Heinrich-Heine-Universität Düsseldorf

Interplay of Quantum Resources in Bell-type Scenarios

Inaugural dissertation

presented to the Faculty of Mathematics and Natural Sciences of the
Heinrich-Heine-Universität Düsseldorf
for the degree of

Doctor of Natural Sciences (Dr. rer. nat.)

by

Lucas Tendick

from Duisburg

Düsseldorf, January 2023

From the Institute for Theoretical Physics III
at the Heinrich-Heine-Universität Düsseldorf

Published by permission of the
Faculty of Mathematics and Natural Sciences at
Heinrich-Heine-Universität Düsseldorf

Supervisor: Prof. Dr. Dagmar Bruß

Co-supervisor: PD Dr. Hermann Kampermann

Date of the oral examination: 23.03.2023

Declaration

Ich versichere an Eides statt, dass die Dissertation von mir selbstständig und ohne unzulässige fremde Hilfe unter Beachtung der “Grundsätze zur Sicherung guter wissenschaftlicher Praxis an der Heinrich-Heine-Universität Düsseldorf” erstellt worden ist.

Düsseldorf, January 23, 2023

Lucas Tendick

To everyone who supported me on this incredible journey.

Abstract

Quantum nonlocality, i.e., the effect that distant quantum systems can exhibit stronger correlations than allowed by any classical theory, is one of the most remarkable features of nature. Beyond its foundational significance, it was realized in recent years that nonlocality is a central resource for information processing tasks such as the reduction of communication complexity, randomness generation, and cryptography.

These practical applications necessitate the systematical analysis and quantification of which resources facilitate quantum nonlocality in a so-called Bell experiment. That is, we need to thoroughly understand which properties of the used quantum states and the performed local measurements enhance quantum nonlocality and how these state and measurement resources result quantitatively in nonlocal correlations.

This thesis is devoted to deepening the understanding of the interplay of quantum resources in Bell-type scenarios that lead to quantum nonlocality using the general framework of quantum resource theories. To that end, we first study some counter-intuitive effects of this interplay of quantum resources leading to Bell nonlocality. Furthermore, we study systematically how hierarchical structures for state and measurement resources result in bounds on the strength of nonlocal correlations.

Regarding the interplay of nonlocality and entanglement, we prove that there exist bound entangled states that are local in any standard Bell experiment. Nevertheless, their nonlocal properties can be activated by local filters in a sequential Bell scenario. Therefore, our results show that hidden nonlocality does not imply entanglement distillability.

Including more state resources in our analysis, we determine analytically the minimal purity necessary to achieve a certain level of nonlocality for any Bell experiment in which the used measurements are fixed. We also discuss the implications of our results for coherence, discord, and entanglement of quantum states using that the purity bounds these quantities. Furthermore, we show that, in general, there is no trade-off between entanglement and incompatibility of quantum measurements, i.e., increasing one resource does not allow for decreasing the other while keeping the desired amount of nonlocality fixed.

Our study of measurement resources is built upon a general framework of distance-based resource quantification that we propose for any convex resource theory of measurements. Using a particular distance-based monotone, relying on the diamond norm, we derive a hierarchy of measurement resources that includes quantum steering and Bell nonlocality. We study instances in which different resources of the hierarchy attain the same value and derive upper and lower bounds on the incompatibility of any set of measurements.

Focusing specifically on the incompatibility of measurements, we show how the incompatibility of a set of measurements is limited through the incompatibility of its subsets. That allows us to bound the maximal incompatibility that can be gained from adding more measurements to an existing measurement scheme. Finally, we discuss the implications of our bounds for the nonlocal correlations of Bell tests with more than two measurements.

Zusammenfassung

Quanten-Nichtlokalität, d.h. der Effekt, dass weit voneinander entfernte Quantensysteme stärkere Korrelationen aufweisen können als jede klassische Theorie es zulässt, ist eine der bemerkenswertesten Eigenschaften der Natur. Über ihre fundamentale Bedeutung hinaus, wurde in den letzten Jahren erkannt, dass Nichtlokalität eine zentrale Ressource für Informationsverarbeitungsaufgaben wie der Reduzierung von Kommunikationskomplexität, Zufallsgenerierung und Kryptographie ist.

Diese praktischen Anwendungen erfordern eine systematische Analyse und Quantifizierung der Ressourcen die Quanten-Nichtlokalität in einem sogenannten Bell-Experiment ermöglichen. Das heißt, es ist notwednig zu verstehen welche Eigenschaften der verwendeten Quantenzustände und der durchgeführten lokalen Messungen die Quanten-Nichtlokalität ermöglichen und wie diese Zustands- und Messressourcen quantitativ zu nichtlokalen Korrelationen führen.

Diese Arbeit widmet sich der Vertiefung des Verständnisses des Zusammenspiels von Quantenressourcen in Bell-Szenarien, die zu Quanten-Nichtlokalität führen, unter Verwendung des allgemeinen Rahmens der Quantenressourcen-Theorien. Zu diesem Zweck untersuchen wir zunächst einige kontraintuitive Effekte des Zusammenspiels von Quantenressourcen, die zu Bell-Nichtlokalität führen. Außerdem untersuchen wir systematisch, wie hierarchische Strukturen für Zustands- und Messressourcen zu einer Begrenzung der Stärke von nichtlokalen Korrelationen führen.

Was das Zusammenspiel von Nichtlokalität und Verschränkung betrifft, beweisen wir, dass es gebunden verschränkte Zustände gibt, die in jedem Standard-Bell-Experiment lokal sind. Dennoch können ihre nichtlokalen Eigenschaften durch lokale Filter in einem sequentiellen Bell-Szenario aktiviert werden. Daher zeigen unsere Ergebnisse dass versteckte Nichtlokalität keine Destillierbarkeit der Verschränkung impliziert.

Indem wir weitere Zustandsressourcen in unsere Analyse einbeziehen, bestimmen wir analytisch die minimale Reinheit, die notwendig ist, um ein bestimmtes Maß an Nichtlokalität für jedes Bell-Experiment zu erreichen, in dem die verwendeten Messungen festgelegt sind. Wir erörtern auch die Auswirkungen unserer Ergebnisse auf Kohärenz, Dissonanz und Verschränkung von Quantenzuständen, unter Verwendung der Tatsache, dass die Reinheit diese Größen beschränkt. Des Weiteren, zeigen wir, dass es im Allgemeinen kein Gleichgewicht zwischen Verschränkung und der Inkompatibilität von Quantenmessungen gibt, d.h. eine Erhöhung der einen

Ressource erlaubt nicht die Verringerung der anderen, während das gewünschte Maß an Nichtlokalität beibehalten wird.

Unsere Untersuchung der Messressourcen basiert auf einem allgemeinen Konzept der distanzbasierten Ressourcenquantifizierung für jede konvexe Ressourcentheorie von Messungen, welches wir vorschlagen. Unter Verwendung eines bestimmten distanzbasierten Monoton, das auf der Diamant-Norm beruht, leiten wir eine Hierarchie von Messressourcen ab, die auch Quantenfernsteuerung und Bell-Nichtlokalität umfasst. Wir untersuchen Instanzen, in denen verschiedene Ressourcen der Hierarchie den gleichen Wert erreichen, und leiten obere und untere Grenzen für die Inkompatibilität einer beliebigen Menge von Messungen her.

Wir konzentrieren uns dann speziell auf die Inkompatibilität von Messungen und zeigen, wie die Inkompatibilität einer Menge von Messungen durch die Inkompatibilität ihrer Teilmengen begrenzt wird. Dies erlaubt uns die maximale Inkompatibilität zu begrenzen, die durch das Hinzufügen weiterer Messungen zu einem bestehenden Messschema gewonnen werden kann. Schließlich diskutieren wir die Auswirkungen unserer Schranken für die nichtlokalen Korrelationen in Bell-Tests mit mehr als zwei Messungen.

Acknowledgements

Before we dive into the fascinating topic of quantum correlations and quantum resources, let us deal with what is possibly the most challenging part of this thesis for me. That is, expressing my sincere gratitude to the many people that, in one way or another, supported me throughout the years. It's the most challenging part for me because it is impossible that these few words can express the gratefulness I feel toward these people. However, to avoid sinking too deep into sentimentality, let's try anyway.

First of all, I would like to thank my supervisor Dagmar Bruß for the opportunity to be part of her group, to learn from her, for the freedom to work on the topics I like, for all the advice, and for all the scientific and non-scientific conversations throughout the years.

No less, I would like to thank Hermann Kampermann for supporting me on my journey to become a scientist and accompanying me as a co-supervisor since I started as a Bachelor's student in this group.

I also wish to thank all my colleagues and friends in the group for the fantastic time in the past years, for all the discussions (we even managed to squeeze in some scientific ones), for all the support, the good time at conferences, and all the dinners, drinks, free-time activities, and laughs we had together. So thank you to Thomas Wagner, Julia Kunzelmann, Sarnava Datta, Federico Grasselli, Gláucia Murta, Nikolai Wyderka, Giacomo Carrara, Justus Neumann, Carlo Liorni, Giulio Gianfelici, Timo Holz, Felix Bischof, Daniel Miller, the "Kliesch guys" Lennart Bittel, Raphael Brieger, Christopher Cedzich, Markus Heinrich, Nikolai Miklin, Mirko Arienzo, Pascal Baßler, Alexander Gresch, Matthias Zipper, Juan Henning, and Salwa Shaglel. Of course, also many thanks to Martin Kliesch for having the opportunity to learn from him and to work with him.

Also, many thanks to all the other members of the Institute, in particular to Jens Bremer, Cornelia Glowacki, and Cordula Hoffjan.

When it comes to this thesis specifically, I would like to thank Julia Kunzelmann, Giacomo Carrara, Nikolai Wyderka, Federico Grasselli, and Gláucia Murta for proofreading parts of the work.

Let me also thank all my other friends for their support and the incredible time I could spend with them so far. I hope this is only the beginning, with many more years and memories to come.

Last but not least, thousand thanks to my family, especially my Mum, for helping me to become the person I am today and enabling me to achieve this goal in the first place.

Contents

1 Introduction

> *Fall in love with some activity, and do it! Nobody ever figures out what life is all about, and it doesn't matter. Explore the world. Nearly everything is really interesting if you go into it deeply enough. Work as hard and as much as you want to on the things you like to do the best. Don't think about what you want to be, but what you want to do. Keep up some kind of a minimum with other things so that society doesn't stop you from doing anything at all.*
>
> — **Richard P. Feynman**

In 2022, the year of writing this thesis, the Nobel Prize in Physics was awarded to Alain Aspect, John F. Clauser, and Anton Zeilinger "for experiments with entangled photons, establishing the violation of Bell inequalities and pioneering quantum information science" [1]. In their reasoning, the Royal Swedish Academy of Sciences argues that the accomplishments of these three scientists "have cleared the way for new technology based upon quantum information."

Indeed, today, the field of quantum information represents a promising field of science that paves the way for potential technological advances in fields like computation [2–5], sensing [6], and secure communication [7–9]. On the other hand, quantum information is a field that aims at answering central questions regarding the foundations of quantum theory, such as the role of measurements and the correlations predicted in experiments that date back to the early works of Einstein [10, 11], Schrödinger [12, 13], and Heisenberg[14].

Arguably, one of the most impactful discoveries for the field of quantum physics, if not all of science, was made by Bell in the '60s [15]. He discovered that the puzzling features of quantum mechanics, and entangled states in particular, can be tested in an experiment that would rule out *local hidden-variables* as an explanation for the strong correlations between distant quantum systems predicted by quantum theory. His proposed test of quantum mechanics, nowadays known as *Bell test* or *Bell experiment*, was first experimentally realized by Clauser [16] and later in a refined experiment by Aspect [17]. Many even more refined Bell tests followed these pioneering experiments in recent years, see e.g., [18–21]. Their results show that quantum theory is indeed *Bell nonlocal*, i.e., that nature cannot be explained by a (more specifically, Bell's) model of *local realism*. Moreover, these works show the importance and the power of the correlations of entangled quantum states. Anton Zeilinger then used these entangled states to demonstrate the feasibility of the important *quantum teleportation* protocol [22], which is the transfer of a quantum state from one system to another over a distance.

Besides their fundamental importance, these results are also milestones for understanding the technological advances that quantum theory promises. Other notable milestones for these technological advances are the works by Bennett and Brassard [23, 24] and Ekert [25]. They showed that quantum theory, particularly entangled states and nonlocal correlations, can be used for secure cryptography. Their works sparked a huge field devoted to developing *device-independent quantum cryptography* protocols [26], which also advanced the field of randomness generation using quantum systems based on Bell's theorem [27–29].

Not only the field of quantum communication promises applications of the foundations of quantum theory, but also the fields of quantum sensing and computation make steady process. Here, the most remarkable result is probably Shor's work [30], which shows that a quantum computer can, in principle, perform prime factoring in polynomial time, which seems unlikely to be possible with a conventional computer.

All these emerging applications of so-called *quantum information processing tasks* have a common theme. They rely on particular quantum states with very distinct quantum properties. The same is true for the measurement schemes that are used to obtain classical information about the quantum system. That is, these quantum information processing tasks rely on *quantum resources* [31], such as entanglement [32]. The field of quantum resource theories is devoted to studying the quantum phenomena that boost quantum technologies. It studies which quantum states and measurements provide an advantage for specific tasks and how to use them optimally. Moreover, it provides methods to quantify this advantage and tools to study the conversion of quantum resources into each other.

Over the years, many promising candidates for quantum advantages, besides entanglement, have been proposed, and these resources have been systematically analyzed using the framework of resource theories (see, e.g., [33–42] for a non-exhaustive list). However, as quantum states and measurements typically possess many of these different quantum resources simultaneously, it becomes increasingly important to understand the interplay of these resources. Understanding this interplay in the context of Bell nonlocality sets the primary motivation for this thesis.

1.1 Motivation and Results

This thesis and the author's doctoral research that preceded it are devoted to deepening the understanding of the role of quantum resources [31] in quantum information processing tasks. Quantum correlation phenomena such as Bell nonlocality [15, 43] and quantum steering [44, 45] are central to this thesis. Our goal is to understand quantitatively which quantum resources of quantum states and measurements are necessary to achieve a certain level of quantum nonlocality. As the respective quantum phenomena themselves are understood already rather well, we focus particularly on the interplay of different state and measurement resources and how they result in nonlocal quantum correlations, which then can be further used as a resource.

That this interplay of resources sets certain challenges can be understood by examining some of the milestones in studying the influence of quantum resources on each other. For instance, Gisin showed that every pure bipartite quantum state that is entangled could also reveal nonlocality if appropriate measurements are chosen for the Bell test [46]. On the hand, Werner and Barrett showed that entanglement and nonlocality are different notions for mixed quantum states [47, 48]. However, a more general Bell test involving the nonlocality activation of states that appear local in the standard Bell test questions, whether nonlocality and entanglement are entirely different phenomena [49–52]. Moreover, there exist quantitative examples for the so-called *anomaly of nonlocality* [53], i.e., the effect that more entanglement does not necessarily imply more nonlocality.

From the viewpoint of measurement resources, it is known that non-jointly measurable measurements are necessary to reveal Bell nonlocal correlations but not sufficient beyond the simplest scenario [54–56]. On the other hand, a one-to-one correspondence exists between the incompatibility of measurements and quantum steering [57–59]. Understanding the interplay of quantum resources becomes even more challenging when resources like purity [60, 61], coherence [35], and discord [36] on the state side and informativeness [40] and coherence [41, 62] on the measurement side are also considered.

To this aim, the author's doctoral research contributes to a better understanding of the interplay of a variety of quantum resources in two different ways. On the one hand, we analyze the hierarchical structure of quantum resources and develop frameworks for identifying the minimal necessary resources for a desired amount of nonlocality. On the other hand, we also study specific instances of the quantum resource interplay to reveal some peculiar effects of quantum nonlocality.

In the first publication [63] (Appendix A), we studied the interplay between entanglement and nonlocality in a hidden nonlocality scenario. In particular, we first showed that there exists a multipartite bound entangled state with a local model for all possible quantum measurements. In a second step, we showed that the nonlocal properties of this seemingly local state could be activated in a sequential Bell scenario using local filters. Hence, we show that bound entangled states can possess genuine hidden nonlocality. Together with the result in [52] our findings imply that hidden nonlocality and distillability of quantum states are entirely different concepts.

In our second publication [64] (Appendix B), we quantified what minimal state resources in terms of purity, coherence, discord, and entanglement are necessary to achieve a certain violation of a given Bell inequality once the measurements are fixed. We show that the minimal purity necessary to achieve said violation can always be determined analytically by explicitly constructing the state that minimizes the purity for a given Bell operator and a fixed violation. Notably, this result is general, i.e., it applies to any Bell inequality, with any number of parties, measurements, and outcomes, in any finite dimension. Using our insights from the resource of purity, we show in the case of two qubits and any full-correlation Bell inequality that there exists a quantum state that simultaneously minimizes the purity, coherence, discord, and entanglement that is necessary to achieve a fixed violation. We show that our results have a counterintuitive consequence for the Clauser-Horne-Shimony-Holt (CHSH) inequality in particular. Namely, we show that for a fixed violation of the CHSH inequality, there are instances where increasing the measurement resources in terms of their incompatibility requires increasing the state resources in terms of the entanglement. That is, there is generally no trade-off between measurement and state resources for a target violation of the CHSH inequality.

In our third manuscript [65] (Appendix C), we developed a distance-based framework for the quantification of measurement resources for any convex resource theory. More specifically, we define distances between sets of measurements and show that they naturally induce resource quantifiers. Based on a specific quantifier, the diamond distance, we derive a hierarchy of different measurement resources, including quantum steering and Bell nonlocality. Furthermore, we analytically derive bounds on several measurement resources, with a special focus on the incompatibility of measurements that are valid for any set of measurements.

In our fourth and final work [66] (Appendix D), we focused specifically on measurement incompatibility. First, we analyze how the incompatibility of subsets of measurements constrains the incompatibility of the whole. That allows us to find bounds on the incompatibility gained from increasing the number of measurements in a particular setup. Moreover, we decompose the total incompatibility in terms of the measurements' incompatibility with respect to substructures like pairwise and genuine triplewise incompatibility. Finally, we present tight examples for most of our bounds and discuss the consequences of our work for the limits that can be set on the correlations in steering and Bell experiments with more than two measurements.

1.2 Thesis Structure

This thesis is structured as follows:

- In Chapter 2, we introduce the theoretical background of this thesis. In particular, we revisit the basic construction of the framework of quantum mechanics. Afterward, we introduce the basic concepts from quantum information theory necessary to follow this thesis. In particular, we introduce geometric and entropic measures for quantum information that will later be useful to quantify quantum resources.

- We introduce the framework of quantum resource theories in Chapter 3. First, we study the general notions of resource theories, such that particular resource theories emerge as a specific case of our framework. Afterward, we review each of the state, measurement, and quantum correlation resources relevant to this thesis. Section 3.4.1 contains a detailed introduction to the Bell scenario and the phenomenon of Bell nonlocality.

- Chapter 4 contains an overview and a discussion of the results obtained during the author's doctoral research.

- In Chapter 5, we conclude this thesis and look out for directions of future research.

- The original works can be found in the appendices A to D.

2 Theoretical Background

> *Do not worry about your difficulties in mathematics, I assure you that mine are greater.*
>
> — **Albert Einstein**

Since its inception in the first half of the 20th century, quantum theory has proven its success time and time again. Indeed, quantum theory is one of our most accurate descriptions of nature so far, as demonstrated by many experiments (see, e.g., [67, 68]). The foundations of quantum mechanics are a few postulates from which the theory can be derived. These postulates were heavily debated in the early days and are still subject to active research (see, e.g., [69]). However, up to this day, these postulates are the minimal set of assumptions necessary to formulate the framework of quantum mechanics. The mathematical language of this framework is linear algebra. Throughout this chapter, we will revisit the postulates of quantum mechanics and use the opportunity to introduce the essential concepts from linear algebra necessary to follow this thesis as we go along. To round up the theoretical background of this work, we will also introduce the concepts of quantum information theory necessary to understand the following chapters.

If not stated otherwise, the mathematical introduction in the following sections follows the same lines as [70–72] and is also inspired by [73–82].

2.1 Hilbert Spaces and State Vectors

The first postulate introduces *state vectors* and the *Hilbert spaces* they live in. In this thesis, we exclusively consider physical systems with a finite number d of degrees of freedom. Thus, we will always deal with finite-dimensional Hilbert spaces. Therefore, the Hilbert spaces we consider always reduce to *inner product spaces* $\mathbb{C}^d$ over the complex field $\mathbb{C}$.

Postulate 1. *To any isolated physical system, there is an associated* state space*, described by a Hilbert Space $\mathcal{H}$ (complex vector space with an inner product) of the system. Any physical state of that system is completely described by a* state vector *, i.e., a unit vector $|\psi\rangle \in \mathcal{H}$ that is unique up to a complex phase factor.*

The first postulate already requires some commenting, as it introduces many of quantum mechanics' structures and includes two implicit assumptions. First, Postulate 1 specifically deals with *isolated physical systems*, which means we assume for now that the system does not interact with other systems and, in particular, any environment. Second, Postulate 1 implicitly states that we are using the Dirac notation for abstract vectors throughout this work, in which the vectors $|\psi\rangle$ are known as *ket*. In principle, any other label can replace the ψ. For instance, $|i\rangle$ with $i = 0, \cdots, d-1$, will often be used to denote the i-th degree of freedom of a d-dimensional system. Note that $|0\rangle$ is an entirely different object than the zero-vector, which we denote by $\mathbf{0}$.

To properly introduce the inner product, we first introduce the dual space $\mathcal{H}^*$ and its elements, the dual vectors $\langle\psi|$.

Definition 2.1.1. *(*Dual space and dual vectors*). Let $\mathcal{H}$ be a Hilbert space over the complex field $\mathbb{C}$. Its corresponding dual space $\mathcal{H}^*$ is the vector space of all linear maps $\langle\psi| : \mathcal{H} \mapsto \mathbb{C}$. The elements $\langle\psi| \in \mathcal{H}^*$ are called* bra*.*

Note that the dual space $\mathcal{H}^*$ is isomorphic to $\mathcal{H}$ due to *Riesz' representation theorem* (see, e.g., [72]). Hence, there is a one-to-one mapping between kets $|\psi\rangle \in \mathcal{H}$ and bras $\langle\psi| \in \mathcal{H}^*$. This mapping is given by the *Hermitian adjoint* or *Hermitian conjugate* $(\cdot)^\dagger$.

Definition 2.1.2. *(*Hermitian adjoint*). Let $|\psi\rangle \in \mathcal{H}$ and $\langle\psi| \in \mathcal{H}^*$. The one-to-one correspondence between $|\psi\rangle$ and $\langle\psi|$ is given by the* Hermitian adjoint $(|\psi\rangle)^\dagger := \langle\psi|$ *that acts such that*

$$(\alpha|\psi_1\rangle + \beta|\psi_2\rangle)^\dagger = \overline{\alpha}\langle\psi_1| + \overline{\beta}\langle\psi_2|, \tag{2.1a}$$

$$(\overline{\alpha}\langle\psi_1| + \overline{\beta}\langle\psi_2|)^\dagger = \alpha|\psi_1\rangle + \beta|\psi_2\rangle, \tag{2.1b}$$

for any complex numbers α, β with complex conjugates $\overline{\alpha}, \overline{\beta}$.

Definition 2.1.3. *(*Inner product*). The inner product of a Hilbert space $\mathcal{H}$ is a mapping $\langle\cdot|\cdot\rangle : \mathcal{H} \times \mathcal{H} \mapsto \mathbb{C}$, which associates to any pair of vectors $|\psi\rangle, |\phi\rangle \in \mathcal{H}$ a complex number and that satisfies the following conditions:*

$$1.\ \langle\phi|\psi\rangle = \overline{\langle\psi|\phi\rangle}, \tag{2.2a}$$

$$2.\ \langle\psi|\psi\rangle \geq 0,\ \textit{and}\ \langle\psi|\psi\rangle = 0 \iff |\psi\rangle = \mathbf{0}, \tag{2.2b}$$

$$3.\ \langle\phi|\alpha\psi_1 + \beta\psi_2\rangle = \alpha\langle\phi|\psi_1\rangle + \beta\langle\phi|\psi_2\rangle,\ \textit{for any}\ \alpha, \beta \in \mathbb{C}. \tag{2.2c}$$

Note that it follows from the conditions (2.2a) and (2.2c) that $\langle\alpha\phi_1 + \beta\phi_2|\psi\rangle = \overline{\alpha}\langle\phi_1|\psi\rangle + \overline{\beta}\langle\phi_2|\psi\rangle$, which means the inner product is a *sesquilinear form*.

We already used in the Definitions 2.1.2 and 2.1.3 the very basic fact that linear combinations of vectors in $\mathcal{H}$ are also contained in $\mathcal{H}$. The very innocent-looking postulation that vectors in a Hilbert space describe quantum mechanics already allows for some of its key features. For instance, the *superposition principle*, which is the foundation for the phenomenon of *coherence* that we explicitly discuss in Section 3.2.2, follows from this simple postulate.

A vector $|\psi\rangle$ is called a *state vector* if it is normalized, i.e., $\||\psi\rangle\| = \sqrt{\langle\psi|\psi\rangle} = 1$. Here, $\|\cdot\|$ is the norm induced by the inner product. Furthermore, the family of states $|\psi^\theta\rangle = \exp(i\theta)|\psi\rangle$ with $\theta \in \mathbb{R}$ are regarded as physically equivalent, as they only differ by a global complex phase which is not observable in any experiment. Besides the notion of a norm, the inner product also gives a sense of orientation between two vectors. Like for two vectors and the Euclidean dot product between them, the vectors $|\psi\rangle$ and $|\phi\rangle$ are said to be orthogonal if and only if $\langle\phi|\psi\rangle = 0$.

Throughout this thesis, we will need some additional concepts from linear algebra to study vector spaces and state vectors in more detail. However, we assume that any potential reader will be familiar with them, so we introduce the concepts in the context of Hilbert spaces without further commenting.

Definition 2.1.4. *(*Spanning set*). A set of vectors $\{|v_i\rangle\}_{i=0}^{n-1}$ is a* spanning set *of a Hilbert space $\mathcal{H}$ if any vector $|v\rangle \in \mathcal{H}$ can be written as a linear combination*

$$|v\rangle = \sum_{i=0}^{n-1} c_i |v_i\rangle, \tag{2.3}$$

with some complex coefficients c_i.

Definition 2.1.5. *(*Linear independence*). A set of non-zero vectors $\{|v_i\rangle\}_{i=0}^{n-1}$ is said to be* linearly independent *if the equation*

$$\sum_{i=0}^{n-1} c_i |v_i\rangle = \mathbf{0}, \tag{2.4}$$

only holds true if $c_i = 0$, for all $i = 0, \cdots, n-1$, and it is called linearly dependent *otherwise.*

Definition 2.1.6. *(*Basis*). Any set $\{|v_i\rangle\}_{i=0}^{d-1}$ of linear independent vectors that spans $\mathcal{H}$ is called a* basis *of $\mathcal{H}$. Any basis of $\mathcal{H}$ contains the same number of elements, which is called the* dimension *d of the Hilbert space. A basis is said to be* orthonormal *if it holds that $\langle v_i|v_j\rangle = \delta_{ij}$ for any $i, j \in \{0, 1, \cdots, d-1\}$, where δ_{ij} is the Kronecker delta, i.e., the vectors are pairwise* orthogonal *and normalized.*

Definition 2.1.7. *(*Mutually unbiased bases*). Two orthonormal bases $\{|v_i\rangle\}_{i=0}^{d-1}$, $\{|w_i\rangle\}_{i=0}^{d-1}$ on a Hilbert space $\mathcal{H}$ are said to be mutually unbiased bases (MUB) if it holds that $|\langle v_i|w_j\rangle| = \frac{1}{\sqrt{d}}$ for all i, j.*

It is often convenient to work with the representation of state vectors in a particular basis. That leads to the concept of coordinate vectors.

Definition 2.1.8. *(*Coordinate vector*). Given an orthonormal basis $\{|v_i\rangle\}_{i=0}^{d-1}$ of $\mathcal{H}$. Than any vector $|v\rangle \in \mathcal{H}$ admits a unique decomposition*

$$|v\rangle = \sum_{i=0}^{d-1} v_i |v_i\rangle, \tag{2.5}$$

i.e., any $|v\rangle$ can be represented by a (column) vector, the so-called coordinate vector $\mathbf{v} := \begin{pmatrix} v_0 \\ v_1 \\ \vdots \\ v_{d-1} \end{pmatrix} \hat{=} |v\rangle$.

We will obey the convention that slightly abuses the notation in the following and also use the symbol $|v\rangle$ for its representation $\mathbf{v}$. Having specified the representation of a particular ket $|v\rangle$ as column vector, we can identify the corresponding bra as row vector, i.e., $\langle v| = (|v\rangle)^\dagger = \begin{pmatrix} \overline{v_0} & \overline{v_1} & \cdots & \overline{v_{d-1}} \end{pmatrix}$, which means that the Hermitian conjugate $(\cdot)^\dagger$, is a conjugate transpose in a particular basis. With that, we can write the inner product of two vectors $|v\rangle$, $|w\rangle$ as

$$\langle v|w\rangle = \sum_{i=0}^{d-1} \overline{v_i} w_i, \tag{2.6}$$

that is, as the matrix product of the row vector $\langle v|$ and the column vector $|w\rangle$.

2.2 Linear Operators

Besides state vectors and Hilbert spaces, additional ingredients are necessary to reconstruct quantum mechanics. Namely, we also need linear operators to act on these states to account for state transformations and measurements in quantum mechanics.

Definition 2.2.1. *(*Linear operator*). Let $\mathcal{H}_1$ and $\mathcal{H}_2$ be two Hilbert spaces. A linear operator A between $\mathcal{H}_1$ and $\mathcal{H}_2$ is a function $A : \mathcal{H}_1 \mapsto \mathcal{H}_2$ that maps any vector $|\psi\rangle \in \mathcal{H}_1$ to a vector $A|\psi\rangle := A(|\psi\rangle) \in \mathcal{H}_2$ such that*

$$A\Big(\sum_i c_i|\psi_i\rangle\Big) = \sum_i c_i A|\psi_i\rangle, \tag{2.7}$$

for any $|\psi_i\rangle \in \mathcal{H}_1$ and any coefficients $c_i \in \mathbb{C}$.

In the following, we will often consider linear operators that map elements of a Hilbert space onto other elements of the *same* Hilbert space, i.e., mappings $A : \mathcal{H} \mapsto \mathcal{H}$, in which case we say A is defined (or acts) on $\mathcal{H}$. We denote by $\mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$ the set of all linear maps from $\mathcal{H}_1$ to $\mathcal{H}_2$ with $\mathcal{L}(\mathcal{H}) := \mathcal{L}(\mathcal{H}, \mathcal{H})$, which itself is a vector space. We will often focus on the set $\mathcal{L}(\mathcal{H})$ for simplicity. One particularly important operator is the identity operator $\mathbb{1}_{\mathcal{H}}$, which acts trivially on any vector in the sense that $\mathbb{1}_{\mathcal{H}}|\psi\rangle = |\psi\rangle$ for any $|\psi\rangle \in \mathcal{H}$.

We already saw that kets $|\psi\rangle \in \mathcal{H}$ can be represented by column vectors and bras $\langle\psi| = (|\psi\rangle)^\dagger \in \mathcal{H}^*$ as row vectors. We will see that linear operators can be represented as matrices in the following. We use the *dyadic* or *outer* product to do so. Consider vectors $|v\rangle, |w\rangle, |z\rangle \in \mathcal{H}$ and the linear operator $|v\rangle\langle w|$ that defines the outer product of $|v\rangle$ and $\langle w| = (|w\rangle)^\dagger$, which acts on $|z\rangle$ such that

$$(|v\rangle\langle w|)|z\rangle := \langle w|z\rangle|v\rangle. \tag{2.8}$$

We can also consider linear combinations of outer products. In particular we consider an orthonormal basis $\{|v_i\rangle\}_{i=0}^{d-1}$ and a vector $|v\rangle = \sum_{i=0}^{d-1} v_i|v_i\rangle$. It follows directly that $\langle v_i|v\rangle = v_i$. This allows us to see that

$$\Big(\sum_{i=0}^{d-1}|v_i\rangle\langle v_i|\Big)|v\rangle = \sum_{i=0}^{d-1}|v_i\rangle\langle v_i|v\rangle = \sum_{i=0}^{d-1} v_i|v_i\rangle = |v\rangle, \tag{2.9}$$

from which we obtain the *completeness relation* $\sum_{i=0}^{d-1}|v_i\rangle\langle v_i| = \mathbb{1}_{\mathcal{H}}$.

One can use the completeness relation to establish the matrix representation of a linear operator $A \in \mathcal{L}(\mathcal{H})$ in an orthonormal basis $\{|v_i\rangle\}_{i=0}^{d-1}$. Namely,

$$\mathbb{1}_{\mathcal{H}}A\mathbb{1}_{\mathcal{H}} = \Big(\sum_{i=0}^{d-1}|v_i\rangle\langle v_i|\Big)A\Big(\sum_{j=0}^{d-1}|v_j\rangle\langle v_j|\Big) = \sum_{i,j=0}^{d-1}\langle v_i|A|v_j\rangle|v_i\rangle\langle v_j| = \sum_{i,j=0}^{d-1}A_{ij}|v_i\rangle\langle v_j|, \tag{2.10}$$

where $A_{ij} = \langle v_i|A|v_j\rangle$ is the matrix element (i-th row and j-th column) of A in the basis $\{|v_i\rangle\}_{i=0}^{d-1}$. Therefore, we denote by (A_{ij}) the matrix representation of A with respect to the basis $\{|v_i\rangle\}_{i=0}^{d-1}$. Similar to the case for state vectors, we synonymously use the symbol A to talk about a particular representation.

There are many important classes of linear operators that we want to define in what follows. However, before we do so, we define some frequently used concepts that come up when we deal with these operators.

Definition 2.2.2. *(*Eigenvalue *and* eigenvector*). Let $A \in \mathcal{L}(\mathcal{H})$ be a linear operator and $|v\rangle \in \mathcal{H}$ a non-zero vector such that*

$$A|v\rangle = \lambda|v\rangle, \tag{2.11}$$

for some scalar $\lambda \in \mathbb{C}$. We say that $|v\rangle$ is an eigenvector *of A and λ its corresponding* eigenvalue*. We say that a normalized* eigenvector *is an* eigenstate *and we call the set of all eigenvalues of A the* spectrum $\lambda(A)$.

Definition 2.2.3. *(*Hermitian adjoint of operators*). Let $A \in \mathcal{L}(\mathcal{H})$ be a linear operator on $\mathcal{H}$. We define the* Hermitian adjoint *operator $A^\dagger \in \mathcal{L}(H^*)$ of A to be the unique linear operator that fulfills*

$$\langle\psi|A^\dagger|\phi\rangle = \overline{\langle\phi|A|\psi\rangle}, \tag{2.12}$$

for any $|\psi\rangle, |\phi\rangle \in \mathcal{H}$. This means in particular that $(A^\dagger)_{ij} = (\overline{A})_{ji}$ for the matrix representation of A in any orthonormal basis. That is, the Hermitian adjoint corresponds to a complex conjugation and a transpose, once a particular basis is fixed. Note that this is consistent with the definition of $(\cdot)^\dagger$ acting on vectors.

Definition 2.2.4. *(*Commutator*). Let $A, B \in \mathcal{L}(\mathcal{H})$ be two linear operators. We say A and B* commute *if the* commutator *defined as*

$$[A, B] := AB - BA, \tag{2.13}$$

vanishes, i.e., if $[A, B] = 0$.

Definition 2.2.5. *(*Normal operators*). An operator $A \in \mathcal{L}(\mathcal{H})$ is said to be* normal *if it commutes with its Hermitian adjoint, that is*

$$A^\dagger A = AA^\dagger. \tag{2.14}$$

The class of normal operators is essential in the context of quantum mechanics because it contains two important cases as subclasses. These classes are the Hermitian operators, which are crucial in the context of observable quantities in quantum mechanics, and the unitary operators, which describe the evolution of quantum systems.

Definition 2.2.6. *(*Hermitian operators*). A linear operator $A \in \mathcal{L}(\mathcal{H})$ is said to be* Hermitian*, if it holds that*

$$A = A^\dagger. \tag{2.15}$$

We denote the set of all Hermitian operators by $\mathrm{Herm}(\mathcal{H}) \subset \mathcal{L}(\mathcal{H})$. Note that it follows directly from Definition 2.2.3 of the Hermitian adjoint, that any operator $A \in \mathrm{Herm}(\mathcal{H})$ can only have real-valued eigenvalues, i.e., $\lambda(A) \in \mathbb{R}$.

There is an additional subset of Hermitian operators that will come up frequently. These are the so-called positive semi-definite operators.

Definition 2.2.7. *(*Positive semi-definite operators*). A Hermitian operator $A \in \mathrm{Herm}(\mathcal{H})$ is said to be* positive semi-definite*, if*

$$\langle v|A|v\rangle \geq 0, \tag{2.16}$$

for any non-zero vector $|v\rangle \in \mathcal{H}$. We will denote this in the following by $A \succeq 0$ ($A \succ 0$ if $\langle v|A|v\rangle > 0$) and the set of all positive semi-definite operators by $\mathrm{Pos}(\mathcal{H}) \subset \mathrm{Herm}(\mathcal{H})$.

Note, it follows directly that any positive semi-definite operator has only non-negative eigenvalues. In fact, this property is an equivalent definition of a positive semi-definite operator.

Definition 2.2.8. *(*Unitary operators*). A linear operator $U \in \mathcal{L}(\mathcal{H})$ is said to be* unitary *if it holds that*

$$UU^\dagger = U^\dagger U = \mathbb{1}_{\mathcal{H}}, \tag{2.17}$$

i.e., $U^\dagger = U^{-1}$ is the inverse of U.

Unitary operators are themselves a subset of the set of linear isometries.

Definition 2.2.9. *(*Isometry*). A linear operator $V \in \mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$ with $d_{\mathcal{H}_2} \geq d_{\mathcal{H}_1}$ is said to be an isometry if*

$$V^\dagger V = \mathbb{1}_{\mathcal{H}_1}. \tag{2.18}$$

An important concept, very similar to eigenvalues, are the so-called singular values and the singular value decomposition, where unitary operators (and a diagonal matrix) are used to decompose an operator A conveniently.

Theorem 2.2.10. *(*Singular values *and the* singular value decomposition*). Let $A \in \mathcal{L}(\mathcal{H})$ be a linear operator. There exist unitary operators $U, V \in \mathcal{L}(\mathcal{H})$ and a diagonal matrix Σ with non-negative entries σ_i, such that A admits a so-called* singular value decomposition *given by*

$$A = U\Sigma V. \tag{2.19}$$

The entries σ_i of Σ are called the singular values *of A. The proof can be found in [70] (Corollary 2.4 therein).*

Definition 2.2.11. *(*Projector*). A linear operator $\Pi \in \mathcal{L}(\mathcal{H})$ is called a* projector *if*

$$\Pi^2 = \Pi. \tag{2.20}$$

Additionally, if $\Pi \in \mathrm{Herm}(\mathcal{H})$, we say Π is an orthogonal projector.

Note that it follows directly that any eigenvalue λ of a projector Π has to be in $\{0, 1\}$. This can be seen from

$$\Pi^2 |v\rangle = \lambda^2 |v\rangle = \lambda |v\rangle = \Pi |v\rangle, \tag{2.21}$$

for any eigenstate $|v\rangle$ of Π.

One example of particular importance is the instance of a *rank-one* orthogonal projector $|\psi\rangle\langle\psi|$. Such an operator maps any vector $|v\rangle \in \mathcal{H}$ onto a 1-dimensional subspace which is spanned by $|\psi\rangle$. Moreover, given an orthonormal basis $\{|b_i\rangle\}_{i=0}^{d-1}$, any orthogonal projector on a subspace S can be written as $\Pi = \sum_{i \in S} |b_i\rangle\langle b_i|$.

Rank-one projectors can, for instance, be used to write any normal operator A in a convenient way. By virtue of the *spectral decomposition* it holds that

$$A = \sum_{i=0}^{d-1} \lambda_i |a_i\rangle\langle a_i|, \tag{2.22}$$

where the $\lambda_i \in \lambda(A)$ are the eigenvalues of A and $|a_i\rangle\langle a_i|$ are the rank-one projectors onto the corresponding eigenstates $|a_i\rangle$.

As a final tool in this section, we introduce the trace of an operator, respectively, of a matrix.

Definition 2.2.12. *(*Trace*). The trace of any operator $A \in \mathcal{L}(\mathcal{H})$ in an orthonormal basis $\{|b_i\rangle\}_{i=0}^{d-1}$ is given by*

$$\mathrm{Tr}[A] := \sum_{i=0}^{d-1} \langle b_i|A|b_i\rangle = \sum_{i=0}^{d-1} A_{ii}, \tag{2.23}$$

i.e., by the sum of its diagonal elements.

Note that the trace is cyclic under products. That means $\mathrm{Tr}[AB] = \mathrm{Tr}[BA]$, which guarantees that the trace $\mathrm{Tr}[A]$ is independent of the particular orthonormal basis. This follows since the change of the representation of A in one orthonormal basis to another is described by a unitary mapping $A \mapsto UAU^\dagger$. That means, in particular, that $\mathrm{Tr}[A] = \sum_{i=0}^{d-1} \lambda_i$ for any normal operator A, where the λ_i are the eigenvalues of A. Finally note that $\mathrm{Tr}[A|\psi\rangle\langle\psi|] = \langle\psi|A|\psi\rangle$, which follows from the completeness relation.

Having the concept of the trace at hand, we can upgrade the vector space of linear operators $\mathcal{L}(\mathcal{H})$ to a Hilbert space by introducing the *Hilbert-Schmidt* inner product $\langle A, B\rangle_{\mathrm{HS}} := \mathrm{Tr}[A^\dagger B]$.

2.3 Density Operators

So far, we have described quantum states via state vectors $|\psi\rangle$. However, there are situations where it is not possible to say in which state of an ensemble $\{|\psi_i\rangle\}_{i=0}^{n-1}$ a quantum state is in. Consider, for instance, a state preparation device that depends on some classical parameter I, that we have no direct access to. Instead, we only observe that the device produces a quantum state $|\psi_i\rangle$ with probability $p_i := p(i) := p(I = i)$. In such a situation, we cannot describe the quantum system by a single state vector $|\psi\rangle$, but instead we have to describe it as a statistical mixture of quantum states

$\{|\psi_i\rangle\}_{i=0}^{n-1}$ that occur with probabilities $\{p_i\}_{i=0}^{n-1}$. This leads to the *density operator* (or *density matrix*) ρ of the system which is per definition given by

$$\rho \coloneqq \sum_i p_i |\psi_i\rangle\langle\psi_i|, \tag{2.24}$$

i.e., by a *convex combination* (i.e., $p_i \geq 0,\ \sum_{i=0}^{n-1} p_i = 1$) of the projectors onto the states $|\psi_i\rangle$. Let us characterize the density operator more formally.

Theorem 2.3.1. *(*Density operator*). An operator $\rho \in \mathcal{L}(\mathcal{H})$ is a density operator of some ensemble of probabilities and states $\{p_i, |\psi_i\rangle\}_{i=0}^{n-1}$ if and only if*

$$\rho = \rho^\dagger,\ \rho \succeq 0,\ \mathrm{Tr}[\rho] = 1, \tag{2.25}$$

which means $\rho \in \mathrm{Pos}(\mathcal{H})$ is a positive semi-definite operator that is normalized to $\mathrm{Tr}[\rho] = 1$*. The proof can be found in [70](Theorem* 2.5 *therein). We denote the set of all density operators acting on $\mathcal{H}$ by $\mathcal{S}(\mathcal{H})$.*

State vectors $|\psi\rangle$ can be described in the density operator formalism as the particular operator $\rho = |\psi\rangle\langle\psi|$, which means there is only one state $|\psi\rangle$ that occurs with probability $p = 1$. From now on, we call state vectors or their corresponding projectors *pure states*. On the other hand, density operators that do not correspond to a single projector are called *mixed states*, reflecting that these states are a statistical mixture of two or more pure states. Clearly, the set $\mathcal{S}(\mathcal{H})$ contains both of these particular cases. Moreover, it can directly be seen from Theorem 2.3.1 that the set of all density operators $\mathcal{S}(\mathcal{H})$ is a convex set, i.e., for any two density matrices $\rho_1, \rho_2 \in \mathcal{S}(\mathcal{H})$ and any $\eta \in [0, 1]$ the operator $\rho = \eta\rho_1 + (1 - \eta)\rho_2$ is again in $\mathcal{S}(\mathcal{H})$. The pure states $\rho = |\psi\rangle\langle\psi|$ describe the extreme points of $\mathcal{S}(\mathcal{H})$, i.e., those states that cannot be written as a convex combination of other states. Note that $\mathcal{S}(\mathcal{H})$ cannot be described by a finite amount of extreme points and that the ensemble a density matrix ρ describes is not unique. Take, for instance, any orthonormal basis $\{|b_i\rangle\}_{i=0}^{d-1}$ of a d-dimensional Hilbert space and let $\{p_i\}_{i=0}^{d-1}$ be distributed uniformly, i.e, $p_i = \frac{1}{d}\ \forall i$. It follows directly that $\rho = \sum_{i=0}^{d-1} p_i |b_i\rangle\langle b_i| = \frac{\mathbb{1}}{d}$. From now on, we say that $\rho = \frac{\mathbb{1}}{d}$ is the *maximally mixed state*, which is a name that we will justify later (see Sections 2.8 and 3.2.1). There is a simple way to see, whether a density operator describes a pure or a mixed state. It holds that $\frac{1}{d} \leq \mathrm{Tr}[\rho^2] \leq 1$ for all $\rho \in \mathcal{S}(\mathcal{H})$ and $\mathrm{Tr}[\rho^2] = 1$ if and only if $\rho = |\psi\rangle\langle\psi|$ is a pure state.

2.3.1 Qubits

So far, we have only described quantum states from an abstract point of view and discussed some generic properties. However, for any particular application, we need a more concrete description of the physical situation. The simplest yet non-trivial system we frequently deal with is the quantum bit or, in short, the qubit. Remember that a classical bit takes on one of the values in $\{0, 1\}$. To describe a qubit, we consider a two-dimensional Hilbert space $\mathbb{C}^2$ and use the so-called *computational basis* described by the states $\{|0\rangle, |1\rangle\}$ with $|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $|1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$.

Multiple physical degrees of freedom can implement a qubit. For instance, the state of a spin-$1/2$ particle, such as an electron, the polarization of a photon, or a two-level transition of an atom. Contrary to a classical bit, a qubit is generally in a superposition state (as a linear combination of two basis vectors of a Hilbert space), i.e., a pure qubit state can be written as

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle, \tag{2.26}$$

where α, β are complex numbers such that $|\alpha|^2 + |\beta|^2 = 1$, due to the normalization of $|\psi\rangle$. Not all of the parameters (four real numbers) within α, β are independent. Due to the normalization and the fact that state vectors are unique only up to a global phase, only two (real) degrees of freedom are necessary to describe a general pure qubit state. Namely,

$$|\psi\rangle = \cos\frac{\vartheta}{2}|0\rangle + e^{-i\varphi}\sin\frac{\vartheta}{2}|1\rangle, \tag{2.27}$$

where $0 \leq \vartheta \leq \pi,\ 0 \leq \varphi < 2\pi$. To describe a general qubit state, including mixed states, we first introduce a convenient basis for for all Hermitian 2×2 matrices. This basis is given in terms of the Pauli matrices

$$\sigma_0 = \mathbb{1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},\ \sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},\ \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix},\ \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{2.28}$$

Note that the Pauli matrices $\{\sigma_i\}$ are unitary, Hermitian, and fulfill $\sigma_i^2 = \mathbb{1}$. Moreover, they are orthogonal with respect to the Hilbert-Schmidt inner product, i.e., $\mathrm{Tr}[\sigma_i\sigma_j] = 2\delta_{ij}$ for all i, j and $\mathrm{Tr}[\sigma_i] = 0$ for $i \in \{1, 2, 3\}$. Let $\boldsymbol{\sigma}$ be the vector that contains the Pauli matrices such that $\boldsymbol{\sigma} = (\sigma_1\ \sigma_2\ \sigma_3)^T$. It is possible to write any qubit state such that

$$\rho = \frac{1}{2}(\mathbb{1} + \mathbf{r} \cdot \boldsymbol{\sigma}), \tag{2.29}$$

where $\mathbf{r} \in \mathbb{R}^3$ is the so-called *Bloch vector* of ρ. Note that the eigenvalues of ρ are given by $\lambda_{1,2} = \frac{1}{2}(1 \pm |\mathbf{r}|)$, which means it is necessary that $|\mathbf{r}| = \sqrt{r_1^2 + r_2^2 + r_3^2} \leq 1$ to describe a quantum state ρ. It follows directly that $|\mathbf{r}| = 1$ corresponds to a pure state.

Describing a qubit by its (unique) Bloch vector allows not only for a compact description of its density operator ρ, but also for a nice geometrical representation.

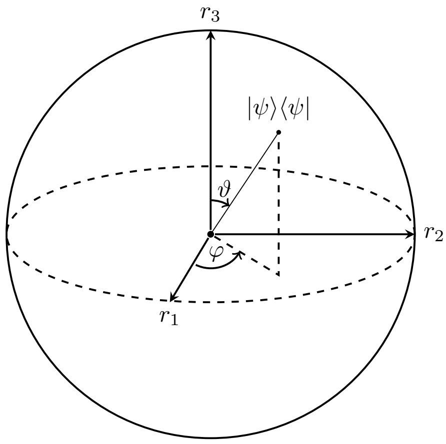

Fig. 2.1.: Representation of qubit states on the Bloch sphere. A pure quantum state $|\psi\rangle\langle\psi|$ corresponds to a Bloch vector $\mathbf{r}$ of length one with components r_1, r_2, r_3. The pure state $|\psi\rangle\langle\psi|$ is completely determined by the two angles φ, ϑ. While the pure states lie on the surface of the Bloch sphere, the mixed states are in its interior.

Namely, all qubit states ρ are contained within a sphere, the *Bloch sphere*, of radius 1. The pure states are on the surface of the Bloch sphere, while the mixed states are in its interior, see also Figure 2.1.

2.4 Composite Systems

Until now, we only considered single quantum systems. However, many of the remarkable features of quantum theory that we discuss in later chapters only become prevalent when considering two or more quantum systems. Here, we recall the mathematical foundations for these phenomena.

Systems composed of multiple systems are mathematically described by the *tensor product* between Hilbert spaces and the *Kronecker product* on the level of matrix representations. We use the symbol $\otimes$ to denote both of these operations.

Postulate 2. *The state space $\mathcal{H}$ of a system composed of n (sub-)systems with associated Hilbert spaces $\mathcal{H}_i$, where $i = 1, \cdots, n$, is described by the tensor product of the single system's Hilbert spaces: $\mathcal{H} := \mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \cdots \otimes \mathcal{H}_n$. Furthermore, if the individual systems are prepared in the states $|\psi_i\rangle$, the joint state of the system is given by $|\psi\rangle = |\psi_1\rangle \otimes |\psi_2\rangle \otimes \cdots \otimes |\psi_n\rangle$, analogously if the individual systems are prepared in the states ρ_i, the joint state of the system is given by $\rho = \rho_1 \otimes \rho_2 \otimes \cdots \otimes \rho_n$.*

Note that we will use the shorthand notations $|\psi_1\rangle \otimes |\psi_2\rangle = |\psi_1\rangle|\psi_2\rangle = |\psi_1\psi_2\rangle$. Furthermore, we call states of the form $|\psi\rangle = |\psi_1\rangle \otimes |\psi_2\rangle \otimes \cdots \otimes |\psi_n\rangle$, respectively $\rho = \rho_1 \otimes \rho_2 \otimes \cdots \otimes \rho_n$, *product states*. We say that a quantum state of n systems, is an n-partite quantum system, shared by n parties in a quantum information processing task. In the following discussions, we will primarily focus on systems composed of two subsystems, i.e., bipartite states.

Consider two Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$ in which we fix a particular orthonormal basis $\{|a_i\rangle\}_{i=0}^{d_A-1}$ and $\{|b_j\rangle\}_{j=0}^{d_B-1}$, respectively. Then $\{|a_i b_j\rangle\}_{ij}$ (we omit the explicit indexing of the set here and in the following whenever it simplifies the notation and does not leave out crucial information) is a basis on the composite Hilbert space $\mathcal{H}_{AB} \coloneqq \mathcal{H}_A \otimes \mathcal{H}_B$. It follows directly that any state $|\psi\rangle \in \mathcal{H}_{AB}$ can be written as

$$|\psi\rangle = \sum_{i,j} c_{ij} |a_i b_j\rangle, \tag{2.30}$$

with some complex coefficients c_{ij}. However, it is often helpful to choose a different decomposition of $|\psi\rangle$ into bases of the subsystems in order to reduce the number of coefficients necessary to describe $|\psi\rangle$. This decomposition is known as *Schmidt decomposition*.

Theorem 2.4.1. *(*Schmidt decomposition*). Let $|\psi\rangle \in \mathcal{H}_{AB}$ be a bipartite state vector of systems with local Hilbert spaces $\mathcal{H}_A, \mathcal{H}_B$ of dimension d_A and d_B such that $\mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B$ and $d_{AB} = d_A d_B$. Then there exist orthonormal bases $\{|i_A\rangle\}_{i=0}^{d_A-1}, \{|i_B\rangle\}_{j=0}^{d_B-1}$ such that*

$$|\psi\rangle = \sum_{i=0}^{d_{\min}-1} \sqrt{c_i} |i_A i_B\rangle, \tag{2.31}$$

where the $\sqrt{c_i}$ are non-negative real numbers such that $\sum_i c_i = 1$, known as Schmidt coefficients*, and $d_{\min} = \min\{d_A, d_B\}$. The number of non-zero Schmidt coefficients is the* Schmidt rank *R of $|\psi\rangle$. A proof can be found in [72] (Proposition* 3.4.1 *therein).*

The Schmidt decomposition, which is entirely based on the *singular value decomposition* (see Theorem 2.2.10), can also be formulated for (density-)operators, which often allows for simplified calculations on mixed states of two systems. The operator Schmidt decomposition [83] guarantees that any state $\rho \in \mathcal{S}(\mathcal{H}_{AB})$ can be decomposed such that

$$\rho = \sum_{i=0}^{d_{\min}^2-1} \gamma_i A_i \otimes B_i, \tag{2.32}$$

where $\gamma_i \geq 0$ for all i and $\{A_i\}_{i=0}^{d_A^2-1}$,$\{B_i\}_{i=0}^{d_B^2-1}$ are orthonormal bases (with respect to the Hilbert-Schmidt inner product) for d_A, d_B dimensional matrices respectively. As before, it holds $d_{\min} = \min\{d_A, d_B\}$.

The Schmidt decomposition plays a crucial role in deciding whether a bipartite pure state $|\psi\rangle$ is a product state $|\psi\rangle = |\psi_1\rangle \otimes |\psi_2\rangle$ or not. Namely, product states are states of Schmidt rank 1. An example of a state that has a higher Schmidt rank is given by $|\Phi^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$, which is clearly already in its Schmidt form and has Schmidt rank 2, the maximum for a two-qubit state. The distinction between pure product states and non-product states leads to the phenomenon of entanglement, which is not only one of the most remarkable features of nature but also a crucial resource for quantum information processing tasks.

Definition 2.4.2. *(*Pure state entanglement*). Let $\mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B$ be the Hilbert space of a bipartite system. A pure quantum state $|\psi\rangle \in \mathcal{H}_{AB}$ is called* entangled *if it cannot be written as a product state, i.e.,*

$$|\psi\rangle \neq |\psi_A\rangle \otimes |\psi_B\rangle, \tag{2.33}$$

for any states $|\psi_A\rangle \in \mathcal{H}_A$, $|\psi_B\rangle \in \mathcal{H}_B$, and it is called separable *otherwise. Equivalently, a pure quantum state $|\psi\rangle$ is called entangled if it has a Schmidt rank larger than 1.*

We will study the phenomenon of entanglement and its role and applications in quantum information processing tasks in depth in Section 3.2.3. Here, we also introduce the concept of entanglement for bipartite mixed states, which is slightly more technical than its pure state counterpart.

Definition 2.4.3. *(*Entanglement*). Let $\mathcal{S}(\mathcal{H}_{AB})$ be the set of all density operators on $\mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B$. A density operator $\rho \in \mathcal{S}(\mathcal{H}_{AB})$ is called* entangled*, if it cannot be written as a probabilistic mixture (that is a convex combination) of product states, i.e.,*

$$\rho \neq \sum_i p_i \rho_i^A \otimes \rho_i^B, \tag{2.34}$$

for any quantum states $\rho_i^A \in \mathcal{H}_A$, $\rho_i^B \in \mathcal{H}_B$, and it is called separable *otherwise.*

2.4.1 Reduced Density Operators

Sometimes when dealing with multipartite quantum states, one is in a situation where only partial information about the system is available. Consider, for instance, a bipartite quantum state $\rho_{AB} \in \mathcal{S}(\mathcal{H}_{AB})$ held by two parties, Alice (A) and Bob (B), where only system A is available to us. That might be the case in some quantum information processing tasks, where one can only control and observe particle A due to experimental limitations. For instance, particle B could be stored and manipulated in a different lab than particle A. Alternatively, system B could describe the environment of particle A, which is not under our control and, therefore, unknown to us.

We want to find a description of the system A that allows us to be ignorant about system B while being consistent with the state ρ_{AB}. The *unique* operation which describes this process is the so-called *partial trace*.

Definition 2.4.4. *(*Partial trace*). Let $V \in \mathcal{L}(\mathcal{H}_A)$ and $W \in \mathcal{L}(\mathcal{H}_B)$ be any two linear operators and $\mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B$. The* partial trace Tr_2 *is a linear mapping* $\mathrm{Tr}_2 : \mathcal{L}(\mathcal{H}_{AB}) \mapsto \mathcal{L}(\mathcal{H}_A)$ *such that*

$$\mathrm{Tr}_2[V \otimes W] = \mathrm{Tr}[W]V. \tag{2.35}$$

Note that it holds in particular that $\mathrm{Tr}_2[|a_1\rangle\langle a_2| \otimes |b_1\rangle\langle b_2|] = \langle b_2|b_1\rangle|a_1\rangle\langle a_2|$ and that the partial trace Tr_1 with respect to system A is defined analogously.

Having the partial trace as a mathematical tool available, we can define the *reduced density operator* of a bipartite quantum state and convince us that the partial trace is the correct (and unique) method to describe reduced density operators.

Definition 2.4.5. *(*Reduced density operator*). Let $\rho_{AB} \in \mathcal{S}(\mathcal{H}_{AB})$ be a density operator and $\mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B$. Then, the reduced density operators $\rho_A \in \mathcal{S}(\mathcal{H}_A)$, and $\rho_B \in \mathcal{S}(\mathcal{H}_B)$ of ρ_{AB} are defined as*

$$\rho_A \coloneqq \mathrm{Tr}_2[\rho_{AB}],\ \rho_B \coloneqq \mathrm{Tr}_1[\rho_{AB}]. \tag{2.36}$$

In particular, for any orthonormal basis $\{|b_i\rangle\}$ on $\mathcal{H}_B$ it holds that $\rho_A = \mathrm{Tr}_2[\rho_{AB}] = \sum_i(\mathbb{1}_A \otimes \langle b_i|)\rho_{AB}(\mathbb{1}_A \otimes |b_i\rangle) =: \sum_i \langle b_i|\rho_{AB}|b_i\rangle$, where $\mathbb{1}_A$ is the identity acting on $\mathcal{H}_A$. The analogous holds for $\rho_B = \mathrm{Tr}_1[\rho_{AB}]$.

While it might not be surprising that the reduced density operators of a mixed quantum state ρ_{AB} can be mixed, it is pretty astonishing that the reduced density operators of a pure state $|\psi_{AB}\rangle\langle\psi_{AB}|$ can also be mixed states. To see that this is indeed the case, we can use the Schmidt-decomposition in Theorem 2.4.1. Let the Schmidt decomposition of $|\psi_{AB}\rangle$ be given by $|\psi_{AB}\rangle = \sum_i \sqrt{\lambda_i}|i_A i_B\rangle$. It follows directly, that the reduced density operators ρ_A, ρ_B are given by

$$\rho_A = \sum_i \lambda_i |i_A\rangle\langle i_A|,\ \rho_B = \sum_i \lambda_i |i_B\rangle\langle i_B|, \tag{2.37}$$

which also directly implies that the spectrum of both reduced states is equivalent, i.e., $\lambda(\rho_A) = \lambda(\rho_B)$. Note that this is a feature of pure states that does clearly not translate to mixed states. We consider again the state $|\Phi^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$ and compute its reduced states. It follows straightforwardly that $\rho_A = \rho_B = \frac{\mathbb{1}}{2}$, which is a mixed state, in particular, the maximally mixed state. This means, there are (pure) quantum states $|\psi_{AB}\rangle$, of which there exist a global (on $\mathcal{H}_{AB}$) complete description, but the local states ρ_A and ρ_B lack information about the system. Moreover, a description of ρ_A and ρ_B is not sufficient to describe the state $|\psi_{AB}\rangle$. In fact, as shown in the particular example above, the local systems can be completely undetermined. This observation plays a crucial role for entanglement, which we discuss in detail in Section 3.2.3.

To conclude this section, we want to discuss why the partial trace is the correct and (unique) operation to define the reduced states.

Theorem 2.4.6. *(*Uniqueness of the partial trace*). Consider any linear operator $X_A \in \mathcal{L}(\mathcal{H}_A)$ and any density operator $\rho_{AB} \in \mathcal{S}(\mathcal{H}_{AB})$ with $\mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B$. Let $f : \mathcal{S}(\mathcal{H}_{AB}) \mapsto \mathcal{S}(\mathcal{H}_A)$ be any map that maps density operators in $\mathcal{S}(\mathcal{H}_{AB})$ onto density operators in $\mathcal{S}(\mathcal{H}_A)$. The partial trace Tr_2 is the unique function that satisfies*

$$\mathrm{Tr}[X_A f(\rho_{AB})] = \mathrm{Tr}[(X_A \otimes \mathbb{1}_B)\rho_{AB}]. \tag{2.38}$$

A proof can be found in [70](page 107, Box 2.6 therein).

That is, the partial trace is the unique operation that gives a consistent description of operations (in particular measurements) on a subsystem of a larger system.

Finally, we want to show that it is possible to write any mixed state $\rho_A \in \mathcal{S}(\mathcal{H}_A)$ as part of a pure quantum state $|\psi\rangle_{AB} \in \mathcal{H}_{AB}$ such that $\mathrm{Tr}_2[|\psi_{AB}\rangle\langle\psi_{AB}|] = \rho_A$. This purely mathematical procedure is known as *purification*.

Definition 2.4.7. *(*Purification*). Any quantum state $|\psi_{AB}\rangle \in \mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B$ that fulfills*

$$\mathrm{Tr}_2[|\psi_{AB}\rangle\langle\psi_{AB}|] = \rho_A, \tag{2.39}$$

is called a purification *of $\rho_A \in \mathcal{S}(\mathcal{H}_A)$.*

Theorem 2.4.8. *(*Existence of a purification*). Consider a quantum state $\rho_A \in \mathcal{S}(\mathcal{H}_A)$ with spectral decomposition $\rho_A = \sum_{i=0}^{d_A-1} \lambda_i |i_A\rangle\langle i_A|$, where d_A is the dimension of $\mathcal{H}_A$. Let $\{|i_B\rangle\}_{i=0}^{d_B-1}$ and $\{|i_{B'}\rangle\}_{i=0}^{d_{B'}-1}$ be orthonormal bases on $\mathcal{H}_B$, $\mathcal{H}_{B'}$ with $d_{B'} \geq d_B \geq d_A$. Then, the two states*

$$|\psi_{AB}\rangle = \sum_{i=0}^{d_A-1} \sqrt{\lambda_i}|i_A i_B\rangle \text{ and } |\psi'_{AB'}\rangle = \sum_{i=0}^{d_A-1} \sqrt{\lambda_i}|i_A i_{B'}\rangle, \tag{2.40}$$

are purifications of ρ_A. Moreover, there exists an isometry $V : \mathcal{L}(\mathcal{H}_B) \mapsto \mathcal{L}(\mathcal{H}_{B'})$ such that $|\psi'_{AB'}\rangle = (\mathbb{1} \otimes V)|\psi_{AB}\rangle$.

Proof. It is readily verified that $\mathrm{Tr}_2[|\psi_{AB}\rangle\langle\psi_{AB}|] = \mathrm{Tr}_2[|\psi_{AB'}\rangle\langle\psi_{AB'}|] = \rho_A$. Since $\{|i_B\rangle\}_{i=0}^{d_B-1}$ and $\{|i_{B'}\rangle\}_{i=0}^{d_{B'}-1}$ are both orthonormal bases, but not necessarily of the same dimension, it follows that the mapping $V : \mathcal{L}(\mathcal{H}_B) \mapsto \mathcal{L}(\mathcal{H}_{B'})$ describes an isometry. See also [72] (Proposition 4.1.1 therein). □

2.5 Evolution of Quantum Systems

Now that we introduced the fundamental properties of quantum states and before we study the measurement process, we discuss the evolution of quantum states, i.e., the possible transformations a quantum state can undergo from a time t_i to a time t_f. The most fundamental transformation is given by a unitary evolution of a closed system.

Postulate 3. *The evolution of a quantum state $|\psi(t_i)\rangle$ at time $t = t_i$ to a quantum state $|\psi'(t_f)\rangle$ at time $t = t_f$ in a closed quantum system is described by a unitary transformation $U(t_i, t_f)$, which is unique up to a complex phase, such that*

$$|\psi'(t_f)\rangle = U(t_i, t_f)|\psi(t_i)\rangle. \tag{2.41}$$

Moreover, a quantum state $\rho(t_i)$ undergoing the same unitary evolution $U(t_i, t_f)$ is transformed to the state $\rho'(t_f) = U(t_i, t_f)\rho(t_i)U(t_i, t_f)^\dagger$. The unitary time evolution is determined by the Hamiltonian H of the system through the Schrödinger equation

$$i\hbar\partial_t|\psi(t)\rangle = H|\psi(t)\rangle, \tag{2.42}$$

where $\hbar$ is the reduced Planck's constant. For a Hamiltonian that is not explicitly time-dependent, the unitary evolution is given by $U(t_i, t_f) = \exp\left(-\frac{i}{\hbar}H(t_f - t_i)\right)$.

Postulate 3 deals with the evolution of closed systems. This requires specifically that the system does not interact with any environment. However, this requirement is, in reality, never met. Therefore, we need to find a framework that describes the evolution of quantum systems interacting with an environment. We will see in the following that there are different ways to represent such an evolution, each of which has its advantages.

2.5.1 Representations of Quantum Channels

The first and maybe most intuitive way to describe a more general physical transformation $\Lambda(\rho) = \rho'$, including the interaction with the environment, is to describe the environment as part of the system. From here on, we will leave out the explicit time dependence of the evolution of a quantum system and only consider the explicit transformation. Consider that our quantum system is prepared in the state $\rho \in \mathcal{S}(\mathcal{H}_S)$ and the environment in some pure state $|\chi\rangle\langle\chi| \in \mathcal{S}(\mathcal{H}_E)$. Note that it is not restrictive to consider the environment to be in a pure state, as we did not specify the dimension of the Hilbert space $\mathcal{H}_E$ and we can make use of purifications if necessary. Let $\{|\chi_i\rangle\}_i$ be an orthonormal basis on $\mathcal{H}_E$. Since we included the environment into our system, the evolution of the system on $\mathcal{H}_{SE} = \mathcal{H}_S \otimes \mathcal{H}_E$ is again governed by unitary operators. This means, we can describe the transformation as

$$\Lambda(\rho) = \mathrm{Tr}_2[U(\rho \otimes |\chi\rangle\langle\chi|)U^\dagger] = \sum_i \langle\chi_i|U(\rho \otimes |\chi\rangle\langle\chi|)U^\dagger|\chi_i\rangle. \tag{2.43}$$

It turns out that this evolution does not only describe a valid transformation from quantum states to quantum states but also that this is the most general way to treat the evolution of quantum systems, as guaranteed by *Stinespring's dilation theorem* (see, e.g., [71]). We say that every such evolution describes a *quantum channel* $\Lambda(\cdot)$. We conceptualize the transformations of quantum states further in the following.

Definition 2.5.1. *(*Completely positive and trace preserving (CPTP) map*). Let $\mathcal{L}(\mathcal{H}_{\mathrm{in}})$ and $\mathcal{L}(\mathcal{H}_{\mathrm{out}})$ be the set of linear operators on the Hilbert spaces $\mathcal{H}_{\mathrm{in}}$ and $\mathcal{H}_{\mathrm{out}}$ and let $\mathbb{1}_n$ be the n-dimensional identity operator. A linear map $\Lambda : \mathcal{L}(\mathcal{H}_{\mathrm{in}}) \mapsto \mathcal{L}(\mathcal{H}_{\mathrm{out}})$ is called* positive *if $\Lambda(X) \in \mathrm{Pos}(\mathcal{H}_{\mathrm{out}})$ for any $X \in \mathrm{Pos}(\mathcal{H}_{\mathrm{in}})$. It is called* completely positive (CP) *if $(\Lambda \otimes \mathbb{1}_n)(X) \in \mathrm{Pos}(\mathcal{H}_{\mathrm{out}})$ for any $X \in \mathrm{Pos}(\mathcal{H}_{\mathrm{in}})$ and any $n \in \mathbb{N}$. Furthermore, the linear map Λ is called* trace preserving (TP) *if $\mathrm{Tr}[\Lambda(X)] = \mathrm{Tr}[X]$ for any $X \in \mathcal{L}(\mathcal{H}_{\mathrm{in}})$. Finally, a linear map $\Lambda : \mathcal{L}(\mathcal{H}_{\mathrm{in}}) \mapsto \mathcal{L}(\mathcal{H}_{\mathrm{out}})$ is said to be CPTP, or simply a quantum channel, if it is CP and TP.*

Let us comment on each of the properties of a CPTP map. To describe the most general transformation Λ of a quantum state $\rho \in \mathcal{S}(\mathcal{H}_{\mathrm{in}})$ to a quantum state $\rho' \in \mathcal{S}(\mathcal{H}_{\mathrm{out}})$, it is necessary that Λ maps positive semi-definite operators to positive semi-definite operators. This should also be true when Λ is only applied to a subsystem of ρ, which means that also $(\Lambda \otimes \mathbb{1}_n)$ has to fulfill the positivity condition. Note, it is enough to check positivity for the case $\mathbb{1}_n = \mathbb{1}_{d_{\mathrm{in}}}$, where d_{in} is the dimension of $\mathcal{H}_{\mathrm{in}}$, to guarantee complete positivity (see, e.g., [71]). Finally, we want that the map Λ preserves the normalization of quantum states, from which the trace preserving property follows. An additional property of linear maps that is frequently used is the notion of unitality. A linear map Λ is said to be *unital* if it preserves the form of the identity, i.e., $\Lambda(\mathbb{1}_{d_{\mathrm{in}}}) = \mathbb{1}_{d_{\mathrm{out}}}$.

So far, we have described how quantum channels act on quantum states, reflecting the so-called *Schrödinger picture*. However, it is also possible to describe quantum channels in terms of their action on observables, which reflects the Heisenberg picture. In the next section, we will discuss this possibility when we introduce quantum measurements. This allows us to generalize the notion of a quantum channel by introducing sub-channels and instruments, which take into account that a quantum channel can be seen as a sum of linear maps that are CP and trace non-increasing.

For now, we focus on the representation of CPTP maps. A very convenient representation of quantum channels can be obtained directly from the previous representation through the environment in Eq. (2.43). Namely, it follows that

$$\Lambda(\rho) = \mathrm{Tr}_2[U(\rho \otimes |\chi\rangle\langle\chi|)U^\dagger] = \sum_i \langle\chi_i|U(\rho \otimes |\chi\rangle\langle\chi|)U^\dagger|\chi_i\rangle = \sum_i K_i \rho K_i^\dagger, \quad (2.44)$$

where $K_i := \langle\chi_i|U|\chi\rangle$ is a so-called *Kraus operator* and the representation

$$\Lambda(\rho) = \sum_i K_i \rho K_i^\dagger,$$

is the *operator sum* or *Kraus* representation of a quantum channel. The Kraus representation is especially convenient, as it has many useful properties, which we formalize in the following theorem.

Theorem 2.5.2. *(*Kraus representation*). Let $\Lambda : \mathcal{L}(\mathcal{H}_{\mathrm{in}}) \to \mathcal{L}(\mathcal{H}_{\mathrm{out}})$ be a linear map between the Hilbert spaces $\mathcal{H}_{\mathrm{in}}$ and $\mathcal{H}_{\mathrm{out}}$. It is CP if and only if it admits a decomposition*

$$\Lambda(\rho) = \sum_i K_i \rho K_i^\dagger,$$

with Kraus operators $\{K_i\}$. Moreover, Λ is TP if $\sum_i K_i^\dagger K_i = \mathbb{1}_{d_{\mathrm{in}}}$ and unital if $\sum_i K_i K_i^\dagger = \mathbb{1}_{d_{\mathrm{out}}}$. The minimal number R of Kraus operators $\{K_i\}_{i=0}^{R-1}$ necessary to decompose a CP map Λ, is called the Kraus rank *and it is upper bounded by $R \leq d_{\mathrm{in}} d_{\mathrm{out}}$. Furthermore, it is always possible to chose R Kraus operators that are orthogonal, i.e., $\mathrm{Tr}[K_i^\dagger K_j] =\propto \delta_{ij}$. A proof can be found in [71] (Theorem 2.1 therein).*

Finally, there is a third powerful representation of quantum channels that uses the Choi-Jamiołkowski isomorphism. Namely, there is a one-to-one correspondence between quantum channels and bipartite quantum states.

Theorem 2.5.3. *(*Choi-Jamiołkowski representation*). Let $\Lambda : \mathcal{L}(\mathcal{H}_{\mathrm{H_{in}}}) \to \mathcal{L}(\mathcal{H}_{\mathrm{H_{out}}})$ be a linear map between the Hilbert spaces $\mathcal{H}_{\mathrm{in}}$ and $\mathcal{H}_{\mathrm{out}}$ and let $|\Phi^+\rangle = \frac{1}{\sqrt{d_{\mathrm{in}}}} \sum_{i=0}^{d_{\mathrm{in}}-1} |ii\rangle$ be a quantum state in $\mathcal{H}_{\mathrm{in}} \otimes \mathcal{H}_{\mathrm{in}}$. The following statements, called the* Choi-Jamiołkowski isomorphism*, defines a one-to-one mapping between linear maps Λ and linear operators $J(\Lambda)$:*

$$J(\Lambda) := (\Lambda \otimes \mathbb{1})(|\Phi^+\rangle\langle\Phi^+|), \;\; \Lambda_J(\rho) := d_{\mathrm{in}}\, \mathrm{Tr}_2[(\mathbb{1}_{\mathrm{out}} \otimes \rho^T) J(\Lambda)].$$

Here, the transpose is taken with respect to the basis $\{|i\rangle\}$. *Moreover it holds:* Λ *is CP* $\iff J(\Lambda) \succeq 0$, Λ *is TP* $\iff \mathrm{Tr}_1[J(\Lambda)] = \mathbb{1}_{\mathrm{in}}/d_{\mathrm{in}}$, Λ *is unital* $\iff \mathrm{Tr}_2[J(\Lambda)] = \mathbb{1}_{\mathrm{out}}/d_{\mathrm{out}}$. *A proof can be found in [71] (Proposition* 2.1 *therein), see also [72].*

An example of a non-unitary quantum channel of particular importance in quantum information theory is the *depolarizing channel* given by

$$\Lambda^{\eta}(\rho) = \eta\rho + (1-\eta)\,\mathrm{Tr}[\rho]\frac{\mathbb{1}}{d}, \tag{2.45}$$

which means it describes a probabilistic mixture of the state ρ with the maximally mixed state $\frac{\mathbb{1}}{d}$ depending on the *noise parameter* $\eta \in [0, 1]$.

2.6 Quantum Measurements

> *An experiment is a question which science poses to nature, and a measurement is the recording of nature's answer. But before an experiment can be performed, it must be planned–the question to nature must be formulated before being posed. Before the result of a measurement can be used, it must be interpreted–nature's answer must be understood properly.*
>
> — **Max Planck**

So far, we have discussed quantum states and their evolution. However, we still miss how to obtain information about a system's state and how to make predictions about a quantum state's observable properties. For this purpose, we introduce the notion of quantum measurements, which will play an essential role in this thesis. Here, we only introduce the basic notions of quantum measurements, focusing mainly on the mathematical description. A more refined discussion about the properties of specific quantum measurements and, in particular, of sets of quantum measurements can be found in Section 3.3. For more background on the theory of quantum measurements, we refer to the book [84].

Quantum measurements have two main remarkable properties that differentiate them from classical ones: First, quantum mechanics only allows for probabilistic statements, i.e., quantum theory is inherently random and it is only possible to predict the probability that a specific outcome occurs. Second, a quantum measurement will generally disturb the quantum state. That means, contrary to classical physics, we need to update the system's state once the measurement took place.

Both properties are summarized in the following pivotal postulate, which tells us how to obtain the probability of a given outcome and how to update the quantum state after the measurement takes place.

Postulate 4. *A quantum measurement is described by a set of k* measurement operators *$\{E_a\}_{a=0}^{k-1} \in \mathcal{L}(\mathcal{H})$, that fulfill the completeness relation $\sum_{a=0}^{k-1} E_a^\dagger E_a = \mathbb{1}_{\mathcal{H}}$, where the label a denotes one of the k possible measurement outcomes. If the measurement is performed on a quantum state $\rho \in \mathcal{S}(\mathcal{H})$, the probability to obtain outcome a is given by the* Born rule

$$p(a) = \mathrm{Tr}[E_a^\dagger E_a \rho].$$

Moreover, immediately after the measurement, the state ρ has to be replaced by the post-measurement state *ρ_a of the system via the* update rule

$$\rho_a = \frac{E_a \rho E_a^\dagger}{\mathrm{Tr}[E_a^\dagger E_a \rho]}.$$

In the following subsections, we first introduce the general notion of positive operator valued measures (POVMs) to describe quantum measurements and afterward focus on the particular case of projective measurements.

2.6.1 POVMs

In some experiments, one might only be interested in the measurement statistics, i.e., the probability distribution $\{p(a)\}_{a=0}^{k-1}$ and not in the post-measurement state. That can be the case, for example, when the post-measurement state is not accessible due to experimental limitations. In that case, we are only interested in operators $M_a \in \mathrm{Pos}(\mathcal{H})$ that map the quantum state ρ onto the probability $p(a)$. Following Postulate 4, this map is given by the Born rule $p(a) = \mathrm{Tr}[M_a \rho]$. It follows directly that the operators M_a, called *measurement effects* or POVM elements, have to fulfill $0 \preceq M_a \preceq \mathbb{1}$ $\forall a$ and, due to the normalization of the distribution $\{p(a)\}_{a=0}^{k-1}$, $\sum_a M_a = \mathbb{1}$. The set $\{M_a\}_{a=0}^{k-1}$ is known as POVM. The connection between effect operators and measurement operators could be seen by setting $E_a = \sqrt{M_a}$, where $\sqrt{M_a}$ is the (unique) positive semi-definite operator such that $\sqrt{M_a}^\dagger \sqrt{M_a} = M_a$. It follows directly that $E_a^\dagger E_a = M_a$. However, the decomposition of the measurement operators M_a into effect operators E_a (which are essentially just Kraus operators) itself is not unique. In particular $E_a' = U_a \sqrt{M_a}$ for any unitary U_a is also a valid decomposition, as $E_a'^\dagger E_a' = M_a$. This means that the post-measurement state ρ_a is not uniquely defined by a given POVM.

Having the concept of measurements at hand, we can further refine our understanding of quantum channels, which on the other hand also offer a way to implement POVMs. Consider a quantum channel Λ that can be decomposed such that $\Lambda = \sum_a \Lambda_a$, where the Λ_a are CP trace non-increasing maps. Any decomposition $\{\Lambda_a\}$ of a quantum channel Λ is called an *instrument* and the maps Λ_a are known as *subchannels*. The interpretation of an instrument is the following: The subchannels $\{\Lambda_a\}_a$ map the state ρ onto the sub-normalized state $\Lambda_a(\rho)$ with probability $\mathrm{Tr}[\Lambda_a(\rho)]$, hence the state of the system is updated to $\rho_a = \Lambda_a(\rho)/\,\mathrm{Tr}[\Lambda_a(\rho)]$. This offers the possibility to see a quantum measurements in the POVM formalism as a particular instrument. The simplest way to realize the POVM $\{M_a\}$ is given by the *Lüders instrument* $\{\Lambda_a(\rho) = \sqrt{M_a}\rho\sqrt{M_a}\}$.

Let us now consider, in the Schrödinger picture, the measurement statistics of any state evolved through the channel Λ, i.e., $\mathrm{Tr}[M_a\Lambda(\rho)]$. We can now introduce the unique Hilbert-Schmidt adjoint map $\Lambda^\dagger$ such that $\mathrm{Tr}[M_a\Lambda(\rho)] = \mathrm{Tr}[\Lambda^\dagger(M_a)\rho]$ for all quantum states ρ and POVM elements M_a. The right hand side of the equation corresponds to the (equivalent) *Heisenberg picture*, where the POVM evolves according to $\Lambda^\dagger$ and the state is left unchanged. It follows directly that a CPTP map Λ in the Schrödinger picture corresponds to a CP and unital map $\Lambda^\dagger$ in the Heisenberg picture.

Finally, a second connection between a quantum measurement and quantum channels is obtained by considering the so-called *measure-and-prepare* channel. A measure-and-prepare channel is given by

$$\Lambda_{\mathcal{M}}(\rho) = \sum_a \mathrm{Tr}[M_a\rho]|a\rangle\langle a|, \tag{2.46}$$

where $\{|a\rangle\}_{a=0}^{k-1}$ is any orthonormal basis. This type of channel corresponds to the situation where we measure the outcome a with probability $p(a) = \mathrm{Tr}[M_a\rho]$ and prepare the *register state* $|a\rangle\langle a|$. Note that the states $|a\rangle\langle a|$ could in principle be replaced by any other states τ_a and that this form of channel corresponds to a situation where the post-measurement state is not relevant, as no information over it is kept.

2.6.2 Projective Measurements

Projective measurements are a special case of POVMs, of particular relevance, introduced in any quantum mechanics textbook. The POVM elements of a projective measurement are given by orthogonal projectors $M_a = \Pi_a$ such that $\Pi_a\Pi_{a'} = \delta_{aa'}\Pi_a$. It follows directly that the measurement operators can be written as $E_a = M_a = \Pi_a$. Given the projective measurement $\{\Pi_a\}$ with outcomes a, we can define the Hermitian operator $A = \sum_a a\Pi_a$ associated to it, which is known as *observable*. With that, the expected outcome of the measurement can be calculated as $\langle A\rangle = \sum_a ap(a) = \sum_a a\,\mathrm{Tr}[\Pi_a\rho] = \mathrm{Tr}[A\rho]$. In the case all projectors are rank-one operators, i.e.,

$\Pi_a = |a\rangle\langle a|$, we say $\{\Pi_a\}$ is a *von Neumann measurement*.

Finally, we want remark that every POVM can be realized by a projective measurement in a higher dimensional Hilbert space, by means of the *Naimark extension* (see, e.g., [71]). This can already be seen from the fact that any POVM can be implemented by a particular quantum instrument $\{\Lambda_a\}$, which itself can be implemented with projective measurements on the environment state and a unitary evolution (see Eq. (2.44)).

2.6.3 Measurements on Composite Systems

In this work, we will often consider local measurements that are performed on a composite quantum state $\rho_{AB} \in \mathcal{S}(\mathcal{H}_{AB})$ with $\mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B$. That is, two parties, Alice and Bob, each hold a quantum measurement described by POVMs $\{M_a\}$ and $\{N_b\}$, respectively. The measurement formalism on a composite system has to fulfill two crucial properties. First, the measurement $\{M_a\}$ on Alice's reduced state $\rho_A = \mathrm{Tr}_2[\rho_{AB}]$ has to agree with the measurement $\{M_a \otimes \mathbb{1}_B\}$ on the global state ρ_{AB}. This is always straightforwardly fulfilled, as $\mathrm{Tr}[M_a\rho_A] = \mathrm{Tr}[(M_a \otimes \mathbb{1}_B)\rho_{AB}]$ (see Theorem 2.4.6). Second, in order to not violate the laws of special relativity, Alice should not be able to signal information about her measurement (specifically her choice of measurement) to Bob, who might hold his part of the state in a distant lab. This means that $\sum_a \mathrm{Tr}_1[(M_a \otimes \mathbb{1}_B)\rho_{AB}] = \rho_B$, which is also naturally fulfilled due to the properties of the trace, using $\sum_a M_a = \mathbb{1}_A$. Note that all arguments hold analogously for Bob's measurements $\{N_b\}$ on his share of ρ_{AB}. This further underpins the partial trace's role in consistently describing observable quantities on composite systems.

2.7 Geometric Measures for Quantum Information

In almost any quantum information processing task, one wishes to prepare a certain target quantum state ρ, measure a certain target POVM $\{M_a\}$, apply a particular target quantum channel Λ, or obtain some specific target probability distribution $\{p(a)\}$. However, as quantum processes are always subjected to some kind of noise or imperfection, we will eventually end up preparing a state ρ', measuring the POVM $\{M'_a\}$, applying the channel Λ' or obtaining the probability distribution $\{p'(a)\}$.

Many questions come immediately to our minds. (i) How does the performance of a particular quantum information processing task change when we deal with the primed objects instead of our targeted objects? (ii) How to quantify this change? (iii) How similar is the primed object to the target object? (iv) How well can we distinguish the primed from the target objects in an experiment?

In order to answer any of these questions, it is essential to find measures/functions that quantify quantum information or, more precisely, quantify the *closeness* or the *distinguishability* of two objects of interest. These functions should fulfill some necessary mathematical conditions and ideally have an operational interpretation. It is not hard to imagine that there will not be a single measure that answers all of these questions for any application. Therefore, it is often helpful to consider a variety of functions, each of which tells us some part of the answers to our questions. One approach to obtain such measures is to look for functions that quantify some distance in state space, for instance. As we will quickly see, many useful functions are actually not a metric (a distance) but they are still valuable for distinguishing quantum objects. This broader class of functions is called *geometric measures*.

2.7.1 Distances

We first focus on "*true*" distance measures in the sense that they induce a *metric* and a *metric space*.

Definition 2.7.1. *(*Metric*). Let $\mathcal{V}$ be a set and $\mathrm{D} : \mathcal{V} \times \mathcal{V} \mapsto \mathbb{R}_{\geq 0}$ a function that maps any two elements $X, Y \in \mathcal{V}$ to the non-negative real numbers. We say $\mathrm{D}(X, Y)$ is a* metric *(or simply a* distance*) if it fulfills the following conditions:*

$$1.\ \mathrm{D}(X,Y) = 0 \iff X = Y \text{ (faithfulness)}, \tag{2.47a}$$

$$2.\ \mathrm{D}(Y,X) = \mathrm{D}(X,Y) \text{ (symmetry)}, \tag{2.47b}$$

$$3.\ \mathrm{D}(X,Z) \leq \mathrm{D}(X,Y) + \mathrm{D}(Y,Z) \text{ (triangle inequality)}, \tag{2.47c}$$

for any elements $X, Y, Z \in \mathcal{V}$. The tuple $(\mathcal{V}, \mathrm{D})$ is called a metric space.

Moreover, we will often employ the distance of a point (i.e., an element $X \in \mathcal{V}$) to a non-empty subset $\mathcal{W} \subset \mathcal{V}$. We define the distance of a point to a set as follows.

Definition 2.7.2. *(*Distance to a subset*). Let $\mathcal{V}$ be a set and $\mathrm{D} : \mathcal{V} \times \mathcal{V} \mapsto \mathbb{R}_{\geq 0}$ a distance function that promotes $\mathcal{V}$ to the metric space $(\mathcal{V}, \mathrm{D})$. Furthermore, let $\mathcal{W} \subset \mathcal{V}$ be a non-empty subset of $\mathcal{V}$. We define the distance of a point $X \in \mathcal{V}$ to the set $\mathcal{W}$ as*

$$\mathrm{D}(X, \mathcal{W}) \coloneqq \inf_{Y \in \mathcal{W}} \mathrm{D}(X, Y). \tag{2.48}$$

Note that the infimum is attained whenever $\mathcal{W}$ is a compact set, which means the infimum can be replaced by a minimum in this instance.

To get an idea of distances for quantum objects, it is helpful to first consider distances of probability distributions before generalizing them to quantum states.

Once we introduce distances for quantum states, we can introduce distances for quantum channels, which helps us find specific distances for quantum measurements.

Consider two probability distributions $\{p(x)\}, \{q(x)\}$ over the same index set x. We can represent them by *probability vectors* $\mathbf{p}, \mathbf{q}$, with entries p_x, q_x, i.e., $\mathbf{p}, \mathbf{q} \in [0,1]^n$, such that $\sum_{x=0}^{n-1} p_x = 1 = \sum_{x=0}^{n-1} q_x$ and $p_x = p(x)$, $q_x = q(x)$ respectively. Let us introduce the ℓ_p norms of a general vector $\mathbf{v} \in \mathbb{C}^n$.

Definition 2.7.3. *(*ℓ_p-norm/distance*). Let* $\mathbf{v} \in \mathbb{C}^d$ *be a general vector and* $\mathrm{p} \in [1, \infty)$. *The* ℓ_p*-norm of* $\mathbf{v}$ *is defined as*

$$\|\mathbf{v}\|_{\ell_\mathrm{p}} := \Big(\sum_{i=0}^{d-1} |v_i|^\mathrm{p}\Big)^{1/\mathrm{p}}, \tag{2.49}$$

with the special case $\|\mathbf{v}\|_{\ell_\infty} := \max_i |v_i|$. *The instance* $\mathrm{p} = 2$ *is the usual* Euclidean norm *and the case* $\mathrm{p} = 1$ *is also known as* Kolmogrov norm *or as classical* trace norm. *Every* ℓ_p*-norm naturally induces a distance on* $\mathbb{C}^d$ *defined as*

$$\mathrm{D}_{\ell_\mathrm{p}}(\mathbf{v}, \mathbf{w}) := \frac{1}{2}\|\mathbf{v} - \mathbf{w}\|_{\ell_\mathrm{p}}. \tag{2.50}$$

Definition 2.7.3 already provides us with an infinite number of possible distances with which we can compare two probability distributions. However, in practice, only a few distances (typically the cases $\mathrm{p} = 1, 2, \infty$) find applications. Here, we focus on $\mathrm{p} = 1$, i.e., the classical trace distance, which is particularly important for us as it has a clear operational meaning. Namely, it can be shown (see, e.g., [70]) that for any two probability distributions $\mathbf{p}$ and $\mathbf{q}$ it holds

$$\mathrm{D}_{\ell_1}(\mathbf{p}, \mathbf{q}) = \max_E \Big| \sum_{x \in E} p_x - \sum_{x \in E} q_x \Big|, \tag{2.51}$$

where E denotes all subsets of the index set $\{x\}$, i.e., E is the optimal event with which $\mathbf{p}$ and $\mathbf{q}$ can be distinguished from each other.

We now seek to generalize the ℓ_p-norms to matrices, such that we obtain a notion of distances between quantum states. In particular, we want to generalize the classical trace distance. To do so, let us introduce the *Schatten* p*-norm* of a linear operator.

Definition 2.7.4. *(*Schatten p-norm/distance*). Let* $X \in \mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$ *be a linear operator between the Hilbert spaces* $\mathcal{H}_1$ *and* $\mathcal{H}_2$ *and* $\mathrm{p} \in [1, \infty)$. *The* Schatten p-norm *of* X *is given by*

$$\|X\|_\mathrm{p} := (\mathrm{Tr}[|X|^\mathrm{p}])^{1/\mathrm{p}}, \tag{2.52}$$

where $|X| := \sqrt{X^\dagger X}$. *Equivalently, we can define the* Schatten p-norm *as* $\|X\|_\mathrm{p} := (\sum_i \sigma_i(X)^\mathrm{p})^{1/\mathrm{p}}$*, where* $\sigma_i(X)$ *is the* i*-th singular value of* X *(ordered in non-increasing order). In particular, we define* $\|X\|_\infty := \sigma_1(X)$ *to be the largest singular value of* X*. The case* $\mathrm{p} = \infty$ *is known as* spectral norm *(sometimes also called* the operator norm*), while* $\mathrm{p} = 1$ *corresponds to the* trace norm *and* $\mathrm{p} = 2$ *to the* Frobenius norm *(which is the norm induced by the Hilbert-Schmidt inner product). Every Schatten* p*-norm naturally induces a distance on* $\mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$ *defined as*

$$\mathrm{D}_\mathrm{p}(X, Y) := \frac{1}{2}\|X - Y\|_\mathrm{p}. \tag{2.53}$$

Schatten p-norms enjoy a lot of useful properties (see, e.g., [82]), which we list here:

- $\|UXV^\dagger\|_\mathrm{p} = \|X\|_\mathrm{p}$ for any $X \in \mathcal{L}(\mathcal{H}_2, \mathcal{H}_3)$ and any isometries $U \in \mathcal{L}(\mathcal{H}_3, \mathcal{H}_4)$, $V \in \mathcal{L}(\mathcal{H}_2, \mathcal{H}_1)$.
- $\|X\|_\mathrm{p} \geq \|X\|_\mathrm{q}$ for any $X \in \mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$ and $1 \leq \mathrm{p} \leq \mathrm{q} \leq \infty$.
- For any $X \in \mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$ and numbers $\mathrm{p}, \mathrm{q} \in [1, \infty)$ such that $\frac{1}{\mathrm{p}} + \frac{1}{\mathrm{q}} = 1$ (with $\frac{1}{\infty} := 0$) it holds: $\|X\|_\mathrm{p} = \sup\{|\mathrm{Tr}[Y^\dagger X]| : Y \in \mathcal{L}(\mathcal{H}_1, \mathcal{H}_2), \|Y\|_\mathrm{q} \leq 1\}$.
- $\|XY\|_1 \leq \|X\|_\mathrm{p}\|X\|_\mathrm{q}$ for any $Y \in \mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$, $X \in \mathcal{L}(\mathcal{H}_2, \mathcal{H}_3)$ and $\mathrm{p}, \mathrm{q} \in [1, \infty)$ such that $\frac{1}{\mathrm{p}} + \frac{1}{\mathrm{q}} = 1$.
- $|\mathrm{Tr}[Y^\dagger X]| \leq \|X\|_\mathrm{p}\|Y\|_\mathrm{q}$ for any $X, Y \in \mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$ and $\mathrm{p}, \mathrm{q} \in [1, \infty)$ such that $\frac{1}{\mathrm{p}} + \frac{1}{\mathrm{q}} = 1$.
- $\|XY\|_\mathrm{p} \leq \|X\|_\mathrm{p}\|Y\|_\mathrm{p}$ for any $Y \in \mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$, $X \in \mathcal{L}(\mathcal{H}_2, \mathcal{H}_3)$.

Similar to the situation for the ℓ_p-distances, not all Schatten p-distances are relevant in typical applications. We will focus here on the *trace distance*, i.e., the case $\mathrm{p} = 1$. The trace distance can be seen as a generalization of the classical trace distance in the following sense. Consider two quantum states $\rho = \sum_i p_i |i\rangle\langle i|$ and $\rho' = \sum_i q_i |i\rangle\langle i|$. It follows directly that

$$\mathrm{D}_1(\rho, \rho') = \mathrm{D}_{\ell_1}(\mathbf{p}, \mathbf{q}), \tag{2.54}$$

where $\mathbf{p}$ and $\mathbf{q}$ are the distributions of the eigenvalues of ρ and ρ' here. The trace distance has an operational interpretation that can be seen as a generalization of its classical counterpart. Namely, it can be shown for any two quantum states $\rho_0, \rho_1 \in \mathcal{S}(\mathcal{H})$ that

$$\mathrm{D}_1(\rho_0, \rho_1) = \max_{0 \preceq M \preceq \mathbb{1}} \mathrm{Tr}[M(\rho_0 - \rho_1)], \tag{2.55}$$

which allows us to connect the trace distance to the optimal distinguishability of two equally likely distributed states ρ_0 and ρ_1 in a *single-shot* (i.e., one round) experiment. To do so, we use a dichotomic POVM $\{M_0, M_1\}$, where M_0 corresponds to guessing the label "0" (corresponding to ρ_0) and M_1 to guessing "1" (corresponding to ρ_1), respectively. It follows directly that the optimal probability to guess the labels correctly is given by

$$p_{1,\text{guess}}^{(\rho_0,\rho_1)} \coloneqq \frac{1}{2}(p(0|\rho_0) + p(1|\rho_1)) = \frac{1}{2}(1 + \mathrm{D}_1(\rho_0, \rho_1)), \tag{2.56}$$

where $p(0|\rho_0)$ is the probability to guess the label "0" provided that ρ_0 was distributed and $p(1|\rho_1)$ is the probability to guess the label "1" provided that ρ_1 was distributed.

Besides this operational interpretation and the basic properties of a metric, the trace distance enjoys an additional property, which is especially useful in the context of quantum information and, in particular, for quantifying quantum resources. That is, the trace distance is also *contractive* under general CPTP maps, i.e., for two quantum states ρ_0, ρ_1 and a general CPTP map Λ it holds

$$\mathrm{D}_1(\rho_0, \rho_1) \geq \mathrm{D}_1(\Lambda(\rho_0), \Lambda(\rho_1)). \tag{2.57}$$

That means there exists no quantum channel that can increase the distinguishability of ρ_0 and ρ_1. Note that in the particular case of the partial trace, it follows that reduced density operators can never be distinguished better than the corresponding global states.

Now that we have a notion of distances in state space, we also want to define distances between quantum channels. In particular, we will see how the trace norm can be generalized to quantum channels through the *diamond norm*. Probably the most intuitive way to distinguish quantum channels based on the trace norm is to introduce the induced norm

$$\|\Lambda\|_{1\to 1} \coloneqq \max_{\rho \in \mathcal{S}(\mathcal{H}_{\text{in}})} \|\Lambda(\rho)\|_1, \tag{2.58}$$

where $\Lambda : \mathcal{L}(\mathcal{H}_{\text{in}}) \mapsto \mathcal{L}(\mathcal{H}_{\text{out}})$ is any linear map between linear operators on $\mathcal{H}_{\text{in}}$ and $\mathcal{H}_{\text{out}}$. The induced distance between CPTP maps Λ_0, Λ_1 is then given by

$$\mathrm{D}_{1\to 1}(\Lambda_0, \Lambda_1) \coloneqq \frac{1}{2}\|\Lambda_0 - \Lambda_1\|_{1\to 1}. \tag{2.59}$$

However, it turns out that this is not the optimal generalization of the trace distance to quantum channels. In particular, quantum channels can be distinguished with a

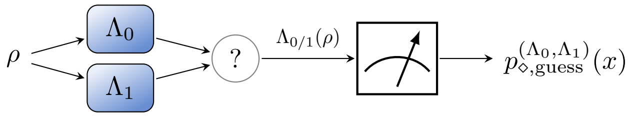

Fig. 2.2.: Representation of the operational interpretation of the diamond norm from [65] (Paper C). Two quantum channels Λ_0, Λ_1 are distinguished by preparing an optimal state ρ and subjecting the states $\Lambda_0(\rho), \Lambda_1(\rho)$ to a dichotomic measurement that distinguishes the states as well as possible. The probability $p_{\diamond,\text{guess}}^{(\Lambda_0,\Lambda_1)}$ signifies the optimal probability to correctly guess the label "0/1".

higher probability if one considers an ancilla space $\mathcal{H}_E$ such that quantum channel acts only on a subsystem of a possibly entangled bipartite quantum state. This leads to the definition of the *diamond norm*.

Definition 2.7.5. *(*Diamond norm/distance*). Let $\Lambda : \mathcal{H}_{\text{in}} \mapsto \mathcal{H}_{\text{out}}$ be a CP map between the Hilbert spaces $\mathcal{H}_{\text{in}}$ and $\mathcal{H}_{\text{out}}$. The* diamond norm *of Λ is defined as*

$$\|\Lambda\|_\diamond := \|\Lambda \otimes \mathbb{1}_{\text{in}}\|_{1\to 1}, \tag{2.60}$$

where $\mathbb{1}_{\text{in}}$ denotes the identity operator with dimension of $\mathcal{H}_{\text{in}}$. The diamond distance *between two CP maps Λ_0, Λ_1 induced by the diamond norm is defined as*

$$\mathrm{D}_\diamond(\Lambda_0, \Lambda_1) := \frac{1}{2}\|\Lambda_0 - \Lambda_1\|_\diamond. \tag{2.61}$$

Since the diamond norm is a natural extension of the trace distance, it enjoys very similar properties. Most importantly, the trace distance's operational interpretation directly translates to the diamond norm. This means that the optimal probability with which two equally likely quantum channels Λ_0, Λ_1 can be distinguished by preparing an optimal quantum state ρ and performing an ideal (dichotomic) measurement is given by

$$p_{\diamond,\text{guess}}^{(\Lambda_0,\Lambda_1)} = \frac{1}{2}(1 + \mathrm{D}_\diamond(\Lambda_0, \Lambda_1)). \tag{2.62}$$

This operational interpretation is schematized in Figure 2.2. The main task we will use the diamond distance for within this thesis, is to compare measure-and-prepare channels (see Eq. (2.46)) with each other. That gives us a notion of distances between measurements. We will specifically use distances between measurements in Sections 4.3 and 4.4, which discuss Publication C and D, to quantify the resources of different sets of measurements.

2.7.2 Robustnesses

Apart from metrics, many more functions give us a notion of *closeness* or *similarity* between different quantum objects. Among the most popular of such functions are the so-called *robustnesses*. These are typically employed to quantify how *noise robust* a given quantum state ρ is. Consider, for instance, the depolarizing channel in Eq. (2.45) which we repeat here for convenience:

$$\Lambda^{\eta}(\rho) = \eta\rho + (1-\eta)\,\mathrm{Tr}[\rho]\frac{\mathbb{1}}{d}. \tag{2.63}$$

As the parameter $(1-\eta) \in [0,1]$ increases, the fraction of the maximally mixed state $\frac{\mathbb{1}}{d}$ becomes more and more dominant. This means, that even a pure state $\rho = |\psi\rangle\langle\psi|$ will eventually end up in the maximally mixed state, hence losing all its useful properties.

We can generalize this notion of susceptibility to noise as follows. First, instead of mixing the state ρ with the maximally mixed state, we allow for mixtures with states $\tau \in \mathcal{T}(\mathcal{H}) \subset \mathcal{S}(\mathcal{H})$, where $\mathcal{T}(\mathcal{H})$ is any subset of all states $\mathcal{S}(\mathcal{H})$. Second, we are generally interested in the robustness with respect to a second set $\mathcal{V}(\mathcal{H}) \subset \mathcal{S}(\mathcal{H})$ instead of the robustness with respect to the maximally mixed state $\frac{\mathbb{1}}{d}$. Finally, we substitute $\eta = \frac{1}{1+r}$. With these generalizations, we can define different robustnesses (see, e.g., [31]) for any choice of $\mathcal{T}(\mathcal{H})$ and $\mathcal{V}(\mathcal{H})$, i.e.,

$$\mathrm{R}^{\mathcal{T}}(\rho) = \inf_{\tau\in\mathcal{T},r}\Big\{r \geq 0 : \frac{\rho + r\tau}{1+r} \in \mathcal{V}\Big\}, \tag{2.64}$$

where we take the infimum over all possible states $\tau \in \mathcal{T}$ that we can mix ρ with and the mixing parameter r. Note that this robustness function is not a metric, as it is not symmetric and does generally not obey the triangle inequality. Moreover, it could be infinite, depending on our choice of the sets $\mathcal{T}$ and $\mathcal{V}$. In practice however, these sets are always chosen such that $\mathrm{R}^{\mathcal{T}}(\rho)$ is finite.

The concept of robustness can be further generalized to any type of set, such that it also applies to probability distributions, measurements, and quantum channels.

Definition 2.7.6. (Robustness). *Let $\mathcal{V}, \mathcal{T} \subset \mathcal{S}$ be two subsets of a set $\mathcal{S}$. The $\mathcal{T}$-*robustness $\mathrm{R}^{\mathcal{T}}(X) : \mathcal{S} \mapsto \mathbb{R}_{\geq 0}$ *of a point $X \in \mathcal{S}$ with respect to the non-empty set $\mathcal{V} \subset \mathcal{S}$ is defined as*

$$\mathrm{R}^{\mathcal{T}}(X) := \inf_{Y\in\mathcal{T},r}\Big\{r \geq 0 : \frac{X + rY}{1+r} \in \mathcal{V}\Big\}. \tag{2.65}$$

There are two specific choices for the set $\mathcal{T}$ that are frequently used. First, in the case $\mathcal{T} = \mathcal{V}$, we say $R^{\mathcal{T}}(X)$ is the *absolute robustness*. Second, in the case $\mathcal{T} = \mathcal{S}$ we say $\mathrm{R}^{\mathcal{T}}(X)$ is the *generalized robustness* [31].

2.7.3 Weight Decompositions

Similar to robustnesses, *weight decompositions* quantify the similarity or closeness of an object with respect to a specific set. However, instead of asking how much noise has to be mixed to a state in order to end up in the specific set, one asks how objects can be decomposed into a general object and a noise component. Take again the example of a depolarized state $\rho^\eta = \eta\rho + (1-\eta)\frac{\mathbb{1}}{d}$. That means ρ^η can be decomposed as a probabilistic mixture of a general state ρ (with probability η) and the maximally mixed state (with probability $(1-\eta)$). We want to generalize this concept to general sets, the same way we generalized the robustness.

Definition 2.7.7. *(*Weight*). Let $\mathcal{V} \subset \mathcal{S}$ be a non-empty subset of a set $\mathcal{S}$. The* weight $\mathrm{W}(X) : \mathcal{S} \mapsto [0,1]$ *of an element $X \in \mathcal{S}$ with respect to $\mathcal{V}$ is defined as*

$$\mathrm{W}(X) := \inf_{Y \in \mathcal{V}, X' \in \mathcal{S}, w} \Big\{ w \geq 0 : X = wX' + (1-w)Y \Big\}. \tag{2.66}$$

Similarly to the robustness, weight functions are not metrics. However, by definition, they are always finite. Moreover, extremal points $X \notin \mathcal{V}$ always have weight $\mathrm{W}(X) = 1$, as they are exactly those points of a set $\mathcal{S} \setminus \mathcal{V}$ that cannot be written as a non-trivial convex combination.

2.8 Entropic Measures for Quantum Information

> *You should call it entropy, for two reasons. In the first place your uncertainty function has been used in statistical mechanics under that name, so it already has a name. In the second place, and more important, nobody knows what entropy really is, so in a debate you will always have the advantage.*
>
> — **John von Neumann**
> In correspondence to Claude Shannon.

As we have already seen, quantum mechanics is an inherently probabilistic theory, meaning generally only probabilistic statements about measurement outcomes are possible. Moreover, when we introduced mixed states, we already saw some uncertainty in our system's description, which is not prevalent for pure states. The mathematical tools that can be used to quantify this uncertainty are entropies. Furthermore, entropies allow for yet another notion of similarity between probability distributions and quantum states. Let us first define the *Shannon entropy* of a probability distribution $\mathbf{p}$ that describes the distribution of the outcomes x of a random variable X.

Definition 2.8.1. *(*Shannon entropy*). Let X be a random variable with outcomes x that are distributed according to the probability distribution $\mathbf{p}$. The* Shannon entropy *of $\mathbf{p}$ is defined as*

$$\mathrm{H}(\mathbf{p}) := -\sum_x p(x)\log p(x), \tag{2.67}$$

where the logarithm is taken with respect to base 2 in this thesis and $0\log(0) := 0$.

The Shannon entropy is a quantifier of the uncertainty or ignorance about the value of X before we learn its outcome x. On the other hand, the Shannon entropy can be seen as the information gained after learning the value of X.

It can be shown that $0 \leq \mathrm{H}(\mathbf{p}) \leq \log n$ for any probability distribution $\mathbf{p}$ having n different outcomes. The bounds $\mathrm{H}(\mathbf{p}) = 0$ and $\mathrm{H}(\mathbf{p}) = \log n$ are particularly important. The lower bound $\mathrm{H}(\mathbf{p}) = 0$ holds if and only if $p(x) = \delta_{x',x}$ for a specific outcome x' of X. That is, one particular outcome occurs with probability 1 while all the other outcomes never occur. This means that $\mathrm{H}(\mathbf{p}) = 0$ signifies that there is no uncertainty about the value of X and that there is no information to obtain when X is measured. On the other hand, $\mathrm{H}(\mathbf{p}) = \log n$ holds if and only if $p(x) = \frac{1}{n}\forall n$, i.e. the outcome probability is uniformly distributed. This represents the fact, that we are maximally uncertain about the value of X, if all of its outcomes are equally likely. This implies that learning the value of a random variable X that is uniformly distributed, lets us obtain the maximal amount of information possible.

2.8.1 Von Neumann Entropy

We now want to extend the concept of entropies to the quantum regime. This leads us to the definition of the *von Neumann entropy*.

Definition 2.8.2. *(*Von Neumann entropy*). The* von Neumann entropy *of a quantum state $\rho \in \mathcal{S}(\mathcal{H})$ is defined as*

$$\mathrm{S}(\rho) := -\operatorname{Tr}[\rho\log\rho]. \tag{2.68}$$

Moreover, let $\lambda(\rho) = \{\lambda_i\}$ be the spectrum of ρ. Then, the von Neumann entropy can be written as

$$\mathrm{S}(\rho) = -\sum_i \lambda_i \log \lambda_i, \tag{2.69}$$

which is the Shannon entropy of the eigenvalues $\lambda(\rho)$.

Similarly to the Shannon entropy, the von Neumann entropy is bounded such that $0 \leq \mathrm{S}(\rho) \leq \log d$, where d is the dimension of $\mathcal{H}$. It holds that $\mathrm{S}(\rho) = 0$ if and only if $\rho = |\psi\rangle\langle\psi|$ is a pure state, which means the state of the system contains no uncertainty. On the other hand, $\mathrm{S}(\rho) = \log d$ if and only if $\rho = \frac{\mathbb{1}}{d}$ is the maximally mixed state. This further justifies the name *maximally mixed state*, as it is the state of maximal uncertainty.

2.8.2 Rényi Entropies

The Shannon entropy is not the *unique* measure of the uncertainty of a random variable and many other entropies can be used instead. One important family consists of the so-called *Rényi entropies*.

Definition 2.8.3. *(*Rényi entropies*). Let X be a random variable with outcomes x that are distributed according to the probability distribution $\mathbf{p}$. The* Rényi entropy *of order α is defined as*

$$\mathrm{H}_\alpha(\mathbf{p}) \coloneqq \frac{1}{1-\alpha} \log\Big(\sum_x p(x)^\alpha\Big), \tag{2.70}$$

for any $\alpha \geq 0$ such that $\alpha \neq 1$. Note that in the limit $\lim_{\alpha\to1} \mathrm{H}_\alpha(\mathbf{p}) = \mathrm{H}(\mathbf{p})$, i.e., the Rényi entropies converges to the Shannon entropy.

Like the Shannon entropy, the Rényi entropies fulfill $0 \leq \mathrm{H}_\alpha(\mathbf{p}) \leq \log n$ for any probability distribution $\mathbf{p}$ and all α. Furthermore, the bound $\mathrm{H}_\alpha(\mathbf{p}) = 0$ is achieved if and only if $p(x) = \delta_{x',x}$ and $\mathrm{H}_\alpha(\mathbf{p}) = \log n$ if and only if $p(x) = \frac{1}{n}$ for all x. The Rényi entropies can also be generalized to quantum states.

Definition 2.8.4. *(*Quantum Rényi entropies*). The* quantum Rényi entropy *of order α of a quantum state $\rho \in \mathcal{S}(\mathcal{H})$ is defined as*

$$\mathrm{S}_\alpha(\rho) \coloneqq \frac{1}{1-\alpha} \log \mathrm{Tr}[\rho^\alpha], \tag{2.71}$$

for all $\alpha \geq 0$ such that $\alpha \neq 1$. Note that in the limit $\lim_{\alpha \to 1} \mathrm{S}_\alpha(\rho) = \mathrm{S}(\rho)$, i.e., the quantum Rényi entropies converge to the von Neumann entropy.

Moreover, let $\lambda(\rho)$ be the spectrum of ρ. Then, the quantum Rényi entropies can be written as

$$\mathrm{S}_\alpha(\rho) = \frac{1}{1-\alpha} \log \Big(\sum_i \lambda_i^\alpha \Big), \tag{2.72}$$

which are the Rényi entropies of the eigenvalues $\lambda(\rho)$.

The quantum Rényi entropies and the von Neumann entropy as a special case, enjoy many useful mathematical properties (see e.g. [78]) which we list in the following:

- $0 \leq \mathrm{S}_\alpha(\rho) \leq \log d$ with $\mathrm{S}_\alpha(\rho) = 0$ if and only if ρ is pure and $\mathrm{S}_\alpha(\rho) = \log d$ if and only if ρ is maximally mixed.

- $\mathrm{S}_\alpha(\rho) = \mathrm{S}_\alpha(V\rho V^\dagger)$ for any isometry $V \in \mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$.

- $\mathrm{S}_\alpha(\rho_1 \otimes \rho_2) = \mathrm{S}_\alpha(\rho_1) + \mathrm{S}_\alpha(\rho_2)$.

- $\mathrm{S}_{\alpha_1}(\rho) \geq \mathrm{S}_{\alpha_2}(\rho)$ for any $1 < \alpha_1 \leq \alpha_2$.

Additionally, the von Neumann entropy $\mathrm{S}(\rho)$ fulfills:

- $\mathrm{S}(\rho_{AB}) \leq \mathrm{S}(\rho_A) + \mathrm{S}(\rho_B)$ for any bipartite state with $\rho_B = \mathrm{Tr}_1[\rho_{AB}]$ and $\rho_A = \mathrm{Tr}_2[\rho_{AB}]$.

- $\mathrm{S}(\sum_i p_i \rho_i) \geq \sum_i p_i \mathrm{S}(\rho_i)$ for any probability distribution $\{p_i\}$ and density operators $\{\rho_i\}$, i.e., the von Neumann entropy is concave.

2.8.3 Relative Entropy

Until now, we focused on entropies that tell us something about the uncertainty or the possible information gain in a quantum state (or a classical random variable). Now, we want to introduce a tool that gives us a notion of similarity for quantum states in terms of entropies. This leads to the concept of the *(quantum) relative entropy*. Note that the quantum relative entropy is a generalization of the classical relative entropy or *Kullback-Leibler divergence*, which we do not discuss here. Instead, we introduce the quantum case directly.

Definition 2.8.5. *(*Relative entropy*). The* relative entropy *of two quantum states $\rho_0, \rho_1 \in \mathcal{S}(\mathcal{H})$ is defined as*

$$\mathrm{S}(\rho_0 \| \rho_1) \coloneqq \mathrm{Tr}[\rho_0(\log \rho_0 - \log \rho_1)], \tag{2.73}$$

with the convention $\mathrm{S}(\rho_0\|\rho_1) \coloneqq \infty$ in the case $\mathrm{supp}(\rho_0) \cap \ker(\rho_1) \neq 0$, where $\mathrm{supp}(\rho_0)$ is the support of ρ_0 and $\ker(\rho_1)$ is the kernel of ρ_1.

From the definition of the relative entropy, it is not clear that it can be used as a tool for distinguishing quantum states. In particular, it is not hard to see that the relative entropy is not only non-symmetric, i.e., $\mathrm{S}(\rho_0\|\rho_1) \neq \mathrm{S}(\rho_1\|\rho_0)$, but also that it does not obey the triangle inequality. That means the relative entropy is not a metric. However, *Klein's inequality* (see e.g. [78]) tells us that

$$\mathrm{S}(\rho_0\|\rho_1) \geq 0, \tag{2.74}$$

with the equality holding if and only if $\rho_0 = \rho_1$. This means, the relative entropy is at least faithful. Moreover, the relative entropy is contractive under CPTP maps. Namely, it holds

$$\mathrm{S}(\rho_0\|\rho_1) \geq \mathrm{S}(\Lambda(\rho_0)\|\Lambda(\rho_1)) \tag{2.75}$$

for any two quantum states $\rho_0, \rho_1 \in \mathcal{S}(\mathcal{H})$ and any CPTP map Λ.

The relevance of the relative entropy is twofold: First, many other entropies which are important in quantum information theory can be formulated as particular instances of the relative entropy. Second, it has a clear (but more complicated) operational interpretation in hypothesis testing through *quantum Stein's Lemma* [85, 86]. Throughout this thesis, it will simply serve as an additional method to compare quantum states and quantify quantum resources.

2.9 Semidefinite Progamming

Many problems in quantum information involve optimizing some quantity of interest over the set of density matrices, quantum measurements, or over the set of CPTP maps. These problems typically have in common that one seeks to minimize a linear function or, more generally, a convex function over some convex domain. Such problems are often tackled via semidefinite programming, which is a special instance of convex optimization for which efficient numerical tools and ready-to-use software exists. However, the framework of semidefinite programming goes beyond its numerical uses, as it can also be used to obtain analytical bounds and, in some instances, even the exact analytical results. Here, we give a short introduction to semidefinite programming in the language of [44], which is itself heavily based on [82]. For a wider overview see also [87]. Note that there are many equivalent formulations, and we will only present one, which is particularly close to the problems in quantum information. A general semidefinite program (SDP) is given by:

$$\begin{aligned}
&\underline{\text{Primal problem:}} \\
&\text{given}: \quad A, B, \Lambda \\
&\text{find } \alpha := \min_X \operatorname{Tr}[AX] \\
&\text{subject to:} \\
&\Lambda(X) = B, \ X \succeq 0,
\end{aligned} \tag{2.76}$$

where $A \in \operatorname{Herm}(\mathcal{H}_1)$ and $B \in \operatorname{Herm}(\mathcal{H}_2)$ are some Hermitian matrices and $\Lambda(\cdot)$: $\operatorname{Herm}(\mathcal{H}_1) \mapsto \operatorname{Herm}(\mathcal{H}_2)$ is a linear mapping between Hermitian matrices. The meaning of the name *primal problem* will become clear later. We say that $\operatorname{Tr}[AX]$ is the *primal objective function* that is to be minimized subject to the *primal constraints* $\Lambda(X) = B, \ X \succeq 0$. Furthermore, we say that every X that fulfills all constraints is a *primal feasible point* and α is the *primal optimal value*.

One can associate every such primal problem with a *Lagrangian* that incorporates the constraints explicitly into the primal objective function by introducing *Lagrange multipliers*. We obtain the Lagrangian given by

$$\begin{aligned}
\mathscr{L}(X, Y, Z) &= \operatorname{Tr}[AX] + \operatorname{Tr}[Y(B - \Lambda(X))] - \operatorname{Tr}[ZX] \\
&= \operatorname{Tr}[X(A - \Lambda^\dagger(Y) - Z)] + \operatorname{Tr}[YB],
\end{aligned} \tag{2.77}$$

where Y, Z are Hermitian Lagrange multipliers (matrices of appropriate size) and $\Lambda^\dagger(Y)$ is the Hilbert-Schmidt conjugate map of Λ such that $\operatorname{Tr}[\Lambda(X)Y] = \operatorname{Tr}[\Lambda^\dagger(Y)X]$. It is convenient to restrict to $Z \succeq 0$ because in that case we have $\alpha \geq \mathscr{L}(X, Y, Z)$ for any primal feasible X. We want to obtain a function from $\mathscr{L}(X, Y, Z)$ that is still a lower bound to α but that is independent of X. This will give us the possibility to get a lower bound for any matrices Y, Z. We define the dual function

$$\mathrm{G}(Y, Z) := \inf_X \mathscr{L}(X, Y, Z). \tag{2.78}$$

Note that the dual function $\mathrm{G}(Y, Z)$ is unbounded from below unless certain constraints (*the dual constraints*) are met. More formally,

$$\mathrm{G}(Y, Z) = \begin{cases} \operatorname{Tr}[YB], & \text{if } A - \Lambda^\dagger(Y) = Z, \ Z \succeq 0, \\ -\infty, & \text{otherwise.} \end{cases} \tag{2.79}$$

Clearly, $\alpha \geq \mathrm{G}(Y, Z)$ for any Y, Z. To find the best lower bound on α, we maximize the dual function $\mathrm{G}(Y, Z)$ over the dual constraints. This means, we obtain the optimization problem:

$$\begin{aligned} &\underline{\text{Dual problem:}} \\ &\text{given}: \quad A, B, \Lambda \\ &\text{find } \beta := \max_{Y,Z} \mathrm{Tr}[YB] \\ &\text{subject to:} \\ &A - \Lambda^\dagger(Y) = Z, \ Z \succeq 0, \end{aligned} \tag{2.80}$$

which is formally called the *dual problem* of our primal problem. However, we can simplify the dual problem by eliminating the variable Z, that only functions as a slack variable here. We finally obtain:

$$\begin{aligned} &\underline{\text{Dual problem:}} \\ &\text{given}: \quad A, B, \Lambda \\ &\text{find } \beta := \max_{Y} \mathrm{Tr}[YB] \\ &\text{subject to:} \\ &A - \Lambda^\dagger(Y) \succeq 0. \end{aligned} \tag{2.81}$$

We way that $\mathrm{Tr}[YB]$ is the *dual objective* and β is the *dual optimal value*. Furthermore, way say that every Y that satisfies the dual constraints is a *dual feasible point*.

From the above derivation, it follows by design that

$$\Delta := \alpha - \beta \geq 0, \tag{2.82}$$

which is referred to as *weak duality* and Δ is called the *duality gap*. Note that weak duality holds for general problems, even if they do not fall under the category of *convex optimization problems*.

Most important for our purposes are the instances where $\Delta = 0$, i.e., $\alpha = \beta$, which we refer to as *strong duality*. In that case, we can use the primal and dual as equivalent formulations of the same problem. While it is, in general, a difficult task to certify that strong duality holds, there is one powerful condition that is often applicable in practice and will help us to prove strong duality for semidefinite programs. This sufficient condition, known as *Slater's condition* (see e.g. [82, 87]), states that strong duality of an SDP problem holds if at least one of the problems (primal or dual) has a strictly feasible point, i.e., a point $X \succ 0$ with $\Lambda(X) = B$ for the primal or a point Y with $A - \Lambda^\dagger(Y) \succ 0$ for the dual. Note that this strictly feasible point does not need to be optimal.

Finally, we want to comment that the above formulation of an SDP can straightforwardly be extended to the case of multiple variables or constraints. Also further inequality constraints can be incorporated using slack variables. Furthermore, SDPs used in practice are often not written in the *standard form* of the primal or dual problem shown above. Instead, they are often formulated in a more complicated form, which can always be reduced to the standard form. Nevertheless, formulating the Lagrangian and obtaining the dual function works also in these more complicated situations.

3 Quantum Resource Theories

> *My favorite things in life don't cost any money. It's really clear that the most precious resource we all have is time.*
>
> — **Steve Jobs**

Studying quantum resources and their interplay in quantum information processing tasks is an integral part of this thesis. Our main goal is to investigate the role of quantum resources in Bell-type scenarios. We study the Bell scenario in depth in Section 3.4.1 after introducing the components necessary to discuss Bell nonlocality and Bell's theorem in detail. These components are the resources of quantum states and quantum measurements. The field of quantum resource theorys (QRTs) adapts many ideas from economics, as it deals in the broader sense with goods of different utility, commonness, demand, and cost. Clearly, the different properties of a good are not necessarily independent. The cost of a good typically correlates to its demand, which, for instance, originates from its utility.

An example of resources in the economic sense are different energy sources. These are highly useful, and the demand is ever-growing, which results in a high cost, even though different energy sources are available nowadays. Some forms of energy are easier to use or transform into other forms of energy than others. We could regard these as more resourceful than forms of energy that are difficult to use or transform. We need functions that quantify the resource to see how valuable an energy source is. For instance, one could quantify the resource of an energy source in terms of its energy density or the amount of electricity that can be generated from it.

Coming back to the quantum case, roughly speaking, a quantum resource is a property of some quantum object, e.g., of a state or a measurement that provides some advantage in a quantum information processing task over objects that do not have said property. QRTs offer a general framework to study and quantify the usefulness of specific resources in a given scenario. In the following, we will study the general structure of resource theories before we study resource theories of quantum states, measurements, and correlations in depth.

3.1 General Structures in Resource Theories

QRTs are used to analyze the properties of various essential objects in quantum information theory, like quantum states, measurements, channels, instruments, or correlations. Historically, resource theories for quantum states have been developed earlier than those of other fields. That is because the entanglement of quantum states has been recognized as a resource first and quickly became the prime example of a resource theory. Therefore, QRTs of quantum states are well-established, and the analysis of the general underlying structures is more developed for states than for other fields of QRTs. However, more recently, the focus of attention shifted toward the resources of quantum channels and measurements. Typically, the framework of resource theories is introduced separately for each of the objects mentioned above. Here, we go a different route and introduce the structures of resource theories more generally by adapting the framework presented in [31], such that the specific cases emerge as a particular instance of our formulation. Then, we will have a more detailed look at the specific resource theories of the particular objects in the upcoming sections.

A general resource theory consists of the following ingredients:

- Sets $\mathcal{S}_1, \mathcal{S}_2$ that are each divided into two sets $\mathcal{V}_i \subset \mathcal{S}_i$ and $\mathcal{R}_i := \mathcal{S}_i \backslash \mathcal{V}_i$ for $i = \{1, 2\}$.
- A set $\mathscr{F}(\mathcal{S}_1, \mathcal{S}_2)$ of linear maps $\Lambda : \mathcal{S}_1 \mapsto \mathcal{S}_2$ such that $\Lambda(X) \in \mathcal{V}_2$ for any $X \in \mathcal{V}_1$.
- Functions $\mathrm{R} : \mathcal{S}_1 \cup \mathcal{S}_2 \mapsto \mathbb{R}_{\geq 0}$ that fulfill $\mathrm{R}(X) = 0$ for any $X \in \mathcal{V}_i$ and $\mathrm{R}(X) \geq \mathrm{R}(\Lambda(X))$ for any $\Lambda \in \mathscr{F}(\mathcal{S}_1, \mathcal{S}_2)$ and any $X \in \mathcal{S}_1$.

The sets $\mathcal{S}_1, \mathcal{S}_2$ define the general objects we deal with. For instance, these could be the set of density operators on Hilbert spaces $\mathcal{H}_1, \mathcal{H}_2$ or sets of POVMs acting on these Hilbert spaces. These general sets are divided into sets $\mathcal{R}_i$ of resource objects and the sets $\mathcal{V}_i$ of free (or void) objects. Considering a general quantum information processing task, the free objects $F \in \mathcal{V}_i$ do not provide any advantage (hence, fail in the task), while the resources $X \in \mathcal{R}_i$ do provide such an advantage. Moreover, free objects F are typically easy to produce or obtain in an experimental situation, while resource objects are more challenging to generate. Typically, the sets $\mathcal{V}_i$ are defined via some condition that unambiguously describes the mathematical form of their elements. This condition is essentially the same for the sets $\mathcal{V}_1$ and $\mathcal{V}_2$. However, they could, for instance, be defined via vector spaces of different dimensions.

The linear maps $\Lambda \in \mathscr{F}(\mathcal{S}_1, \mathcal{S}_2)$ do map free objects $F \in \mathcal{V}_1$ to free objects $\Lambda(F) \in \mathcal{V}_2$, i.e., they cannot create resources from nothing and are therefore called

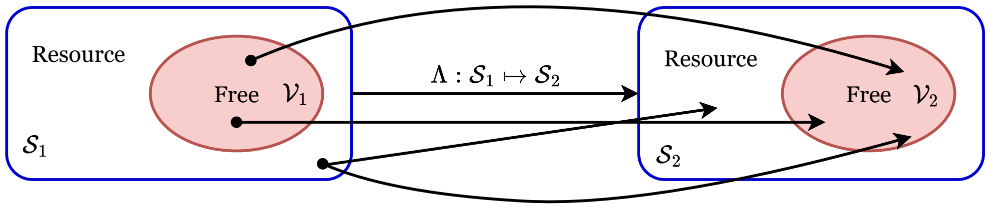

Fig. 3.1.: Sketch of a resource theory. The resource theory is determined by the sets of free objects $\mathcal{V}_1 \subset \mathcal{S}_1$, $\mathcal{V}_2 \subset \mathcal{S}_2$ and the set $\mathscr{F}(\mathcal{S}_1, \mathcal{S}_2)$ of free maps $\Lambda : \mathcal{S}_1 \mapsto \mathcal{S}_2$ that map resource free objects to resource free objects.

free maps. On the other hand, maps $\Lambda : \mathcal{S}_1 \mapsto \mathcal{S}_2$ such that $\Lambda(X) \notin \mathcal{V}_2$ for at least one $X \in \mathcal{V}_1$ are called resource creating maps. Free maps are typically those transformations that are easier to perform in practice.

The functions $\mathrm{R} : \mathcal{S}_1 \cup \mathcal{S}_2 \mapsto \mathbb{R}_{\geq 0}$ quantify the given resource. Hence, they give us a tool to go beyond the dichotomy of free and resource objects by telling us how useful a given object is (for a particular task). To do so, they have to fulfill two fundamental properties. First, they should assign zero value to free objects, i.e., $\mathrm{R}(F) = 0$ for any $F \in \mathcal{V}_i$. Second, they should be monotonous under free operations i.e., $\mathrm{R}(X) \geq \mathrm{R}(\Lambda(X))$ for any $\Lambda \in \mathscr{F}(\mathcal{S}_1, \mathcal{S}_2)$ and any $X \in \mathcal{S}_1$, which captures the fact that free operations cannot increase resources. In particular, this implies that $\mathrm{R}(\Lambda(F)) = 0$ for any $F \in \mathcal{V}_1$. We explicitly assumed that R is defined on both the domains $\mathcal{S}_1, \mathcal{S}_2$, however, in practice it is more convenient to restrict it to specific domain of R in each separate case. To simplify the notation, we consider in the following that $\mathcal{S} = \mathcal{S}_1 = \mathcal{S}_2$, hence $\mathcal{V} = \mathcal{V}_1 = \mathcal{V}_2$. Note that this is often assumed implicitly for resource theories in the literature. However, in practice, we can always achieve this situation by embedding vector spaces of smaller dimensions into higher-dimensional ones.

Let us formalize the above ingredients into definitions so we can use them throughout this thesis. In the remainder of this section, we use the term *resource theory* instead of QRT, as the following definitions are general and could, in principle, apply to other concepts outside of quantum theory.

Definition 3.1.1. *(*Resource theory for general sets*). Let $\mathcal{S}$ be a set with the subset $\mathcal{V} \subset \mathcal{S}$ and let $\mathscr{Q}$ be the set of all linear maps $\Lambda : \mathcal{S} \mapsto \mathcal{S}$. A resource theory for general sets is defined by the tuple $Q := (\mathcal{V}, \mathscr{F})$, where $\mathscr{F} \subset \mathscr{Q}$ is a subset of all maps $\mathscr{Q}$ such that:*

1. *The set $\mathscr{F}$ contains the identity map $\mathrm{id}_{\mathcal{S}}$ with $\mathrm{id}_{\mathcal{S}}(X) = X$ for any $X \in \mathcal{S}$.*
2. *For any two maps $\Lambda_1, \Lambda_2 \in \mathscr{F}$, their composition $\Lambda_1 \circ \Lambda_2$ is contained in $\mathscr{F}$.*
3. *For any $F \in \mathcal{V}$ and any $\Lambda \in \mathscr{F}$, it holds $\Lambda(F) \in \mathcal{V}$.*

The set $\mathcal{V}$ is the set of free objects *and $\mathscr{F}$ is called the set of* free operations. *The set $\mathcal{S}\backslash\mathcal{V}$ is the so-called set of* resource objects *and $\mathscr{Q}\backslash\mathscr{F}$ is the set of* resource creating operations.

Typically, we deal with the situation that a probabilistic mixture of any free resources is a free resource again. Hence, the set of free objects $\mathcal{V}$ is often convex. The same holds for a probabilistic mixture of free operations, which implies that also the set of free operations $\mathscr{F}$ is usually convex. We say that a resource theory is convex if the free set $\mathcal{V}$ and the set of free operations $\mathscr{F}$ are convex, i.e., $\eta F_1 + (1-\eta)F_2 \in \mathcal{V}$ for any $\eta \in [0,1]$ and any $F_1, F_2 \in \mathcal{V}$. Similarly, it has to hold $\eta\Lambda_1 + (1-\eta)\Lambda_2 \in \mathscr{F}$ for all free maps $\Lambda_1, \Lambda_2 \in \mathscr{F}$. Note that the free objects play a more important role than the set of free operations in this thesis. Therefore, we will also call any resource theory convex if the set $\mathcal{V}$ is convex and implicitly assume that we take the closure of the convex hull of $\mathscr{F}$ if it is not convex already.

Definition 3.1.1 captures two important points. First, doing nothing is always free. Hence, the identity map has to be a free operation. Second, one can employ free operations in any order and any number of times. They can never create resources from free objects. Condition $3.$ is sometimes referred to as the *golden rule* of QRTs [31] as it captures the main feature of resource theories, i.e., some operations cannot create resources from resource-free objects.

Instead of defining the set of free objects and free operations simultaneously, in some situations, it is desirable to define either the free operations or free objects first and adjust the respective counterpart accordingly. This leads to the *maximal set of free operations* and the *minimal set of free objects*.

Definition 3.1.2. *(*Maximal set of free operations*). Let $\mathcal{V} \subset \mathcal{S}$ be a set of free objects F. The* maximal set of free operations *$\mathscr{F}_{\max}$ consistent with $\mathcal{V}$ contains all maps $\Lambda : \mathcal{S} \mapsto \mathcal{S}$ such that*

$$\Lambda(F) \in \mathcal{V}(\mathcal{S}), \tag{3.83}$$

for any $F \in \mathcal{V}$. That is, $\mathscr{F}_{\max}$ is the largest set that contains free operations Λ, such that $Q = (\mathcal{V}, \mathscr{F}_{\max})$ is a resource theory.

Definition 3.1.3. *(*Minimal set of free objects*). Let $\mathscr{F}$ be a set of free operations. The associated minimal set of free objects $\mathcal{V}_{\min}$ is defined as*

$$\mathcal{V}_{\min} := \Big\{F \;:\; \forall\, X \in \mathcal{S}\;\exists\;\Lambda \in \mathscr{F} \text{ such that } F = \Lambda(X)\Big\}, \tag{3.84}$$

which means every $F \in \mathcal{V}_{\min}$ can be generated by operations in $\mathscr{F}$ from any other object $X \in \mathcal{S}$.

A fixed set $\mathscr{F}$ of free operations Λ imposes a *preorder* on the set $\mathcal{S}$. We write $X \xrightarrow{\mathscr{F}} Y$ if there exist a map $\Lambda \in \mathscr{F}$ such that $\Lambda(X) = Y$. In the case $X \xrightarrow{\mathscr{F}} Y$ and $Y \xrightarrow{\mathscr{F}} X$, we write $X \overset{\mathscr{F}}{\approx} Y$. If $X \xrightarrow{\mathscr{F}} Y$ holds, X is at least as resourceful as Y, since it can be transformed into it for free. To see that $\xrightarrow{\mathscr{F}}$ is a preorder, note that $X \xrightarrow{\mathscr{F}} Y$ and $Y \xrightarrow{\mathscr{F}} Z$ clearly imply $X \xrightarrow{\mathscr{F}} Z$. We can use the notion of a preorder to establish more properties of the set of minimal free states $\mathcal{V}_{\min}$.

First, note that for every other set of free objects $\mathcal{V}$ consistent with the free operations $\mathscr{F}$ such that $Q = (\mathcal{V}, \mathscr{F})$ is a resource theory, it holds $\mathcal{V}_{\min} \subset \mathcal{V}$. This shows that $\mathcal{V}_{\min}$ is indeed the minimal set of free objects. To see this, note that for every $F \in \mathcal{V}$ and any $G \in \mathcal{V}_{\min}$ it holds by definition $F \xrightarrow{\mathscr{F}} G$, which implies that $G \in \mathcal{V}$, since G can be obtained from a free object and free operations. Next, we want to see in which situation $\mathcal{V}_{\min}$ is the only set of free objects consistent with the free operations $\mathscr{F}$. This is the case, when any two free objects $F, G \in \mathcal{V}$ can be transformed into each other, i.e., $F \overset{\mathscr{F}}{\approx} G$. The statement follows from $F \xrightarrow{\mathscr{F}} G$ for any $F \in \mathcal{V}_{\min}$ and $G \in \mathcal{V}$ and the fact that $X \xrightarrow{\mathscr{F}} F$ for any $X \in \mathcal{S}$ by definition. Therefore, $X \xrightarrow{\mathscr{F}} G$, which implies $G \in \mathcal{V}_{\min}$.

3.1.1 Quantification

Definition 3.1.4. *(*Resource monotones*). Let $Q = (\mathcal{V}, \mathscr{F})$ be a resource theory of objects $\mathcal{S}$ and let $\mathrm{R} : \mathcal{S} \mapsto \mathbb{R}_{\geq 0}$ be a function from the set $\mathcal{S}$ to the non-negative numbers. Any function R that obeys the following two conditions is a* resource monotone *for the resource theory associated to Q.*

1. *Vanishing for free objects:* $F \in \mathcal{V} \implies \mathrm{R}(F) = 0$.

2. *Monotonicity:* $\mathrm{R}(X) \geq \mathrm{R}(\Lambda(X))$ *for any $X \in \mathcal{S}$ and any free map $\Lambda \in \mathscr{F}$.*

Moreover, a resource monotone R is said to be faithful *if $\mathrm{R}(X) = 0 \iff X \in \mathcal{V}$, i.e., it is zero if and only if it is evaluated on a resource free object. We will use the terms resource monotone and* resource quantifier *interchangeably.*

Note that $\mathrm{R}(X) > \mathrm{R}(Y)$ directly implies that there is no free transformation $\Lambda \in \mathscr{F}$ such that $X = \Lambda(Y)$. However, it does not generally hold that $\mathrm{R}(X) \geq \mathrm{R}(Y)$ implies $X \xrightarrow{\mathscr{F}} Y$.

A resource monotone $\mathrm{R} : \mathcal{S} \mapsto \mathbb{R}_{\geq 0}$ is said to be a *convex resource monotone* if it holds that

$$\mathrm{R}\Big(\sum_i p_i X_i\Big) \leq \sum_i p_i \mathrm{R}(X_i), \tag{3.85}$$

for any ensemble of objects $X_i \in \mathcal{S}$ and any probability vector $\mathbf{p}$. Note that we implicitly assumed here that $\mathcal{S}$ is a convex set, i.e., $\sum_i p_i X_i \in \mathcal{S}$.

There are several classes of relevant resource monotones. The first class are resource monotones based on *contractive distances*, i.e., distance functions $\mathrm{D}(X, Y)$ between elements $X, Y \in \mathcal{S}$ that satisfy

$$\mathrm{D}(X, Y) \geq \mathrm{D}(\Lambda(X), \Lambda(Y)), \tag{3.86}$$

for any allowed map $\Lambda \in \mathscr{Q}$. An example of such a distance is the trace distance between quantum states, that is contractive under all CPTP maps (see Eq. (2.57)).

Theorem 3.1.5. *(*Distance based resource monotones*). Let $Q = (\mathcal{V}, \mathscr{F})$ be a resource theory and let $\mathrm{D} : \mathcal{S} \times \mathcal{S} \mapsto \mathbb{R}_{\geq 0}$ be a contractive distance. It follows that*

$$\mathrm{R}_{\text{dist}}(X) := \mathrm{D}(X, \mathcal{V}) = \inf_{F \in \mathcal{V}} \mathrm{D}(X, F), \tag{3.87}$$

is a faithful resource monotone for the resource theory Q induced by the distance D.

Proof. The faithfulness follows trivially. The proof of the monotonicity follows directly from the contractiveness of the distance D under any map $\Lambda \in \mathscr{Q}$ and the fact that set $\mathcal{V}$ is closed under free transformations. Namely,

$$\begin{aligned}\mathrm{R}_{\text{dist}}(X) = \inf_{F \in \mathcal{V}} \mathrm{D}(X, F) &\geq \inf_{F \in \mathcal{V}} \mathrm{D}(\Lambda(X), \Lambda(F)) \\ &\geq \inf_{F' \in \mathcal{V}} \mathrm{D}(\Lambda(X), F') = \mathrm{R}_{\text{dist}}(\Lambda(X)).\end{aligned} \tag{3.88}$$

□

Note that the infimum can be replaced by a minimum whenever $\mathcal{V}$ is a compact set. Furthermore, it follows directly that R_{dist} is a convex function whenever $\mathcal{V}$ is convex and the underlying distance $\mathrm{D}(X, Y)$ is *jointly convex*, i.e., it holds

$$\mathrm{D}(pX_1 + (1-p)X_2, pY_1 + (1-p)Y_2) \leq p\mathrm{D}(X_1, Y_1) + (1-p)\mathrm{D}(X_2, Y_2), \tag{3.89}$$

for any $p \in [0, 1]$ and any $X_1, X_2, Y_1, Y_2 \in \mathcal{S}$. In particular, all distances based on norms are jointly convex, as norms are convex functions.

As we did not use specifically all properties of a distance, we can extend the above statement to all faithful and contractive functions. This includes functions that are not a metric, like the relative entropy (see Eq. (2.75)) in the case of quantum states.

The next resource quantifier is based on robustnesses with respect to a set. We focus here on the generalized robustness, the most commonly used robustness.

Theorem 3.1.6. *(*Robustness based resource monotones*). Let $Q = (\mathcal{V}, \mathscr{F})$ be a resource theory. It follows that*

$$\mathrm{R}_{\mathrm{rob}}(X) := \inf_{Y \in \mathcal{S}, r} \Big\{ r \geq 0 : \frac{X + rY}{1 + r} \in \mathcal{V} \Big\}. \tag{3.90}$$

is a faithful resource monotone for the resource theory Q.

Proof. The faithfulness follows again trivially. The proof of the monotonicity follows directly from the fact that we can apply any linear map $\Lambda \in \mathscr{F}$ onto the left hand side in the definition of $\mathrm{R}_{\mathrm{rob}}(X)$. Therefore,

$$\frac{\Lambda(X) + r\Lambda(Y)}{1 + r} = \frac{\Lambda(X) + rY'}{1 + r} \in \mathcal{V}, \tag{3.91}$$

is a valid convex combination of $\Lambda(X)$ and an other element $Y' \in \mathcal{S}$ into a free resource with the robustness $\mathrm{R}_{\mathrm{rob}}(X) = r$. However, this convex combination does not need to be optimal, which implies $\mathrm{R}_{\mathrm{rob}}(X) \geq \mathrm{R}_{\mathrm{rob}}(\Lambda(X))$. Note that by following the ideas in [88], it can also be shown that $\mathrm{R}_{\mathrm{rob}}(X)$ is convex, if and only if $\mathcal{V}$ is convex. □

Finally, we can define a resource monotone based on the weight decompositions.

Theorem 3.1.7. *(*Weight based resource monotones*). Let $Q = (\mathcal{V}, \mathscr{F})$ be a resource theory. It follows that*

$$\mathrm{R}_{\mathrm{weight}}(X) := \inf_{F \in \mathcal{V}, Y \in \mathcal{S}, w} \Big\{ w \geq 0 : X = wY + (1 - w)F \Big\}, \tag{3.92}$$

is a faithful resource monotone for the resource theory Q.

Proof. The faithfulness holds trivially. The proof of the monotonicity follows from the fact that $\Lambda(X)$ admits the decomposition

$$\Lambda(X) = w\Lambda(Y) + (1 - w)\Lambda(F) = wY' + (1 - w)F' \tag{3.93}$$

with $Y' \in \mathcal{S}$ and $F' \in \mathcal{V}$. Since this decomposition for $\Lambda(X)$ does not need to be optimal, it follows $\mathrm{R}_{\mathrm{weight}}(X) \geq \mathrm{R}_{\mathrm{weight}}(\Lambda(X))$. By adapting the ideas from [34] one can show that $\mathrm{R}_{\mathrm{weight}}(X)$ is convex as well for convex sets $\mathcal{V}$. □

3.2 Resource Theories for Quantum States

Having the general structures and all the above resource monotones at hand, we are in the position to study the most relevant QRTs within the context of this thesis. We start by looking at resource theories for quantum states.

Here, the set $\mathcal{S}$ of all objects is replaced by the set $\mathcal{S}(\mathcal{H})$ of all density matrices ρ, where we denote by d the dimension of $\mathcal{H}$. Furthermore, the set $\mathscr{Q}$ now denotes the set of all CPTP maps from $\mathcal{S}(\mathcal{H})$ to itself.

3.2.1 Purity

The first resource theory of quantum states we look at is the resource theory of purity, which is probably the simplest resource theory for quantum states. In a certain sense, that we justify later (see Section 3.2.4), it is also the most fundamental resource of a quantum state. The resource theory of purity, which originated from studying resources in the context of thermodynamics [89–94], follows the simple idea that pure states are more valuable than mixed states in quantum information processing tasks. We already saw in Sections 2.3 and 2.8 that mixed states come with an inherent drawback, that is, it is only probabilistically determined in which pure state our system is in. The resource theory of purity captures this idea and aims at quantifying how *non-mixed* a certain quantum state is. Within this section, we follow the same lines as [61] and we also refer to [60] for a detailed overview.

In the resource theory of purity, the set of free states is particularly simple, as it only contains a single state. More precisely, the only free state is the maximally mixed state

$$\mathcal{V}_{\text{Purity}} := \frac{\mathbb{1}}{d}. \tag{3.94}$$

It is now easy to identify the *maximal set of free operations* (see Definition 3.1.2) as the set of all unital CPTP maps. This means $\mathscr{F}_{\text{max}} = \mathscr{F}_{\text{U}}$ contains all quantum channels that fulfill

$$\Lambda_U\Big(\frac{\mathbb{1}}{d}\Big) = \frac{\mathbb{1}}{d}. \tag{3.95}$$

Note that $\mathscr{F}_{\text{U}}$ includes two frequently used subsets of free operations. First, the set $\mathscr{F}_{\text{NO}}$ of *noisy operations*, whose elements are given by channels of the form

$$\Lambda_{\text{NO}}(\rho) = \text{Tr}_2[U\Big(\rho \otimes \frac{\mathbb{1}_E}{d_E}\Big)U^\dagger], \tag{3.96}$$

for any unitary U and any maximally mixed state $\frac{\mathbb{1}_E}{d_E}$ on a d_E-dimensional ancilla (environment) space $\mathcal{H}_E$. Second, the set $\mathscr{F}_{\mathrm{MU}}$ of channels that can be written as a mixture of unitaries, i.e.,

$$\Lambda_{\mathrm{MU}}(\rho) = \sum_i p_i U_i \rho U_i^\dagger, \tag{3.97}$$

with $\mathbf{p}$ being a probability vector and the U_i being unitary operators.

It can be shown that the (strict) set inclusion

$$\mathscr{F}_{\mathrm{MU}} \subset \mathscr{F}_{\mathrm{NO}} \subset \mathscr{F}_{\mathrm{U}} \tag{3.98}$$

holds. However, interestingly, the state conversion abilities of all sets are equivalent, in the sense that $\rho \xrightarrow{\mathscr{F}_{\mathrm{U}}} \sigma$ implies $\rho \xrightarrow{\mathscr{F}_{\mathrm{MU}}} \sigma$. The other relations follow trivially. For more details, see [60].

Whether a state transformation $\rho \xrightarrow{\mathscr{F}_{\mathrm{U}}} \sigma$ is possible can be elegantly determined through the majorization (see, e.g., [70]) of quantum states. This is a consequence of the fact that unital channels are a generalization of classical bistochastic maps. A quantum state ρ is said to majorize the state σ if

$$\sum_{i=0}^{k} \lambda_i^{\downarrow}(\rho) \geq \sum_{i=0}^{k} \lambda_j^{\downarrow}(\sigma), \tag{3.99}$$

for all $0 \leq k \leq d-1$, where $\lambda_i^{\downarrow}(\rho)$ is the i-th largest eigenvalue of ρ. We use the shorthand notation $\rho \succeq_m \sigma$ to signify that ρ majorizes σ.

Theorem 3.2.1. *(*State transformation and majorization*). Let $\rho, \sigma \in \mathcal{S}(\mathcal{H})$ be two quantum states. The state ρ can be transformed to the state σ by a unital CPTP map Λ_U if and only if ρ majorizes σ, i.e, $\rho \succeq_m \sigma$.*

Proof. The proof can be found in [95] and we also refer to [82] (Theorem 4.32 therein). $\square$

According to [60] and [61], a resource quantifier P in the resource theory $Q_{\mathrm{Purity}} = (\mathcal{V}_{\mathrm{Purity}}, \mathscr{F}_{\mathrm{U}})$ of purity should fulfill the following four conditions, of which the first two are the typical (faithful) monotone requirements:

(P1) Faithfulness: $\mathrm{P}(\rho) \geq 0$ for any $\rho \in \mathcal{S}(\mathcal{H})$ and $\mathrm{P}(\rho) = 0 \iff \rho = \dfrac{\mathbb{1}}{d}$.

(P2) Monotonicity: $\mathrm{P}(\rho) \geq \mathrm{P}(\Lambda_U(\rho))$ for any $\Lambda_U \in \mathscr{F}_{\mathrm{U}}$ and any $\rho \in \mathcal{S}(\mathcal{H})$.

(P3) Additivity: $\mathrm{P}(\rho \otimes \sigma) = \mathrm{P}(\rho) + \mathrm{P}(\sigma)$ for any $\rho, \sigma \in \mathcal{S}(\mathcal{H})$.

(P4) Normalization: $\mathrm{P}(|\Psi\rangle\langle\Psi|) = \log_2(d)$ for any $|\Psi\rangle \in \mathcal{H}$.

By design, any of the general resource monotones introduced in Section 3.1.1 fulfill the conditions (P1) and (P2). However, these monotones do not fulfill the conditions (P3) and (P4) in general. Note that these conditions could be seen as conditions coming more from a mathematical than a physical point of view. An important family of resource quantifiers that fulfills all four conditions is given by the *Rényi α-purities* (see [61] and Section 2.8.4)

$$\mathrm{P}_\alpha(\rho) = \log_2(d) - \mathrm{S}_\alpha(\rho), \tag{3.100}$$

where the $\mathrm{S}_\alpha(\rho)$ are the Rényi α-entropies as defined in Eq. (2.71). Note that in the case $\lim_{\alpha\to 1} \mathrm{P}_\alpha(\rho) = \log_2(d) - \mathrm{S}(\rho) = \mathrm{S}(\rho\|\frac{\mathbb{1}}{d})$ we obtain the relative entropy of purity, where $\mathrm{S}(\rho)$ denotes the von Neumann Entropy (see Definition 2.8.2) of ρ. Additionally, $\mathrm{P}_\alpha(\rho)$ is convex for $0 \le \alpha \le 1$. Besides the case of the relative entropy, two additional Rényi purities are of particular importance. First, the case $\alpha = 2$ leads to the Rényi 2-purity

$$\mathrm{P}_2(\rho) = \log_2(d\,\mathrm{Tr}[\rho^2]), \tag{3.101}$$

which is a monotonic function of the *linear purity* $\mathrm{Tr}[\rho^2]$. We want to emphasize here that the term *(linear) purity*, often used for the expression $\mathrm{Tr}[\rho^2]$ in the literature, is misleading as it does not fulfill the requirements $(\mathrm{P1}) - (\mathrm{P4})$. Furthermore, it also does not capture an important point. The dimension of a quantum state is part of its purity resource. That is, any pure state fulfills $\mathrm{Tr}[\rho^2] = 1$ but higher dimensional pure states are generally more useful than lower dimensional states, which is exactly captured by the Rényi α-purities in Eq. (3.100).

The final important instance is the case

$$\lim_{\alpha\to\infty} \mathrm{P}_\alpha(\rho) = \log_2(d\lambda_1(\rho)), \tag{3.102}$$

where $\lambda_1(\rho)$ is the largest eigenvalue of ρ. It is noteworthy that the case $\lim_{\alpha\to\infty} \mathrm{P}_\alpha(\rho)$ is particularly easy to study and compute, since it does only depend on a single eigenvalue of ρ. Furthermore, it was shown in [96] that $\lambda_1(\rho)$ also completely determines the generalized robustness of purity (see Definition 3.1.6) as $\mathrm{P}_{\mathrm{rob}}(\rho) = d\lambda_1(\rho) - 1$. We use the fact that there exists a purity quantifier that only depends on $\lambda_1(\rho)$ in Section 4.2, respectively in Publication B.

3.2.2 Coherence

The purity of a quantum state can hardly count as a genuine quantum property. It is more like a classical resource a quantum state has to have to be useful. Quantum coherence [35], on the other hand, is a phenomenon that divides quantum and classical physics. Quantum coherence describes the effect, respectively the possibility, that quantum states can be in a superposition of basis states. Originating from the superposition principle and being responsible for effects like interference, quantum coherence has a longstanding research history, starting from the coherence of optical fields [97–99]. The modern day notion of coherence, commonly used in quantum information theory today, was first introduced by Åberg [100] before it was further developed in [101]. Nowadays, quantum coherence represents a mature research field which attracts a lot of attention. In the following, we will follow the notions of [35] and [101]. Moreover, we generally refer to the review [35] for more details.

The resource theory of coherence is inherently basis-dependent. Typically, one chooses the energy-eigenbasis of a given Hamiltonian, the eigenbasis of a specific observable, or simply the computational basis. Basis states of the fixed basis $\{|i\rangle\}_{i=0}^{d-1}$ are considered as free, while superpositions of basis states are considered resourceful. More generally, including mixed states, the set of free or incoherent states is given by

$$\mathcal{V}_{\text{Coherence}} := \Big\{\rho_I = \sum_{i=0}^{d-1} p_i |i\rangle\langle i|\Big\}, \tag{3.103}$$

where $\mathbf{p}$ is a probability vector. Clearly, $\mathcal{V}_{\text{Coherence}}$ is convex by design. An alternative definition of the free set can be given via the *dephasing operation*

$$\Delta(\rho) = \sum_{i=0}^{d-1} \langle i|\rho|i\rangle |i\rangle\langle i|, \tag{3.104}$$

which destroys the coherence of any quantum state and acts on an incoherent state such that $\Delta(\rho_I) = \rho_I$. In general, a resource theory with a CPTP map that maps any quantum state to a free state, while keeping any free state invariant is called a resource theory with a *resource destroying map* [31].

Several classes of free operations are typically considered, each motivated by a different physical context. Formally, one can define the maximal set of free operations, called the set of *maximally incoherent operations (MIO)* as the set containing all channels $\Lambda_{\text{MIO}} \in \mathscr{Q}$ that fulfill

$$\Lambda_{\text{MIO}}(\rho_I) \in \mathcal{V}_{\text{Coherence}} \tag{3.105}$$

for any $\rho_I \in \mathcal{V}_{\text{Coherence}}$. However, this set of operations is not very well characterized, and in many cases, the more relevant set of free operations is the set $\mathscr{F}_{\text{IO}}$ of *incoherent operations (IO)*. The set $\mathscr{F}_{\text{IO}}$ contains all CPTP maps $\Lambda_{\text{IO}}(\rho) =$

$\sum_a \Lambda_a(\rho) = \sum_a K_a \rho K_a^\dagger$ that can be implemented via an instrument $\{\Lambda_a\}$ such that each of the subchannels Λ_a is incoherent in the sense that

$$\frac{K_a \rho_I K_a^\dagger}{\mathrm{Tr}[K_a \rho_I K_a^\dagger]} \in \mathcal{V}_{\mathrm{Coherence}}. \tag{3.106}$$

That means IO cannot create coherence, even in a probabilistic sense. Alternatively, we can say that IO are those operations which can be written in terms of *incoherent Kraus operators* K_a, such that $K_a|i\rangle \sim |j\rangle$, where $|i\rangle$ and $|j\rangle$ are (incoherent) basis states.

A third relevant set $\mathscr{F}_{\mathrm{SIO}}$ of free operations contains the *strictly incoherent operations (SIO)*. These are given by those CPTP maps that can be implemented via instruments with incoherent Kraus operators K_a such that $K_a^\dagger$ is an incoherent Kraus operator as well. For further relevant classes of free operations see [35].

It can be shown [102, 103] that the set of free operations obey the strict set inclusions

$$\mathscr{F}_{\mathrm{SIO}} \subset \mathscr{F}_{\mathrm{IO}} \subset \mathscr{F}_{\mathrm{MIO}}. \tag{3.107}$$

The state conversion abilities of coherent states under free operations are less clear than in the case of purity. We refer, however, to the notable results in [104], where the possible transformations between pure states are characterized in terms of a majorization criterion, and the results in [105], where the state transformation abilities for single-qubit states are solved entirely. Furthermore, it was shown in [101] that states of the form

$$|\Psi\rangle = \frac{1}{\sqrt{d}} \sum_{j=0}^{d-1} \exp\left(i\varphi_j\right)|j\rangle, \tag{3.108}$$

can be converted to any other d-dimensional state ρ via IO (and therefore also via MIO). Therefore, states of the form in Eq. (3.108) are called *maximally coherent states*. The notion of a maximally coherent state was generalized in [61] to mixed states for a fixed spectrum. To understand what a maximally coherent mixed state is, we must understand how coherence is typically quantified.

According to [100], respectively [101], a coherence quantifier C in the resource theory $Q_{\mathrm{Coherence}} = (\mathcal{V}_{\mathrm{Coherence}}, \mathscr{F}_{\mathrm{IO}})$ of coherence, should fulfill the following four conditions. The first two are our familiar (faithful) monotone requirements:

(C1) Faithfulness: $\mathrm{C}(\rho) \geq 0$ for any $\rho \in \mathcal{S}(\mathcal{H})$ and $\mathrm{C}(\rho) = 0 \iff \rho \in \mathcal{V}_{\mathrm{Coherence}}$.

(C2) Monotonicity: $\mathrm{C}(\rho) \geq \mathrm{C}(\Lambda_{\mathrm{IO}}(\rho))$ for any $\Lambda_{\mathrm{IO}} \in \mathscr{F}_{\mathrm{IO}}$ and any $\rho \in \mathcal{S}(\mathcal{H})$.

(C3) Strong monotonicity: $\mathrm{C}(\rho) \geq \sum_i p_i \mathrm{C}(\sigma_i)$ for any $\sigma_i = \frac{K_i \rho K_i^\dagger}{p_i}$ with $p_i = \mathrm{Tr}[K_i \rho K_i^\dagger]$ and incoherent Kraus operators $\{K_i\}$ and any $\rho \in \mathcal{S}(\mathcal{H})$.

(C4) Convexity: $\mathrm{C}\Big(\sum_i q_i \rho_i\Big) \leq \sum_i q_i \mathrm{C}(\rho_i)$ for any probability vector $\mathbf{q}$ and any $\rho_i \in \mathcal{S}(\mathcal{H})$.

Note that convexity is typically a requirement for a quantifier in the resource theory of coherence, contrary to the resource theory of purity, where convexity is typically regarded as an additional feature. More importantly, in the resource theory of coherence one requires typically that a quantifier behaves also monotonic on average under the selective IO as captured by condition $(\mathrm{C3})$. In [35], two additional requirements are proposed: a uniqueness condition for pure states and the additivity under tensor products. Finally, note that the conditions $(\mathrm{C2})$ and $(\mathrm{C3})$ can trivially be adapted to other sets of incoherent operations.

Clearly all resource monotones introduced in Section 3.1.1 fulfill the conditions $(\mathrm{C1})$, $(\mathrm{C2})$, and $(\mathrm{C4})$ by design. However, not all of them will fulfill condition $(\mathrm{C3})$. In particular, the coherence monotone based on the trace distance does not fulfill condition $(\mathrm{C3})$, as shown in [106]. However, it was shown in [107] that the generalized robustness of coherence does indeed fulfill the conditions $(\mathrm{C1}) - (\mathrm{C4})$ for the class of IO operations. In the following, we discuss two important coherence monotones in more detail. Both of them obey the conditions $(\mathrm{C1}) - (\mathrm{C4})$ and admit a closed form expression which does not require any optimization [35].

The first quantifier is defined via the entrywise ℓ_1-distance between a quantum state ρ and the incoherent states $\rho_I \in \mathcal{V}_{\mathrm{Coherence}}$ and it is given by

$$\mathrm{C}_{\ell_1}(\rho) = \min_{\rho_I \in \mathcal{V}_{\mathrm{Coherence}}} \|\rho - \rho_I\|_{\ell_1} = \sum_{i \neq j} |\rho_{ij}|, \tag{3.109}$$

where the closed form expression on the right-hand-side was shown in [101]. This makes the ℓ_1-distance of coherence not only particularly easy to compute but also reflects the intuitive view that the magnitude of the off-diagonal elements determines a state's coherence. Interestingly, this intuitive quantifier violates the monotonicity condition $(\mathrm{C2})$ for the class of MIO.

The second quantifier is based on the relative entropy and is given by

$$\mathrm{C}_{\mathrm{rel}}(\rho) = \min_{\rho_I \in \mathcal{V}_{\mathrm{Coherence}}} \mathrm{S}(\rho \| \rho_I) = \mathrm{S}(\Delta(\rho)) - \mathrm{S}(\rho), \tag{3.110}$$

where the closed form expression on the right-hand-side, in terms of the von-Neumann entropy $\mathrm{S}(\rho)$, was also shown in [101]. The relative entropy of coherence also obeys the condition $(\mathrm{C2})$ for MIO. Moreover, it coincides with the *distillable coherence* [35], hence it has a well-defined operational meaning.

Coming back to the question of what a maximally coherent mixed state is, it was shown in [61] that for a fixed spectrum $\lambda(\rho)$ the state

$$\rho_{\max} = \sum_{\tilde{n}=0}^{d-1} \lambda_{\tilde{n}} |\tilde{n}\rangle\langle\tilde{n}|, \tag{3.111}$$

maximizes any coherence monotone (under MIO), where $\{|\tilde{n}\rangle\}$ is a mutually unbiased basis to the incoherent basis $\{|i\rangle\}$, i.e., $|\langle i|\tilde{n}\rangle| = \frac{1}{\sqrt{d}}$ for all $i, \tilde{n}$ (see also Definition 2.1.7). Note that the maximally coherent pure states in Eq. 3.108 are exactly defined such that they are diagonal in a basis that is mutually unbiased to $\{|i\rangle\}$.

Finally, let us comment on the coherence of multipartite states. Coherence is inherently a local property and can straightforwardly be defined for a multipartite state by treating the multiple particles as one system. However, the more exciting scenario occurs if one considers the coherence of a multipartite state with respect to an incoherent product basis. Namely, a bipartite incoherent state can be written in the form

$$\rho_I = \sum_i p_i |i_A\rangle\langle i_A| \otimes |i_B\rangle\langle i_B|, \tag{3.112}$$

which can directly be generalized to n-partite systems. The coherence of multipartite states with respect to product bases was first studied in [108, 109] and was also further investigated in [61, 110]. In our Publication B, we deal with a specific case of coherence with respect to incoherent product bases. We consider the minimal coherence with respect to all possible product bases, which is equivalent [61, 111] to

$$\mathrm{C}_{\min}(\rho) = \min_{U_A \otimes U_B} \mathrm{C}((U_A \otimes U_B)\rho(U_A \otimes U_B)^\dagger), \tag{3.113}$$

where U_A, U_B are unitaries on system A and system B, respectively. This introduces not only a basis independent form of quantum coherence but also coincides with the symmetric discord [36, 112] of a quantum state ρ. Quantum discord [38, 113, 114] was introduced as a measure of quantum correlations in addition to entanglement, which can be non-zero for separable states and quantifies the difference between the total and classical correlations in a quantum state. For a detailed discussion and review on discord, see [36, 112]. While discord can be viewed as a resource for certain quantum information processing tasks, it also has certain drawbacks. In particular, the set of zero-discord states is non-convex as it contains (in the case of symmetric discord) all states of the form

$$\rho_{(c-c)} = \sum_i p_i |\phi_i\rangle\langle\phi_i| \otimes |\psi_i\rangle\langle\psi_i|, \tag{3.114}$$

with $\langle\phi_i|\phi_j\rangle = \delta_{ij}$ and similarly $\langle\psi_i|\psi_j\rangle = \delta_{ij}$. These states are also known as *classical-classical states*. Moreover, the set of free operations is not well characterized for discord. So far, it is only known that discord is invariant under local unitary operations, which follows in the case of symmetric discord directly from Eq. (3.113).

3.2.3 Entanglement

> *I would not call that one but rather the characteristic trait of quantum mechanics, the one that enforces its entire departure from classical lines of thought.*
>
> — **Erwin Schrödinger**
> About the role of entanglement.

Entanglement [32] is by far one of the most puzzling and, at the same time, astonishing features of nature. While it led to many philosophical questions in the early days of quantum theory, it was later realized that there are specific quantum information processing tasks in which entanglement plays a crucial role. In fact, it was realized that entanglement is necessary for certain cryptographic [8, 25], communication [7, 115] and computation tasks [30, 116]. Furthermore, it is a key ingredient for the existence of quantum correlations like EPR-steering [44, 45] and Bell nonlocality [43, 117] (see also Section 3.4). All these findings led to the development of the resource theory of entanglement [118, 119] as the first application of resource theories to quantum theory. Nowadays, it is the prime example of the field of quantum resource theories. Here, we give an overview of the essential concepts from entanglement theory relevant to this thesis. For a review and more details on entanglement, see [32, 120, 121].

We already formally defined the concept of entanglement in Section 2.4 (see Definition 2.4.2 and 2.4.3) which we repeat here for convenience. A pure bipartite state $|\Psi\rangle \in \mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B$ is called separable if it can be written as $|\Psi\rangle = |\Psi_A\rangle \otimes |\Psi_B\rangle$ and it is called entangled otherwise. The canonical example for entangled pure states are the two-qubit *Bell states*

$$|\Phi^{\pm}\rangle = \frac{1}{\sqrt{2}}(|00\rangle \pm |11\rangle), \quad |\Psi^{\pm}\rangle = \frac{1}{\sqrt{2}}(|01\rangle \pm |10\rangle). \tag{3.115}$$

We have already established a method to check whether a pure state $|\Psi\rangle$ is entangled through the Schmidt decomposition (see Theorem 2.4.1). Namely, by decomposing a pure state such that

$$|\psi\rangle = \sum_{i=0}^{R-1} \sqrt{c_i}|i_A i_B\rangle, \tag{3.116}$$

one can identify separable pure states as those of having a Schmidt-rank of $R = 1$. Entangled states, on the other hand, have a Schmidt-rank $R \geq 2$. As the Schmidt decomposition relies entirely on the sigular value decomposition, the problem of deciding whether a pure state is entangled or separable can be solved efficiently (in the Hilbert space dimension). Note that quantum states of Schmidt rank $R = d$, (where $d = \dim(\mathcal{H}_A) = \dim(\mathcal{H}_B)$ in the following) with $\sqrt{c}_i = \frac{1}{\sqrt{d}}\ \forall\ i$ are called *maximally entangled* states. We will justify this name later within this section (see the discussion of Theorem 3.2.4). The Bell states in Eq. (3.115) are examples for maximally entangled states.

A general mixed state $\rho \in \mathcal{S}(\mathcal{H}_{AB})$ is separable if it can be written as

$$\rho = \sum_i p_i \rho_{A_i} \otimes \rho_{B_i}, \tag{3.117}$$

and it is called entangled otherwise. From here on, we denote by $\mathrm{Sep}(\mathcal{H}_{AB}) \subset \mathcal{S}(\mathcal{H}_{AB})$ the set of all separable states on $\mathcal{H}_{AB}$. To decide whether a given mixed state ρ is entangled or separable is much harder than for pure states. In fact, it has been shown to be NP-hard [122].

There are, however, a variety of sufficient entanglement criteria (for a detailed review, see [32, 121]), out of which the positive partial transpose (PPT) criterion [123] is probably the most famous one. The PPT criterion is a powerful and easy-to-compute criterion that relies on the fact that separable quantum states are mapped to (separable) quantum states when the transposition map is applied only to one of the subsystems of ρ. More precisely, the partial transpose ρ^{T_B} (with respect to Bob's, respectively the second subsystem) of a general quantum state $\rho = \sum_{i,j,k,l} c_{ij,kl}|i\rangle\langle j| \otimes |k\rangle\langle l|$ is defined as

$$\begin{aligned}\rho^{T_B} = (\mathbb{1} \otimes T)(\rho) &\coloneqq \sum_{i,j,k,l} c_{ij,kl}|i\rangle\langle j| \otimes (|k\rangle\langle l|)^T \\ &= \sum_{i,j,k,l} c_{ij,kl}|i\rangle\langle j| \otimes |l\rangle\langle k| = \sum_{i,j,k,l} c_{ij,lk}|i\rangle\langle j| \otimes |k\rangle\langle l|,\end{aligned} \tag{3.118}$$

where $(\cdot)^T$ denotes the usual transposition. This means that the partial transposition acts on the basis element $|i\rangle\langle j| \otimes |k\rangle\langle l|$ such that $(|i\rangle\langle j| \otimes |k\rangle\langle l|)^{T_B} = |i\rangle\langle j| \otimes |l\rangle\langle k|$. Note that all definitions can trivially be adapted to Alice's (the first) subsystem. Furthermore, we want to emphasize that the form of ρ^{T_B} depends on the basis we transpose in, however, the eigenvalues of ρ^{T_B} do not depend on the particular basis. The PPT criterion now states the following:

Theorem 3.2.2. *(*PPT criterion*). Let $\rho \in \mathcal{S}(\mathcal{H}_{AB})$ be a quantum state. If $\rho \in \mathrm{Sep}(\mathcal{H}_{AB})$ is a separable state, it follows that $\rho^{T_B} \succeq 0$.*

Proof. The proof [123] follows directly from inspection of

$$\rho^{T_B} = \sum_i p_i \rho_{A_i} \otimes (\rho_{B_i})^T, \tag{3.119}$$

which has to be non-negative since $(\rho_{B_i})^T$ is in itself a density operator. Note that $\rho^{T_B} \succeq 0 \implies \rho^{T_A} \succeq 0$, which follows by simply transposing ρ^{T_B} (with respect to both subsystems). □

The power of the PPT criterion stems from the fact that $\rho^{T_B} \not\succeq 0$ for some entangled states. That means there is at least one negative eigenvalue of ρ^{T_B}. In fact, violation of the PPT criterion is even necessary for entanglement in systems with dimensions d_A, d_B such that $d_A d_B \leq 6$ [124]. However, in higher dimensions, there exist entangled quantum states with PPT [125] (see also down below).

The underlying reason why the PPT criterion is able to detect entanglement is that the transposition is a positive map but no CP map (see Definition 2.5.1). As shown in [124], the separability of a quantum state ρ can be rephrased in the sense that

$$\rho \in \mathrm{Sep}(\mathcal{H}_{AB}) \iff (\mathbb{1}_A \otimes \Lambda)(\rho) \succeq 0, \tag{3.120}$$

for all positive maps Λ.

While entanglement criteria based on positive but not CP maps are beneficial for detecting entanglement from a mathematical point of view, they are problematic from a practical or experimental view. The problem with these entanglement criteria is that they rely on non-physical transformations. Therefore, they are hard to exploit experimentally, as one typically relies on state tomography. There is, however, a related way to detect entanglement, which is also highly usable from an experimental perspective. That leads us to the entanglement detection via entanglement witnesses [124, 126].

Definition 3.2.3. *(*Entanglement witness*). Let $W \in \mathrm{Herm}(\mathcal{H}_{AB})$ be a Hermitian operator. It is called an* entanglement witness *if it fulfills the following two conditions:*

1. $\mathrm{Tr}[W\rho_{\mathrm{sep}}] \geq 0$ *for all separable states* $\rho_{\mathrm{sep}} \in \mathrm{Sep}(\mathcal{H}_{AB})$.

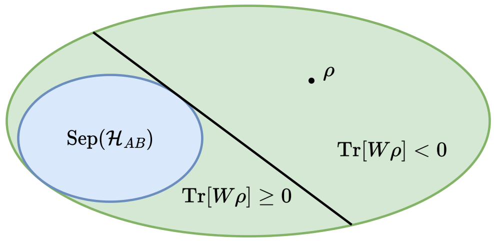

Fig. 3.2.: Geometrical representation of an entanglement witness. The entanglement witness W divides the space of all quantum states into two half-spaces. While all separable quantum states $\rho_{\text{Sep}} \in \text{Sep}(\mathcal{H}_{AB})$ lead to $\mathrm{Tr}[W\rho_{\text{Sep}}] \geq 0$ there exists an entangled state ρ such that $\mathrm{Tr}[W\rho] < 0$, which means W detects the entanglement of ρ.

2. $\mathrm{Tr}[W\rho] < 0$ *for (at least) one state* $\rho \in \mathcal{S}(\mathcal{H}_{AB})$.

The concept of entanglement witnesses is based on the *Hahn-Banach*, respectively the *separating hyperplane* theorem (see, e.g., [87]). The separating hyperplane theorem states that for any two disjoint convex and compact sets, there exists a hyperplane (actually even two parallel ones separated by a non-zero gap, see, e.g., [127]) that separates the two sets into different half-spaces. For a geometric representation, see Figure 3.2. As the set of separable quantum states is compact and convex and a single point is convex and compact as well, it follows that there exist observables W that lead to a positive expectation value for any separable state. In contrast, at least one entangled state ρ leads to a negative expectation value, which means the witness W detects the entangled state ρ. Besides this geometrical interpretation, entanglement witnesses have the advantage that they can be measured in an experiment, even without full knowledge of the quantum state that is to be detected. For further details on entanglement witnesses and ways to construct them, see [121].

Several relevant classes of operations cannot create entanglement from separable states. The maximal set of free operations $\mathscr{F}_{\text{Sep}} \subset \mathscr{Q}$ contains all CPTP maps Λ_{sep} that admit the separable decomposition [128, 129]

$$\Lambda_{\text{sep}}(\rho) = \sum_i (A_i \otimes B_i)\rho(A_i \otimes B_i)^\dagger. \tag{3.121}$$

The most important subset of all separable CPTP maps is given by those channels that can be implemented via local operations and classical communication (LOCC) [115]. These operations often describe the physical situation in a lab. For example, Alice and Bob can perform local operations in their respective lab (including measurements) and send bits via a classical channel. Then, conditioned on the information received from the other party, they can perform additional local operations and restart the process by sending further bits to their counterpart. However, they are not allowed or able to send quantum information. Hence, they cannot

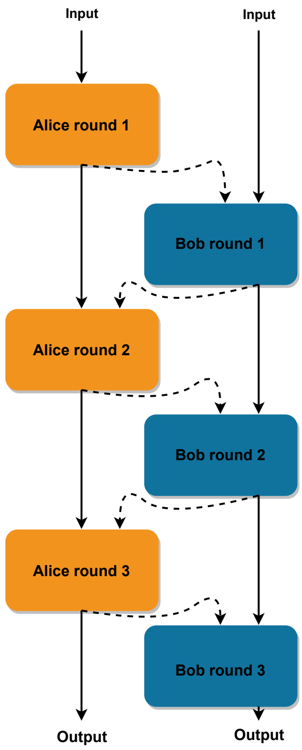

Fig. 3.3.: LOCC paradigm scheme. Alice and Bob perform in each round of the protocol some local operation on their quantum system which can be based on the information they received from the other party via a classical channel (dotted lines).

send quantum states from one lab to the other. For a pictorial representation of the LOCC paradigm, see Figure 3.3. Unfortunately, the mathematical structure of LOCC is rather complicated [130], which is why the larger and easier-to-handle set of separable operations is often used instead.

An important variant of LOCC are those operations that only succeed with a certain probability, respectively rely on post-processing. Formally, this enlarges the class of LOCC to the class of stochastic local operations and classical communication (SLOCC). The most prominent example for these stochastic operations are *local filters* [49]. Local filters are operations from the set of SLOCC described by Kraus operators F_A, F_B such that $F_A F_A^\dagger \preceq \mathbb{1}$ and $F_B F_B^\dagger \preceq \mathbb{1}$ which map a state ρ to the state

$$\rho_F = \frac{(F_A \otimes F_B)\rho(F_A \otimes F_B)^\dagger}{\mathrm{Tr}[(F_A \otimes F_B)\rho(F_A \otimes F_B)^\dagger]}, \tag{3.122}$$

with probability $p = \mathrm{Tr}[(F_A \otimes F_B)\rho(F_A \otimes F_B)^\dagger]$. Local filters can be understood via the concept of subchannels or they can be regarded as quantum measurement with four outcomes (see, e.g., [131]) given by $\{F_A \otimes F_B, F_A \otimes \bar{F}_B, \bar{F}_A \otimes F_B, \bar{F}_A \otimes \bar{F}_B\}$, where $\bar{F}_A$ and $\bar{F}_B$ are operators such that $F_A F_A^\dagger + \bar{F}_A \bar{F}_A^\dagger = \mathbb{1}$ $(F_B F_B^\dagger + \bar{F}_B \bar{F}_B^\dagger = \mathbb{1})$. Classical communication allows Alice and Bob to keep only the quantum state corresponding to the first outcome, which is equivalent to obtaining the state in

Eq. (3.122).

The concept of LOCC allows us to justify that pure states with Schmidt coefficients $c_i = \frac{1}{\sqrt{d}}$ for all $i = 0, \cdots, d-1$ are termed *maximally entangled*. The canonical example for such a state is given by

$$|\Phi_d^+\rangle = \frac{1}{\sqrt{d}} \sum_{i=0}^{d-1} |ii\rangle. \tag{3.123}$$

The name is justified, as states of this form can deterministically be transformed to any other d-dimensional pure states via LOCC. More precisely, Nielsen [132] proved the following theorem.

Theorem 3.2.4. *(*Pure state conversion via LOCC*). Let $|\psi\rangle$ and $|\phi\rangle$ be two quantum states in the Hilbert space $\mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B$. The state transformation $|\psi\rangle \xrightarrow{\mathscr{F}_{\mathrm{LOCC}}} |\phi\rangle$ is possible if and only if*

$$\mathrm{Tr}_2[|\phi\rangle\langle\phi|] \succeq_m \mathrm{Tr}_2[|\psi\rangle\langle\psi|], \tag{3.124}$$

which means $|\psi\rangle$ can be transformed via LOCC into $|\phi\rangle$ exactly when the reduced state $\rho_A(|\psi\rangle\langle\psi|) = \mathrm{Tr}_2[|\psi\rangle\langle\psi|]$ is majorized by $\rho_A(|\phi\rangle\langle\phi|) = \mathrm{Tr}_2[|\phi\rangle\langle\phi|]$.

Theorem 3.2.4 together with the fact that states in Eq. (3.123) lead to a spectrum $\lambda(\rho_A) = \{1/d, \cdots, 1/d\}$ allows us to conclude that maximally entangled states can be transformed to any d-dimensional pure state.

To study the conversion possibilities $\rho \xrightarrow{\mathscr{F}_{\mathrm{LOCC}}} \sigma$ for general mixed states is a more challenging task, due to the complexity of the LOCC paradigm and the structure of entanglement itself. See, however, the results [133] for more details regarding advanced (pure) state conversion schemes, including probabilistic transformations. Note that local filters can be used to transform a pure entangled state $|\psi\rangle = a|00\rangle + b|11\rangle$ with $a > b > 0$ into the maximally entangled state $|\Phi_2^+\rangle$ with non-zero probability.

The transformation of an entangled quantum state ρ via local filtering into a maximally entangled state of the same Hilbert space dimension can be seen as the simplest form of *entanglement distillation* [134]. However, local filters are generally not enough to achieve this task. In general, entanglement distillation describes the task to transform n-copies of a state ρ into m copies of a maximally-entangled two-qubit state $|\Phi_2^+\rangle$ via all possible *distillation protocols*, i.e., via all possible LOCC acting on $\rho^{\otimes n} = \underbrace{\rho \otimes \rho \otimes \cdots \otimes \rho}_{n \text{ times}}$. It is convenient to write $m = rn$, which introduces the distillation rate r. The optimal distillation rate is now described via the *distillable entanglement* (see e.g. [32, 120]) defined as

$$\mathrm{E}_D(\rho) = \sup \Big\{ r : \lim_{n\to\infty} \Big(\inf_{\Lambda \in \mathscr{F}_{\mathrm{LOCC}}} \|\Lambda(\rho^{\otimes n}) - (|\Phi_2^+\rangle\langle\Phi_2^+|)^{\otimes rn}\|_1 \Big) = 0 \Big\}, \tag{3.125}$$

in the limit of infinitely many copies of ρ. The distillable entanglement is physically very relevant, however, it is hard to compute in practice due to its optimization over all possible distillation protocols. For an overview over some distillation protocols see [32].

While it is obvious that entanglement is necessary for a quantum state to be distillable, it is less obvious whether it is also sufficient. Indeed, there exist entangled quantum states with $\mathrm{E}_D(\rho) = 0$ [135]. Hence, these states cannot be distilled to the maximally entangled state. This particularly weak form of entanglement has been termed *bound entanglement*. The characterization of bound entangled states is a famously complex problem. However, a very powerful sufficient criterion exists for a state to be bound entangled. That is, if a state is entangled but its entanglement cannot be detected via the PPT criterion, it is bound entangled [135]. Whether there exists also bound entangled states with a negative partial transpose is one of the most famous open problems of entanglement theory [136].

Despite their weak entanglement, bound entangled states where found to be useful for several quantum information processing tasks [137, 138], including the generation of nonlocal correlations [139, 140]. In our Publication A, we demonstrate nonlocality in bound entanglement in a hidden nonlocality scenario, which answers a question close to the Peres conjecture [141].

While the distillable entanglement quantifies the entanglement in terms of its usefulness for a particular set of protocols, it is also possible to follow the more axiomatic approach. In the following, we present this axiomatic approach, which states that an entanglement monotone has to fulfill the following conditions:

(E1) Vanishing for separable states: $\rho \in \mathrm{Sep}(\mathcal{H}_{AB}) \implies \mathrm{E}(\rho) = 0$.

(E2) Monotonicity: $\mathrm{E}(\rho) \geq \mathrm{E}(\Lambda_{\mathrm{LOCC}}(\rho))$ for any $\Lambda_{\mathrm{LOCC}} \in \mathscr{F}_{\mathrm{LOCC}}$ and any $\rho \in \mathcal{S}(\mathcal{H}_{AB})$.

(E3) Strong monotonicity: $\mathrm{E}(\rho) \geq \sum_i p_i \mathrm{E}(\sigma_i)$ for any $\sigma_i = \dfrac{K_i \rho K_i^\dagger}{p_i}$ with $p_i = \mathrm{Tr}[K_i \rho K_i^\dagger]$ and Kraus operators $\{K_i\}$ describing some LOCC channel and any $\rho \in \mathcal{S}(\mathcal{H}_{AB})$.

(E4) Convexity: $\mathrm{E}\Big(\sum_i q_i \rho_i\Big) \leq \sum_i q_i \mathrm{E}(\rho_i)$ for any probability vector $\mathbf{q}$ and any $\rho_i \in \mathcal{S}(\mathcal{H}_{AB})$.

Due to the hardness of the separability problem, entanglement monotones are typically not required to be faithful. Prominent examples of non-faithful monotones are the distillable entanglement introduced above, and the negativity [142]. Like for coherence, the typical monotonicity condition (E2) is sometimes replaced by the stronger condition (E3). However, it is often considered as an additional property and not as a necessary requirement. Clearly, all the monotones introduced in Section 3.1.1 satisfy at least the conditions (E1), (E2), and (E4). For a more general overview over entanglement monotones and discussions about additional requirements for a proper entanglement quantifier, see [32].

We want to finish the discussion about entanglement by describing the structural differences between multipartite and bipartite entanglement. However, we will focus here on the tripartite case since we need it to discuss Publication A. A tripartite quantum state $\rho_{ABC} \in \mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C)$ is called *(fully) separable* if it can be written as convex combination of product states, i.e.,

$$\rho_{ABC} = \sum_i p_i \rho_{A_i} \otimes \rho_{B_i} \otimes \rho_{C_i}, \tag{3.126}$$

and it is entangled otherwise. Besides the notion of full separability, tripartite entanglement can also be defined with respect to certain bipartitions. For instance, the state

$$\rho_{ABC} = |0\rangle\langle 0|_A \otimes |\Phi^+\rangle\langle\Phi^+|_{BC}, \tag{3.127}$$

is clearly entangled. However, the entanglement is only prevalent due to the entanglement in the BC system and it is separable with respect to the bipartition $A|BC$. More generally, a quantum state ρ_{ABC} is called *biseparable* if it can be written as convex combination of biseparable states with respect to the different bipartitions, i.e,

$$\rho_{ABC} = p_{A|BC}\ \rho_{A|BC} + p_{B|AC}\ \rho_{B|AC} + p_{C|AB}\ \rho_{C|AB}, \tag{3.128}$$

with probabilities $p_{A|BC}$, $p_{B|AC}$, $p_{C|AB}$ such that $p_{A|BC} + p_{B|AC} + p_{C|AB} = 1$. If a quantum state is not biseparable, it is called *genuinely multipartite entangled*, which is the strongest form of multipartite entanglement.

The most important examples of genuinely multipartite entangled states are the GHZ [143] and the W state [144], given by

$$|\mathrm{GHZ}\rangle = \frac{1}{\sqrt{2}}(|000\rangle + |111\rangle), \quad |\mathrm{W}\rangle = \frac{1}{\sqrt{3}}(|001\rangle + |010\rangle + |100\rangle), \tag{3.129}$$

which both can be regarded as maximally entangled tripartite states. However, the notion of maximally entangled states is much more complex for multipartite states, as different SLOCC orbits of entangled states exist. This has the consequence that no SLOCC protocol exists that can convert the GHZ state to the W state and vice versa.

3.2.4 Relations Between State Resources: A Hierarchy

In general, the different resources of a quantum state are not independent of each other. Therefore, the framework of resource theories is advantageous as the general structures allow us to compare different resources and establish relations among them. In the following, we present a hierarchy of quantum resources presented in [61]. Consider the purity $\mathrm{P}(\rho)$, the coherence with respect to some product basis $\mathrm{C}(\rho)$, the symmetric discord $\mathrm{D}(\rho)$, and the entanglement $\mathrm{E}(\rho)$ of a general quantum state $\rho \in \mathcal{S}(\mathcal{H}_{AB})$ quantified by any of the resource quantifiers (by varying the free set for a fixed quantification method) presented in Section 3.1.1. It holds that

$$\mathrm{P}(\rho) \geq \mathrm{C}(\rho) \geq \mathrm{D}(\rho) \geq \mathrm{E}(\rho). \tag{3.130}$$

The hierarchy in Eq. (3.130) establishes the purity $\mathrm{P}(\rho)$ as the most fundamental resource as all other resources are upper bounded by it. As noted in [61] the hierarchy also holds in the form

$$\mathrm{P}(\rho) = \sup_U \mathrm{C}(U\rho U^\dagger) \geq \sup_U \mathrm{D}(U\rho U^\dagger) \geq \sup_U \mathrm{E}(U\rho U^\dagger), \tag{3.131}$$

that is, even when the resources of coherence, discord, and entanglement are maximized over arbitrary unitaries applied on the state ρ. The bound between coherence and purity was proven to be tight using the concept of the maximally coherent state in Eq. (3.111).

The proof of the hierarchy in Eq. (3.130) is surprisingly simple. It follows directly from the fact that each of the quantifiers in Section 3.1.1 is defined as the minimum of some distance-like function (like the robustness or the relative entropy) with respect to the set of free states. As the separable states (Eq. (3.117)) contain all classical-classical states (Eq. (3.114)), which themselves contain the incoherent states (Eq. (3.103)) with the maximally mixed state $\frac{\mathbb{1}}{d}$ as particular instance, the hierarchy follows directly from the nested structure of the free states.

3.3 Resource Theories for Quantum Measurements

Quantum measurements have been recognized to fit into the framework of QRTs only recently. However, certain (sets) of quantum measurements were known to be more valuable than others in specific tasks already earlier, for instance, for the violation of Bell inequalities [56, 145], and especially in the field of state discrimination [146–148].

Resource theories for quantum measurements have many similarities with those of states and quantum channels [149–151] but also some critical differences. For instance, we typically consider the resource of one state, while, in the case of measurements, the resource of a set of measurements is of particular interest. Some attempts to give a global overview on resources of quantum measurements were made in [41, 152]. However, the works focus either on the notion of measurement simulability [153] or consider only a single measurement and particular resource theories. On the other hand, the general quantification of convex measurement resources in the context of robustness and weight quantifiers was studied, for instance, in [154–156].

In the context of measurements, the set $\mathcal{S}$ (see Definition 3.1.1) of all objects will be replaced by the set $\mathcal{A}_{(m,d,k)}$, which contains itself all sets of m measurements acting on a d-dimensional Hilbert space $\mathcal{H}$, each containing k outcomes. Given that a POVM is defined by a set of effects $\{M_a\}_a$, where the index a denotes the outcome, we denote by $\mathcal{M} = \{M_{a|x}\}_{a,x}$ a set of multiple POVMs, where x denotes the measurement setting (the respective POVM). From here on, we call $\mathcal{M}$ a *measurement assemblage*. Furthermore, the set of all transformations $\mathscr{Q}^\dagger$ denotes the set of all unital maps (acting on the POVM elements) on the Hilbert space $\mathcal{H}$.

3.3.1 Informativeness

The resource theory of measurement informativeness was recently introduced in [40] to quantify how valuable a given measurement is in obtaining information about a quantum state. It formalized and translated earlier ideas (see, e.g., [157–160]) to the language of resource theories and proposed a framework for *a single POVM*. We will review the resource theory of measurement informativeness as introduced in [40] in the following and present our generalization to measurement assemblages in Section 4.3, respectively our Publication C.

In the resource theory of measurement informativeness, the set of free POVMs is given by

$$\mathcal{V}_{\mathrm{Info}} := \{\{q(a)\mathbb{1}\}_a\}, \tag{3.132}$$

where $\mathbf{q}$ is a probability distribution that contains the probabilities $q(a)$ of obtaining the outcome a. These measurements are termed *trivial* or *uninformative*, as they are unable to reveal any information about a quantum state. Instead, random outcomes are obtained according to the distribution $\mathbf{q}$.

The convexity of the set $\mathcal{V}_{\mathrm{Info}}$ can be easily checked. Moreover, it is closed under all unital maps $\Lambda^\dagger \in \mathscr{Q}^\dagger$, since unital maps preserve the identity by definition. This means that the set $\mathscr{F}$ of free channels $\Lambda^\dagger$ coincides with the set of all channels $\mathscr{Q}^\dagger$. We want to point out that this is a distinction to resource theories of states. Such a case would immediately deem the resource theory trivial by using quantum channels that simply prepare a quantum state ρ.

Furthermore, we emphasize that unital channels $\Lambda^\dagger$ were actually not considered as a free resource in [40]. Instead, the authors considered *measurement simulations* ξ [153]. A POVM $\mathcal{M}$ simulates the measurement $\mathcal{N} = \xi(\mathcal{M})$ if there exists conditional probabilities such that:

$$N_b = \sum_a p(b|a) M_a. \tag{3.133}$$

This means that, upon obtaining the outcome a from measuring the POVM $\mathcal{M}$, we output the new outcome b with the probability $p(b|a)$. This way, actually measuring $\mathcal{M}$ gives us access to the measurement statistics of $\mathcal{N}$.

Measurement simulability plays an essential role in resource theories for measurements, and measurement simulations can be seen as an additional type of free operation besides unital channels. We will consider them in more detail in the Section 3.3.3 and in our Publication C. For a detailed overview on the topic see, [153]. To capture both types of free operations appropriately for resource theories of measurements, we will define measurement resource theories via the tuple $Q = (\mathcal{V}, \mathscr{F}, \mathbb{S})$. Here, the set $\mathbb{S}$ contains all free measurement simulations ξ. Typically, and for all measurement resources presented in this thesis, $\mathbb{S}$ coincides with the set of all possible measurement simulations, meaning that any possible measurement simulation is also considered to be free.

An informativeness quantifier $\mathrm{IF}(\mathcal{M})$ should fulfill the following conditions (see also [40]):

(IF1) Faithfulness: $\mathrm{IF}(\mathcal{M}) = 0 \iff \mathcal{M} \in \mathcal{V}_{\mathrm{Info}}$.

(IF2) Monotonicity under free channels: $\mathrm{IF}(\mathcal{M}) \geq \mathrm{IF}(\Lambda^\dagger(\mathcal{M}))$ for any $\Lambda^\dagger \in \mathscr{Q}^\dagger$.

(IF3) Monotonicity under simulations: $\mathrm{IF}(\mathcal{M}) \geq \mathrm{IF}(\xi(\mathcal{M}))$ for any measurement simulation of the form in Eq. (3.133).

(IF4) Convexity: $\mathrm{IF}(\sum_i q_i \mathcal{M}_i) \leq \sum_i q_i \mathrm{IF}(\mathcal{M}_i)$ for any POVMs $\mathcal{M}_i$ and any probability distribution $\mathbf{q}$.

Note that $\sum_i q_i \mathcal{M}_i$ denotes a POVM whose respective effects are convex combinations of those of the $\mathcal{M}_i$ with the weights q_i.

In [40], the authors focused on the quantifier known as the generalized robustness of informativeness, given by

$$\mathrm{IF}_{\mathrm{rob}}(\mathcal{M}) \coloneqq \inf_{\mathcal{N},r,\mathbf{q}} \left\{ r \geq 0 : \frac{M_a + rN_a}{1+r} = q(a)\mathbb{1} \ \forall a \right\}, \tag{3.134}$$

where $\mathbf{q}$ is any probability distribution and $\mathcal{N}$ is a general POVM of the same dimension and number of outcomes as $\mathcal{M}$. The robustness $\mathrm{IF}_{\mathrm{rob}}(\mathcal{M})$ clearly fulfills the conditions (IF1), (IF2), and (IF4) due to our general considerations in Section 3.1.1. The monotonicity under simulations in condition (IF3) was shown in [40] and follows from similar ideas.

Moreover, it was shown that $\mathrm{IF}_{\mathrm{rob}}(\mathcal{M})$ admits the closed form expression

$$\mathrm{IF}_{\mathrm{rob}}(\mathcal{M}) = \sum_a \|M_a\|_\infty - 1, \tag{3.135}$$

in terms of the spectral norm $\|\cdot\|_\infty$. The expression in Eq. (3.135) was derived by exploiting the SDP formulation of the generalized robustness of informativeness. A straightforward analysis leads to the bound

$$\mathrm{IF}_{\mathrm{rob}}(\mathcal{M}) \leq \min(k,d) - 1, \tag{3.136}$$

for a measurement with k outcomes. That means for a fixed dimension d, the maximal informativeness $\mathrm{IF}_{\mathrm{rob}}(\mathcal{M}) = d-1$ can only be obtained for measurements with at least d outcomes. This maximal informativeness can easily be seen to be achieved by any POVM with effects that are proportional to rank-1 projectors.

Finally, it was shown that $\mathrm{IF}_{\mathrm{rob}}(\mathcal{M})$ admits an operational interpretation in terms of a state discrimination task (following the general results in [41, 155, 156]) and also in terms of the *accessible min-information* of the measurement, if it is treated as a measure-and-prepare channel (see Eq. (2.46)). We refer to the original work [40] for more details.

3.3.2 Coherence of Measurements

The second measurement resource that we consider is the coherence of measurements. The coherence of measurements was considered recently [41, 62, 161] and follows the ideas of coherence for quantum states. The idea of coherence for measurements comes also from the practical fact that coherence in a quantum state is only useful, if it can be detected by measurements. Beyond the case of measurements, coherence has also been considered as a resource for quantum operations [150]. Here, we follow the lines in [41, 62] to introduce the concept of coherence for a *single POVM*. In Section 4.3, which discusses Publication C, we present our approach to quantifying the coherence of a measurement assemblage. See also [161] for an alternative, basis-independent, approach to the coherence of a set.

In the resource theory of measurement coherence, the set of free POVMs is given by those POVMs that are diagonal in a fixed basis $\{|i\rangle\}$. Namely,

$$\mathcal{V}_{\mathrm{Coh}} := \Big\{\{M_a = \sum_{i=0}^{d-1} \alpha_{i|a} |i\rangle\langle i|\}_a\Big\}, \tag{3.137}$$

where $\alpha_{i|a} = \langle i|M_a|i\rangle$. This definition not only mirrors the situation for quantum states but is also equivalent to the set of measurements for which it holds that

$$\mathrm{Tr}[M_a \rho] = \mathrm{Tr}[M_a \Delta(\rho)], \tag{3.138}$$

for any state ρ and all outcomes a. Here, $\Delta(\rho) = \sum_{i=0}^{d-1} \langle i|\rho|i\rangle |i\rangle\langle i|$ is the fully dephased version (see also Eq. (3.104)) of ρ. Therefore, incoherent measurements are those that cannot distinguish between a state ρ and its incoherent version $\Delta(\rho)$ in an experiment.

The set $\mathcal{V}_{\mathrm{Coh}}$ is convex, as convex combinations of diagonal operators remain diagonal. Essentially for the same reason, measurement simulations (see Eq. (3.133)) also preserve the incoherence of measurements. Similarly to the situation for the coherence of quantum states, there are multiple options for classes of incoherent quantum channels. For instance, in [41] all incoherent channels were considered. This means that all unital channels $\Lambda^\dagger$ that preserve diagonal operators are free. On the contrary, in [62] only SIO, i.e., channels for which the Kraus operators K_a and their adjoint $K_a^\dagger$ are incoherent Kraus operators, were considered. It was argued that this is the best choice, since these operators cannot create coherence on the measurement and on the state side. For simplicity, we will consider here the set $\mathscr{F}_{\mathrm{IO}}$, which contains all unital channels $\Lambda^\dagger$ that map diagonal operators (in the basis $\{|i\rangle\}$) to diagonal operators, following [41].

A quantifier of measurement coherence for the resource theory $Q = (\mathcal{V}_{\mathrm{Coh}}, \mathscr{F}_{\mathrm{IO}}, \mathbb{S})$, where $\mathbb{S}$ denotes as before the set of all measurement simulations, should fulfill the requirements:

(C1) Faithfulness: $\mathrm{C}(\mathcal{M}) = 0 \iff \mathcal{M} \in \mathcal{V}_{\mathrm{Coh}}$.

(C2) Monotonicity under free channels: $\mathrm{C}(\mathcal{M}) \geq \mathrm{C}(\Lambda^\dagger(\mathcal{M}))$ for any $\Lambda^\dagger \in \mathscr{F}_{\mathrm{IO}}$.

(C3) Monotonicity under simulations: $\mathrm{C}(\mathcal{M}) \geq \mathrm{C}(\xi(\mathcal{M}))$ for any measurement simulation of the form in Eq. (3.133).

(C4) Convexity: $\mathrm{C}(\sum_i q_i \mathcal{M}_i) \leq \sum_i q_i \mathrm{C}(\mathcal{M}_i)$ for any POVMs $\mathcal{M}_i$ and any probability distribution $\mathbf{q}$.

Note that in [62], a stronger form of the monotonicity was considered in terms of *selective operations*. However, this condition does not play a role in our discussions, therefore we refer to the original work at this point.

Several candidates for measurement coherence quantifiers have been proposed in [62]. Among them are quantifiers based on the relative entropy (see also [162]), the classical trace distance, and most notably matrix norm and robustness quantifiers. We focus on the latter here, which has also been discussed in [41]. The generalized robustness of coherence is defined as

$$\mathrm{C}_{\mathrm{rob}}(\mathcal{M}) := \inf_{\mathcal{N},r,\alpha_{i|a}} \Big\{ r \geq 0 : \frac{M_a + rN_a}{1+r} = \sum_{i=0}^{d-1} \alpha_{i|a} |i\rangle\langle i| \ \forall a \Big\}, \tag{3.139}$$

where $\alpha_{i|a}$ are some positive coefficients and $\mathcal{N}$ is a general POVM of the same dimension and number of outcomes as $\mathcal{M}$. The generalized robustness of measurement coherence fulfills all the conditions $(\mathrm{C1}) - (\mathrm{C4})$, which can be seen by applying the general methods from Section 3.1.4.

Analogously to the informativeness, it was shown in [41] that the generalized robustness of measurement coherence is bounded such that

$$\mathrm{C}_{\mathrm{rob}}(\mathcal{M}) \leq \min(k, d) - 1. \tag{3.140}$$

Notably, this is the exact same upper bound as for the informativeness in Eq. (3.136). In Section 4.3, which discusses Publication C, we show that this is no coincidence but a consequence of a nested structure between the uninformative and incoherent measurements. The bound in Eq. (3.140) can be tight. In particular, it can be achieved by measurements defined in the *Fourier basis* with respect to the incoherent basis $\{|i\rangle\}$. We show in our Publication C that a similar result holds for the generalization to *sets of measurements*.

3.3.3 Measurement Incompatibility

Measurement incompatibility is arguably the most prominent resource of quantum measurements. In recent years, there has been extensive research on the incompatibility of quantum devices and measurement incompatibility in particular. The importance of measurement incompatibility can be understood from different perspectives. From a quantum foundations point of view, it generalizes Heisenberg's and Robertson's studies about the uncertainty of observables. From the standpoint of quantum correlations, measurement incompatibility is an essential prerequisite for Bell nonlocality and for quantum steering, which has particularly close connections to measurement incompatibility. Also, the connection to quantum contextuality has been investigated in recent works. Finally, from a prepare-and-measure perspective, measurement incompatibility is an important resource as it provides advantages in state discrimination tasks and for quantum random access codes.

For a general overview of measurement incompatibility, including a list of the recent literature mentioned above, we refer to the reviews [163] and [39], where we follow the notions of the latter here.

The most popular notion of incompatible measurements, concerns the statistics of two observables A, B that do not share a common set of eigenstates, i.e., that do not commute. It captures the fact that $[A, B] \neq 0$ results in a restriction of the precision that can be obtained from A and B simultaneously, as described by the Heisenberg-Robertson uncertainty relation [164]:

$$\Delta_\psi A \cdot \Delta_\psi B \geq \frac{1}{2} |\langle [A, B] \rangle_\psi|. \tag{3.141}$$

Here, $\Delta_\psi X = \sqrt{\langle X^2 \rangle_\psi - \langle X \rangle_\psi^2}$ denotes the standard deviation of X and $\langle X \rangle_\psi = \langle \psi | X | \psi \rangle$ is the expectation value of X with respect to the pure state $|\psi\rangle$.

Despite its fundamental importance, the uncertainty relation in Eq. (3.141) does not capture the whole notion of measurement incompatibility and comes with certain drawbacks. First, it is a state-dependent statement about the incompatibility of two observables in terms of their non-commutativity. Second, it is written as a product of the standard deviations, which can make it trivial in specific scenarios. For an overview of modern formulations of generalized uncertainty relations, see [165].

More importantly, the non-commutativity of two observables only captures the incompatibility of projective measurements. That it does capture the commutativity properties of projective measurements follows directly from the spectral decomposition, see Eq. (2.22). Generally, we say a measurement assemblage $\mathcal{M}$ describes *commuting measurements* if

$$[M_{a|x}, M_{a'|x'}] = 0 \; \forall \; a, a' \text{ and } x \neq x'. \tag{3.142}$$

To include also general POVMs into the framework, we have to introduce a more general notion of *measurement simulations*. We already saw in Eq. (3.133) how one POVM $\mathcal{M}$ can simulate a different POVM via classical post-processing i.e.,

$$N_b = \sum_a p(b|a) M_a. \tag{3.143}$$

This notion of a measurement simulation can be generalized in many different ways [42, 153]. In particular, a POVM $\{G_\lambda\}$ can be used to simulate the statistics of more than one measurement. More specifically, $\{G_\lambda\}$ simulates the measurement assemblage $\mathcal{F}$ if it holds that

$$F_{a|x} = \sum_\lambda p(a|x, \lambda) G_\lambda \; \forall \; a, x. \tag{3.144}$$

That means, the measurement statistics of the assemblage $\mathcal{F}$ can be obtained by performing the measurement $\{G_\lambda\}$ and post-processing its outcome distribution $p(\lambda) = \mathrm{Tr}[G_\lambda \rho]$ via the probabilities $p(a|x,\lambda)$. Since all the measurements from $\mathcal{F}$ can be measured by only performing a single measurement, it is called *jointly measurable* or compatible. Assemblages that are not jointly measurable are called *incompatible*. This clearly generalizes the idea of commuting projective measurements to general POVMs.

At this point, we have to comment on a few points. First, there are more notions apart from non-joint measurability that can be seen as a generalization of measurement incompatibility to POVMs. For instance, there are the notions of non-disturbance [166] and coexistence [167]. However, joint measurability is by far the most studied and important notion and we will use the term *measurement incompatibility* interchangeably here. Second, in order to decide whether there exist a joint measurement $\{G_\lambda\}$ for the assemblage $\mathcal{F}$, also known as *parent POVM*, it is useful to realize that all the randomness in the post-processing $p(a|x,\lambda)$ can by hidden inside the parent POVM $\{G_\lambda\}$. This means that one can restrict to deterministic post-processing strategies $v(a|x,\lambda)$ such that

$$F_{a|x} = \sum_\lambda v(a|x,\lambda) G_\lambda \;\forall\; a, x, \tag{3.145}$$

where the $v(a|x,\lambda)$ are vertices of a probability polytope, taking on only the values 0 or 1. This limits the numbers of outcomes needed to describe $\{G_\lambda\}$, as there are only $N_{\mathrm{det}} = k^m$ deterministic assignements for an assemblage $\mathcal{F}$ of m measurements with k outcomes.

The deterministic post-processing in Eq. (3.145) can also be seen from a different perspective. Namely, the joint measurability of $\mathcal{F}$ can equivalently be written as

$$F_{a|x} = \sum_{\vec{a}\backslash a_x} G_{\vec{a}}, \tag{3.146}$$

where $G_{\vec{a}}$ is a parent POVM with differently labeled outcome sets and the sum over all elements $\vec{a}\backslash a_x$ refers to all elements of $\vec{a}$ but a_x [168].

To clarify the notation and to show the differences between commutativity and joint measurability, we consider the following canonical example. Let $M^\eta_{i|1} = \{\frac{1}{2}(\mathbb{1} \pm \eta\sigma_x)\}_{i=\pm1}$ and $M^\eta_{j|2} = \{\frac{1}{2}(\mathbb{1} \pm \eta\sigma_z)\}_{j=\pm1}$ be two measurements corresponding to noisy Pauli-x and z measurements. We want to find out when the measurements become jointly measurable, depending on the noise parameter η. It is relatively easy to see that

$$G^\eta_{i,j} = \Big\{\frac{1}{4}(\mathbb{1} + \eta(i\sigma_x + j\sigma_z))\Big\}_{i,j=\pm1} \tag{3.147}$$

is a good candidate for a parent POVM, as it is complete, i.e., $\sum_{i,j=\pm1} G^\eta_{i,j} = \mathbb{1}$ and reproduces the measurements via $M^\eta_{i|1} = \sum_j G^\eta_{i,j}$ and $M^\eta_{j|2} = \sum_i G^\eta_{i,j}$. However,

$G^{\eta}_{i,j}$ is only a positive operator for $\eta \in [0, \frac{1}{\sqrt{2}}]$. This means that $M^{\eta}_{i|1}$ and $M^{\eta}_{j|2}$ are jointly measurable for $\eta \leq \frac{1}{\sqrt{2}}$. Indeed, it can be shown that the measurements are incompatible for $\eta > \frac{1}{\sqrt{2}}$, which means the construction based on $\{G^{\eta}_{i,j}\}_{i,j}$ is optimal. Notably, except for the case $\eta = 0$, the measurements $M^{\eta}_{i|1}$ and $M^{\eta}_{j|2}$ do not commute but are still jointly measurable for $\eta \leq 1/\sqrt{2}$. To see that commutativity implies joint measurability, it is enough to notice that $\{G_{i,j} = M_{i|1}M_{j|2}\}_{i,j}$ is a valid parent POVM provided that $\{M_{i|1}\}_i$ and $\{M_{j|2}\}_j$ are commuting measurements (see, e.g., [169]). The generalization to pairwise commuting measurements is straight forward.

Let us finally comment on more general measurement simulations. In Eq. (3.144) we saw how one measurement can simulate many. In general, an assemblage $\mathcal{M}$ can simulate the assemblage $\mathcal{N}$ if there exist probabilities $p(x|y)$ and $p(b|y,x,a)$ such that:

$$N_{b|y} = \sum_{a,x} p(x|y)p(b|y,x,a)M_{a|x}. \tag{3.148}$$

That means, we choose measurement x of the assemblage $\mathcal{M}$ with the probability $p(x|y)$, given that we want to simulate setting y of $\mathcal{N}$. Finally, upon receiving the outcome a we declare the outcome b with probability $p(b|y,x,a)$.

Measurement simulations cannot generate incompatibility from compatible measurements. Indeed, by applying the simulation in Eq. (3.148) onto a compatible assemblage $\mathcal{F}$, we obtain

$$F'_{b|y} = \sum_{a,x} p(x|y)p(b|y,x,a)F_{a|x} = \sum_{\lambda}\sum_{a,x} p(x|y)p(b|y,x,a)p(a|x,\lambda)G_{\lambda}, \tag{3.149}$$

which means $\{G_{\lambda}\}$ is a parent POVM for $\mathcal{F}'$ with the post processing

$$p(b|y,\lambda) = \sum_{a,x} p(x|y)p(b|y,x,a)p(a|x,\lambda), \tag{3.150}$$

which defines a well-defined probability distribution.

Also quantum channels cannot be used to create incompatible measurements from compatible ones. Indeed, by simply using linearity one can show that:

$$\Lambda^{\dagger}(F_{a|x}) = \sum_{\lambda} p(a|x,\lambda)\Lambda^{\dagger}(G_{\lambda}),$$

which is a valid decomposition of the assemblage $\Lambda^{\dagger}(\mathcal{F})$ through the parent POVM $\{\Lambda^{\dagger}(G_{\lambda})\}$.

This means that all quantum channels $\Lambda^{\dagger} \in \mathcal{Q}^{\dagger}$ and all classical simulations $\xi \in \mathbb{S}$ are free operations in a resource theory of measurement incompatibility. Together with the convex set of jointly measurable assemblages $\mathcal{V}_{\text{JM}}$, we obtain the resource theory of measurement incompatibility $Q_{\text{Incomp}} = (\mathcal{V}_{\text{JM}}, \mathscr{Q}^{\dagger}, \mathbb{S})$.

Before we can discuss ways to quantify incompatibility, we have to find ways to certify that a given assemblage is incompatible or jointly measurable in the first place. This can be done via the following SDP:

$$\begin{aligned} \text{given}: \quad & \mathcal{M} = \{M_{a|x}\} \\ \underset{\{G_\lambda\}}{\text{maximize}} \quad & \mu \\ \text{subject to:} \quad & M_{a|x} = \sum_\lambda v(a|x,\lambda) G_\lambda \; \forall\, a, x, \\ & G_\lambda \succeq \mu \mathbb{1} \; \forall\, \lambda, \sum_\lambda G_\lambda = \mathbb{1}, \end{aligned} \tag{3.151}$$

which will result in a value $\mu \geq 0$ for any jointly measurable assemblage $\mathcal{M}$. On the other hand, incompatible assemblages lead to a value $\mu < 0$. Therefore, the incompatibility decision problem is efficiently solvable by an SDP. Note, however, that the number $N_{\text{det}} = k^m$ of deterministic assignments $v(a|x,\lambda)$ grows exponential in the number of settings m, which typically represents the bottle neck in applications.

There exist also various analytical criteria to certify joint measurability, respectively incompatibility and several methods for the explicit construction of parent POVMs for certain types of measurements have been proposed. We refer to the review [39] for more details.

The quantification of measurement incompatibility has been studied in various works and is very much subject of active research. In Section 4.3, which discusses Publication C, we discuss a distance-based approach to measurement incompatibility. Here, we follow similar lines as [169] and require that an incompatibility quantifier for the resource theory on incompatibility $Q_{\text{Incomp}} = (\mathcal{V}_{\text{JM}}, \mathscr{Q}^\dagger, \mathbb{S})$ has to fulfill the typical four conditions:

(I1) Faithfulness: $\mathrm{I}(\mathcal{M}) = 0 \iff \mathcal{M} \in \mathcal{V}_{\text{JM}}$.

(I2) Monotonicity under free channels: $\mathrm{I}(\mathcal{M}) \geq \mathrm{I}(\Lambda^\dagger(\mathcal{M}))$ for any $\Lambda^\dagger \in \mathscr{Q}^\dagger$.

(I3) Monotonicity under simulations: $\mathrm{I}(\mathcal{M}) \geq \mathrm{I}(\xi(\mathcal{M}))$ for any measurement simulation of the form in Eq. (3.148).

(I4) Convexity: $\mathrm{I}(\sum_i q_i \mathcal{M}_i) \leq \sum_i q_i \mathrm{I}(\mathcal{M}_i)$ for any POVMs $\mathcal{M}_i$ and any probability distribution $\mathbf{q}$.

Two classes of incompatibility quantifiers were subject to extensive research. These two classes are the incompatibility weight [170] and the robustness-based quantifiers (see, e.g., [169]). It is an interesting open problem to find the most incompatible assemblage in a dimension d for a certain quantifier and a fixed number of measurements m. Especially the role of MUB in that context has gained some attention recently [171, 172].

Let us finish this section by discussing more detailed structures of incompatibility in the case of $m = 3$ measurements. Similarly to the case of multipartite entanglement, incompatibility reveals more complex structures when we go beyond the simplest non-trivial case. More precisely, in [173] two additional forms of incompatibility for $m = 3$ measurements were discussed. Beyond the case of full incompatibility in which the assemblage $\mathcal{M} = \{\mathcal{M}^1, \mathcal{M}^2, \mathcal{M}^3\}$ can be obtained from a single parent POVM, there exist also the notions of *pairwise incompatibility* and *genuine triplewise incompatibility*. Pairwise incompatibility describes the effect that the measurement pairs $\mathcal{M}^{(1,2)} = \{\mathcal{M}^1, \mathcal{M}^2\}$, $\mathcal{M}^{(1,3)} = \{\mathcal{M}^1, \mathcal{M}^3\}$, and $\mathcal{M}^{(2,3)} = \{\mathcal{M}^2, \mathcal{M}^3\}$ are compatible, but $\mathcal{M}$ could possibly be incompatible.

Finally, an assemblage $\mathcal{M}$ is genuine triplewise incompatible if it cannot be decomposed as a convex combination of the form

$$M_{a|x} = p_{(1,2)} F^{(1,2)}_{a|x} + p_{(2,3)} F^{(2,3)}_{a|x} + p_{(1,3)} F^{(1,3)}_{a|x}, \tag{3.152}$$

where $F^{(s,t)}_{a|x}$ is an assemblage of $m = 3$ measurements, where the measurement s and t are compatible.

Note that this definition of triplewise incompatibility mirrors the definition of genuinely multipartite entanglement. An illustrative example for the different incompatibility structures is given by using the three projective measurements corresponding to the Pauli operators $\sigma_x, \sigma_y, \sigma_z$ and subject them to white noise. Namely, we consider the measurements given such that

$$M^{\eta}_{a|x} = \eta \Pi_{a|x} + (1-\eta)\,\mathrm{Tr}[\Pi_{a|x}]\frac{\mathbb{1}}{2}, \ \forall\, a, x \tag{3.153}$$

where $\Pi_{a|x}$ are the projectors corresponding to the Pauli measurements. The assemblage $\mathcal{M}^{\eta}$ is fully jointly measurable for $\eta \leq \frac{1}{\sqrt{3}}$, pairwise compatible for $\eta \leq \frac{1}{\sqrt{2}}$ and genuinely triplewise incompatible for $\eta > \frac{\sqrt{2}+1}{3}$ [173].

In Section 4.4, which discusses Publication D, we study these incompatibility structures in more detail and analyze how the incompatibility of an assemblage is constrained through the incompatibility of its subsets.

3.4 Resource Theories for Quantum Correlations

In this section, we introduce resources of quantum correlations. In particular, we discuss that Bell nonlocal correlations are stronger than *local* or *classical correlations* and how nonlocality can be analyzed and quantified on the level of the statistics obtained from a Bell experiment. Furthermore, we look at the quantum correlations and their resources from steering experiments, representing an intermediate level between entanglement and nonlocality. In a certain sense, these two quantum resources are the product of the formerly discussed state and measurement resources.

Steering and nonlocality rely on entangled states and incompatible measurements as prerequisite resources for the respective experiments. We will also see in more detail how these different resources are connected. Note that other forms of quantum correlations, e.g., contextuality, are not discussed within this thesis. Finally, note that the resources of quantum entanglement and discord are also frequently denoted as quantum correlations in the literature. However, within this thesis, we have already classified these resources as state resources.

3.4.1 Bell Nonlocality

> *Bell's theorem is the most profound discovery of science.*
>
> — **Henry Stapp**

Bell nonlocality is arguably one of the most fascinating facets of quantum theory. Bell's seminal work [15], as his answer to the infamous EPR-paradox [174], and his theorem that shows that nature is not compatible with (his model of) local realism are cornerstones for the foundations of quantum physics. In 2022, the importance of Bell nonlocality and the experiments that verify Bell's predictions were finally also acknowledged by the *Royal Swedish Academy of Sciences* [1]. Not only does Bell nonlocality contradict our everyday experience, but it is also a key component for many applications in quantum information theory, such as randomness generation [27–29] and cryptography based on the device-independent paradigm [26, 175], see also [176]. Here, we introduce the key features of quantum nonlocality and Bell's theorem and refer for more details to the literature, see especially the review [43] and the references therein. For a historical perspective on the topic, see also [177].

On a very basic level, Bell nonlocality makes statements about the behavior, and the possible underlying causal structure, of two parties and their systems which are subjected to a series of measurements in an experiment. Let us call the parties (their systems) Alice (A) and Bob (B). Imagine that Alice and Bob each hold a measurement device that has m buttons as input and k light bulbs that represent the output or outcome of the measurement device. In each round of the experiment, Alice and Bob are allowed to press one of their m buttons, which will lead to one of the k outcomes for each of them. Let us denote by x the input of Alice and by y the input of Bob. Further, let a and b be the outcome of Alice, respectively, Bob.

Our aim is to make statements about the nature of the underlying correlations that Alice and Bob could observe while making as few assumptions as possible. Therefore, their measurement devices can be regarded as black boxes, as depicted in Figure 3.4. The only assumption we make, for now, is that Alice and Bob are placed in distant labs such that the measurement events corresponding to the same round of the experiment are space-like separated. That means Alice cannot see what button Bob pushes in each round of the experiment and vice versa. Even more, every signal (which is limited by the speed of light in the vacuum) going from A to B or B to A needs a longer traveling time than the time one experiment round took. By performing enough experiment rounds, Alice and Bob collect the outcomes' frequencies, given their measurement setting. This means Alice can approximate the probabilities $p(a|x)$ and Bob the probabilities $p(b|y)$. Moreover, they can meet after the experiment to approximate the probabilities $p(ab|xy)$. For the sake of simplicity,

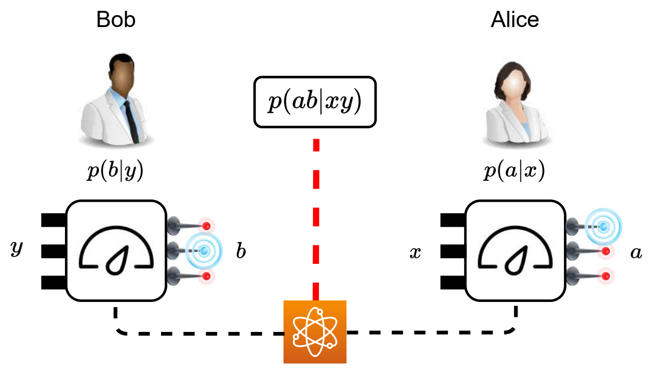

Fig. 3.4.: Sketch of the Bell scenario. Alice and Bob have each access to a measurement device in a spacelike separated setup (indicated by the dashed red line). In each round of the Bell experiment, they get an input to their device from a source, which could, for instance, distribute a bipartite quantum state. Alice and Bob perform a measurement and treat their devices as a black box, i.e., the experiment does not rely on any specific implementation or the assumption that quantum mechanics describes the experiment. Upon performing the measurement x and y, Alice and Bob obtain the outcome a and b, respectively. The central object of interest is the behavior $\mathbf{p}$ containing the probabilities $p(ab|xy)$.

we assume here that the respective distributions are obtained exactly, i.e., we work in the limit of infinitely many measurement rounds.

The question now is, which underlying model could explain their observed distribution and its potential correlations? Let us refer to $\mathbf{p} = \{p(ab|xy)\}_{a,b,x,y}$ as a behavior or distribution. We can see $\mathbf{p}$ alternatively as a vector in probability space that describes the experiment. Therefore, abusing notations slightly, we use the boldface symbol $\mathbf{p}$ for both notions interchangeably here. We say that the behavior $\mathbf{p}$ is uncorrelated if

$$p(ab|xy) = p(a|x)p(b|y) \;\forall\; a, b, x, y. \tag{3.154}$$

However, Alice and Bob generally observe that their behavior is correlated. That is nothing unexpected, as our everyday experience tells us that the past interaction of systems A and B could be responsible for the underlying correlations. To speak about any kind of model that is in accordance with causality, respectively special relativity, a probability distribution $\mathbf{p}$ has to obey the so-called *no-signaling principle* that states that the marginals $\{p(a|x)\}$, $\{p(b|y)\}$ of Alice and Bob should be independent of the actions (and in particular the measurement setting) of the other party. Note that the no-signaling principle is imposed due to the assumption of spacelike separated events here. More formally, a no-signaling behavior $\mathbf{p}$ has to obey

$$\sum_b p(ab|xy) = \sum_b p(ab|xy') = p(a|x) \;\forall a, x, y, y', \tag{3.155a}$$

$$\sum_a p(ab|xy) = \sum_a p(ab|x'y) = p(b|y) \;\forall b, y, x, x'. \tag{3.155b}$$

This means that Alice and Bob cannot use their black-box measurement device to signal their measurement setting to the other party (which would allow for superluminal communication), and it assures that the marginals of Alice and Bob are well-defined and consistent with the global distribution $\mathbf{p}$. We denote the set of all behaviors that obey the no-signaling conditions by $\mathcal{NS}$. The set $\mathcal{NS}$ is a (convex) polytope in the space of all behaviors, as it is fully described by an intersection of half-spaces. This means that the set $\mathcal{NS}$ can equivalently be described by a finite set of extremal points, namely the vertices of the no-signaling polytope.

What could be an explanation for the correlations that Alice and Bob observe? One very plausible explanation would be that the systems A and B obey a local model, or more precisely, a local hidden-variable model (LHV). To motivate it, we follow the same lines as in [43], which gives a modern-day motivation for the local model Bell proposed in his seminal paper [15]. In an LHV, we assume that a variable λ fully accounts for the correlations between Alice and Bob. It could be that the description of λ is not accessible to us but it influences the behavior $\mathbf{p}$ on a statistical level. It could especially contain information about the past interaction of the systems A and B. Therefore, the parameter λ is a type of *hidden* influence on the experiment that is fully responsible for the observed correlations. The hidden variable is *local* if it allows for a factorization of the global distribution into its marginals, i.e.,

$$p(ab|xy, \lambda) = p(a|x, \lambda)p(b|y, \lambda). \tag{3.156}$$

This means that the systems A and B are uncorrelated if the influence of the parameter λ is taken into account. In general, the hidden variable λ does not need to be constant during the time of the experiment. Moreover, λ could involve parameters that cannot be perfectly controlled in an experiment. This means, it could vary according to a probability distribution $\pi(\lambda)$. Taken this factor into account, a behavior $\mathbf{p}$ that admits an LHV can be decomposed such that

$$p(ab|xy) = \int_\Lambda \pi(\lambda)p(a|x, \lambda)p(b|y, \lambda)d\lambda,\ \forall\, a, b, x, y, \tag{3.157}$$

where Λ denotes the parameter space the variable λ lives in. Note that it is enough to consider one variable λ, instead of a set of local variables, by enlarging the parameter space Λ if necessary. Note further that we specifically assume in the LHV in Eq. (3.157) that $\pi(\lambda) = \pi(\lambda|x, y)$, i.e., that λ is *statistically independent* of the measurement settings x and y.

The model in Eq.(3.157) is undoubtedly a reasonable explanation for the origin of correlations in an experiment, and it clearly matches our intuition from everyday life. However, it cannot be an explanation for all possible no-signaling correlations, and the central point of Bell's theorem is that it cannot be in accordance with quantum mechanics. In other words, Bell showed that an LHV cannot explain quantum mechanics and how this can be witnessed.

We denote the set of all behaviors admitting an LHV according to Eq. (3.157) by $\mathcal{L}$ and we say that a behavior is (Bell-)local if $\mathbf{p} \in \mathcal{L}$ and it is (Bell-)nonlocal otherwise. Note that Eq. (3.157) can be rewritten as

$$p(ab|xy) = \sum_{\lambda} p(\lambda)p(a|x,\lambda)p(b|y,\lambda), \tag{3.158}$$

for any finite number of measurement settings for Alice and Bob, where $\{p(\lambda)\}$ is the probability distribution of λ in this case. The set $\mathcal{L}$ is, similarly to the no-signaling set, a convex polytope. The extremal points of the local polytope $\mathcal{L}$ are given by local-deterministic strategies (vertices), i.e., by distributions $p(ab|xy) = v(a|x)v(b|y)$, where $v(a|x), v(b|y)$ denote deterministic input-output distributions for Alice and Bob [178].

Before we see how quantum mechanics' incompatibility with a local hidden variable model can be proven, let us quickly define quantum behaviors formally. Quantum behaviors are those behaviors $\mathbf{p}$ that admit a quantum implementation in terms of a quantum state $\rho \in \mathcal{H}_A \otimes \mathcal{H}_B$ and measurement assemblages $\mathcal{M}$ acting on $\mathcal{H}_A$ and $\mathcal{N}$ acting on $\mathcal{H}_B$ such that

$$p(ab|xy) = \mathrm{Tr}[(M_{a|x} \otimes N_{b|y})\rho_{AB}], \tag{3.159}$$

where the dimension of the Hilbert spaces $\mathcal{H}_A, \mathcal{H}_B$ can be possibly be infinite. Since we do not constraint the dimension of the underlying Hilbert space, all the measurements and the state in Eq. (3.159) can be assumed to be projective, respectively pure. We denote the set of all quantum behaviors by $\mathcal{Q}$. Note that for infinite dimensional systems, the behaviors defined via the tensor product structure in Eq. (3.159) have been shown (the paper actually still needs to be peer-reviewed at the time of writing this thesis) to not coincide with behaviors based on commuting strategies [179]. However, this does not play a role for our discussion here.

The concept of a *Bell inequality* was introduced to show that not all no-signaling, and especially not all quantum behaviors, can be explained by an LHV. Since the set $\mathcal{L}$ is a convex and compact subset in the space of all behaviors, it follows via the separating hyperplane theorem (see, e.g., [87], see also Section 3.2.3) that every behavior $\mathbf{q} \notin \mathcal{L}$ can be detected (witnessed) by a hyperplane that separates $\mathbf{q}$ and $\mathcal{L}$ into different half-spaces. Every such *Bell inequality* can be written as

$$F = \sum_{a,b,x,y} C_{abxy}p(ab|xy) \leq L, \tag{3.160}$$

where the $\{C_{abxy}\}_{a,b,x,y}$ are real coefficients and $L \geq \max_{\mathbf{q}\in\mathcal{L}} F$ is the local bound that is obeyed by every behavior $\mathbf{q} \in \mathcal{L}$. A Bell inequality with $L = \max_{\mathbf{q}\in\mathcal{L}} F$ touches the local polytope at some point and a Bell inequality that describes a facet of the local polytope is called a *facet inequality*. See also Figure 3.5 for a depiction of the different cases.

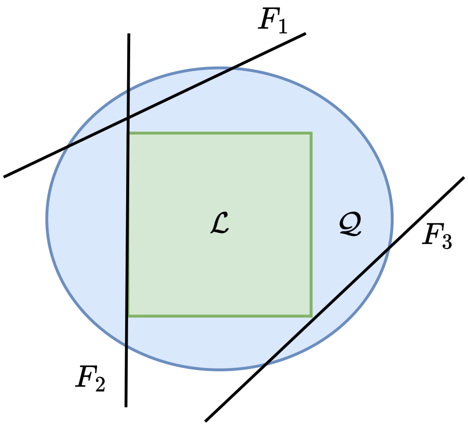

Fig. 3.5.: Depiction of Bell inequalities. Behaviors outside the local polytope $\mathcal{L}$ can be witnessed by different types of Bell inequalities. The Bell inequality F_1 does not touch the local polytope, while Bell inequality F_3 touches it in one point. The Bell inequality F_2 corresponds to a facet of the local polytope.

The description of the local polytope in terms of deterministic vertices is equivalent to the description in terms of its facets. This conversion is computationally costly but can be performed for simple Bell scenarios. In the simplest non-trivial scenario of two inputs per party, i.e., $m = 2$ and dichotomic outputs, i.e., $k = 2$, there exists only one facet Bell inequality up to relabelings of the inputs, outputs, and parties.

This inequality is the famous *CHSH inequality* [180], which is the by far most studied Bell inequality in the literature. It is given by

$$\langle F_{\text{CHSH}} \rangle = \langle A_0 B_0 \rangle + \langle A_0 B_1 \rangle + \langle A_1 B_0 \rangle - \langle A_1 B_1 \rangle \leq 2, \tag{3.161}$$

where $\langle A_x B_y \rangle = \sum_{a,b} ab\ p(ab|xy)$ with $a, b \in \{\pm 1\}$. In quantum theory, it holds $\langle A_x \otimes B_y \rangle = \operatorname{Tr}[(A_x \otimes B_y)\rho]$, where A_x, B_y are the observables of Alice and Bob, that are defined by $A_x = M_{+1|x} - M_{-1|x}$ and analogously for Bob's observables B_y. The operator $F_{\text{CHSH}} = A_0 \otimes B_0 + A_1 \otimes B_0 + A_0 \otimes B_1 - A_1 \otimes B_1$ is known as the *Bell operator* corresponding to the CHSH inequality. More generally, $F_{\text{op}} = \sum_{a,b,x,y} C_{abxy}(M_{a|x} \otimes M_{b|y})$ is the Bell operator corresponding to the Bell inequality defined by the coefficients $\{C_{abxy}\}$ and it holds $\langle F_{\text{op}} \rangle = \operatorname{Tr}[\rho F_{\text{op}}]$.

By measuring the observables

$$A_0 = \sigma_x,\ A_1 = \sigma_z,\ B_0 = \frac{\sigma_x + \sigma_z}{\sqrt{2}},\ B_1 = \frac{\sigma_x - \sigma_z}{\sqrt{2}}, \tag{3.162}$$

on the maximally entangled state $\rho = |\Phi^+\rangle\langle\Phi^+|$ with $|\Phi^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$, it can be shown that $\langle F_{\text{CHSH}} \rangle = 2\sqrt{2} > 2$. This shows, that not all quantum behaviors can be explained by an LHV, i.e., $\mathcal{L} \neq \mathcal{Q}$. That $\mathcal{Q}$ contains indeed all local correlations, i.e., that $\mathcal{L} \subset \mathcal{Q}$ holds was, for instance, shown in [181].

Moreover, using a separable state $\rho = \sum_\lambda p(\lambda) \rho_{A_\lambda} \otimes \rho_{B_\lambda}$ and performing arbitrary local measurements $\{M_{a|x}\}_{a,x}$ and $\{M_{b|y}\}_{b,y}$ on it, always results in a local model of the form

$$p(ab|xy) = \sum_{\lambda} p(\lambda) \operatorname{Tr}[M_{a|x}\rho_{A_\lambda}] \operatorname{Tr}[M_{b|y}\rho_{B_\lambda}], \tag{3.163}$$

as a special instance of obtaining local behaviors. This implies that entanglement is a pivotal resource for the violation of Bell inequalities.

The value of $\langle F_{\mathrm{CHSH}}\rangle = 2\sqrt{2}$ is special for the CHSH inequality, as it is the maximal value that can be attained by quantum theory [182]. In general, the quantum bound of a Bell inequality, i.e., the maximal expectation value of its Bell operator (optimized over all possible quantum measurements), is known as *Tsirelson bound*.

To show that $\mathcal{Q} \subset \mathcal{NS}$, let us first check that every quantum behavior is indeed no-signaling. This follows by realizing that

$$\begin{aligned} \sum_a p(ab|xy) &= \operatorname{Tr}[((\sum_a M_{a|x}) \otimes M_{b|y})\rho] \\ &= \operatorname{Tr}[(\mathbb{1} \otimes M_{b|y})\rho] \\ &= \operatorname{Tr}[M_{b|y}\rho_B] \\ &= p(b|y), \end{aligned} \tag{3.164}$$

where $\rho_B = \operatorname{Tr}_1[\rho]$ is the reduced state of Bob. The calculation for Alice's marginal distribution is analogous. Note that this property was already implicitly discussed in Section 2.6.3 and follows from the properties of the (partial) trace and the underlying tensor product structure.

Finally, consider the behavior $\mathbf{p}_{\mathrm{PR}}$ defined via

$$p_{\mathrm{PR}}(ab|xy) = \begin{cases} 1/2, & \text{if } a \oplus b = xy \\ 0, & \text{otherwise} \end{cases}, \tag{3.165}$$

where $\oplus$ denotes the addition modulo two. The behavior $\mathbf{p}_{\mathrm{PR}}$ is known as a PR-box [183]. It is straight forward to see that $\mathbf{p}_{\mathrm{PR}} \in \mathcal{NS}$. However, it follows for the CHSH inequality that $\langle F_{\mathrm{CHSH}}\rangle(\mathbf{p}_{\mathrm{PR}}) = 4$. Hence, $\mathbf{p}_{\mathrm{PR}} \notin \mathcal{Q}$ and it follows that

$$\mathcal{L} \subset \mathcal{Q} \subset \mathcal{NS}. \tag{3.166}$$

It is an essential question of why quantum theory is weaker than general no-signaling theories and which other well-motivated physical principles could lead to the description of quantum correlations. The question has been studied in various works which we refer to for the interested reader, see, e.g., [184, 185].

In the following, we want to discuss the relationship between measurement and state resources on the one side and Bell nonlocality on the other. We already saw in Eq. (3.163) that entanglement is necessary for Bell nonlocality. For pure states, it is also sufficient, as shown by Gisin [46]. However, as Werner showed [47], there exist mixed entangled states that cannot show Bell nonlocality, even if an infinite number of projective measurements is employed. Later, this result was even extended to arbitrary POVMs [48]. Surprisingly, this is not the whole story. Popescu [186] and Gisin [50] showed that states that cannot lead to any Bell inequality violation in the standard setup can still lead to Bell nonlocality in a sequential scenario. Take for instance the state

$$\rho = q|\Psi^-\rangle\langle\Psi^-| + (1-q)|0\rangle\langle 0| \otimes \frac{\mathbb{1}}{2}, \tag{3.167}$$

where $|\Psi^-\rangle = (|01\rangle - |10\rangle)/\sqrt{2}$ is the singlet state and $0 \leq q \leq 1$. The state was proven in [51] to admit an LHV for all projective measurements if $q \leq 1/2$ while it is entangled for any $q > 0$. However, if the local filters (see Eq. (3.122))

$$F_A = \epsilon|0\rangle\langle 0| + |1\rangle\langle 1|, \; F_B = \delta|0\rangle\langle 0| + |1\rangle\langle 1|, \tag{3.168}$$

with $\delta = \epsilon/\sqrt{q}$ are applied onto ρ, one obtains the state

$$\rho_F = \frac{(F_A \otimes F_B)\rho(F_A \otimes F_B)^\dagger}{\mathrm{Tr}[(F_A \otimes F_B)\rho(F_A \otimes F_B)^\dagger]}, \tag{3.169}$$

by appropriately selecting one of the four outcomes in the filtering measurement prior to the Bell test. Evaluating the expression above, one obtains the state

$$\rho_F = \sqrt{q}|\Psi^-\rangle\langle\Psi^-| + (1-\sqrt{q})\frac{|01\rangle\langle 01| + |10\rangle\langle 10|}{2} + \mathcal{O}(\epsilon^2), \tag{3.170}$$

that can violate the CHSH inequality up to a value of $\langle F_{\mathrm{CHSH}}\rangle = 2\sqrt{1+q}$ in the limit of $\epsilon \to 0$. This effect is called *hidden nonlocality* and it can be extended to *genuine hidden nonlocality* [51] for states that even admit a local model for all POVMs but exhibit Bell nonlocality after local filtering.

In Section 4.1, that discusses Publication A, we show that there exist bound entangled states that admit an LHV for general POVMs but can nevertheless be activated to show Bell nonlocality in a sequential scenario. This proves that genuine hidden nonlocality does not imply entanglement distillability.

Note that hidden nonlocality via local filters is only one method for activating the nonlocality of an entangled state that admits a local model in the usual Bell scenario. For a more detailed overview, we refer to [43].

From the measurement resource side, it is also easy to see that measurement incompatibility is necessary for Bell nonlocality. Indeed, let Alice's measurements be jointly measurable, i.e., $M_{a|x} = \sum_\lambda p(a|x,\lambda)G_\lambda$. It follows that

$$\begin{aligned} p(ab|xy) &= \mathrm{Tr}[(M_{a|x} \otimes M_{b|y})\rho] \\ &= \sum_\lambda p(a|x,\lambda)\,\mathrm{Tr}[(G_\lambda \otimes M_{b|y})\rho] \\ &= \sum_\lambda p(\lambda)p(a|x,\lambda)\,\mathrm{Tr}[M_{b|y}\rho_\lambda] \\ &= \sum_\lambda p(\lambda)p(a|x,\lambda)p(b|y,\lambda), \end{aligned} \tag{3.171}$$

which is a particular implementation of an LHV. Therefore, incompatible measurements are a prerequisite for Bell nonlocality. In the CHSH scenario, i.e., two dichotomic measurements per party, measurement incompatibility is also sufficient for nonlocality [56]. However, for general assemblages, it was shown that measurement incompatibility is not sufficient for exhibiting Bell nonlocality [54, 55]. Further relations between measurement incompatibility and Bell nonlocality have been established for instance in [173, 187].

In the Sections 4.2 and 4.3 that discuss the Publications B and C, we follow these lines by quantifying which state resources are necessary to violate a Bell inequality given that certain measurement resources are fixed. Furthermore, we establish a hierarchy of measurement resources that, for instance, upper bounds the amount of nonlocality that can be obtained from a given measurement assemblage.

Let us conclude this section by commenting on the resource theory of nonlocality of behaviors $\mathbf{p}$. The resource theory of nonlocality was first studied in [188] and further developed in [189, 190]. Most notably in [191] the set of free operations on the level of the behaviors has been identified. Here, we follow the lines of [33], as it gives a general overview of the topic. However, as there are many different aspects of the resource theory of nonlocality, we will restrict to a very general and high-level overview and refer to the above references for more details. Also, we restrict to a single copy scenario, i.e., to a scenario where only one copy (often called a box) of a behavior $\mathbf{p}$ is accessible, and protocols like nonlocality distillation or wirings do not play a role.

In the resource theory of nonlocality, the free set is given by the set $\mathcal{L}$ of local behaviors, and all allowed objects are elements of the no-signaling set $\mathcal{NS}$. Our goal is to establish a framework for the quantification of the nonlocality of a behavior $\mathbf{p}$.

There are several classes of free operations[33], of which the four relevant for us are :

1. Relabelings $\mathscr{R}$ of outcomes and settings.

2. Mixing operations $\mathscr{M}$ with local behaviors.

3. Output merging or splitting operations $\mathscr{O}$.

4. Input shortening and input enlarging operations $\mathscr{I}$.

The relabeling operations $\mathscr{R}$ exchange the labels of the outcomes and inputs by some permutation. Of course, it makes only sense to relabel outcomes to other outcomes and inputs to other inputs. However, as these operations merely permute the order of entries within a behavior $\mathbf{p}$, they cannot change the nonlocality of behaviors, which implies that relabeling operations cannot create nonlocality.

The probabilistic mixing operations $\mathscr{M}$ of a given behavior $\mathbf{p}$ with a local behavior $\mathbf{p}_L$ (that can always be created from Alice and Bob by preparing marginal distributions and correlating them via shared-randomness) such that $\mathbf{p}' = \eta\mathbf{p} + (1-\eta)\mathbf{p}_L$ also cannot increase the nonlocality of any behavior. Note that this is intuitively clear from the perspective of a convex resource quantifier.

Output merging and splitting operations $\mathcal{O}$ are operations that, in principle, can change the number of outputs of a behavior. Therefore, these operations can be used to compare behaviors with a different number of outputs. The operations $\mathscr{O}$ consist of merging two outcomes a and a' together to a new outcome a''. This operation could depend on the setting x. Similarly, Bob could merge his outcomes b in a certain setting y. Reversely, an outcome a could also be split into new outcomes a_1 and a_2. These operations convert local distributions $\mathbf{p}_L$ to other local behaviors $\mathbf{p}'_L$ in a possibly different Bell scenario (concerning the number of outcomes) and are therefore free.

Finally, input shorting and input enlarging operations $\mathscr{I}$ represent the action of Alice and Bob to either ignore a specific input or add more inputs. First, ignoring some entries from a behavior $\mathbf{p}$ clearly preserves locality. Second, Alice and Bob could decide to use for the measurement setting $m+1$ a desired marginal distribution that is uncorrelated to the output of the other party. Finally, both parties have the choice to copy the input-output statistics of a particular setting into a new one. However, also these operations were shown to preserve locality in the new Bell scenario (with a possibly different number of settings than the initial one).

Following the above discussion, we require the following conditions from a nonlocality quantifier $\mathrm{N}(\mathbf{p})$:

(N1) Faithfulness: $\mathrm{N}(\mathbf{p}) = 0 \iff \mathbf{p} \in \boldsymbol{\mathcal{L}}$.

(N2) Monotonicity: $\mathrm{N}(\mathbf{p}) \geq \mathrm{N}(\mathbf{p}')$, where $\mathbf{p}'$ is any behavior obtained from relabeling $\mathscr{R}$, local mixing $\mathscr{M}$, output merging or splitting operations $\mathscr{O}$, and input shortening and input enlarging operations $\mathscr{I}$ applied on $\mathbf{p}$.

(N3) Convexity: $\mathrm{N}(\sum_i \eta_i \mathbf{p}_i) \leq \sum_i \eta_i \mathrm{N}(\mathbf{p}_i)$, for any behaviors $\mathbf{p}_i$ and any probability distribution $\boldsymbol{\eta}$.

Many candidates for resource quantifiers have been considered, including the nonlocal weight, or EPR2 decomposition [192], robustness measures (see. e.g., [187]) or distance-based quantifiers, like the classical trace distance [193] (see also Section 2.7.1) to the local polytope. Also the violation of a singular Bell inequality has been used widely. However, as it was pointed out, for instance in [33], using the violation of a fixed Bell inequality as nonlocality quantifier is flawed, as it is not monotonic under free operations.

Here, we present the distance-based approach from [193]. The nonlocality of a behavior $\mathbf{q}$ can be quantified by its distance to the local polytope $\boldsymbol{\mathcal{L}}$. More precisely, the nonlocality of $\mathbf{q}$ as quantified by the trace-distance is given by

$$\mathrm{N}_1(\mathbf{q}) = \frac{1}{2} \min_{\mathbf{t} \in \boldsymbol{\mathcal{L}}} \sum_{a,b,x,y} p(x,y) |q(a,b|x,y) - t(a,b|x,y)|, \tag{3.172}$$

where the $p(x,y)$ are the probabilities that Alice chooses setting x and Bob setting y.

3.4.2 Steering

The final quantum resource that we consider is the resource of quantum steering. The phenomenon of quantum steering dates back to the works of Einstein [11] and Schrödinger [12, 13], following the already mentioned EPR-paper [174]. In a nutshell, quantum steering describes the possibility of Alice to remotely *steer* the state of Bob's particle into different sets of conditional states by measuring her share of an entangled state. While she cannot transfer any information to Bob through this process, she can convince Bob that he shares an entangled state with her. That is, his observed conditional states, respectively, the correlations of the experiment, cannot be explained by pre-determined states on Bob's side. Even though the initial idea of quantum steering is already rather old, it was only more recently [194] that the field gained popularity. From a quantum information point of view, steering describes a scenario between entanglement and nonlocality. While both parties have complete control over their measurement devices in an entanglement setup, they treat their device as black boxes in a nonlocality scenario. In quantum steering, one of the parties, Alice, treats her device as a black box, while Bob has full control over his particle and can perform state tomography in principle. That makes steering inherently asymmetric. See also Figure 3.6 for a distinction of the three scenarios.

Like Bell nonlocality in the device-independent paradigm, steering as a one-sided device-independent setup has found various applications in quantum information processing tasks. For instance, in randomness certification [195, 196], cryptography

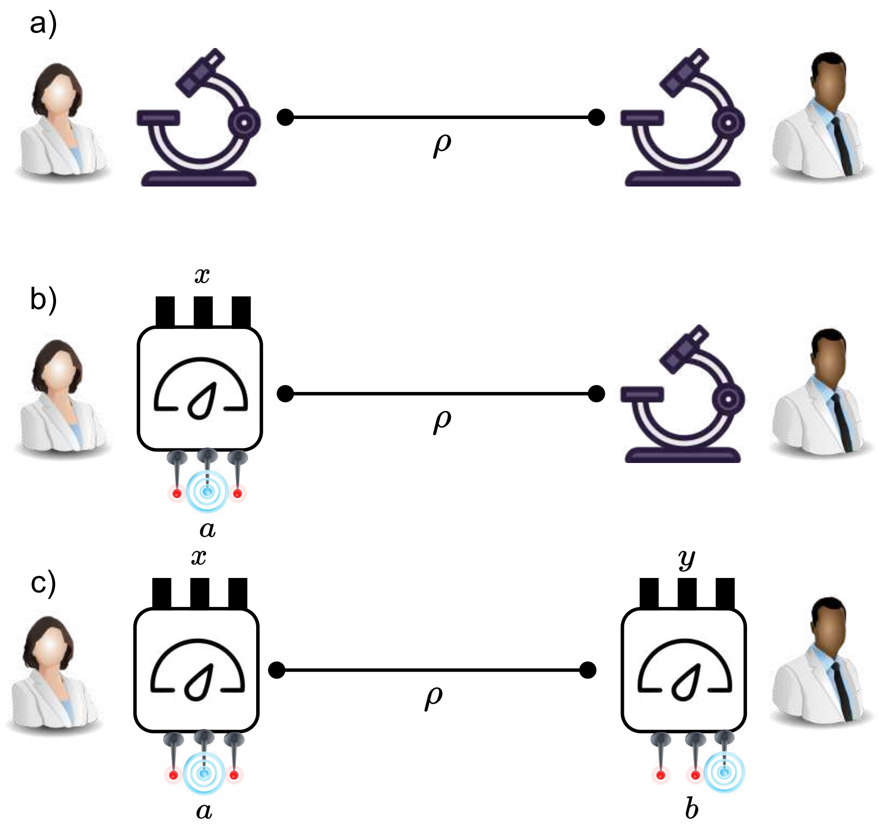

Fig. 3.6.: Comparison of three forms of inseparability. In scenario a), Alice and Bob certify entanglement with their trusted measurement devices that, in principle, are able to perform state tomography. In the nonlocality scenario c), both parties treat their measurement devices as black box, i.e., they only use the outcome statistics. In the steering scenario b), Alice treats her device as a black box, while Bob trusts his measurement device like in the entanglement scenario.

[197] or sub-channel discrimination [198]. Additionally, its close connection to measurement incompatibility [39] makes quantum steering an exciting field for research.

As the research on quantum steering gained much momentum in the last 15 years, a wide variety of literature exists. We will follow in this introduction the lines of the reviews [44, 45] but would also like to mention the overview in [199]. Furthermore, we will skip the historical motivation and arguments here and directly start with the modern-day formulation. For the interested reader, we refer to Section V.M. of reference [45] for a historical overview of the topic.

In steering, Alice has a black box measurement device with m measurements, where we assume that each of the settings x has k outcomes for simplicity. Upon receiving the outcome a with probability $p(a|x)$, the state on Bob's side is updated (transformed) to the state $\rho_{a|x}$. This means, the information one has available in steering is given by the collection $(\{\rho_{a|x}\}_{a,x}, \{p(a|x)\}_{a,x})$ of conditional probabilities and states. Usually, this information is conveniently summarized by considering the steering assemblage $\vec{\sigma} = \{\sigma_{a|x}\}_{a,x}$, where $\sigma_{a|x} = p(a|x)\rho_{a|x}$ are the sub-normalized conditional states on Bob's side after Alice's measurement.

Invoking the quantum formalism, the conditional states $\sigma_{a|x}$ are given by

$$\sigma_{a|x} = \mathrm{Tr}_1[(M_{a|x} \otimes \mathbb{1})\rho], \tag{3.173}$$

where, in general, the measurement assemblage $\mathcal{M} = \{M_{a|x}\}_{a,x}$ and the state ρ are assumed to be unknown. Even the Hilbert space dimension of Alice's system is considered to be unknown.

Every steering assemblage must fulfill three constraints corresponding to the positivity of probabilities, no-signaling, and normalization. More formally, any assemblage $\vec{\sigma}$ has to obey

$$\sigma_{a|x} \succeq 0 \;\forall\, a, x,\; \sum_a \sigma_{a|x} = \sum_a \sigma_{a|x'} = \rho_B \;\forall\, x, x', \tag{3.174}$$
$$\mathrm{Tr}[\sum_a \sigma_{a|x}] = 1 \;\forall\, x,$$

where $\rho_B = \mathrm{Tr}_1[\rho]$ is the reduced state of Bob in the quantum formalism.

We want to point out that it is a priori unclear whether all assemblages $\vec{\sigma}$ on Bob's side admit a quantum realization in terms of a bipartite quantum state ρ and measurements $\mathcal{M}$. That is, also for steering, there could exist a discrepancy between general no-signaling theories and quantum theory. However, this is not the case in the simple bipartite scenario, as it was shown in [200] by using the *HJW theorem* [201]. That means, every steering assemblage admitting the constraints in Eq. (3.174) can be regarded as having a quantum origin. However, this is no longer true for multipartite systems [200] and more involved bipartite scenarios [202].

The central question regarding quantum steering is whether an assemblage $\vec{\sigma}$ can be explained by local means (i.e., classically) or not. In the case it cannot be explained by a so-called local hidden-state model (LHS), Bob has to believe that Alice can steer his conditional states by measuring on an entangled state. Steering can be seen as a weaker notion of nonlocality than Bell nonlocality. We will shortly see that an LHS is a particular form of an LHV. More formally, an assemblage $\vec{\sigma}$ is said to admit an LHS, following the ideas from Bell nonlocality, if it can be written such that

$$\sigma_{a|x} = \int_\Lambda \pi(\lambda) p(a|x,\lambda)\rho_\lambda d\lambda \;\forall\, a, x, \tag{3.175}$$

and it is called *steerable* otherwise. The LHS in Eq. (3.175) can be understood as follows. If an assemblage $\vec{\sigma}$ admits an LHS, the conditional states $\sigma_{a|x}$ that Bob observes can be explained by the existence of a source that sends a classical message or variable λ to Alice, who announces after performing the measurement x the outcome a with probability $p(a|x,\lambda)$. Furthermore, the source sends a corresponding state ρ_λ to Bob. If a local model of the form in Eq. (3.175) fails to explain the observed assemblage $\vec{\sigma}$, Bob has to conclude that Alice can steer his system and that they have shared an entangled state.

To see that entanglement is indeed necessary for steering, we first simplify our notation, as we will always work in the scenario of finitely many measurements on Alice's side. That means, it is always possible to write the LHS in Eq. (3.175) as

$$\sigma_{a|x} = \sum_\lambda p(\lambda)p(a|x,\lambda)\rho_\lambda = \sum_\lambda p(a|x,\lambda)\sigma_\lambda, \tag{3.176}$$

where the σ_λ are sub-normalized hidden states such that $\sum_\lambda \sigma_\lambda = \rho_B$. Moreover, the probabilities $p(a|x,\lambda)$ can always be replaced with deterministic response functions $v(a|x,\lambda)$ by possibly enlarging the number of hidden states [44]. We want to highlight here, that the definition of an LHS resembles very much the joint measurability condition in Eq. (3.144).

Now, if Alice applies measurements from a measurement assemblage $\mathcal{M}$ on a separable state $\rho = \sum_\lambda p(\lambda)\rho_{A_\lambda} \otimes \rho_{B_\lambda}$, we obtain an assemblage of the form

$$\begin{aligned} \sigma_{a|x} &= \mathrm{Tr}_1[(M_{a|x} \otimes \mathbb{1})\rho] \\ &= \sum_\lambda p(\lambda)\,\mathrm{Tr}[M_{a|x}\rho_{A_\lambda}]\rho_{B_\lambda} \\ &= \sum_\lambda p(\lambda)p(a|x,\lambda)\rho_{B_\lambda}, \end{aligned} \tag{3.177}$$

which is clearly *unsteerable*, i.e., it admits an LHS of the form in Eq. (3.176). To see that steering is arranged between entanglement and nonlocality, let us apply measurements $\mathcal{N}$ on an unsteerable state assemblage held by Bob. We obtain that the distribution

$$\begin{aligned} p(ab|xy) &= \sum_\lambda p(\lambda)p(a|x,\lambda)\,\mathrm{Tr}[M_{b|y}\rho_{B_\lambda}] \\ &= \sum_\lambda p(\lambda)p(a|x,\lambda)p(b|y,\lambda), \end{aligned} \tag{3.178}$$

admits an LHV. More precisely, it admits a special type of LHV in which Bob's response function is of the form $p(b|y,\lambda) = \mathrm{Tr}[M_{b|y}\rho_{B_\lambda}]$. Therefore, we can conclude that nonlocality implies steerability of the underlying assemblage, and steerability implies entanglement of the underlying state. However, the converse does not hold. Not every mixed entangled state can lead to steering, and not all steerable assemblages can be used to reveal Bell nonlocality, see also Figure 3.7. Similarly to the situation for Bell nonlocality, there exist also entangled quantum states that cannot exhibit steering for all possible measurements [203].

Interestingly, quantum steering is also directional, i.e., there exists states that are steerable from Alice to Bob but not vice versa. This phenomenon is known as *one-way steerability* [204]. Also phenomenona like hidden nonlocality and more general nonlocality activation translate to the steering scenario, i.e., to *hidden steering* [51, 203] and activation of steering in general [45].

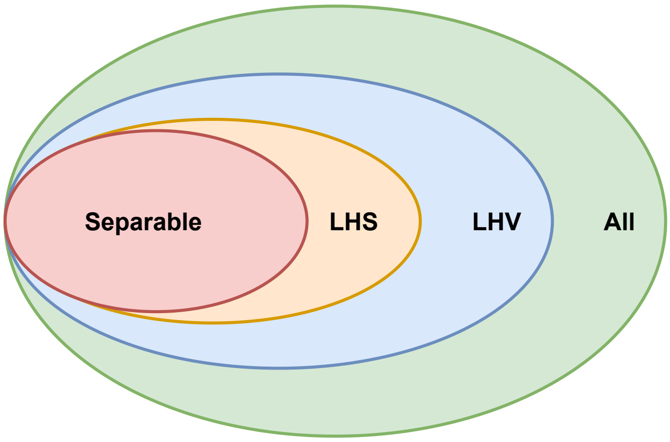

Fig. 3.7.: Scheme of subsets in state space. The states that admit an LHV for all POVMs are a convex subset of all quantum states. Some of the local quantum states even admit an LHS, which means that they are not only Bell local but also unsteerable. The set of separable states is a proper subset of all unsteerable states.

Regarding measurement resources, steering has a particularly close connection to measurement incompatibility. Let us first check that measurement incompatibility is necessary for quantum steering. Let Alice's measurements be jointly measurable, i.e, $M_{a|x} = \sum_\lambda p(a|x,\lambda)G_\lambda$. It follows that

$$\begin{aligned}\sigma_{a|x} &= \mathrm{Tr}_1[(M_{a|x} \otimes \mathbb{1})\rho] \\ &= \sum_\lambda p(a|x,\lambda)\,\mathrm{Tr}_1[(G_\lambda \otimes \mathbb{1})\rho] \\ &= \sum_\lambda p(\lambda)p(a|x,\lambda)\rho_\lambda,\end{aligned} \tag{3.179}$$

where we used that $\mathrm{Tr}_1[(G_\lambda \otimes \mathbb{1})\rho] = p(\lambda)\rho_\lambda$. Since the jointly measurable assemblage $\mathcal{M}$ leads to an unsteerable assemblage for any state ρ, incompatible measurements are necessary for steering.

The other way around, measurement incompatibility also implies steerability, when the assemblage $\vec{\sigma}$ is optimized over all possible states. This might not come as a big surprise, if we compare the definitions in Eq. (3.176) and Eq. (3.144). The simplest way [58] to show that measurement incompatibility can always lead to steering is by using the maximally entangled state $|\Phi^+\rangle = \frac{1}{\sqrt{d}}\sum_{i=0}^{d-1}|ii\rangle$ of dimension d. Any steering assemblage coming from measurements $\mathcal{M}$ on $|\Phi^+\rangle$ will evaluate to

$$\sigma_{a|x} = \mathrm{Tr}_1[(M_{a|x} \otimes \mathbb{1})|\Phi^+\rangle\langle\Phi^+|] = \frac{M_{a|x}^T}{d}, \tag{3.180}$$

where the transpose is with respect to the computational basis. By comparing the definitions of joint measureability and local assemblages, it follows directly that $\left\{\frac{M_{a|x}^T}{d}\right\}_{a,x}$ is steerable if and only if $\mathcal{M}$ is incompatible. The above one-to-one correspondence can be extended from the maximally entangled state to all Schmidt-rank d states $|\psi\rangle\langle\psi|$ [59].

A similar idea allows a mapping of steerability to incompatibility problems and vice versa [57, 205]. Given any steering assemblage $\vec{\sigma}$ such that $\rho_B = \sum_a \sigma_{a|x}$ is the reduced state of Bob. One can define the so-called *steering equivalent measurement assemblage* $\mathcal{N}$ via

$$N_{b|y} = \rho_B^{-1/2} \sigma_{a|x} \rho_B^{-1/2}, \tag{3.181}$$

which is certainly well-defined if ρ_B has full rank. In the other cases one restricts ρ_B to its support, i.e., one applies a pseudo-inverse. It now follows, that $\vec{\sigma}$ is steerable if and only if $\mathcal{N}$ is incompatible.

So far, we have not yet discussed how steering is detected, i.e., how one practically decides whether an assemblage $\vec{\sigma}$ is steerable or not. Due to the similarities to measurement incompatibility and Bell nonlocality, we can use very similar methods. Similarly, as for measurement incompatibility, the steerability problem can be decided by looking at the following SDP:

$$\begin{aligned} \text{given :} \quad & \vec{\sigma} \\ \underset{\{\sigma_\lambda\}}{\text{maximize}} \quad & \mu \\ \text{subject to:} \quad & \sigma_{a|x} = \sum_\lambda v(a|x,\lambda)\sigma_\lambda \ \forall \ a, x, \\ & \sigma_\lambda \succeq \mu \mathbb{1} \ \forall \ \lambda, \end{aligned} \tag{3.182}$$

which returns a non-negative value if $\vec{\sigma}$ is unsteerable, while it returns a negative value for steerable assemblages. The corresponding dual problem introduces the concept of *steering inequalities*. The dual is given by

$$\begin{aligned} \text{given :} \quad & \vec{\sigma} \\ \underset{\{F_{a|x}\}}{\text{minimize}} \quad & \sum_{a,x} \mathrm{Tr}[F_{a|x}\sigma_{a|x}] \\ \text{subject to:} \quad & \sum_{a,x} F_{a|x} v(a|x,\lambda) \succeq 0 \ \forall \ \lambda, \\ & \mathrm{Tr}[\sum_{a,x,\lambda} F_{a|x} v(a|x,\lambda)] = 1, \end{aligned} \tag{3.183}$$

which can be understood as a steering inequality, analog to a Bell inequality, in the following sense. The matrices $F_{a|x}$ correspond to the Bell coefficients, while the local bound of the inequality was fixed to $L = 0$. Indeed, it follows from the first constraint that

$$\sum_{a,x} \mathrm{Tr}[F_{a|x} \sum_\lambda v(a|x,\lambda)\sigma_\lambda] \geq 0, \tag{3.184}$$

for any collection of hidden states σ_λ. However, steerable assemblages might lead to a value $\mu = \sum_{a,x} \mathrm{Tr}[F_{a|x}\sigma_{a|x}] < 0$, which witnesses the steerability. The second constraint in the dual SDP simply fixes the scale of the optimization, such that it is bounded. The idea behind steering inequalities is essentially the same as for Bell nonlocality and entanglement witnesses, as it relies on the separating hyperplane theorem. Besides the SDP method, many analytical and numerical methods on the level of states, assemblages, and correlations have been developed. We refer for an overview to the reviews [44, 45].

The resource theory of steering was proposed in [34], and several resource monotones have been identified. The free objects in the resource theory of steering are the unsteerable assemblages $\vec{\tau} \in \mathrm{LHS}$. The class of operations that do not create steering from unsteerable assemblages is given by the class of one-way LOCC operations. The party allowed to perform local operations is Bob (the trusted party), who is afterward allowed to communicate classically to Alice. After receiving the message, Alice can apply classical pre-and post-processing to her measurement device. Since the concrete structure is rather complicated, we will simply denote such operations by maps $\Lambda_{1W-\mathrm{LOCC}}$ and refer to the literature [34] for more details. For a steering quantifier $\mathrm{S}(\vec{\sigma})$, we require the typical conditions:

(S1) Faithfulness: $\mathrm{S}(\vec{\sigma}) = 0 \iff \vec{\sigma} \in \mathrm{LHS}$.

(S2) Monotonicity: $\mathrm{S}(\vec{\sigma}) \geq \mathrm{S}(\Lambda_{1W-\mathrm{LOCC}}(\vec{\sigma}))$ for any one-way LOCC map $\Lambda_{1W-\mathrm{LOCC}}$ and any assemblage $\vec{\sigma}$.

(S3) Convexity: $\mathrm{S}(\sum_i p_i\vec{\sigma}_i) \leq \sum_i p_i \mathrm{S}(\vec{\sigma}_i)$, for any steering assemblages $\vec{\sigma}_i$ and any probability distribution $\mathbf{p}$.

The steerability can than be quantified by typical weight [206] and robustness quantifiers [198], as well as entropic [34] and distance-based monotones [207]. We focus on the latter here. The distance-based steering quantifier is given by

$$\mathrm{S}_1(\vec{\sigma}) = \min_{\vec{\tau} \in \mathrm{LHS}} \frac{1}{2} \sum_{a,x} p(x) \|\sigma_{a|x} - \tau_{a|x}\|_1, \tag{3.185}$$

where $p(x)$ denotes the probability that Alice decides to perform the measurement x on her share of the bipartite system. We analyze and relate the quantifier $\mathrm{S}_1(\vec{\sigma})$ to measurement resources within a resource hierarchy in Section 4.3. Furthermore, we explicitly formulate it as an SDP and relate it to the violation of an optimal steering inequality in our Publication C. For more details, we refer to the original work [207].

4 Overview of the Results

In this chapter, we summarize the main results of this thesis, based on the scientific publications prepared during the time of the author's doctoral research. The original publications can be found in the appendices: Paper A to Paper D.

4.1 Activation of Nonlocality in Bound Entanglement (Paper A)

In the work [63], we analyzed the interplay of entanglement and Bell nonlocality in the sequential hidden nonlocality scenario (see Section 3.4.1). We focused on the particularly weak form of entanglement, called bound entanglement (see Section 3.2.3). Due to their weak entanglement, Peres conjectured that bound entangled states can never lead to nonlocal correlations [141]. However, this conjecture was recently proven to be false, i.e., even bound entangled states can lead to violations of Bell inequalities [140, 208].

Our contribution advances scientific research in two directions. First, despite the Peres conjecture, it was never proven that there actually exist bound entangled states with an LHV for all POVMs. We show that this is the case by employing an SDP method proposed in [209, 210] that is in principle able to find local models for arbitrary quantum states without relying on special symmetries. Even more, we show by explicit construction that the Bell nonlocal properties of this seemingly local bound entangled state can be revealed in the hidden nonlocality scenario, i.e., we prove that local filters can activate its nonlocality. This proves that genuine hidden nonlocality does not imply entanglement distillability. Our results are summarized in Figure 4.1 and Figure 4.2, from the original publication [63].

Fig. 4.1.: Enlarging the set of nonlocal bound entangled states from [63] (Paper A). Our work shows that the set of nonlocal bound entangled states is enlarged in the hidden nonlocality (HNL) scenario. To show potential nonlocality for all bound entangled states, methods like superactivation (SA) and more asymptotic scenarios, involving infinitely many copies of a state and local operations, might be necessary.

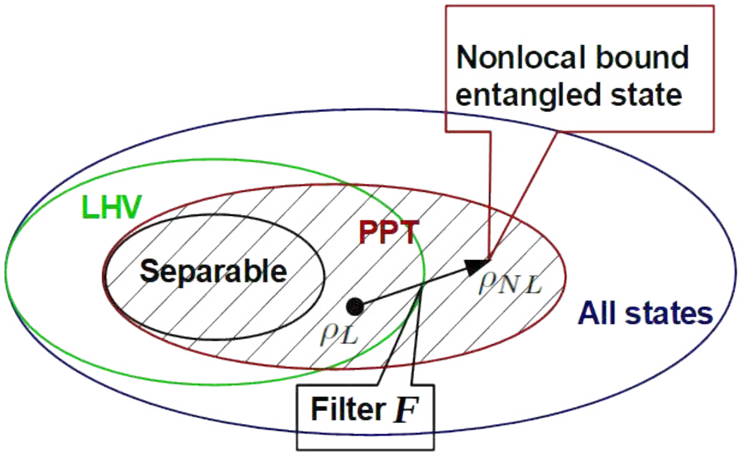

Fig. 4.2.: Scheme of nonlocality activation by local filters from [63] (Paper A). In our work, we show that there exists a state ρ_L that is bound entangled and that admits a local model for all POVMs. Next, we show that local filters can reveal the nonlocal properties of ρ_L.

Going more into the details, we focused on a three-qubit bound entangled state. Specifically, we considered in a first step the nonlocal state (respresented in the computational basis $\{|000\rangle, |001\rangle, |010\rangle, ..., |111\rangle\}_{ABC}$)

$$\rho_{NL} = (r_{ij})_{1\leq i,j\leq 8}, \tag{4.186}$$

where the entries r_{ij} are given by

$$\begin{aligned} & r_{11} = 0.0290, r_{12} = r_{13} = r_{15} = -0.0098, \\ & r_{14} = r_{16} = r_{17} = r_{23} = r_{25} = r_{35} = -0.0083, \\ & r_{18} = r_{27} = r_{36} = r_{45} = 0.0646, \\ & r_{22} = r_{33} = r_{55} = 0.0412, \\ & r_{24} = r_{26} = r_{34} = r_{37} = r_{56} = r_{57} = -0.0335, \\ & r_{28} = r_{38} = r_{46} = r_{47} = r_{58} = r_{67} = -0.0598, \\ & r_{44} = r_{66} = r_{77} = 0.1352, \\ & r_{48} = r_{68} = r_{78} = 0.0102, r_{88} = 0.4418, \end{aligned} \tag{4.187}$$

with all other entries being fixed by the Hermiticity of ρ.

By design, the state is invariant under partial transpose with respect to any party, and it is also invariant under the exchange of any of the subsystems. This makes the state not only PPT but also *fully biseparable*, i.e., biseparable with respect to any bipartiton. Hence, it is bound entangled. Nevertheless, the state can be shown to violate a Bell inequality, in particular, Sliwa's inequality number 5 [211]. Note that this state is similar to the states considered in [140]. Note further that we used numerical methods based on SDPs and the see-saw algorithm [212] to find the state ρ_{NL}, its description as given in Eq. (4.187) is, however, exact.

In a second step, we showed that there exist invertible local filters F_A, F_B, F_C and a local state ρ_L such that

$$\rho_{NL} = \frac{F_A \otimes F_B \otimes F_C\ \rho_L\ F_A^\dagger \otimes F_B^\dagger \otimes F_C^\dagger}{\mathrm{Tr}(F_A \otimes F_B \otimes F_C\ \rho_L\ F_A^\dagger \otimes F_B^\dagger \otimes F_C^\dagger)}, \tag{4.188}$$

with the local filters being given by

$$F_A = \begin{bmatrix} 0.4310 & -0.2971 \\ -0.2488 & 0.7291 \end{bmatrix},$$
$$F_B = \begin{bmatrix} 0.0342 & -0.0808 \\ -0.3664 & 0.8688 \end{bmatrix},$$
$$F_C = \begin{bmatrix} 0.3268 & -0.1873 \\ -0.1773 & 0.6440 \end{bmatrix}.$$

Note that this directly implies that ρ_L is entangled and PPT with respect to any bipartiton, since (invertible) local filters map PPT states onto PPT states [32] and preserve the non-separability. To prove that the state ρ_L is indeed local for all POVMs, we used the algorithmic SDP method in [209, 210], that is able to generate LHV for an infinite set of measurements by approximating it from the inside and adding potentially noise on the targeted state. For more details on the specific use of this method, see also our original publication [63].

The locality of ρ_L, the nonlocality of ρ_{NL}, and their connection via the local filters F_A, F_B, and F_C prove that there exist local bound entangled states of which the nonlocal nature can be revealed in the hidden nonlocality scenario. Even more, we actually employed an LHS for ρ_L, which implies the LHV. That means we also show that the activation of genuine hidden steerability in bound entanglement is possible.

4.2 Quantifying Necessary Quantum Resources for Nonlocality (Paper B)

Our work [64] analyzes what quantum state resources are necessary to achieve a specific Bell inequality violation, given that the measurements of Alice and Bob that determine the corresponding Bell operator are fixed. We focus first on the minimal purity necessary to achieve the targeted Bell inequality violation. As purity is the most fundamental resource of a quantum state (see Section 3.2.4) that bounds its other resources, this has consequences for the minimal necessary coherence, discord, and entanglement. Our setup is summarized in Figure 4.3.

In a second step, we show that our method is also helpful in determining the minimal coherence, discord, and entanglement necessary to achieve a specific violation of

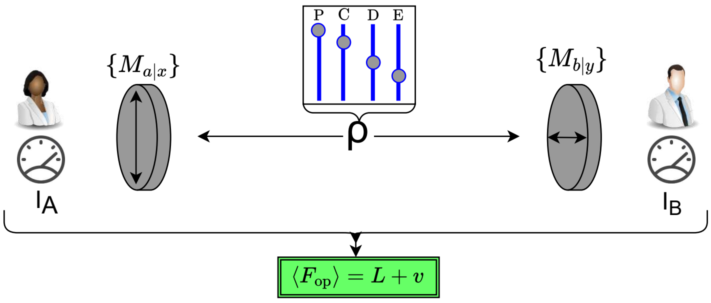

Fig. 4.3.: State resources versus measurement resources in a Bell experiment from [64] (Paper B). In our work, we consider a state's necessary purity P, coherence C, discord D, and entanglement E, given that Alice and Bob apply measurements with incompatibility I_A, respectively I_B. The goal of the Bell experiment is to achieve the violation v, beyond the local bound L.

any two-qubit full-correlation Bell inequality. In particular, we show that there exists a two-qubit quantum state ρ_{opt} that simultaneously minimizes all necessary state resources for a given full-correlation Bell inequality and a targeted Bell inequality violation.

Surprisingly, this has a counter-intuitive consequence for the CHSH-inequality. Namely, we demonstrate that more measurement resources, in terms of incompatibility, do not always allow for fewer state resources, for instance, in terms of the entanglement needed to achieve a fixed violation. In particular, sometimes it is possible to achieve the same (fixed) Bell violation while decreasing the amount of entanglement and incompatibility used to achieve this violation simultaneously.

Going more into the details, we first fix a Bell operator

$$F_{\text{op}} = \sum_{a,b,x,y} C_{abxy}\, M_{a|x} \otimes M_{b|y}, \tag{4.189}$$

corresponding to a Bell inequality defined by the Bell coefficients $\{C_{abxy}\}_{a,b,x,y}$ and the local bound L, such that $\sum_{a,b,x,y} C_{abxy}\, p(ab|xy) \leq L$, for any behavior $\mathbf{p} \in \boldsymbol{\mathcal{L}}$. Our main objective is it to find the minimal resource

$$\mathrm{R}^* = \min_{\rho}\{\mathrm{R}(\rho) \;:\; \langle F_{\text{op}} \rangle = \mathrm{Tr}(\rho F_{\text{op}}) = L + v\}, \tag{4.190}$$

for a given violation $v > 0$ and some state resource $\mathrm{R}(\rho)$. For the state resources of purity P, coherence C, discord D, and entanglement E, we considered the respective generalized robustness

$$\mathrm{R}_{\text{rob}}(\rho) := \inf_{r, \tau \in \mathcal{S}(\mathcal{H})} \left\{ r \geq 0 : \frac{\rho + r\tau}{1 + r} \in \mathcal{V} \right\}, \tag{4.191}$$

where $\mathcal{V}$ denotes the set of free resource states for the above considered resources. See also Section 2.7.2 and Section 3.1.1.

In our first main result, we show that it is always possible to determine the min-

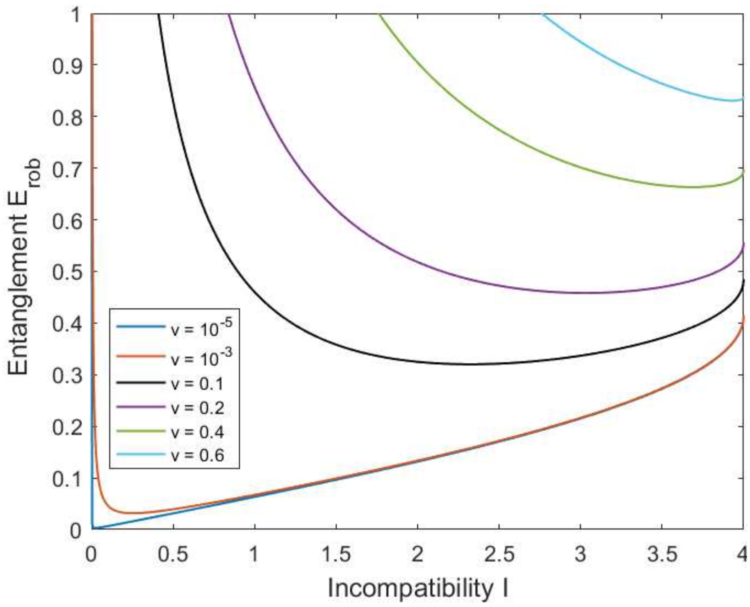

Fig. 4.4.: Entanglement versus incompatibility for the CHSH inequality from [64] (Paper B). We consider the minimal entanglement $\mathrm{E}_{\mathrm{rob}}$ a state has to contain to achieve the violation v, given the incompatibility I. Surprisingly, for every level of violation v, there is a region where more measurement resources also require higher entanglement. This effect is especially strong for small violations. The curves diverge, if there does not exist a state that achieves the violation for a given parameter I.

imal generalized robustness of purity necessary, independently of the dimension, number of measurements, outcomes, parties or the considered Bell operator. The generalized robustness of purity was previously shown [96] to be fully characterized by the largest eigenvalue $\lambda_1(\rho)$, as it holds $\mathrm{P}_{\mathrm{rob}}(\rho) = d\lambda_1(\rho) - 1$. Using the spectral decomposition (see Eq. (2.22)), we were able to determine the minimal $\lambda_1(\rho)$, such that the constraints in Eq. (4.190) hold.

Our second main result extends this finding to the resources of coherence with respect to product bases, discord, and entanglement for two-qubit full-correlation Bell inequalities, i.e., for inequalities with Bell operators of the form

$$F_{\mathrm{op}} = \sum_{x,y} g_{xy}\, A_x \otimes B_y. \tag{4.192}$$

Here, the g_{xy} are real-valued coefficients and $A_x = \vec{a}_x \cdot \vec{\sigma}$, respectively $B_y = \vec{b}_y \cdot \vec{\sigma}$ are the local observables of Alice and Bob, determined by the Bloch vectors $\vec{a}_x$ and $\vec{b}_y$. More precisely, we show by construction that there exists a quantum state ρ_{opt} that simultaneously minimizes all state resources under the constraint in Eq. (4.190). That is, we construct analytically from the eigenstates and eigenvalues of F_{op} the state ρ_{opt} that minimizes all required state resources for a given violation.

As a surprising application of our results, we show that for the CHSH inequality, there is no general trade-off between measurement and state resources. Consider the CHSH operator

$$F_{\text{CHSH}} = A_1 \otimes B_1 + A_1 \otimes B_2 + A_2 \otimes B_1 - A_2 \otimes B_2, \tag{4.193}$$

and the quantifier $\mathrm{I} = \mathrm{I}_A \mathrm{I}_B = \|[A_1, A_2]\|_\infty \cdot \|[B_1, B_2]\|_\infty$, where I_A and I_B are the measurement incompatibility of Alice's and Bob's observables. Note that the quantities I_A and I_B are related to the noise robustness of the underlying projective measurements, as it was shown in [213]. A large value of I allows for potentially large violations of the CHSH-inequality, as it determines the eigenvalues $\{\mu_i\}$ of F_{CHSH}, such that

$$\mu_{1/4} = \pm\sqrt{4 + \mathrm{I}}, \ \mu_{2/3} = \pm\sqrt{4 - \mathrm{I}}. \tag{4.194}$$

Now, one might intuitively think that a higher value of the measurement resource I allows for fewer state resources, while keeping the Bell value $\langle F_{\text{CHSH}} \rangle = L + v$ fixed. That this is actually not the case for the minimal necessary resources, is our final result as shown in Figure 4.4. For more details, see our work [64].

4.3 Distance-based Resource Quantification for Sets of Quantum Measurements (Paper C)

In our work [65], we provide a distance-based framework to quantify and compare the resources of measurement assemblages in any convex resource theory. We use this framework to introduce a hierarchy for measurement resources that is similar to the hierarchy for states presented in [61] (see also Section 3.2.4). We focus mainly on a specific resource monotone based on the diamond distance between measure-and-prepare channels. We show that said quantifier fulfills all required monotone conditions and enjoys many additional properties, such as continuity, an operational interpretation, and an efficient way to compute it utilizing SDPs. Based on this resource quantifier, we identify scenarios when specific measurement resources of a given measurement assemblage obtain the same value. For the resource theory of measurement incompatibility specifically, we derive general upper and lower bounds on the incompatibility of any measurement assemblage. Finally, we show that these bounds are tight for certain measurement assemblages based on MUB.

Our work extends the existing literature in several ways. First, our work provides the first method to use a proper distance (in the sense of a metric) as a resource quantifier for measurement assemblages. Previous methods relied chiefly on weight and robustness quantifiers to circumvent the problem of a missing metric for measurement assemblages (see, e.g., [40, 154–156, 169, 170, 187]). Especially properties such as continuity and the obeyed triangle inequality, which have not been available for previous quantifiers, could prove very useful in the future.

Second, we extend the hierarchy relations found in [187] and [161] and compare the resource hierarchies for states and measurements. Furthermore, we prove analytically instances in which the different resources of an assemblage coincide to an equal amount. Finally, for the resource theory of measurement incompatibility, we provide results that accompany previously established results [169, 171] and show that MUB lead to asymptotic (in the Hilbert space dimension d) maximal incompatibility among all rank-1 projective measurements.

Going more into the details, we first define contractive distances between *weighted measurement assemblages*. A measurement assemblage $\mathcal{M}$ is weighted by a probability distribution $\mathbf{p}$ that contains the probabilities $p(x)$ with which a particular measurement setting x is performed. This captures, for instance, the spirit of Bell and steering scenarios. A weighted measurement assemblage is simply defined as the tuple $\mathcal{M}_{\mathbf{p}} = (\mathcal{M}, \mathbf{p})$. Now, any distance (see Section 2.7.1) $\mathrm{D}(\mathcal{M}_{\mathbf{p}}, \mathcal{N}_{\mathbf{p}})$ between two weighted measurement assemblages $\mathcal{M}_{\mathbf{p}}, \mathcal{N}_{\mathbf{p}}$ is called contractive (see also Eq. (3.86)), if it holds that

$$\mathrm{D}(\mathcal{M}_{\mathbf{p}}, \mathcal{N}_{\mathbf{p}}) \geq \mathrm{D}(\Lambda^{\dagger}(\mathcal{M})_{\mathbf{p}}, \Lambda^{\dagger}(\mathcal{N})_{\mathbf{p}}), \tag{4.195a}$$

$$\mathrm{D}(\mathcal{M}_{\mathbf{p}}, \mathcal{N}_{\mathbf{p}}) \geq \mathrm{D}(\xi(\mathcal{M}_{\mathbf{p}}), \xi(\mathcal{N}_{\mathbf{p}})), \tag{4.195b}$$

for any CP unital quantum channel $\Lambda^{\dagger}$ and any classical measurement simulation ξ. Our first result shows that any contractive and jointly convex distance induces a faithful convex resource monotone for a resource theory $Q = (\mathcal{V}, \mathscr{F}, \mathbb{S})$, with a convex (and compact) free set $\mathcal{V}$, free operations $\Lambda^{\dagger} \in \mathscr{F}$, and free classical simulations $\xi \in \mathbb{S}$, via the construction

$$\mathrm{R}(\mathcal{M}_{\mathbf{p}}) \coloneqq \min_{\mathcal{F} \in \mathcal{V}} \mathrm{D}(\mathcal{M}_{\mathbf{p}}, \mathcal{F}_{\mathbf{p}}). \tag{4.196}$$

The quantifier we then focus on, is induced by the distance

$$\mathrm{D}_{\diamond}(\mathcal{M}_{\mathbf{p}}, \mathcal{N}_{\mathbf{p}}) \coloneqq \sum_{x} p(x)\, \mathrm{D}_{\diamond}(\Lambda_{\mathcal{M}_x}, \Lambda_{\mathcal{N}_x}), \tag{4.197}$$

where $\mathrm{D}_{\diamond}(\Lambda_1, \Lambda_2)$ is the diamond distance between two channels as introduced in Definition 2.7.5 and $\Lambda_{\mathcal{M}_x}(\rho) = \sum_a \mathrm{Tr}[M_{a|x}\rho] |a\rangle\langle a|$ is the measure-and-prepare channel associated to the POVM $\{M_{a|x}\}_a$. Due to the operational interpretation of the diamond distance, our resource quantifier $\mathrm{R}_{\diamond}(\mathcal{M}_{\mathbf{p}}) = \min_{\mathcal{F} \in \mathcal{V}} \mathrm{D}_{\diamond}(\mathcal{M}_{\mathbf{p}}, \mathcal{F}_{\mathbf{p}})$ has a similar operational interpretation in terms of hypothesis testing in a single-shot experiment.

As an example, the incompatibility of $\mathcal{M}_{\mathbf{p}}$ in our distance-based framework is now quantified by

$$\mathrm{I}_{\diamond}(\mathcal{M}_{\mathbf{p}}) = \min_{\mathcal{F}\in\mathcal{V}_{\mathrm{JM}}} \sum_x p(x)\, \mathrm{D}_{\diamond}(\Lambda_{\mathcal{M}_x}, \Lambda_{\mathcal{F}_x}). \tag{4.198}$$

On the other hand, the informativeness is now given by

$$\mathrm{IF}_{\diamond}(\mathcal{M}_{\mathbf{p}}) = \min_{\mathcal{F}\in\mathcal{V}_{\mathrm{Info}}} \sum_x p(x)\, \mathrm{D}_{\diamond}(\Lambda_{\mathcal{M}_x}, \Lambda_{\mathcal{F}_x}), \tag{4.199}$$

generalizing the informativeness of a single measurement to the average informativeness of the assemblage $\mathcal{M}$. Note that we generalize the coherence of a single POVM in the same way to the average coherence of an assemblage.

We show in our second result, based on the diamond distance quantifier $\mathrm{R}_{\diamond}(\mathcal{M}_{\mathbf{p}})$, that there exists a resource hierarchy for measurements, similar to the in [61] for quantum states. Namely, we show that the following chain of inequalities holds:

$$\mathrm{IF}_{\diamond}(\mathcal{M}_{\mathbf{p}_A}) \geq \mathrm{C}_{\diamond}(\mathcal{M}_{\mathbf{p}_A}) \geq \mathrm{I}_{\diamond}(\mathcal{M}_{\mathbf{p}_A}) \geq \mathrm{S}_1(\vec{\sigma}_{\mathbf{p}_A}) \geq \mathrm{N}_1(\mathbf{q}_{\mathbf{p}}), \tag{4.200}$$

where $\mathrm{IF}_{\diamond}(\mathcal{M}_{\mathbf{p}_A})$, $\mathrm{C}_{\diamond}(\mathcal{M}_{\mathbf{p}_A})$, and $\mathrm{I}_{\diamond}(\mathcal{M}_{\mathbf{p}_A})$ are the informativeness, the coherence, and the incompatibility of Alice's weighted measurement assemblage $\mathcal{M}_{\mathbf{p}_A} = (\mathcal{M}, \mathbf{p}_A)$. As mentioned above, this extends the notions of informativeness and coherence from one POVM (as introduced in the Sections 3.3.1 and 3.3.2) to an assemblage, by taking the average informativeness, respectively, coherence of the assemblage. Furthermore, $\mathrm{S}_1(\vec{\sigma}_{\mathbf{p}_A})$ is the steerability of any weighted steering assemblage $\vec{\sigma}_{\mathbf{p}_A} = (\sigma, \mathbf{p}_A)$ obtained from the weighted assemblage $\mathcal{M}_{\mathbf{p}_A}$ via $\sigma_{a|x} = \mathrm{Tr}_1[(M_{a|x} \otimes \mathbb{1})\rho]$ and quantified via the distance-based quantifier defined in Eq. (3.185). Finally, $\mathrm{N}_1(\mathbf{q}_{\mathbf{p}})$ with $\mathbf{q}_{\mathbf{p}} = (\mathbf{q}, \mathbf{p})$ is the nonlocality, as quantified by the monotone defined in Eq. (3.172), of a behavior $\mathbf{q}$ obtained via $q(ab|xy) = \mathrm{Tr}[(M_{a|x} \otimes N_{b|y})\rho]$, where $\mathcal{N}_{\mathbf{p}_B}$ is any weighted measurement assemblage on Bob's side and the distribution $\mathbf{p}$ is obtained via $p(x, y) = p_A(x)p_B(y)$.

The proof idea relies on the nested structure of the problem. For instance, all uninformative assemblages are also incoherent, as their POVM effects are proportional to the identity by definition. Further, incoherent measurements commute pairwise, i.e., they are also jointly measurable. Finally, we use that incompatibility is necessary for steering and that steering is necessary for Bell nonlocality. See also Figure 4.5 for a depiction of the proof idea.

In our work, we then ask in which instances the equalities $\mathrm{IF}_{\diamond}(\mathcal{M}_{\mathbf{p}_A}) = \mathrm{C}_{\diamond}(\mathcal{M}_{\mathbf{p}_A})$ and $\mathrm{I}_{\diamond}(\mathcal{M}_{\mathbf{p}_A}) = \mathrm{S}_1(\vec{\sigma}_{\mathbf{p}_A})$ can be obtained. We show that certain measurement assemblages based on MUB can be used to achieve this equality. Furthermore, we conjecture that the strict inequalities $\mathrm{C}_{\diamond}(\mathcal{M}_{\mathbf{p}_A}) > \mathrm{I}_{\diamond}(\mathcal{M}_{\mathbf{p}_A})$ and $\mathrm{S}_1(\vec{\sigma}_{\mathbf{p}_A}) > \mathrm{N}_1(\mathbf{q}_{\mathbf{p}})$ hold for non-trivial scenarios, i.e., whenever the involved quantities are non-zero. For more details see our work [65].

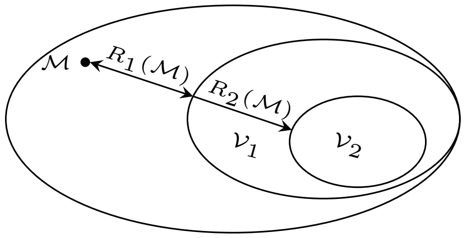

Fig. 4.5.: Nested structure of measurement resources from [65] (Paper C). Part of the proof idea of the hierarchy in Eq. (4.200) relies on the nested structure of the respective free sets. Consider two resources associated to the free sets $\mathcal{V}_1$ and $\mathcal{V}_2$. If it holds that $\mathcal{V}_2 \subset \mathcal{V}_1$, it follows directly that $\mathrm{R}_2(\mathcal{M}) \geq \mathrm{R}_1(\mathcal{M})$ for the resources of the measurement assemblage $\mathcal{M}$, in the case that the respective resource monotones are induced by the same distance.

Finally, based on the SDP formulation of the quantifier $\mathrm{I}_\diamond(\mathcal{M}_\mathbf{p})$, we derive general upper and lower bounds on the incompatibility of any measurement assemblage in dimension d that is weighted with a distribution $\mathbf{p}$, such that $p(x) = \frac{1}{m}$, where m is the number of measurements. More formally, we prove that the following bounds hold:

$$\mathrm{I}_\diamond(\mathcal{M}_\mathbf{p}) \geq \frac{1}{md}\sum_{a,x}\mathrm{Tr}[M_{a|x}^2] - \frac{1}{m}\Big(\max_{a,x}\|M_{a|x}\|_\infty + \tag{4.201a}$$

$$(m-1)\max_{a,a',x,x'\neq x}\|M_{a|x}^{1/2}M_{a'|x'}^{1/2}\|_\infty\Big),$$

$$\mathrm{I}_\diamond(\mathcal{M}_\mathbf{p}) \leq \frac{m-1}{(d+1)m^2}\sum_x\|d\mathbb{1} - \sum_a \mathrm{Tr}[M_{a|x}]M_{a|x}\|_\infty. \tag{4.201b}$$

In the special case of measurements based on MUB and in the limit $d \to \infty$ both bounds collapse to $\mathrm{I}_\diamond(\mathcal{M}_\mathbf{p}) \approx 1 - \frac{1}{m}$. Now, since the upper bound is completely general, this shows that MUB lead to maximally incompatible assemblages of m measurements among all measurement assemblages consisting of rank-1 projective measurements in the limit of large dimensions.

Indeed, notice that the upper bound in Eq. (4.201b) is the same for all measurements with $\mathrm{Tr}[M_{a|x}] = 1$ for all a, x. This includes, in particular, all rank-1 projective measurements. Therefore, we conclude that MUB lead asymptotically (for $d \to \infty$) to maximally incompatible assemblages of m measurements among all assemblages consisting of rank -1 projective measurements.

4.4 Distribution of Quantum Incompatibility Across Subsets of Measurements (Paper D)

In this work [66], we continue to study measurement resources by focusing specifically on measurement incompatibility. We analyze how the incompatibility of a measurement assemblage with $m > 2$ settings depends on the incompatibility of subsets of its measurements. In particular, we are interested in the incompatibility gain, i.e., the amount of incompatibility obtained by adding further measurements to an existing measurement scheme.

We show that this incompatibility gain is bounded by the incompatibility of the parent POVMs that are associated with the closest jointly measurable assemblages. More precisely, the incompatibility of these different parent POVMs (associated to the different subsets) with each other, sets a limit on the possible incompatibility gain. More generally, we show how to bound the incompatibility of an assemblage with multiple measurements using the incompatibilities of subsets of measurements. We also analyze how to decompose the total incompatibility of an assemblage into different components associated with different incompatibility structures [173], such as pairwise or genuine triplewise incompatibility (see also Section 3.3.3). We provide tight examples for most of our bounds by using measurements based on MUB to prove the relevance of our bounds. Finally, we show how to apply our methods to steering and Bell nonlocality.

Our work promises further advance of scientific research in many directions, as it provides new tools to find optimal measurement resources and illustrates potentially crucial differences between nonlocality on the one hand and measurement incompatibility on the other hand. Furthermore, our work provides a foundation for comparing the power of specific protocols using a different number of measurements, such as the six-state protocol [214] and the BB84 protocol [23].

Going more into the details, we write a measurement assemblage within this work as an ordered list $\mathcal{M}_{(1,2,\cdots,m)} = (\mathcal{M}_1, \mathcal{M}_2, \cdots, \mathcal{M}_m)$ of m POVMs $\mathcal{M}_1$ to $\mathcal{M}_m$. Focusing on the case $m = 3$, this allows us to write an assemblage $\mathcal{M}_{(1,2,3)}$ as $\mathcal{M}_{(1,2,3)} = (\mathcal{M}_1, \mathcal{M}_2, \mathcal{M}_3)$. On the other hand, it also allows us to consider assemblages $\mathcal{M}_{(1,2,2)} = (\mathcal{M}_1, \mathcal{M}_2, \mathcal{M}_2)$ where the second and the third measurement are equal. Adding a measurement $\mathcal{M}_3$ to an assemblage $\mathcal{M}_{(1,2)} = (\mathcal{M}_1, \mathcal{M}_2)$ is now described by the concatenation $+\!\!\!+$ of ordered lists. That is, $\mathcal{M}_{(1,2,3)}$ is given such that

$$\mathcal{M}_{(1,2,3)} = \mathcal{M}_{(1,2)} +\!\!\!+ \mathcal{M}_3 = (\mathcal{M}_1, \mathcal{M}_2, \mathcal{M}_3). \tag{4.202}$$

The two main questions that we try to answer are the following: How much incompatibility can be gained by concatenating the third measurement to $\mathcal{M}_{(1,2)}$? How does the incompatibility of $\mathcal{M}_{(1,2,3)}$ depend quantitatively on the incompatibility of its subsets?

The incompatibility quantifier which is best suited to answer these questions is the quantifier based on the diamond distance, as introduced in our work [65] (Paper C), see also Eq. (4.198). That means the incompatibility of a weighted measurement assemblage $\mathcal{M}_{\mathbf{p}} = (\mathcal{M}, \mathbf{p})$ (see also Section 4.3) is given by

$$\mathrm{I}_\diamond(\mathcal{M}_{\mathbf{p}}) = \min_{\mathcal{F} \in \mathrm{JM}} \sum_x p(x)\, \mathrm{D}_\diamond(\Lambda_{\mathcal{M}_x}, \Lambda_{\mathcal{F}_x}), \tag{4.203}$$

where, as before, $\Lambda_{\mathcal{M}_x}(\rho) = \sum_a \mathrm{Tr}[M_{a|x}\rho] |a\rangle\langle a|$ is the measure-and-prepare channel associated to the POVM $\mathcal{M}_x$ and $\mathrm{D}_\diamond(\Lambda_1, \Lambda_2)$ is the diamond distance between two channels Λ_1 and Λ_2. In order to be closer to the original work, we use the notation JM instead of $\mathcal{V}_{\mathrm{JM}}$ for the set of jointly measurable assemblages within this section.

The diamond distance quantifier is especially well-suited for answering our questions, as it inherits all metric properties and is written as a convex combination over individual settings. For the sake of simplicity, we consider here the scenario that $\mathbf{p} = \{p(x)\}$ is always uniformly distributed, i.e., $p(x) = \frac{1}{m}\ \forall\, x$ and we simply use the symbol $\mathcal{M}$ for the weighted assemblage in that case. However, our results [66] are general, i.e., they can be adapted to general distributions $\mathbf{p}$.

We denote by $\mathcal{M}^{\#}_{(1,2,\cdots,m)}$ the closest jointly measurable approximation of $\mathcal{M}_{(1,2,\cdots,m)}$, i.e., the *arg-min* on the RHS in Eq.(4.203). Moreover, we denote by $\mathcal{M}^{\#(1,2,\cdots,n)}_{(1,2,\cdots,m)} = \mathcal{M}^{\#}_{(1,2,\cdots,n)} \mathbin{+\!\!\!+} \mathcal{M}_{n+1} \mathbin{+\!\!\!+} \cdots \mathbin{+\!\!\!+} \mathcal{M}_m$ an assemblage in which the subset of first $n < m$ measurements are replaced by their respective closest jointly measurable assemblage.

As depicted in Figure 4.6, $m = 3$ measurements allow for more incompatibility structures beyond standard incompatibility. For instance, the sets $\mathrm{JM}^{(s,t)}$ with $s, t \in \{1, 2, 3\}$ and $s \neq t$ contain assemblages in which the pair (s, t) are jointly measurable. Their intersection $\mathrm{JM}^{\mathrm{pair}} := \mathrm{JM}^{(1,2)} \cap \mathrm{JM}^{(1,3)} \cap \mathrm{JM}^{(2,3)}$ contains all pairwise compatible measurements, and, as a proper subset, the set JM. Moreover, $\mathrm{JM}^{\mathrm{conv}} := \mathrm{Conv}(\mathrm{JM}^{(1,2)}, \mathrm{JM}^{(1,3)}, \mathrm{JM}^{(2,3)})$ denotes the convex hull of the sets $\mathrm{JM}^{(1,2)}$, $\mathrm{JM}^{(1,3)}$, and $\mathrm{JM}^{(2,3)}$. Finally, as defined in [173] and also explained in Section 3.3.3, assemblages $\mathcal{M}_{(1,2,3)} \notin \mathrm{JM}^{\mathrm{conv}}$ are called genuine triplewise incompatible.

We define the incompatibility gain from adding the third measurement $\mathcal{M}_3$ to the assemblage $\mathcal{M}_{(1,2)}$ as

$$\Delta \mathrm{I}_{(1,2)\to(1,2,3)} := \mathrm{I}_\diamond(\mathcal{M}_{(1,2,3)}) - \mathrm{I}_\diamond(\mathcal{M}_{(1,2)}). \tag{4.204}$$

Our main objective is to find upper bounds on the gained incompatibility, i.e., to understand how much resources can maximally be gained. Using the triangle inequality with respect to JM and exploiting the fact that $\mathrm{I}_\diamond(\mathcal{M}_{(1,2,3)}) = \mathrm{I}_\diamond(\mathcal{M}_{(1,2,1,3,2,3)})$, where each of the measurements is now treated as two copies of itself, occurring with probability $p(x) = \frac{1}{6}$, we obtain our first main result.

Namely, assuming that $\mathrm{I}_\diamond(\mathcal{M}_{(1,2)}) \geq \max\{\mathrm{I}_\diamond(\mathcal{M}_{(1,3)}), \mathrm{I}_\diamond(\mathcal{M}_{(2,3)})\}$ the incompatibility gain is bounded such that

$$\Delta \mathrm{I}_{(1,2)\to(1,2,3)} \leq \mathrm{I}_\diamond(\mathcal{N}) \leq \mathrm{I}_\diamond(\mathcal{G}), \tag{4.205}$$

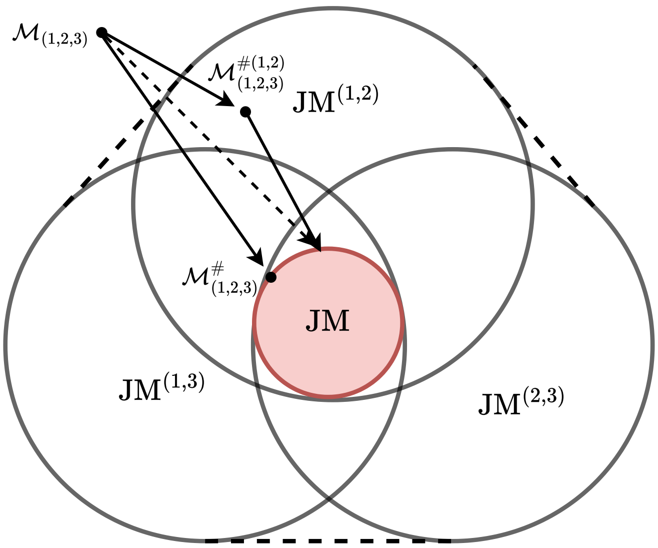

Fig. 4.6.: Incompatibility structures for $m = 3$ measurements from [66] (Paper D). Three measurements allow for incompatibility structures, such as the sets $\mathrm{JM}^{(s,t)}$ in which the measurements s and t are compatible. Their intersection, the set $\mathrm{JM}^{\mathrm{pair}}$ of pairwise jointly measurable assemblages, contains as a proper subset the set of all jointly measurable assemblages JM. Assemblages $\mathcal{M}_{(1,2,3)}$ that are not contained in the convex hull $\mathrm{JM}^{\mathrm{conv}}$ of $\mathrm{JM}^{(1,2)}$, $\mathrm{JM}^{(1,3)}$, and $\mathrm{JM}^{(2,3)}$ are genuinely triplewise incompatible. Using the triangle inequality with respect to JM via the assemblage $\mathcal{M}^{\#(1,2)}_{(1,2,3)}$, it is possible to upper bound the incompatibility of $\mathcal{M}_{(1,2,3)}$.

where $\mathcal{N} = \mathcal{M}^{\#}_{(1,2)} \mathbin{+\!\!\!+} \mathcal{M}^{\#}_{(1,3)} \mathbin{+\!\!\!+} \mathcal{M}^{\#}_{(2,3)}$ is the assemblage that contains the closest jointly measurable approximations of each of the two-measurement subsets of $\mathcal{M}_{(1,2,3)}$. Moreover, $\mathcal{G} = G(\mathcal{M}^{\#}_{(1,2)}) \mathbin{+\!\!\!+} G(\mathcal{M}^{\#}_{(1,3)}) \mathbin{+\!\!\!+} G(\mathcal{M}^{\#}_{(2,3)})$ is the assemblage that contains the three parent POVMs associated to these closest jointly measurable approximations. Note that $\mathrm{I}_\diamond(\mathcal{N}) \leq \mathrm{I}_\diamond(\mathcal{G})$ holds since $\mathcal{N}$ is a classical simulation (see Eq. (3.148)) of $\mathcal{G}$ by the definition of a parent POVM (see Eq. (3.144)). Similarly, $\mathrm{I}_\diamond(\mathcal{M}_{(1,2,3)}) = \mathrm{I}_\diamond(\mathcal{M}_{(1,2,1,3,2,3)})$ holds as $\mathcal{M}_{(1,2,3)}$ can simulate $\mathcal{M}_{(1,2,1,3,2,3)}$ classically and vice versa. Note further that the assumption $\mathrm{I}_\diamond(\mathcal{M}_{(1,2)}) \geq \max\{\mathrm{I}_\diamond(\mathcal{M}_{(1,3)}), \mathrm{I}_\diamond(\mathcal{M}_{(2,3)})\}$ is not relevant for all practical purposes, as one can relabel the measurements such that $\mathcal{M}_{(1,2)}$ contains the maximal incompatibility among the different subsets.

Our result reveals the polygamous nature of measurement incompatibility: for a high incompatibility of $\mathcal{M}_{(1,2,3)}$, each of the subsets, as well as the assemblage containing the parent POVMs associated to the respective closest jointly measurable approximations, have to be highly incompatible. Moreover, we show in our work [66] that the first bound in Eq. (4.205) is tight for noisy Pauli measurements.

Besides bounding the incompatibility gain, we also derive more general bounds on the incompatibility of $\mathcal{M}_{(1,2,3)}$ using information about the subsets. That is, we derive the bound

$$\mathrm{I}_\diamond(\mathcal{M}_{(1,2,3)}) \leq \frac{2}{3}\,\mathrm{I}_\diamond(\mathcal{M}_{(1,2)}) + \mathrm{I}_\diamond(\mathcal{M}^{\#(1,2)}_{(1,2,3)}). \tag{4.206}$$

This means, we show that $\mathrm{I}_\diamond(\mathcal{M}_{(1,2,3)})$ is upper bounded by the incompatibility of $\mathcal{M}_{(1,2)}$ weighted by the probability $p = \frac{2}{3}$ that one of the measurements $(1,2)$ is chosen among the available measurements $(1,2,3)$ plus the incompatibility $\mathrm{I}_\diamond(\mathcal{M}^{\#(1,2)}_{(1,2,3)})$. The incompatibility of $\mathcal{M}^{\#(1,2)}_{(1,2,3)}$ can be understood as a new notion of incompatibility of $\mathcal{M}_{(1,2,3)}$. That is, all other incompatibility contributions apart those that are present due to the measurement $\mathcal{M}_3$ are omitted since $\mathcal{M}^{\#}_{(1,2)}$ is jointly measurable by itself. Hence, it gives us an idea of the contribution of $\mathcal{M}_3$ toward the whole incompatibility of $\mathcal{M}_{(1,2,3)}$. We show [66] that Eq. (4.206) can straightforwardly be generalized to

$$\mathrm{I}_\diamond(\mathcal{M}_{(1,2,\cdots,m)}) \leq \frac{|C|}{m}\,\mathrm{I}_\diamond(\mathcal{M}_C) + \mathrm{I}_\diamond(\mathcal{M}^{\#C}_{1,2,\cdots,m)}), \tag{4.207}$$

for any assemblage $\mathcal{M}_{(1,2,\cdots,m)}$ with a subset of measurements C of cardinality $|C|$. Moreover, we prove the tightness of Eq. (4.206) analytically for noisy Pauli measurements and also provide analytical proofs for more general scenarios covered by Eq. (4.207) based on MUB in any dimension d. Besides this, we also discuss generalizations of Eq. (4.205) in our work.

Moreover, we also show that our methods and results can be adapted to any probability distribution $\mathbf{p}$ that weights the assemblage $\mathcal{M}$ and also derive lower bounds on the incompatibility of any assemblage in terms of its subsets' incompatibility.

We also introduce a bound that uses the decomposition of the total incompatibility $\mathrm{I}_\diamond(\mathcal{M}_{(1,2,3)})$ in terms of its different incompatibility components. More precisely, we show that

$$\mathrm{I}_\diamond(\mathcal{M}_{(1,2,3)}) \leq \mathrm{I}_\diamond{}^{\mathrm{gen}}(\mathcal{M}_{(1,2,3)}) + \mathrm{I}_\diamond{}^{\mathrm{pair}}(\mathcal{M}_{(1,2,3)}) + \mathrm{I}_\diamond{}^{\mathrm{hol}}(\mathcal{M}_{(1,2,3)}), \tag{4.208}$$

with the genuine triplewise incompatibility $\mathrm{I}_\diamond{}^{\mathrm{gen}}(\mathcal{M}_{(1,2,3)})$, the pairwise incompatibility $\mathrm{I}_\diamond{}^{\mathrm{pair}}(\mathcal{M}_{(1,2,3)})$ and the hollow incompatibility $\mathrm{I}_\diamond{}^{\mathrm{hol}}(\mathcal{M}_{(1,2,3)})$. Note that we define $\mathrm{I}_\diamond{}^{\mathrm{gen}}(\mathcal{M}_{(1,2,3)})$ to be the minimal distance of $\mathcal{M}_{(1,2,3)}$ to an assemblage $\mathcal{M}^{\mathrm{conv}}_{(1,2,3)} \in \mathrm{JM}^{\mathrm{conv}}$. Similarly, we define $\mathrm{I}_\diamond{}^{\mathrm{pair}}(\mathcal{M}_{(1,2,3)})$ as the minimal distance of $\mathcal{M}^{\mathrm{conv}}_{(1,2,3)}$ to an assemblage $\mathcal{M}^{\mathrm{pair}}_{(1,2,3)} \in \mathrm{JM}^{\mathrm{pair}}$ and $\mathrm{I}_\diamond{}^{\mathrm{hol}}(\mathcal{M}) := \mathrm{I}_\diamond(\mathcal{M}^{\mathrm{pair}}_{(1,2,3)})$. Again, we show in our work [66] analytically that the decomposition in Eq. (4.208) is tight for Pauli measurements and give numerical evidence that this is generally the case for measurements based on MUB for any dimension $d > 2$.

Let us finally discuss how to apply our results and methods to the Bell nonlocality that could be gained from increasing the number of measurements in a Bell experiment. First, due to the similarities between quantum steering and incompatibility, our results can also be adapted directly to the steering distance introduced in [207] and also discussed in Eq. (3.185).

For the bounds on the nonlocality of a distribution $\mathbf{q}_{(1,2,3)} = \{q(ab|xy)\}$ in which Alice uses $m_A = 3$ measurements, we use a variant of the nonlocality distance introduced in [193] (see also Eq. (3.172)). That is, we quantify the nonlocality via the quantifier

$$\mathrm{N}_1(\mathbf{q}) = \frac{1}{2} \min_{\mathbf{t} \in \mathcal{CL}} \sum_{a,b,x,y} \frac{1}{m_A m_B} |q(a,b|x,y) - t(a,b|x,y)|, \tag{4.209}$$

where we set the input probabilities $p(x,y) = \frac{1}{m_A m_B}$ to be uniformly distributed and we assume $m_B = 2$ for simplicity in the following. Moreover $\mathbf{t} \in \mathcal{CL}$ denotes a local probability distribution that is consistent with $\mathbf{q}$, i.e., $\mathbf{t}$ has the same marginal distributions for Alice and Bob as $\mathbf{q}$. Using the same methods as for the incompatibility before, we derive the bound

$$\mathrm{N}_1(\mathbf{q}_{(1,2,3)}) \leq \frac{1}{3}[\mathrm{N}_1(\mathbf{q}_{(1,2)}) + \mathrm{N}_1(\mathbf{q}_{(1,3)}) + \mathrm{N}_1(\mathbf{q}_{(2,3)})] + \mathrm{N}_1(\mathbf{r}), \tag{4.210}$$

where $\mathbf{q}_{(1,2)}$,$\mathbf{q}_{(1,3)}$, and $\mathbf{q}_{(2,3)}$ are the respective distributions when Alice uses only the measurements (s,t). Moreover $\mathbf{r} = \mathbf{q}^{\#}_{(1,2)} \mathbin{+\!\!+} \mathbf{q}^{\#}_{(1,3)} \mathbin{+\!\!+} \mathbf{q}^{\#}_{(2,3)}$ is a probability distribution in which Alice has 6 measurements, containing the closest local-consistent distributions to these subset distributions. Surprisingly, the bound on the nonlocality in Eq. (4.210) shows a fundamental difference between nonlocality and measurement incompatibility or steering. While all two-measurement subsets could be maximally incompatible for an assemblage of $m = 3$ measurements, this is generally impossible for the corresponding nonlocality terms. That is, it generally holds that $\frac{1}{3}[\mathrm{N}_1(\mathbf{q}_{(1,2)}) + \mathrm{N}_1(\mathbf{q}_{(1,3)}) + \mathrm{N}_1(\mathbf{q}_{(2,3)})] < \mathrm{N}_1(\mathbf{q}^{\max})$, where $\mathbf{q}^{\max}$ is the distribution that contains the maximal possible nonlocality using two measurements on Alice's side. The reason for this is that there generally do not exist $m = 3$ measurements for Alice, out of which every pair of two measurements can be used to violate a Bell inequality maximally while the state and Bob's measurements remain unchanged. We show this explicitly for dichotomic measurements by showing that there does not exist a combination of three versions of the CHSH inequality involving two of the three measurements, such that they are all maximally violated simultaneously. Now crucially, in the case of two inputs and outputs for Alice and Bob $\mathrm{N}_1(\mathbf{q})$ is directly linked to the violation of a CHSH inequality [193]. The insight that the upper bounds on nonlocality behave qualitatively different than for measurement incompatibility could prove helpful for future applications, as it is an open question [193, 215] whether increasing the number of measurements provides any advantage for the maximal nonlocality in a Bell test at all.

5 Conclusion and Outlook

This thesis was devoted to the interplay of different quantum resources and their impact on quantum nonlocality. With the advent of more and more technologies based upon quantum information, it is crucial to understand which properties of quantum systems are essential for a *quantum advantage* over conventional technologies. The field of quantum resource theories [31] is particularly concerned with understanding and quantifying those properties of quantum systems, the quantum resources, that lead to an advantage in quantum information processing tasks. Due to the variety of different quantum phenomena that play a role as a quantum resource, it is essential to understand how they influence each other and to find quantitative dependencies among them.

The goal of this thesis and the preceding doctoral research [63–66] was to solidify further and extend this understanding of different quantum resources. We went two different routes in doing so. On the one hand, we studied some peculiar effects of the interplay of quantum resources, showing the counter-intuitive behavior of quantum theory. On the other hand, we developed general frameworks for quantifying quantum resources and understanding minimal required resources for a certain level of Bell nonlocality.

From the side of peculiar quantum effects, we showed that there exists bound entangled states that admit a local model for all measurements in the standard Bell scenario. Nevertheless, they display nonlocality in a sequential Bell scenario [63]. That shows that even very weakly entangled states, like bound entangled states that appear to provide no advantage for obtaining nonlocal correlations, can be used as a quantum resource if one uses them in the appropriate scenario.

Moreover, we showed that for the CHSH inequality, despite maximal entanglement and maximally incompatible measurements being necessary for a maximal violation of it, there is, in general, no trade-off between measurement and state resources [64]. We analytically proved that increasing the available measurement resource generally does not allow fewer state resources to achieve a fixed violation of the CHSH inequality. Contrary to our intuition, it is sometimes necessary to increase the state resources if the measurement resources, in terms of their incompatibility, are increased while keeping the targeted violation constant. That shows that more entangled states and more incompatible measurements do not always result in stronger quantum nonlocality.

In our work [64], we also provided general results concerning the minimal purity that is necessary to achieve a certain violation of any Bell inequality with given measurements. Our insights from the resource theory of purity and the use of hierarchical structures of quantum resources [61] allowed us to prove that a simultaneous minimization of quantum state resources is possible in the context of two-qubit full correlation Bell inequalities and any level of Bell violation.

Following the lines of resource quantification, we developed in our work [65] a general distance-based framework to quantify resources of quantum measurements in any convex resource theory. We established the notion of distances between measurement assemblages and focused on the diamond distance, which has a clear operational interpretation in terms of single-shot distinguishability. Based on the diamond distance, we established a hierarchy of measurement resources, including quantum steering and Bell nonlocality, that establishes a parallel to the hierarchy [61] from the measurement side. Furthermore, our work provides insights into specific resource theories, such as measurement incompatibility [39]. In particular, we obtained bounds for the incompatibility of any measurement assemblage. We showed that measurements based on MUB take a special role in the resource theory of incompatibility because they are asymptotically (in the Hilbert space dimension d) maximally incompatible.

Our final work [66] concerned the constraints that the incompatibility of subsets of measurements set on the whole assemblage. We showed that our distance-based way of quantifying incompatibility allows for bounding the incompatibility of an assemblage, using the information of the subset incompatibilities. This allowed us to bound the incompatibility that can be gained from adding measurements to an existing measurement scheme. Moreover, we showed that the incompatibility of any assemblage can be upper bound by decomposing its total incompatibility in different components, regarding incompatibility structures like pairwise and genuine triplewise incompatibility. Finally, we showed that our methods also apply to steering and nonlocality and discussed the implications for these resources.

Summing up, this thesis provides a general overview of (many of) the essential quantum resources for quantum nonlocality and discusses how their interplay leads to nonlocal correlations. In addition to the publications, this thesis provides a general framework to apply the machinery of quantum resource theories to more quantum phenomena in the future.

Despite the results in this thesis, there are many directions for future research. Especially the analysis of measurement resources offers a variety of different paths. For instance, our framework of distance-based resource quantification could be applied to quantum resources like projective simulability [42] or to a resource theory of imaginarity [216] that has to be developed for quantum measurements in the first place. Also, the nonlocality revealing properties of quantum measurements, i.e., which measurements can lead to a Bell inequality violation, needs more structural analysis beyond the results in [54–56]. That also includes the possibility of analyzing measurement resources in sequential Bell scenarios.

For the resource theory of measurement incompatibility, it is also necessary to study and quantify more complex incompatibility structures, i.e., incompatible assemblages where subsets of measurements are jointly measurable [173, 217–219]. In particular, a good characterization of sets like pairwise (n-wise) jointly measurable assemblages could prove useful in the future.

An even more promising route to study quantum measurement resources lies in considering measurements on more than one system. In this work and most of the literature, resources like measurement incompatibility [39] are considered as a local resource, i.e., acting on a single Hilbert space. However, considering measurements on more systems would require new notions of measurement incompatibility, similar to multipartite steering [44, 45]. Due to the advent of quantum networks [220–230] as a technology for communication between distant parties, this is a promising path in understanding the measurement resources that are necessary for desired nonclassical effects. In a sense, measurements of joint systems combine effects from entanglement and incompatibility theory. It would be interesting to study the notions of nonclassicality that arise in such scenarios and how different quantum phenomena interplay.

Bibliography

[1] *The Nobel Prize in Physics 2022. NobelPrize.org. Nobel Prize Outreach AB 2022.* https://www.nobelprize.org/prizes/physics/2022/summary/, Accessed: 2022-10-19 (cit. on pp. 1, 79).

[2] J. Preskill, Quantum computing 40 years later, https://doi.org/10.48550/arXiv.2106.10522 (2021) (cit. on p. 1).

[3] A. Steane, Quantum computing, Rep. Prog. Phys. **61**, 117 (1998) (cit. on p. 1).

[4] T. D. Ladd, F. Jelezko, R. Laflamme, Y. Nakamura, C. Monroe, and J. L. O'Brien, Quantum computers, Nature **464**, 45 (2010) (cit. on p. 1).

[5] *40 years of quantum computing*, https://www.nature.com/collections/djhfiebiig/, Accessed: 2022-11-09 (cit. on p. 1).

[6] C. L. Degen, F. Reinhard, and P. Cappellaro, Quantum sensing, Rev. Mod. Phys. **89**, 035002 (2017) (cit. on p. 1).

[7] N. Gisin and R. Thew, Quantum communication, Nature Photonics **1**, 165 (2007) (cit. on pp. 1, 59).

[8] S. Pirandola, U. L. Andersen, L. Banchi, et al., Advances in quantum cryptography, Adv. Opt. Photon. **12**, 1012 (2020) (cit. on pp. 1, 59).

[9] G. Murta, F. Grasselli, H. Kampermann, and D. Bruß, Quantum Conference Key Agreement: A Review, Advanced Quantum Technologies **3**, 2000025 (2020) (cit. on p. 1).

[10] A. Einstein, B. Podolsky, and N. Rosen, Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?, Phys. Rev. **47**, 777 (1935) (cit. on p. 1).

[11] A. Einstein, Quanten-Mechanik und Wirklichkeit, Dialectica **2**, 320 (1948) (cit. on pp. 1, 88).

[12] E. Schrödinger, Discussion of Probability Relations between Separated Systems, Mathematical Proceedings of the Cambridge Philosophical Society **31**, 555 (1935) (cit. on pp. 1, 88).

[13] E. Schrödinger, Probability relations between separated systems, Mathematical Proceedings of the Cambridge Philosophical Society **32**, 446 (1936) (cit. on pp. 1, 88).

[14] W. Heisenberg, Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik, Zeitschrift für Physik **43**, 172 (1927) (cit. on p. 1).

[15]J. S. Bell, On the Einstein Podolsky Rosen paradox, Physics Physique Fizika **1**, 195 (1964) (cit. on pp. 1, 3, 79, 81).

[16]S. J. Freedman and J. F. Clauser, Experimental Test of Local Hidden-Variable Theories, Phys. Rev. Lett. **28**, 938 (1972) (cit. on p. 1).

[17]A. Aspect, J. Dalibard, and G. Roger, Experimental Test of Bell's Inequalities Using Time-Varying Analyzers, Phys. Rev. Lett. **49**, 1804 (1982) (cit. on p. 1).

[18]W. Rosenfeld, D. Burchardt, R. Garthoff, K. Redeker, N. Ortegel, M. Rau, and H. Weinfurter, Event-Ready Bell Test Using Entangled Atoms Simultaneously Closing Detection and Locality Loopholes, Phys. Rev. Lett. **119**, 010402 (2017) (cit. on p. 1).

[19]M. Giustina, M. A. M. Versteegh, S. Wengerowsky, et al., Significant-Loophole-Free Test of Bell's Theorem with Entangled Photons, Phys. Rev. Lett. **115**, 250401 (2015) (cit. on p. 1).

[20]B. Hensen, H. Bernien, A. E. Dréau, et al., Loophole-free Bell inequality violation using electron spins separated by 1.3 kilometres, Nature **526**, 682 (2015) (cit. on p. 1).

[21]L. K. Shalm, E. Meyer-Scott, B. G. Christensen, et al., Strong Loophole-Free Test of Local Realism, Phys. Rev. Lett. **115**, 250402 (2015) (cit. on p. 1).

[22]D. Bouwmeester, J.-W. Pan, K. Mattle, M. Eibl, H. Weinfurter, and A. Zeilinger, Experimental quantum teleportation, Nature **390**, 575 (1997) (cit. on p. 2).

[23]C. H. Bennett and G. Brassard, Quantum cryptography: Public key distribution and coin tossing, Theoretical Computer Science **560**, 7 (2014) (cit. on pp. 2, 104).

[24]C. H. Bennett, G. Brassard, and N. D. Mermin, Quantum cryptography without Bell's theorem, Phys. Rev. Lett. **68**, 557 (1992) (cit. on p. 2).

[25]A. K. Ekert, Quantum cryptography based on Bell's theorem, Phys. Rev. Lett. **67**, 661 (1991) (cit. on pp. 2, 59).

[26]A. Acín, N. Brunner, N. Gisin, S. Massar, S. Pironio, and V. Scarani, Device-Independent Security of Quantum Cryptography against Collective Attacks, Phys. Rev. Lett. **98**, 230501 (2007) (cit. on pp. 2, 79).

[27]R. Colbeck, Quantum And Relativistic Protocols For Secure Multi-Party Computation, `https://doi.org/10.48550/arXiv.0911.3814` (2009) (cit. on pp. 2, 79).

[28]S. Pironio, A. Acín, S. Massar, A. B. de la Giroday, D. N. Matsukevich, P. Maunz, S. Olmschenk, D. Hayes, L. Luo, T. A. Manning, and C. Monroe, Random numbers certified by Bell's theorem, Nature **464**, 1021 (2010) (cit. on pp. 2, 79).

[29]R. Colbeck and A. Kent, Private randomness expansion with untrusted devices, J. Phys. A: Math. Theor. **44**, 095305 (2011) (cit. on pp. 2, 79).

[30]P. W. Shor, Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer, SIAM Journal on Computing **26**, 1484 (1997) (cit. on pp. 2, 59).

[31]E. Chitambar and G. Gour, Quantum resource theories, Rev. Mod. Phys. **91**, 025001 (2019) (cit. on pp. 2, 3, 35, 46, 48, 55, 109).

[32]R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, Quantum entanglement, Rev. Mod. Phys. **81**, 865 (2009) (cit. on pp. 2, 59, 60, 65, 66, 97).

[33] J. I. de Vicente, On nonlocality as a resource theory and nonlocality measures, J. Phys. A: Math. Theor. **47**, 424017 (2014) (cit. on pp. 2, 86, 88).

[34] R. Gallego and L. Aolita, Resource Theory of Steering, Phys. Rev. X **5**, 041008 (2015) (cit. on pp. 2, 51, 94).

[35] A. Streltsov, G. Adesso, and M. B. Plenio, Colloquium: Quantum coherence as a resource, Rev. Mod. Phys. **89**, 041003 (2017) (cit. on pp. 2, 3, 55–57).

[36] A. Bera, T. Das, D. Sadhukhan, S. S. Roy, A. Sen(De), and U. Sen, Quantum discord and its allies: a review of recent progress, Rep. Prog. Phys. **81**, 024001 (2017) (cit. on pp. 2, 3, 58).

[37] V. Veitch, S. A. H. Mousavian, D. Gottesman, and J. Emerson, The resource theory of stabilizer quantum computation, New Journal of Physics **16**, 013009 (2014) (cit. on p. 2).

[38] H. Ollivier and W. H. Zurek, Quantum Discord: A Measure of the Quantumness of Correlations, Phys. Rev. Lett. **88**, 017901 (2001) (cit. on pp. 2, 58).

[39] O. Gühne, E. Haapasalo, T. Kraft, J.-P. Pellonpää, and R. Uola, Incompatible measurements in quantum information science, `https://doi.org/10.48550/arXiv.2112.06784` (2021) (cit. on pp. 2, 73, 76, 89, 110, 111).

[40] P. Skrzypczyk and N. Linden, Robustness of Measurement, Discrimination Games, and Accessible Information, Phys. Rev. Lett. **122**, 140403 (2019) (cit. on pp. 2, 3, 68–70, 100).

[41] M. Oszmaniec and T. Biswas, Operational relevance of resource theories of quantum measurements, Quantum **3**, 133 (2019) (cit. on pp. 2, 3, 68, 70–72).

[42] M. Oszmaniec, L. Guerini, P. Wittek, and A. Acín, Simulating Positive-Operator-Valued Measures with Projective Measurements, Phys. Rev. Lett. **119**, 190501 (2017) (cit. on pp. 2, 73, 110).

[43] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner, Bell nonlocality, Rev. Mod. Phys. **86**, 419 (2014) (cit. on pp. 3, 59, 79, 81, 85).

[44] D Cavalcanti and P Skrzypczyk, Quantum steering: a review with focus on semidefinite programming, Rep. Prog. Phys. **80**, 024001 (2016) (cit. on pp. 3, 41, 59, 89, 91, 94, 111).

[45] R. Uola, A. C. S. Costa, H. C. Nguyen, and O. Gühne, Quantum steering, Rev. Mod. Phys. **92**, 015001 (2020) (cit. on pp. 3, 59, 89, 91, 94, 111).

[46] N. Gisin, Bell's inequality holds for all non-product states, Physics Letters A **154**, 201 (1991) (cit. on pp. 3, 84).

[47] R. F. Werner, Quantum states with Einstein-Podolsky-Rosen correlations admitting a hidden-variable model, Phys. Rev. A **40**, 4277 (1989) (cit. on pp. 3, 84).

[48] J. Barrett, Nonsequential positive-operator-valued measurements on entangled mixed states do not always violate a Bell inequality, Phys. Rev. A **65**, 042302 (2002) (cit. on pp. 3, 85).

[49] S. Popescu, Bell's Inequalities and Density Matrices: Revealing "Hidden" Nonlocality, Phys. Rev. Lett. **74**, 2619 (1995) (cit. on pp. 3, 63).

[50]N. Gisin, Hidden quantum nonlocality revealed by local filters, Physics Letters A **210**, 151 (1996) (cit. on pp. 3, 85).

[51]F. Hirsch, M. T. Quintino, J. Bowles, and N. Brunner, Genuine Hidden Quantum Nonlocality, Phys. Rev. Lett. **111**, 160402 (2013) (cit. on pp. 3, 85, 91).

[52]F. Hirsch, M. T. Quintino, J. Bowles, T. Vértesi, and N. Brunner, Entanglement without hidden nonlocality, New J. Phys. **18**, 113019 (2016) (cit. on pp. 3, 4).

[53]A. A. Methot and V. Scarani, An anomaly of non-locality, `https://doi.org/10.48550/arXiv.quant-ph/0601210` (2006) (cit. on p. 3).

[54]F. Hirsch, M. T. Quintino, and N. Brunner, Quantum measurement incompatibility does not imply Bell nonlocality, Phys. Rev. A **97**, 012129 (2018) (cit. on pp. 3, 86, 110).

[55]E. Bene and T. Vértesi, Measurement incompatibility does not give rise to Bell violation in general, New J. Phys. **20**, 013021 (2018) (cit. on pp. 3, 86, 110).

[56]M. M. Wolf, D. Perez-Garcia, and C. Fernandez, Measurements Incompatible in Quantum Theory Cannot Be Measured Jointly in Any Other No-Signaling Theory, Phys. Rev. Lett. **103**, 230402 (2009) (cit. on pp. 3, 68, 86, 110).

[57]R. Uola, C. Budroni, O. Gühne, and J.-P. Pellonpää, One-to-One Mapping between Steering and Joint Measurability Problems, Phys. Rev. Lett. **115**, 230402 (2015) (cit. on pp. 3, 93).

[58]R. Uola, T. Moroder, and O. Gühne, Joint Measurability of Generalized Measurements Implies Classicality, Phys. Rev. Lett. **113**, 160403 (2014) (cit. on pp. 3, 92).

[59]M. T. Quintino, T. Vértesi, and N. Brunner, Joint Measurability, Einstein-Podolsky-Rosen Steering, and Bell Nonlocality, Phys. Rev. Lett. **113**, 160402 (2014) (cit. on pp. 3, 93).

[60]G. Gour, M. P. Müller, V. Narasimhachar, R. W. Spekkens, and N. Y. Halpern, The resource theory of informational nonequilibrium in thermodynamics, Physics Reports **583**, 1 (2015) (cit. on pp. 3, 52, 53).

[61]A. Streltsov, H. Kampermann, S. Wölk, M. Gessner, and D. Bruß, Maximal coherence and the resource theory of purity, New J. Phys. **20**, 053058 (2018) (cit. on pp. 3, 52–54, 56, 58, 67, 100, 102, 110, 143).

[62]K. Baek, A. Sohbi, J. Lee, J. Kim, and H. Nha, Quantifying coherence of quantum measurements, New J. Phys. **22**, 093019 (2020) (cit. on pp. 3, 70–72).

[63]L. Tendick, H. Kampermann, and D. Bruß, Activation of Nonlocality in Bound Entanglement, Phys. Rev. Lett. **124**, 050401 (2020) (cit. on pp. 3, 95–97, 109, 133).

[64]L. Tendick, H. Kampermann, and D. Bruß, Quantifying necessary quantum resources for nonlocality, Phys. Rev. Research **4**, L012002 (2022) (cit. on pp. 4, 97–100, 109, 143).

[65]L. Tendick, M. Kliesch, H. Kampermann, and D. Bruß, Distance-based resource quantification for sets of quantum measurements, `https://doi.org/10.48550/arXiv.2205.08546` (2022) (cit. on pp. 4, 34, 100, 102–104, 109, 110, 159, 187).

[66] L. Tendick, H. Kampermann, and D. Bruß, Distribution of quantum incompatibility across subsets of measurements, https://doi.org/10.48550/arXiv.2301.08670 (2023) (cit. on pp. 4, 103, 105–107, 109, 110, 187).

[67] D. Hanneke, S. Fogwell, and G. Gabrielse, New Measurement of the Electron Magnetic Moment and the Fine Structure Constant, Phys. Rev. Lett. **100**, 120801 (2008) (cit. on p. 7).

[68] T. Sailer, V. Debierre, Z. Harman, F. Heiße, C. König, J. Morgner, B. Tu, A. V. Volotka, C. H. Keitel, K. Blaum, and S. Sturm, Measurement of the bound-electron g-factor difference in coupled ions, Nature **606**, 479 (2022) (cit. on p. 7).

[69] L. Masanes, T. D. Galley, and M. P. Müller, The measurement postulates of quantum mechanics are operationally redundant, Nature Communications **10**, 10.1038/s41467-019-09348-x (2019) (cit. on p. 7).

[70] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information: 10th Anniversary Edition* (Cambridge University Press, 2010) (cit. on pp. 7, 14, 16, 22, 31, 53).

[71] M. M. Wolf, Quantum channels & operations: Guided tour, **5** (2012) (cit. on pp. 7, 24–26, 29).

[72] J. M. Renes, Quantum Information Theory, (2015) (cit. on pp. 7, 8, 19, 23, 26).

[73] F. Grasselli, „Quantum Cryptography: from Key Distribution to Conference Key Agreement“, PhD thesis (Heinrich Heine University Düsseldorf, Düsseldorf, Sept. 2020) (cit. on p. 7).

[74] T. Holz, „Device-independent Quantum Key Distribution and the Role of Bell Nonlocality“, PhD thesis (Heinrich Heine University Düsseldorf, Düsseldorf, May 2020) (cit. on p. 7).

[75] G. Gianfelici, „Quantum resources in systems with many degrees of freedom“, PhD thesis (Heinrich Heine University Düsseldorf, Düsseldorf, July 2021) (cit. on p. 7).

[76] N. Wyderka, „Learning from correlations: what parts of quantum states tell about the whole“, PhD thesis (Universität Siegen, 2020) (cit. on p. 7).

[77] T. Kraft, „Aspects of quantum resources: coherence, measurements, and network correlations“, PhD thesis (Universität Siegen, 2020) (cit. on p. 7).

[78] I. Bengtsson and K. Życzkowski, *Geometry of quantum states: an introduction to quantum entanglement* (Cambridge university press, 2017) (cit. on pp. 7, 39, 40).

[79] D. Bruß and G. Leuchs, *Lectures on quantum information* (Wiley, 2007) (cit. on p. 7).

[80] R. Renner, Quantum Information Theory, (2013) (cit. on p. 7).

[81] M. M. Wilde, From classical to quantum Shannon theory, https://doi.org/10.1017/9781316809976.001 (2011) (cit. on p. 7).

[82] J. Watrous, *The theory of quantum information* (Cambridge university press, 2018) (cit. on pp. 7, 32, 41, 42, 53).

[83] M. A. Nielsen, Quantum information theory, https://doi.org/10.48550/arXiv.quant-ph/0011036 (2000) (cit. on p. 19).

[84] P. Busch, P. Lahti, J.-P. Pellonpää, and K. Ylinen, *Quantum Measurement* (Springer International Publishing, 2016) (cit. on p. 26).

[85] F. Hiai and D. Petz, The proper formula for relative entropy and its asymptotics in quantum probability, Communications in Mathematical Physics **143**, 99 (1991) (cit. on p. 40).

[86] H. Nagaoka and T. Ogawa, Strong converse and Stein's lemma in quantum hypothesis testing, IEEE Transactions on Information Theory **46**, 2428 (2000) (cit. on p. 40).

[87] S. Boyd and L. Vandenberghe, *Convex Optimization* (Cambridge University Press, 2004) (cit. on pp. 41, 42, 62, 82).

[88] M. Steiner, Generalized robustness of entanglement, Phys. Rev. A **67**, 054305 (2003) (cit. on p. 51).

[89] F. G. S. L. Brandão, M. Horodecki, J. Oppenheim, J. M. Renes, and R. W. Spekkens, Resource Theory of Quantum States Out of Thermal Equilibrium, Phys. Rev. Lett. **111**, 250404 (2013) (cit. on p. 52).

[90] M. Horodecki and J. Oppenheim, Fundamental limitations for quantum and nanoscale thermodynamics, Nature communications **4**, 1 (2013) (cit. on p. 52).

[91] P. Skrzypczyk, A. J. Short, and S. Popescu, Work extraction and thermodynamics for individual quantum systems, Nature communications **5**, 1 (2014) (cit. on p. 52).

[92] L. Masanes and J. Oppenheim, A general derivation and quantification of the third law of thermodynamics, Nature communications **8**, 1 (2017) (cit. on p. 52).

[93] C. Sparaciari, L. Del Rio, C. M. Scandolo, P. Faist, and J. Oppenheim, The first law of general quantum resource theories, Quantum **4**, 259 (2020) (cit. on p. 52).

[94] F. Brandao, M. Horodecki, N. Ng, J. Oppenheim, and S. Wehner, The second laws of quantum thermodynamics, PNAS **112**, 3275 (2015) (cit. on p. 52).

[95] M. A. Nielsen, An introduction to majorization and its applications to quantum mechanics, Lecture Notes, Department of Physics, University of Queensland, Australia (2002) (cit. on p. 53).

[96] Y. Luo and Y. Li, Robustness of purity: an operational approach to resource theory of purity, The European Physical Journal D **73**, `10.1140/epjd/e2019-90693-y` (2019) (cit. on pp. 54, 99).

[97] R. J. Glauber, The Quantum Theory of Optical Coherence, Phys. Rev. **130**, 2529 (1963) (cit. on p. 55).

[98] E. C. G. Sudarshan, Equivalence of Semiclassical and Quantum Mechanical Descriptions of Statistical Light Beams, Phys. Rev. Lett. **10**, 277 (1963) (cit. on p. 55).

[99] L. Mandel and E. Wolf, Coherence Properties of Optical Fields, Rev. Mod. Phys. **37**, 231 (1965) (cit. on p. 55).

[100] J. Åberg, Quantifying Superposition, arXiv: quant-ph/0612146 (2006) (cit. on pp. 55, 56).

[101] T. Baumgratz, M. Cramer, and M. B. Plenio, Quantifying Coherence, Phys. Rev. Lett. **113**, 140401 (2014) (cit. on pp. 55–57).

[102]E. Chitambar and G. Gour, Comparison of incoherent operations and measures of coherence, Phys. Rev. A **94**, 052336 (2016) (cit. on p. 56).

[103]J. I. de Vicente and A. Streltsov, Genuine quantum coherence, J. Phys. A: Math. Theor. **50**, 045301 (2016) (cit. on p. 56).

[104]A. Winter and D. Yang, Operational Resource Theory of Coherence, Phys. Rev. Lett. **116**, 120404 (2016) (cit. on p. 56).

[105]A. Streltsov, S. Rana, P. Boes, and J. Eisert, Structure of the Resource Theory of Quantum Coherence, Phys. Rev. Lett. **119**, 140402 (2017) (cit. on p. 56).

[106]X.-D. Yu, D.-J. Zhang, G. F. Xu, and D. M. Tong, Alternative framework for quantifying coherence, Phys. Rev. A **94**, 060302 (2016) (cit. on p. 57).

[107]C. Napoli, T. R. Bromley, M. Cianciaruso, M. Piani, N. Johnston, and G. Adesso, Robustness of Coherence: An Operational and Observable Measure of Quantum Coherence, Phys. Rev. Lett. **116**, 150502 (2016) (cit. on p. 57).

[108]T. R. Bromley, M. Cianciaruso, and G. Adesso, Frozen Quantum Coherence, Phys. Rev. Lett. **114**, 210401 (2015) (cit. on p. 58).

[109]A. Streltsov, U. Singh, H. S. Dhar, M. N. Bera, and G. Adesso, Measuring Quantum Coherence with Entanglement, Phys. Rev. Lett. **115**, 020403 (2015) (cit. on p. 58).

[110]C. Radhakrishnan, M. Parthasarathy, S. Jambulingam, and T. Byrnes, Distribution of Quantum Coherence in Multipartite Systems, Phys. Rev. Lett. **116**, 150504 (2016) (cit. on p. 58).

[111]Y. Yao, X. Xiao, L. Ge, and C. P. Sun, Quantum coherence in multipartite systems, Phys. Rev. A **92**, 022112 (2015) (cit. on p. 58).

[112]K. Modi, A. Brodutch, H. Cable, T. Paterek, and V. Vedral, The classical-quantum boundary for correlations: Discord and related measures, Rev. Mod. Phys. **84**, 1655 (2012) (cit. on p. 58).

[113]W. Zurek, Einselection and decoherence from an information theory perspective, Annalen der Physik **512**, 855 (2000) (cit. on p. 58).

[114]L Henderson and V Vedral, Classical, quantum and total correlations, J. Phys. A: Math. Gen. **34**, 6899 (2001) (cit. on p. 58).

[115]C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, and W. K. Wootters, Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels, Phys. Rev. Lett. **70**, 1895 (1993) (cit. on pp. 59, 62).

[116]R. Jozsa and N. Linden, On the Role of Entanglement in Quantum-Computational Speed-Up, Proceedings: Mathematical, Physical and Engineering Sciences **459**, 2011 (2003) (cit. on p. 59).

[117]J. S. Bell, On the Einstein Podolsky Rosen paradox, Physics Physique Fizika **1**, 195 (1964) (cit. on p. 59).

[118]C. H. Bennett, H. J. Bernstein, S. Popescu, and B. Schumacher, Concentrating partial entanglement by local operations, Phys. Rev. A **53**, 2046 (1996) (cit. on p. 59).

[119]V. Vedral, M. B. Plenio, M. A. Rippin, and P. L. Knight, Quantifying Entanglement, Phys. Rev. Lett. **78**, 2275 (1997) (cit. on p. 59).

[120] M. B. Plenio and S. S. Virmani, An introduction to entanglement theory, in *Quantum information and coherence* (Springer, 2014), pp. 173–209 (cit. on pp. 59, 65).

[121] O. Gühne and G. Tóth, Entanglement detection, Physics Reports **474**, 1 (2009) (cit. on pp. 59, 60, 62).

[122] L. Gurvits, „Classical deterministic complexity of Edmonds' Problem and quantum entanglement“, in Proceedings of the thirty-fifth ACM symposium on Theory of computing - STOC '03 (2003) (cit. on p. 60).

[123] A. Peres, Separability Criterion for Density Matrices, Phys. Rev. Lett. **77**, 1413 (1996) (cit. on pp. 60, 61).

[124] M. Horodecki, P. Horodecki, and R. Horodecki, Separability of mixed states: necessary and sufficient conditions, Physics Letters A **223**, 1 (1996) (cit. on p. 61).

[125] P. Horodecki, Separability criterion and inseparable mixed states with positive partial transposition, Physics Letters A **232**, 333 (1997) (cit. on p. 61).

[126] B. M. Terhal, Bell inequalities and the separability criterion, Physics Letters A **271**, 319 (2000) (cit. on p. 61).

[127] A. Dax, The distance between two convex sets, Linear Algebra and its Applications **416**, 184 (2006) (cit. on p. 62).

[128] E. M. Rains, Entanglement purification via separable superoperators, `https://doi.org/10.48550/arXiv.quant-ph/9707002` (1997) (cit. on p. 62).

[129] V. Vedral and M. B. Plenio, Entanglement measures and purification procedures, Phys. Rev. A **57**, 1619 (1998) (cit. on p. 62).

[130] A. Grudka, M. Horodecki, P. Horodecki, R. Horodecki, and M. Piani, Entanglement-swapping boxes and their communication properties, Phys. Rev. A **77**, 060307 (2008) (cit. on p. 63).

[131] F. Hirsch, „Hidden Nonlocality“, PhD thesis (Ph. D. thesis, University of Geneva, 2013) (cit. on p. 63).

[132] M. A. Nielsen, Conditions for a Class of Entanglement Transformations, Phys. Rev. Lett. **83**, 436 (1999) (cit. on p. 64).

[133] M. A. Nielsen and G. Vidal, Majorization and the interconversion of bipartite states, Quantum Inf. Comput. **1**, 76 (2001) (cit. on p. 64).

[134] C. H. Bennett, G. Brassard, S. Popescu, B. Schumacher, J. A. Smolin, and W. K. Wootters, Purification of Noisy Entanglement and Faithful Teleportation via Noisy Channels, Phys. Rev. Lett. **76**, 722 (1996) (cit. on p. 64).

[135] M. Horodecki, P. Horodecki, and R. Horodecki, Mixed-State Entanglement and Distillation: Is there a “Bound” Entanglement in Nature?, Phys. Rev. Lett. **80**, 5239 (1998) (cit. on p. 65).

[136] P. Horodecki, Ł. Rudnicki, and K. Życzkowski, Five open problems in quantum information theory, PRX Quantum **3**, 010101 (2022) (cit. on p. 65).

[137] K. G. H. Vollbrecht and M. M. Wolf, Activating Distillation with an Infinitesimal Amount of Bound Entanglement, Phys. Rev. Lett. **88**, 247901 (2002) (cit. on p. 65).

[138]L. Masanes, All Bipartite Entangled States Are Useful for Information Processing, Phys. Rev. Lett. **96**, 150501 (2006) (cit. on p. 65).

[139]T. Vértesi and N. Brunner, Disproving the Peres conjecture by showing Bell nonlocality from bound entanglement, Nature communications **5**, 1 (2014) (cit. on p. 65).

[140]T. Vértesi and N. Brunner, Quantum Nonlocality Does Not Imply Entanglement Distillability, Phys. Rev. Lett. **108**, 030403 (2012) (cit. on pp. 65, 95, 96).

[141]A. Peres, Foundations of Physics **29**, 589 (1999) (cit. on pp. 65, 95).

[142]K. Życzkowski, P. Horodecki, A. Sanpera, and M. Lewenstein, Volume of the set of separable states, Physical Review A **58**, 883 (1998) (cit. on p. 66).

[143]D. M. Greenberger, M. A. Horne, and A. Zeilinger, Going Beyond Bell's Theorem, `https://doi.org/10.48550/arXiv.0712.0921` (2007) (cit. on p. 66).

[144]W. Dür, G. Vidal, and J. I. Cirac, Three qubits can be entangled in two inequivalent ways, Phys. Rev. A **62**, 062314 (2000) (cit. on p. 66).

[145]A. Fine, Joint distributions, quantum correlations, and commuting observables, J. Math. Phys. **23**, 1306 (1982) (cit. on p. 68).

[146]I. Ivanovic, How to differentiate between non-orthogonal states, Physics Letters A **123**, 257 (1987) (cit. on p. 68).

[147]D. Dieks, Overlap and distinguishability of quantum states, Physics Letters A **126**, 303 (1988) (cit. on p. 68).

[148]A. Peres, How to differentiate between non-orthogonal states, Physics Letters A **128**, 19 (1988) (cit. on p. 68).

[149]Y. Liu and X. Yuan, Operational resource theory of quantum channels, Phys. Rev. Research **2**, 012035 (2020) (cit. on p. 68).

[150]T. Theurer, D. Egloff, L. Zhang, and M. B. Plenio, Quantifying Operations with an Application to Coherence, Phys. Rev. Lett. **122**, 190405 (2019) (cit. on pp. 68, 70).

[151]L. Li, K. Bu, and Z.-W. Liu, Quantifying the resource content of quantum channels: An operational approach, Phys. Rev. A **101**, 022335 (2020) (cit. on p. 68).

[152]T. Guff, N. A. McMahon, Y. R. Sanders, and A. Gilchrist, A resource theory of quantum measurements, J. Phys. A: Math. Theor. **54**, 225301 (2021) (cit. on p. 68).

[153]L. Guerini, J. Bavaresco, M. T. Cunha, and A. Acín, Operational framework for quantum measurement simulability, J. Math. Phys. **58**, 092102 (2017) (cit. on pp. 68, 69, 73).

[154]A. F. Ducuara and P. Skrzypczyk, Operational Interpretation of Weight-Based Resource Quantifiers in Convex Quantum Resource Theories, Phys. Rev. Lett. **125**, 110401 (2020) (cit. on pp. 68, 100).

[155]R. Uola, T. Kraft, J. Shang, X.-D. Yu, and O. Gühne, Quantifying Quantum Resources with Conic Programming, Phys. Rev. Lett. **122**, 130404 (2019) (cit. on pp. 68, 70, 100).

[156]R. Takagi and B. Regula, General Resource Theories in Quantum Mechanics and Beyond: Operational Characterization via Discrimination Tasks, Phys. Rev. X **9**, 031053 (2019) (cit. on pp. 68, 70, 100).

[157] S. Massar and S. Popescu, Amount of information obtained by a quantum measurement, Phys. Rev. A **61**, 062303 (2000) (cit. on p. 68).

[158] F. Buscemi, M. Hayashi, and M. Horodecki, Global Information Balance in Quantum Measurements, Phys. Rev. Lett. **100**, 210504 (2008) (cit. on p. 68).

[159] M. Dall'Arno, G. M. D'Ariano, and M. F. Sacchi, Informational power of quantum measurements, Phys. Rev. A **83**, 062304 (2011) (cit. on p. 68).

[160] A. S. Holevo, Information capacities of quantum measurement channels, Physica Scripta **T153**, 014034 (2013) (cit. on p. 68).

[161] S. Designolle, R. Uola, K. Luoma, and N. Brunner, Set Coherence: Basis-Independent Quantification of Quantum Coherence, Phys. Rev. Lett. **126**, 220404 (2021) (cit. on pp. 70, 101).

[162] H.-J. Kim and S. Lee, Relation between quantum coherence and quantum entanglement in quantum measurements, Phys. Rev. A **106**, 022401 (2022) (cit. on p. 72).

[163] T. Heinosaari, T. Miyadera, and M. Ziman, An invitation to quantum incompatibility, J. Phys. A: Math. Theor. **49**, 123001 (2016) (cit. on p. 73).

[164] H. P. Robertson, The Uncertainty Principle, Phys. Rev. **34**, 163 (1929) (cit. on p. 73).

[165] P. J. Coles, M. Berta, M. Tomamichel, and S. Wehner, Entropic uncertainty relations and their applications, Rev. Mod. Phys. **89**, 015002 (2017) (cit. on p. 73).

[166] T. Heinosaari and M. M. Wolf, Nondisturbing quantum measurements, J. Math. Phys. **51**, 092201 (2010) (cit. on p. 74).

[167] P. Busch, J. Kiukas, and P. Lahti, On the notion of coexistence in quantum mechanics, Mathematica Slovaca **60**, 665 (2010) (cit. on p. 74).

[168] S. T. Ali, C. Carmeli, T. Heinosaari, and A. Toigo, Commutative POVMs and Fuzzy Observables, Foundations of Physics **39**, 593 (2009) (cit. on p. 74).

[169] S. Designolle, M. Farkas, and J. Kaniewski, Incompatibility robustness of quantum measurements: a unified framework, New J. Phys. **21**, 113053 (2019) (cit. on pp. 75, 76, 100, 101).

[170] M. F. Pusey, Verifying the quantumness of a channel with an untrusted device, J. Opt. Soc. Am. B **32**, A56 (2015) (cit. on pp. 76, 100).

[171] S. Designolle, P. Skrzypczyk, F. Fröwis, and N. Brunner, Quantifying Measurement Incompatibility of Mutually Unbiased Bases, Phys. Rev. Lett. **122**, 050402 (2019) (cit. on pp. 76, 101).

[172] J. Bavaresco, M. T. Quintino, L. Guerini, T. O. Maciel, D. Cavalcanti, and M. T. Cunha, Most incompatible measurements for robust steering tests, Phys. Rev. A **96**, 022110 (2017) (cit. on p. 76).

[173] M. T. Quintino, C. Budroni, E. Woodhead, A. Cabello, and D. Cavalcanti, Device-Independent Tests of Structures of Measurement Incompatibility, Phys. Rev. Lett. **123**, 180401 (2019) (cit. on pp. 77, 86, 104, 105, 111).

[174] A. Einstein, B. Podolsky, and N. Rosen, Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?, Phys. Rev. **47**, 777 (1935) (cit. on pp. 79, 88).

[175] A. K. Ekert, Quantum cryptography based on Bell's theorem, Phys. Rev. Lett. **67**, 661 (1991) (cit. on p. 79).

[176] G Murta, S. B. van Dam, J Ribeiro, R Hanson, and S Wehner, Towards a realization of device-independent quantum key distribution, Quantum Sci. Technol. **4**, 035011 (2019) (cit. on p. 79).

[177] H. M. Wiseman, From Einstein's theorem to Bell's theorem: a history of quantum non-locality, Contemporary Physics **47**, 79 (2006) (cit. on p. 79).

[178] A. Fine, Hidden Variables, Joint Probability, and the Bell Inequalities, Phys. Rev. Lett. **48**, 291 (1982) (cit. on p. 82).

[179] Z. Ji, A. Natarajan, T. Vidick, J. Wright, and H. Yuen, MIP*=RE, 10.48550/ARXIV.2001.04383 (2020) (cit. on p. 82).

[180] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, Proposed Experiment to Test Local Hidden-Variable Theories, Phys. Rev. Lett. **23**, 880 (1969) (cit. on p. 83).

[181] I. Pitowsky, The range of quantum probability, J. Math. Phys. **27**, 1556 (1986) (cit. on p. 83).

[182] B. S. Cirel'son, Quantum generalizations of Bell's inequality, Letters in Mathematical Physics **4**, 93 (1980) (cit. on p. 84).

[183] S. Popescu and D. Rohrlich, Quantum nonlocality as an axiom, Foundations of Physics **24**, 379 (1994) (cit. on p. 84).

[184] M. Pawłowski, T. Paterek, D. Kaszlikowski, V. Scarani, A. Winter, and M. Żukowski, Information causality as a physical principle, Nature **461**, 1101 (2009) (cit. on p. 84).

[185] M. Navascués and H. Wunderlich, A glance beyond the quantum model, Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences **466**, 881 (2009) (cit. on p. 84).

[186] S. Popescu, Bell's Inequalities and Density Matrices: Revealing "Hidden" Nonlocality, Phys. Rev. Lett. **74**, 2619 (1995) (cit. on p. 85).

[187] D. Cavalcanti and P. Skrzypczyk, Quantitative relations between measurement incompatibility, quantum steering, and nonlocality, Phys. Rev. A **93**, 052112 (2016) (cit. on pp. 86, 88, 100, 101).

[188] J. Barrett, N. Linden, S. Massar, S. Pironio, S. Popescu, and D. Roberts, Nonlocal correlations as an information-theoretic resource, Phys. Rev. A **71**, 022101 (2005) (cit. on p. 86).

[189] N. S. Jones and L. Masanes, Interconversion of nonlocal correlations, Phys. Rev. A **72**, 052312 (2005) (cit. on p. 86).

[190] J. Barrett and S. Pironio, Popescu-Rohrlich Correlations as a Unit of Nonlocality, Phys. Rev. Lett. **95**, 140401 (2005) (cit. on p. 86).

[191] R. Gallego, L. E. Würflinger, A. Acín, and M. Navascués, Operational Framework for Nonlocality, Phys. Rev. Lett. **109**, 070401 (2012) (cit. on p. 86).

[192] A. C. Elitzur, S. Popescu, and D. Rohrlich, Quantum nonlocality for each pair in an ensemble, Physics Letters A **162**, 25 (1992) (cit. on p. 88).

[193] S. G. A. Brito, B. Amaral, and R. Chaves, Quantifying Bell nonlocality with the trace distance, Phys. Rev. A **97**, 022111 (2018) (cit. on pp. 88, 108).

[194] H. M. Wiseman, S. J. Jones, and A. C. Doherty, Steering, Entanglement, Nonlocality, and the Einstein-Podolsky-Rosen Paradox, Phys. Rev. Lett. **98**, 140402 (2007) (cit. on p. 88).

[195] Y. Z. Law, L. P. Thinh, J.-D. Bancal, and V. Scarani, Quantum randomness extraction for various levels of characterization of the devices, J. Phys. A: Math. Theor. **47**, 424028 (2014) (cit. on p. 88).

[196] E. Passaro, D. Cavalcanti, P. Skrzypczyk, and A. Acín, Optimal randomness certification in the quantum steering and prepare-and-measure scenarios, New J. Phys. **17**, 113010 (2015) (cit. on p. 88).

[197] C. Branciard, E. G. Cavalcanti, S. P. Walborn, V. Scarani, and H. M. Wiseman, One-sided device-independent quantum key distribution: Security, feasibility, and the connection with steering, Phys. Rev. A **85**, 010301 (2012) (cit. on p. 89).

[198] M. Piani and J. Watrous, Necessary and Sufficient Quantum Information Characterization of Einstein-Podolsky-Rosen Steering, Phys. Rev. Lett. **114**, 060404 (2015) (cit. on pp. 89, 94).

[199] M. D. Reid, P. D. Drummond, W. P. Bowen, E. G. Cavalcanti, P. K. Lam, H. A. Bachor, U. L. Andersen, and G. Leuchs, Colloquium: The Einstein-Podolsky-Rosen paradox: From concepts to applications, Rev. Mod. Phys. **81**, 1727 (2009) (cit. on p. 89).

[200] A. B. Sainz, N. Brunner, D. Cavalcanti, P. Skrzypczyk, and T. Vértesi, Postquantum Steering, Phys. Rev. Lett. **115**, 190403 (2015) (cit. on p. 90).

[201] L. P. Hughston, R. Jozsa, and W. K. Wootters, A complete classification of quantum ensembles having a given density matrix, Physics Letters A **183**, 14 (1993) (cit. on p. 90).

[202] A. B. Sainz, M. J. Hoban, P. Skrzypczyk, and L. Aolita, Bipartite Postquantum Steering in Generalized Scenarios, Phys. Rev. Lett. **125**, 050404 (2020) (cit. on p. 90).

[203] M. T. Quintino, T. Vértesi, D. Cavalcanti, R. Augusiak, M. Demianowicz, A. Acín, and N. Brunner, Inequivalence of entanglement, steering, and Bell nonlocality for general measurements, Phys. Rev. A **92**, 032107 (2015) (cit. on p. 91).

[204] J. Bowles, T. Vértesi, M. T. Quintino, and N. Brunner, One-way Einstein-Podolsky-Rosen Steering, Phys. Rev. Lett. **112**, 200402 (2014) (cit. on p. 91).

[205] J. Kiukas, C. Budroni, R. Uola, and J.-P. Pellonpää, Continuous-variable steering and incompatibility via state-channel duality, Phys. Rev. A **96**, 042331 (2017) (cit. on p. 93).

[206] P. Skrzypczyk, M. Navascués, and D. Cavalcanti, Quantifying Einstein-Podolsky-Rosen Steering, Phys. Rev. Lett. **112**, 180404 (2014) (cit. on p. 94).

[207] H.-Y. Ku, S.-L. Chen, C. Budroni, A. Miranowicz, Y.-N. Chen, and F. Nori, Einstein-Podolsky-Rosen steering: Its geometric quantification and witness, Phys. Rev. A **97**, 022338 (2018) (cit. on pp. 94, 108).

[208] T. Vértesi and N. Brunner, Disproving the Peres conjecture by showing Bell nonlocality from bound entanglement, Nature Communications **5** (2014) (cit. on p. 95).

[209] F. Hirsch, M. T. Quintino, T. Vértesi, M. F. Pusey, and N. Brunner, Algorithmic Construction of Local Hidden Variable Models for Entangled Quantum States, Phys. Rev. Lett. **117**, 190402 (2016) (cit. on pp. 95, 97).

[210] D. Cavalcanti, L. Guerini, R. Rabelo, and P. Skrzypczyk, General Method for Constructing Local Hidden Variable Models for Entangled Quantum States, Phys. Rev. Lett. **117**, 190401 (2016) (cit. on pp. 95, 97).

[211] C. Śliwa, Symmetries of the Bell correlation inequalities, Physics Letters A **317**, 165 (2003) (cit. on p. 96).

[212] R. F. Werner and M. M. Wolf, Bell inequalities and Entanglement, `https://doi.org/10.48550/arXiv.quant-ph/0107093` (2001) (cit. on p. 96).

[213] T. Heinosaari, J. Kiukas, and D. Reitzner, Noise robustness of the incompatibility of quantum measurements, Phys. Rev. A **92**, 022115 (2015) (cit. on p. 100).

[214] D. Bruß, Optimal Eavesdropping in Quantum Cryptography with Six States, Phys. Rev. Lett. **81**, 3018 (1998) (cit. on p. 104).

[215] M. Araújo, F. Hirsch, and M. T. Quintino, Bell nonlocality with a single shot, Quantum **4**, 353 (2020) (cit. on p. 108).

[216] K.-D. Wu, T. V. Kondra, S. Rana, C. M. Scandolo, G.-Y. Xiang, C.-F. Li, G.-C. Guo, and A. Streltsov, Operational Resource Theory of Imaginarity, Phys. Rev. Lett. **126**, 090401 (2021) (cit. on p. 110).

[217] R. Kunjwal, C. Heunen, and T. Fritz, Quantum realization of arbitrary joint measurability structures, Phys. Rev. A **89**, 052126 (2014) (cit. on p. 111).

[218] T. Heinosaari, D. Reitzner, and P. Stano, Notes on Joint Measurability of Quantum Observables, Foundations of Physics **38**, 1133 (2008) (cit. on p. 111).

[219] Y.-C. Liang, R. W. Spekkens, and H. M. Wiseman, Specker's parable of the overprotective seer: A road to contextuality, nonlocality and complementarity, Physics Reports **506**, 1 (2011) (cit. on p. 111).

[220] M.-O. Renou, E. Bäumer, S. Boreiri, N. Brunner, N. Gisin, and S. Beigi, Genuine Quantum Nonlocality in the Triangle Network, Phys. Rev. Lett. **123**, 140401 (2019) (cit. on p. 111).

[221] A. Pozas-Kerstjens, R. Rabelo, L. Rudnicki, R. Chaves, D. Cavalcanti, M. Navascués, and A. Acín, Bounding the Sets of Classical and Quantum Correlations in Networks, Phys. Rev. Lett. **123**, 140503 (2019) (cit. on p. 111).

[222] N. Gisin, J.-D. Bancal, Y. Cai, P. Remy, A. Tavakoli, E. Z. Cruzeiro, S. Popescu, and N. Brunner, Constraints on nonlocality in networks from no-signaling and independence, Nature Communications **11**, `10.1038/s41467-020-16137-4` (2020) (cit. on p. 111).

[223] P. Contreras-Tejada, C. Palazuelos, and J. I. de Vicente, Genuine Multipartite Nonlocality Is Intrinsic to Quantum Networks, Phys. Rev. Lett. **126**, 040501 (2021) (cit. on p. 111).

[224] A. Pozas-Kerstjens, N. Gisin, and A. Tavakoli, Full Network Nonlocality, Phys. Rev. Lett. **128**, 010403 (2022) (cit. on p. 111).

[225] A. Tavakoli, A. Pozas-Kerstjens, M.-X. Luo, and M.-O. Renou, Bell nonlocality in networks, Rep. Prog. Phys. **85**, 056001 (2022) (cit. on p. 111).

[226] M. Navascués, E. Wolfe, D. Rosset, and A. Pozas-Kerstjens, Genuine Network Multipartite Entanglement, Phys. Rev. Lett. **125**, 240505 (2020) (cit. on p. 111).

[227] T. Kraft, S. Designolle, C. Ritz, N. Brunner, O. Gühne, and M. Huber, Quantum entanglement in the triangle network, Phys. Rev. A **103**, L060401 (2021) (cit. on p. 111).

[228] K. Hansenne, Z.-P. Xu, T. Kraft, and O. Gühne, Symmetries in quantum networks lead to no-go theorems for entanglement distribution and to verification techniques, Nature Communications **13**, `10.1038/s41467-022-28006-3` (2022) (cit. on p. 111).

[229] O. Makuta, L. T. Ligthart, and R. Augusiak, No graph states can be prepared in quantum networks with bipartite sources, `https://doi.org/10.48550/arXiv.2208.12099` (2022) (cit. on p. 111).

[230] Y.-X. Wang, Z.-P. Xu, and O. Gühne, Quantum networks cannot generate graph states with high fidelity, `https://doi.org/10.48550/arXiv.2208.12100` (2022) (cit. on p. 111).

List of Figures

List of Acronyms

Index

A Paper: Activation of Nonlocality in Bound Entanglement

Title:	Activation of Nonlocality in Bound Entanglement
Authors:	Lucas Tendick, Hermann Kampermann, and Dagmar Bruß
Journal:	Physical Review Letters
Impact factor:	9.185 (2021)
Date of submission:	24 April 2019
Publication status:	Published
Contribution by LT:	First author (input approx. 85%)

This publication corresponds to the paper [63]. The summary of the results is presented in Section 4.1.

This work was a continuation of the studies in my Master's thesis. I had the initial idea to work on hidden nonlocality. Together with HK and DB, I collected some ideas for states that might be interesting to investigate, with bound entangled states among them. Then, I checked the literature for papers with appropriate methods to use. HK suggested relevant sources to me to get familiar with SDPs. In the following, I conducted the central part of the research, i.e., the numerical search for local bound entangled states that reveal hidden nonlocality. In the meantime, I had several discussions with DB and especially with HK. After obtaining the numerical results, I showed that it is possible to retrieve an analytical description of the state from the numerics. The results were discussed together with HK and DB. Finally, I wrote the whole manuscript, which my co-authors proofread. I improved the manuscript based on my co-authors' comments on several drafts of the paper.

Activation of Nonlocality in Bound Entanglement

Lucas Tendick,* Hermann Kampermann, and Dagmar Bruß
Institute for Theoretical Physics III, Heinrich-Heine-Universität Düsseldorf, D-40225 Düsseldorf, Germany

(Received 24 April 2019; revised manuscript received 26 September 2019; accepted 10 January 2020; published 3 February 2020)

We discuss the relation between entanglement and nonlocality in the hidden nonlocality scenario. Hidden nonlocality signifies nonlocality that can be activated by applying local filters to a particular state that admits a local hidden-variable model in the Bell scenario. We present a fully biseparable three-qubit bound entangled state with a local model for the most general (nonsequential) measurements. This proves for the first time that bound entangled states can admit a local model for general measurements. We furthermore show that the local model breaks down when suitable local filters are applied. Our results demonstrate the first example of activation of nonlocality in bound entanglement. Hence, we show that genuine hidden nonlocality does not imply entanglement distillability.

DOI: 10.1103/PhysRevLett.124.050401

Performing local measurements on certain entangled quantum states can lead to the phenomenon of quantum nonlocality. That is, the correlations obtained from the measurements are not compatible with the principle of local realism, witnessed by the violation of a so-called Bell inequality [1]. Although entanglement and nonlocality were extensively studied since the foundation of quantum theory [2,3], the relation between both is still not fully understood.

After the seminal work by Bell [1] as an answer to the EPR-Gedankenexperiment [4], it was widely believed that entanglement and nonlocality are just two different notions of the inseparability of quantum states. Indeed, for pure entangled states nonlocality is a generic feature [5,6]. However, Werner [7] showed that there exist mixed entangled states (so-called Werner states) which admit a local model for projective measurements. Later, Barrett [8] extended this result by showing that certain Werner states admit a local model even when positive-operator valued measures (POVMs), i.e., most general nonsequential measurements are considered. This displays the inequivalence of entanglement and nonlocality in the Bell scenario.

It was first noticed by Popescu [9] and more recently by Hirsch *et al.* [10] that some local entangled states can violate a Bell inequality when the observers apply judicious local filters as probabilistic preselection before the Bell test. This phenomenon is referred to as hidden nonlocality, or as genuine hidden nonlocality when one considers an entangled quantum state ρ with a local model even for POVMs. However, it was shown that genuine hidden nonlocality is not a general feature [11]. For example, a particular two-qubit Werner state remains local even after arbitrary local filtering.

Note that hidden nonlocality is not the only extension of the Bell scenario. For instance, nonlocality can also be superactivated [12,13] by allowing the parties to perform joint measurements on multiple copies of a local entangled state. An even more general concept is that of entanglement distillation [3]. In this scenario the parties have access to both local filters and multiple copies of a given state, with the goal to probabilistically obtain pure entangled states. Distillable states can therefore always be seen as nonlocal resource in the so-called asymptotic scenario [14]. However, there exist entangled states which are not distillable to pure entangled states. These states build the famous set of bound entangled states [15], which were the subject of various scientific works in the past [16–20]. Studying the nonlocal properties of bound entangled states will approach the answer of the fundamental open question of whether all entangled states are nonlocal resources. Since bound entanglement is the weakest form of entanglement, it was conjectured by Peres [21] that bound entangled states cannot lead to any nonlocal correlations at all. However, nowadays we know that the Peres conjecture is false [22,23]: bound entangled states can violate a Bell inequality. Despite these results and more advanced scenarios [24], nothing is known about the activation of local bound entanglement.

In this Letter, we answer the open question of whether bound entangled states with genuine hidden nonlocality exist in the affirmative. Specifically, we show that a certain three-qubit bound entangled state with a local model for POVMs can violate a Bell inequality when local filters are applied. This proves that genuine hidden quantum nonlocality does not imply entanglement distillability. Our results and possible extensions are visualized in Fig. 1.

Preliminaries.—Consider three distant parties Alice, Bob, and Charlie sharing an entangled quantum state ρ. The parties can perform local measurements via the positive semidefinite operators $M_{a|x}$, $M_{b|y}$, and $M_{c|z}$ with the settings x, y, z and the outcomes a, b, c. These operators form POVMs, as they satisfy the completeness relation

FIG. 1. Abstract overview of our results. We show that the set of nonlocal bound entangled states (BE states) can be enlarged in the hidden nonlocality scenario (HNL). This is the first step towards a possible equivalence of all BE states and all nonlocal BE states. Further enlargements of the set of nonlocal BE states could be provided by superactivation (SA) and the asymptotic scenario (Asymp.), similar to the case for distillable states. It is also an open question, whether the set can be enlarged to all BE states in such scenarios.

$\sum_a M_{a|x} = \mathbb{1}$ (and analogously for Bob and Charlie), where $\mathbb{1}$ denotes the identity operator. The resulting statistics is given by

$$p(abc|xyz) = \mathrm{Tr}[(M_{a|x} \otimes M_{b|y} \otimes M_{c|z})\rho]. \quad (1)$$

The state ρ is said to be local (for $\{M_{a|x}\}$, $\{M_{b|y}\}$, and $\{M_{c|z}\}$) if the distribution (1) admits a local decomposition of the following form:

$$p(abc|xyz) = \int \pi(\lambda) p(a|x\lambda) p(b|y\lambda) p(c|z\lambda) d\lambda. \quad (2)$$

That is, the statistics can be explained by a local hidden-variable model (LHV), where $\lambda \in \mathbb{R}$ is the shared local hidden variable, distributed according to the density $\pi(\lambda)$ such that $\int \pi(\lambda) d\lambda = 1$. The probability distributions $p(a|x\lambda)$, $p(b|y\lambda)$, and $p(c|z\lambda)$ are typically called local response functions in this context. A state ρ with such a decomposition *for all* possible measurements cannot violate any Bell inequality; otherwise it does violate (at least) one Bell inequality.

A concept which is easier to handle and necessary for Bell nonlocality is the concept of quantum steering [25]. The steering scenario is an asymmetric scenario where one or more parties remotely steer the state of the remaining parties by performing measurements on their part of the state. Here, we focus on the so-called one-sided steering scenario where Alice tries to steer Bob and Charlie. We say a state ρ demonstrates steering if its probability distribution does not admit a decomposition of the form

$$p(abc|xyz) = \int \pi(\lambda) p(a|x\lambda) \mathrm{Tr}(M_{b|y}\sigma_\lambda^B) \mathrm{Tr}(M_{c|z}\sigma_\lambda^C) d\lambda. \quad (3)$$

That is, the statistics can be explained by a so-called local hidden-state model (LHS), where the local response functions of Bob and Charlie are obtained from measurements on predetermined quantum states σ_λ^B and σ_λ^C, respectively. The set of (unnormalized) conditional states $\{\sigma_{a|x}^{BC}\}$ that Alice can prepare for Bob and Charlie, the so-called assemblage, is given by

$$\sigma_{a|x}^{BC} = \mathrm{Tr}_A[(M_{a|x} \otimes \mathbb{1} \otimes \mathbb{1})\rho], \quad (4)$$

where Tr_A denotes the partial trace and $\mathrm{Tr}(\sigma_{a|x}^{BC}) = p(a|x)$ is the probability that Alice obtains outcome a. Here, the measurement sets of Bob and Charlie $\{M_{b|y}\}$ and $\{M_{c|z}\}$ are assumed as tomographically complete. Further, note that any LHS can be considered as an LHV, while the converse does not hold [26]. An assemblage is said to demonstrate steering if it does not admit the decomposition

$$\sigma_{a|x}^{BC} = \int \pi(\lambda) p(a|x\lambda) \rho_\lambda^{BC}, \quad (5)$$

here ρ_λ^{BC} is a separable quantum state shared by Bob and Charlie.

We present now the hidden nonlocality scenario in the spirit of [10]. In this scenario the parties perform a probabilistic preselection according to a desired outcome before the Bell test. Hence, they apply a sequence of measurements on the shared state ρ_L which can lead to nonlocal correlations even if ρ_L admits an LHV for POVMs. In particular, this idea can be implemented by the use of local filters given by arbitrary Kraus operators F_x, fulfilling $F_x^\dagger F_x \leq \mathbb{1}, x \in \{A, B, C\}$ and acting on the respective local Hilbert space of the observers. The state which the parties share after filtering is given by

$$\rho = \frac{F_A \otimes F_B \otimes F_C \rho_L F_A^\dagger \otimes F_B^\dagger \otimes F_C^\dagger}{\mathrm{Tr}(F_A \otimes F_B \otimes F_C \rho_L F_A^\dagger \otimes F_B^\dagger \otimes F_C^\dagger)}, \quad (6)$$

where the success probability of the filtering is given by the normalization factor. We say that a state ρ_L possesses genuine hidden nonlocality if it admits an LHV for POVMs but the state ρ for some judiciously chosen filters F_A, F_B, F_C violates a Bell inequality. Note that local invertible filters do not change the entanglement character of a given state [3], i.e., bound entangled states remain bound entangled. Nevertheless the filters can increase the amount of entanglement (probabilistically) between the parties [27], which gives an intuitive reason why local filters can be useful. Further, by bound entangled states we mean entangled states with positive partial transpose (PPT).

Methods.—In order to derive our results, we will solve two main tasks: we show that the filtered state does violate a Bell inequality and that the state before filtering admits a local model for POVMs. The first task can be solved efficiently by an iterative sequence of semidefinite programs (SDPs) [28], using the so-called seesaw [29] method.

Consider a Bell inequality of the form

$$I = \sum_{a,b,c,x,y,z} c_{abc|xyz} p(abc|xyz) \leq L, \tag{7}$$

with given (real) coefficients $c_{abc|xyz}$ and a local bound L. The Bell operator according to this inequality is then given by

$$\mathcal{B} = \sum_{a,b,c,x,y,z} c_{abc|xyz} M_{a|x} \otimes M_{b|y} \otimes M_{c|z}. \tag{8}$$

The goal is to maximize the quantum value $Q = \mathrm{Tr}(\mathcal{B}\rho)$ for PPT entangled states ρ. Optimizing such an expression over all local measurements and the state is a problem, which cannot be solved by an SDP in general. However, the seesaw method provides a solution: we fix the measurements for two of the parties for a given state ρ, such that the problem becomes linear in the remaining party, let us say Alice. We maximize the expression Q subject to the constraints $M_{a|x} \geq 0$, $\sum_a M_{a|x} = \mathbb{1}$, which leads us to the optimal measurements of Alice. This strategy is iteratively applied over the individual parties and the state, to optimize the quantum value Q, without being guaranteed that it is a global maximum.

The second task is more difficult to solve. Even though there exist analytical constructions for LHVs, they mostly restrict to certain classes of states with high symmetry or they are restricted to projective measurements. Recently in [30,31] a method was presented to algorithmically construct local models, again making use of SDPs. Here, we only point out the main use of this construction (for details see [30,31]). Consider a discrete set of measurements $\{M_{a|x}\}$ associated with a so-called shrinking factor $0 \leq \eta \leq 1$ and the target state ρ_L. Further, consider the following SDP:

$$\begin{aligned}
&\text{given } \rho_L, \{M_{a|x}\}, \eta \\
&\quad \text{find } q^* = \max q \\
&\quad \text{s.t. } \mathrm{Tr}_A[(M_{a|x} \otimes \mathbb{1} \otimes \mathbb{1})\chi] = \sum_\lambda D_\lambda(a|x)\sigma_\lambda^{BC}, \quad \forall\ a, x \\
&\qquad \sigma_\lambda^{BC} \geq 0, (\sigma_\lambda^{BC})^{T_B} \geq 0 \quad \forall\ \lambda \\
&\qquad \eta\chi + (1-\eta)\frac{\mathbb{1}}{d_A} \otimes \mathrm{Tr}_A(\chi) = q\rho_L + (1-q)\frac{\mathbb{1}}{d_A d_B d_C},
\end{aligned} \tag{9}$$

where the Hermitian matrices χ and σ_λ^{BC} are the optimization variables. The SDP can be understood as follows. The first constraint ensures that (not necessarily positive-semidefinite quasistate) χ does admit an LHS for the finite set of measurements $\{M_{a|x}\}$, where $D_\lambda(a|x)$ are the deterministic strategies corresponding to Alice's set of inputs and outputs. More specifically, $D_\lambda(a|x) = \delta_{a,\lambda_x}$, where $\lambda = \lambda_1\lambda_2 \cdots \lambda_{m_A}$ is a string of length m_A, where m_A is the number of Alice's settings. The (subnormalized) states σ_λ^{BC} have to be separable between Bob and Charlie which is in general a nontrivial task, but for two qubits can simply be enforced by the partial transpose constraint $(\sigma_\lambda^{BC})^{T_B} \geq 0$ [32]. The last constraint contains the shrinking factor $0 \leq \eta \leq 1$ and ensures that also a noisy version of the target state ρ_L admits an LHS, but this time for the continuous set of measurements $\mathcal{M}$ (e.g., four-outcome POVMs) which was approximated by the discrete set $\{M_{a|x}\} \subset \mathcal{M}$.

The SDP is based on the fact that the statistics from noisy measurements on a noiseless state are equal to the statistics of a noisy state with noiseless measurements, i.e.,

$$\mathrm{Tr}_A[(M_a^\eta \otimes \mathbb{1} \otimes \mathbb{1})\chi] = \mathrm{Tr}_A[(M_a \otimes \mathbb{1} \otimes \mathbb{1})\rho_L], \tag{10}$$

where the target state is defined by

$$\rho_L = \eta\chi + (1-\eta)\frac{\mathbb{1}}{d_A} \otimes \mathrm{Tr}_A(\chi), \tag{11}$$

and the noisy measurements are given by

$$M_a^\eta = \eta M_a + (1-\eta)\mathrm{Tr}(M_a)\frac{\mathbb{1}}{d_A}, \tag{12}$$

for any $M_a \in \mathcal{M}$.

Note that because χ admits an LHS for the discrete set $\{M_{a|x}\}$, by convexity it admits also a local model for the noisy measurements M_a^η. From the equality in (10) it follows that ρ_L does also admit an LHS for a set of continuous noiseless measurements.

Here, the shrinking factor η is the largest number such that all noisy measurements M_a^η can be written as a convex mixture of elements from the discrete set $\{M_{a|x}\}$, i.e.,

$$M_a^\eta = \sum_x p_x M_{a|x}, \tag{13}$$

with $\sum_x p_x = 1$ and $p_x \geq 0\ \forall\ x$.

The shrinking factor can only be obtained analytically in the case of qubit projective measurements, but for general measurements it can be obtained by an SDP [31].

Results.—We now display our main result by first presenting a nonlocal three-qubit bound entangled state and in a second step show that this state originates from local filtering of a different state with an LHS model for POVMs. Note that the following results were recovered from the numerical data and are therefore exact in an analytical sense, unless indicated differently. Consider the (real-valued) density matrix in the basis $\{|000\rangle, |001\rangle, |010\rangle, \ldots, |111\rangle\}_{ABC}$ given by

$$\rho_{\mathrm{NL}} = (r_{ij})_{1\leq i,j\leq 8}, \tag{14}$$

with the following defining entries:

$$r_{11} = 0.0290, \qquad r_{12} = r_{13} = r_{15} = -0.0098,$$
$$r_{14} = r_{16} = r_{17} = r_{23} = r_{25} = r_{35} = -0.0083,$$
$$r_{18} = r_{27} = r_{36} = r_{45} = 0.0646,$$
$$r_{22} = r_{33} = r_{55} = 0.0412,$$
$$r_{24} = r_{26} = r_{34} = r_{37} = r_{56} = r_{57} = -0.0335,$$
$$r_{28} = r_{38} = r_{46} = r_{47} = r_{58} = r_{67} = -0.0598,$$
$$r_{44} = r_{66} = r_{77} = 0.1352,$$
$$r_{48} = r_{68} = r_{78} = 0.0102, \qquad r_{88} = 0.4418.$$

Note that ρ_{NL} is invariant under partial transpose with respect to any party, as well as invariant under permutation of parties, by construction. Therefore, the state is PPT and also biseparable with respect to any bipartite cut [23,33]. Note further that ρ_{NL} has the same symmetry properties as the family of states in [23] without being a member of this family. Nevertheless, using the seesaw method it can be shown to violate Śliwa's inequality number 5 [34] (which implies ρ_{NL} is entangled), which reads

$$I = \langle \mathrm{sym}[A_1 + A_1B_2 - A_2B_2 - A_1B_1C_1 - A_2B_1C_1 + A_2B_2C_2] \rangle \leq 3, \tag{15}$$

where sym$[X]$ denotes the symmetrization of X over the three parties, e.g., $\mathrm{sym}[A_1B_2] = A_1B_2 + A_1C_2 + A_2B_1 + A_2C_1 + B_1C_2 + B_2C_1$. Here, $A_j = B_j = C_j, j \in \{1,2\}$, and $A_j = M_{1|j} - M_{2|j}$. We choose $A_1 = -0.7909\sigma_z - 0.6119\sigma_x$, $A_2 = -0.2344\sigma_z + 0.9721\sigma_x$, which leads to a quantum violation $Q \approx 3.0152 > 3$ of inequality (15). Note that the maximal quantum value achievable by PPT states only allows violations up to $Q \approx 3.0187$ [35].

Next, we show that ρ_{NL} can originate from a local state by filtering. Consider the state ρ_L defined via the relation

$$\rho_{\mathrm{NL}} = \frac{F_A \otimes F_B \otimes F_C \rho_L F_A^\dagger \otimes F_B^\dagger \otimes F_C^\dagger}{\mathrm{Tr}(F_A \otimes F_B \otimes F_C \rho_L F_A^\dagger \otimes F_B^\dagger \otimes F_C^\dagger)}, \tag{16}$$

with the local filters

$$F_A = \begin{bmatrix} 0.4310 & -0.2971 \\ -0.2488 & 0.7291 \end{bmatrix},$$
$$F_B = \begin{bmatrix} 0.0342 & -0.0808 \\ -0.3664 & 0.8688 \end{bmatrix},$$
$$F_C = \begin{bmatrix} 0.3268 & -0.1873 \\ -0.1773 & 0.6440 \end{bmatrix}.$$

For more details, see the Supplemental Material [36]. Note that it is immediately clear that there exists a valid quantum state ρ_L fulfilling the above equation. This can be seen by using the fact that the above local filters are invertible and the only constraint $F^\dagger F \leq \mathbb{1}$ can always be achieved, since the filters F and cF map onto the same state for any $c \in \mathbb{C}\backslash\{0\}$.

In order to finally show that ρ_L possesses genuine hidden nonlocality, we need to show that it admits a local model for all POVMs. Therefore, we use the same parametrization as in [31] for Alice's finite set of measurements $\{M_{a|x}\}$. It consists of all relabellings of $\{P_+, P_-, 0, 0\}$ where P_+ is a projector onto a vertex of an icosahedron in the Bloch sphere and P_- onto the opposite direction, as well as all relabellings of the trivial set $\{\mathbb{1}, 0, 0, 0\}$. This leads to a set of 76 elements with a shrinking factor of $\eta \approx 0.673$. Note that it is sufficient to consider only extremal POVMs, which for qubits have at most four outcomes [38]. The optimization for the LHS, according to (9) results in $q^* = 1$. The precision of this result is subject to the standard precision of MATLAB [39] as well as the SDP solvers SeDuMi [40] and Mosek [41] for Yalmip [42]. Hence, ρ_L admits a local model for POVMs without the need of additional noise. For a graphical illustration of our main results, see Fig. 2.

Conclusions and outlook.—In the present Letter, we have shown that a fully biseparable bound entangled state of three qubits can admit a local model for POVMs, but can give rise to nonlocal correlations when local filters were applied before the Bell test. Hence, we have shown that bound entangled states can possess genuine hidden nonlocality. This is the first example of activation of nonlocality in bound entanglement. Furthermore, this is also the first example of an LHV of a bound entangled state for all POVMs, while previous models were restricted to projective measurements [31,43]. One important conclusion of our results is that genuine hidden nonlocality (since it also exists for bound entangled states) does not imply entanglement distillability. Together with the result of [11] it shows that genuine hidden nonlocality and entanglement

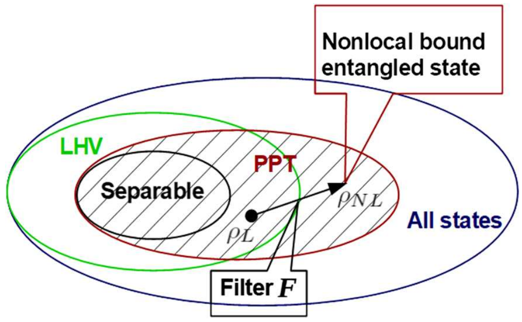

FIG. 2. Schematic overview over the relevant sets of states. The states in the shaded area are undistillable. Our results confirm the existence of bound entangled states with an LHV for POVMs. However, (invertible) local filters F are able to reveal the hidden nonlocality of these states. They map a state ρ_L from the set of states admitting an LHV onto a nonlocal state ρ_{NL}.

distillation are inequivalent. Note that since the local model we have constructed is an LHS model, our results are also relevant for the steering scenario.

It would be interesting to know whether there exist also bound entangled states without hidden nonlocality. Even though we could not prove the existence of such states, we found a bipartite bound entangled state with a local model for POVMs in the so-called filter normal form [27], which seems to play an important role for hidden nonlocality. We think, therefore, that this state is a good candidate to show bound entanglement without hidden nonlocality. For further details, see the Supplemental Material [36]. In the future, one should investigate the potential of bound entangled states in the superactivation or even in the asymptotic scenario. Even 20 years after the Peres conjecture [21], we still learn what bound entangled states are useful for. In the spirit of these developments it seems to be well motivated to state an "inverse Peres conjecture": all bound entangled states are nonlocal resources in the asymptotic case, see Fig. 1.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 675662 and support from the Federal Ministry of Education and Research (BMBF).

*Lucas.Tendick@hhu.de

[1] J. S. Bell, On the Einstein-Podolsky-Rosen paradox, Physics **1**, 195 (1964).
[2] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner, Bell nonlocality, Rev. Mod. Phys. **86**, 419 (2014).
[3] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, Quantum entanglement, Rev. Mod. Phys. **81**, 865 (2009).
[4] A. Einstein, B. Podolsky, and N. Rosen, Can quantum-mechanical description of physical reality be considered complete?, Phys. Rev. **47**, 777 (1935).
[5] N. Gisin, Bell's inequality holds for all non-product states, Phys. Lett. **154A**, 201 (1991).
[6] S. Popescu and D. Rohrlich, Generic quantum nonlocality, Phys. Lett. **166A**, 293 (1992).
[7] R. F. Werner, Quantum states with Einstein-Podolsky-Rosen correlations admitting a hidden-variable model, Phys. Rev. A **40**, 4277 (1989).
[8] J. Barrett, Nonsequential positive-operator-valued measurements on entangled mixed states do not always violate a Bell inequality, Phys. Rev. A **65**, 042302 (2002).
[9] S. Popescu, Bell's Inequalities and Density Matrices: Revealing "Hidden" Nonlocality, Phys. Rev. Lett. **74**, 2619 (1995).
[10] F. Hirsch, M. Quintino, J. Bowles, and N. Brunner, Genuine Hidden Quantum Nonlocality, Phys. Rev. Lett. **111**, 160402 (2013).
[11] F. Hirsch, M. Quintino, J. Bowles, T. Vértesi, and N. Brunner, Entanglement without hidden nonlocality, New J. Phys. **18**, 113019 (2016).
[12] C. Palazuelos, Superactivation of Quantum Nonlocality, Phys. Rev. Lett. **109**, 190401 (2012).
[13] D. Cavalcanti, A. Acín, N. Brunner, and T. Vértesi, All quantum states useful for teleportation are nonlocal resources, Phys. Rev. A **87**, 042104 (2013).
[14] L. Masanes, Asymptotic Violation of Bell Inequalities and Distillability, Phys. Rev. Lett. **97**, 050503 (2006).
[15] M. Horodecki, P. Horodecki, and R. Horodecki, Mixed-State Entanglement and Distillation: Is There a Bound Entanglement in Nature?, Phys. Rev. Lett. **80**, 5239 (1998).
[16] P. Horodecki, M. Horodecki, and R. Horodecki, Bound Entanglement Can Be Activated, Phys. Rev. Lett. **82**, 1056 (1999).
[17] P. W. Shor, J. A. Smolin, and A. V. Thapliyal, Superactivation of Bound Entanglement, Phys. Rev. Lett. **90**, 107901 (2003).
[18] J. T. Barreiro, P. Schindler, O. Gühne, T. Monz, M. Chwalla, C. F. Roos, M. Hennrich, and R. Blatt, Experimental multiparticle entanglement dynamics induced by decoherence, Nat. Phys. **6**, 943 (2010).
[19] J. Zhang, Continuous-variable multipartite unlockable bound entangled Gaussian states, Phys. Rev. A **83**, 052327 (2011).
[20] X. Jia, J. Zhang, Y. Wang, Y. Zhao, C. Xie, and K. Peng, Superactivation of Multipartite Unlockable Bound Entanglement, Phys. Rev. Lett. **108**, 190501 (2012).
[21] A. Peres, All the Bell inequalities, Found Phys. **29**, 589 (1999).
[22] T. Vértesi and N. Brunner, Disproving the Peres conjecture: Bell nonlocality from bipartite bound entanglement, Nat. Commun. **5**, 5297 (2014).
[23] T. Vértesi and N. Brunner, Quantum Nonlocality Does Not Imply Entanglement Distillability, Phys. Rev. Lett. **108**, 030403 (2012).
[24] Y. C. Liang, L. Masanes, and D. Rosset, All entangled states display some hidden nonlocality, Phys. Rev. A **86**, 052115 (2012).
[25] D. Cavalcanti and P. Skrzypczyk, Quantum steering: A review with focus on semidefinite programming, Rep. Prog. Phys. **80**, 024001 (2017).
[26] M. Quintino, T. Vértesi, D. Cavalcanti, R. Augusiak, M. Demianowicz, A. Acín, and N. Brunner, Inequivalence of entanglement, steering, and Bell nonlocality for general measurements, Phys. Rev. A **92**, 032107 (2015).
[27] F. Verstraete, J. Dehaene, and B. DeMoor, Local filtering operations on two qubits, Phys. Rev. A **64**, 010101 (2001).
[28] L. Vandenberghe and S. Boyd, Semidefinite programming, SIAM Rev. **38**, 49 (1996).
[29] R. F. Werner and M. M. Wolf, Bell inequalities and entanglement, arXiv:quant-ph/0107093.
[30] D. Cavalcanti, L. Guerini, R. Rabelo, and P. Skrzypczyk, General Method for Constructing Local Hidden Variable Models for Entangled Quantum States, Phys. Rev. Lett. **117**, 190401 (2016).
[31] F. Hirsch, M. Quintino, T. Vértesi, M. F. Pusey, and N. Brunner, Algorithmic Construction of Local Hidden Variable Models for Entangled Quantum States, Phys. Rev. Lett. **117**, 190402 (2016).
[32] A. Peres, Separability Criterion for Density Matrices, Phys. Rev. Lett. **77**, 1413 (1996).

[33] B. Kraus, J. I. Cirac, S. Karnas, and M. Lewenstein, Separability in $2 \times n$ composite quantum systems, Phys. Rev. A **61**, 062302 (2000).

[34] C. Śliwa, Symmetries of the Bell correlation inequalities, Phys. Lett. A **317**, 165 (2003).

[35] T. Moroder, J. D. Bancal, Y. C. Liang, M. Hofmann, and O. Gühne, Device-Independent Entanglement Quantification and Related Applications, Phys. Rev. Lett. **111**, 030501 (2013).

[36] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.124.050401 for more details on the local state in the main text, as well as for the local bipartite bound entangled state in the filter normal form. This includes Ref. [37].

[37] A. C. Doherty, P. A. Parrilo, and F. M. Spedalieri, Complete family of separability criteria, Phys. Rev. A **69**, 022308 (2004).

[38] G. M. D'Ariano, P. Lo Presti, and P. Perinotti, Classical randomness in quantum measurements, J. Phys. A **38**, 5979 (2005).

[39] MATLAB, Version 8.2.0.701 (R2013b) (The MathWorks Inc., Natick, Massachusetts, 2010).

[40] Jos F. Sturm, Using Sedumi 1.02, a MATLAB toolbox for optimization over symmetric cones, Optimi. Methods and Softw. **11**, 625 1999.

[41] MOSEK ApS, The MOSEK optimization toolbox for MATLAB manual, Version 8.0.0.60, 2019, https://docs.mosek.com/8.0/toolbox/index.html.

[42] J. Löfberg, Yalmip: A toolbox for modeling and optimization in MATLAB, in *Proceedings of the CACSD Conference, Taipei, Taiwan* (2004).

[43] G. Tóth and T. Vértesi, Quantum States with a Positive Partial Transpose are Useful for Metrology, Phys. Rev. Lett. **120**, 020506 (2018).

Supplemental Material for "Activation of nonlocality in bound entanglement "

September 21, 2019

Details on the local state ρ_L.—In order to give a useful representation of the local state ρ_L from (16) in the main text, one has to understand how to obtain this state. Naturally, there is no hint which states one should investigate in order to try to prove their locality or whether they possess genuine hidden nonlocality. However, it becomes immediately clear when one inverts the problem and tries to find a local state after we applied local filters on a nonlocal state. Since we choose the filters to be invertible, we can easily find filters which map the local state onto the nonlocal state. The nonlocal state obtained by the see-saw algorithm has by construction a high amount of symmetry, which we decrease by the local filters and then apply the SDP techniques to find an LHS. Afterwards, the inverted filters increase the symmetry of the state again. Therefore, ρ_L is simply given by

$$\rho_L = \frac{G_A \otimes G_B \otimes G_C \ \rho_{NL} \ G_A^\dagger \otimes G_B^\dagger \otimes G_C^\dagger}{\mathrm{Tr}(G_A \otimes G_B \otimes G_C \ \rho_{NL} \ G_A^\dagger \otimes G_B^\dagger \otimes G_C^\dagger)}, \tag{S1}$$

with the local invertible filters

$$G_A = \begin{bmatrix} 0.7291 & 0.2971 \\ 0.2488 & 0.4310 \end{bmatrix},$$
$$G_B = \begin{bmatrix} 0.8688 & 0.0808 \\ 0.3664 & 0.0342 \end{bmatrix},$$
$$G_C = \begin{bmatrix} 0.6440 & 0.1873 \\ 0.1773 & 0.3268 \end{bmatrix}.$$

and the nonlocal state ρ_{NL} defined in Eq. (14) in the main text.

Local bound entanglement in the filter normal form.—Here, we want to extend our outlook by presenting a bipartite bound entangled state which admits an LHS for POVMs and is a good candidate to show bound entanglement without hidden nonlocality, as we will argue below. An important feature of this state is that the state is already in the filter normal form [1], which means all single-party reduced density matricies are maximally mixed. The filter normal form does play an important role when it comes to hidden nonlocality. For example, the filter normal form does maximize the violation of the CHSH inequality for two-qubits, as well as entanglement monotones [1]. Further, in [2] it was shown that certain Werner states admit an LHS model, even after arbitrary local filtering. Werner states are also already in the filter normal form.

Intuitively, there is no obvious reason why local filters would still be able to activate the nonlocality of such states because they cannot distinguish the *useful* part of a state from white noise. Consider the state, in filter normal form given by

$$\sigma = \frac{\mathbb{1}}{d_A d_B} + \sum_{k=1}^{d_A^2-1} \xi_A H_k^A \otimes H_k^B \tag{S2}$$

with $d_A = 2$, $d_B = 4$, the coefficients ξ_k, and the traceless mutually orthonormal matricies H_k^A, H_k^B. Specifically, we choose

$$\xi_1 = \xi_2 = 1.3219,\ \ \xi_3 = 1.1348,$$

and the matricies

$$\begin{aligned} H_1^A &= \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix},\ H_2^A = \begin{pmatrix} 0 & -1 \\ 0 & 0 \end{pmatrix}, \\ H_3^A &= \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 \\ 0 & -\frac{1}{\sqrt{2}} \end{pmatrix}, \end{aligned}$$

for Alice's subsystem, as well as

$$\begin{aligned} H_1^B &= \begin{pmatrix} 0 & 0 & 0 & -0.0983 \\ -0.6393 & 0 & 0 & 0 \\ 0 & -0.4158 & 0 & 0 \\ 0 & 0 & -0.6393 & 0 \end{pmatrix}, \\ H_2^B &= \begin{pmatrix} 0 & 0.6393 & 0 & 0 \\ 0 & 0 & 0.4158 & 0 \\ 0 & 0 & 0 & 0.6393 \\ 0.0983 & 0 & 0 & 0 \end{pmatrix}, \\ H_3^B &= \begin{pmatrix} -0.4859 & 0 & 0 & 0 \\ 0 & -0.5137 & 0 & 0 \\ 0 & 0 & 0.5137 & 0 \\ 0 & 0 & 0 & 0.4859 \end{pmatrix}, \end{aligned}$$

for Bob's side. As one can quickly verify, σ is a PPT state. Nevertheless, it can be shown to be entangled by the SDP techniques presented in [3]. With the methods described in the main text, we were able to show that σ does admit an LHS model for general POVMs on Alice's side.

As argued above, this state is a good candidate to show bound entanglement without hidden nonlocality. However, it is quite complicated to prove our conjecture, due to the fact that many degrees of freedom are involved. If our conjecture turns out to be true, other scenarios like the superactivation or the asymptotic scenario have to be considered. If it turns out that σ can show hidden nonlocality, it would be the first example of a nonlocal bound entangled state in the lowest possible dimension for two parties. So far the smallest dimension for examples of nonlocal bound entangled states is 3×3 [4].

References

[1] F. Verstraete, J. Dehaene, and B. DeMoor. Local filtering operations on two qubits. *Phys. Rev. A*, 64:010101, Jun 2001.

[2] F. Hirsch, M. Quintino, J. Bowles, T. Vértesi, and N. Brunner. Entanglement without hidden nonlocality. *New Journal of Physics*, 18(11):113019, 2016.

[3] A. C. Doherty, P. A. Parrilo, and F. M. Spedalieri. Complete family of separability criteria. *Phys. Rev. A*, 69:022308, Feb 2004.

[4] T. Vértesi and N. Brunner. Disproving the Peres conjecture: Bell nonlocality from bipartite bound entanglement. *Nature Communications 5, 5297 (2014)*, 2014.

B Paper: Quantifying necessary quantum resources for nonlocality

Title:	Quantifying necessary quantum resources for nonlocality
Authors:	Lucas Tendick, Hermann Kampermann, and Dagmar Bruß
Journal:	Physical Review Research
Impact factor:	No official impact factor yet
Date of submission:	22 February 2021
Publication status:	Published
Contribution by LT:	First author (input approx. 85%)

This publication corresponds to the paper [64]. The summary of the results can be found in Section 4.2.

I initiated this project after finding the counter-intuitive behavior of the measurement and state resources for the CHSH inequality numerically (see Figure 4.4). After some initial discussions with HK, I obtained the first analytical results to explain this behavior. In following discussions with HK and DB, we realized that our results apply more generally to the resource theory of purity. HK and DB suggested that I should look for relevant connections to the paper [61]. After that, I derived the main results of the paper analytically. After more discussions with HK about the obtained results, I started to write the whole manuscript. The manuscript was improved based on my co-authors' comments on several drafts. Especially for some proofs in the Supplemental Material, HK pointed out ways to improve earlier versions of them.

PHYSICAL REVIEW RESEARCH **4**, L012002 (2022)

Letter

Quantifying necessary quantum resources for nonlocality

Lucas Tendick,* Hermann Kampermann, and Dagmar Bruß
Institute for Theoretical Physics III, Heinrich-Heine-Universität Düsseldorf, D-40225 Düsseldorf, Germany

(Received 22 February 2021; accepted 18 November 2021; published 14 January 2022)

Nonlocality is one of the most important resources for quantum information protocols. The observation of nonlocal correlations in a Bell experiment is the result of appropriately chosen measurements and quantum states. We quantify the minimal purity to achieve a certain Bell value for any Bell operator. Since purity is the most fundamental resource of a quantum state, this enables us also to quantify the necessary coherence, discord, and entanglement for a given violation of two-qubit correlation inequalities. Our results shine a light on the Clauser-Horne-Shimony-Holt inequality by showing that for a fixed Bell violation an increase in the measurement resources does not always lead to a decrease of the minimal state resources.

DOI: 10.1103/PhysRevResearch.4.L012002

It is arguably one of the most astonishing features of quantum theory that local measurements performed on certain quantum states can lead to the phenomenon of quantum nonlocality [1]. That is, the measurement statistics cannot be explained classically as they are not compatible with the principle of local realism. Mathematically this can be witnessed by the violation of a so-called Bell inequality [2]. Even though nonlocality [3] has been studied ever since the foundations of quantum theory [4], it is not yet completely understood.

Especially its connection to the properties of the used states and measurements remain challenging. On a qualitative level it is well understood that the resources entanglement and measurement incompatibility are necessary but not sufficient for nonlocality [5–7]; on a quantitative level things are much less clear. One particular example for open challenges is the anomaly of nonlocality [8,9] i.e., the effect that partially entangled states can lead to more nonlocality than the maximally entangled state. The situation becomes even more unclear when we include the influence of the state resources purity [10], coherence [11,12], and discord [13] which all found growing attention recently [14–19]. If one wants to analytically analyze the resources within quantum states and measurements and study their influence on nonlocal correlations, it is most natural to use the full description of the involved physical systems. Resources like purity, entanglement, and coherence are defined naturally in this so-called device-dependent (DD) formalism. An alternative approach to study nonlocality is the device-independent formalism which makes minimal assumptions on the involved systems and usually relies on numerical hierarchies [20,21]. We address in this Letter the following fundamental question in the DD scenario: What are the required properties of a quantum state and its measurements to exhibit nonlocality? In other words, we quantify the interplay between the resource of nonlocal correlations and other resources like purity, coherence, discord, and most famously entanglement on the state side and measurement incompatibility [22] on the measurement side. The physical situation we are going to consider is illustrated in Fig. 1. We derive from the spectrum of any given Bell operator an analytical expression for the minimal purity of a quantum state that is needed to achieve some fixed amount of nonlocality in terms of a Bell inequality violation. This result is general, i.e., it holds for any dimension, any number of parties, measurement settings, and outcomes. In a second step, we show that this criterion also provides the minimal amount of coherence, discord, and entanglement needed for the violation of an inequality with any Bell-diagonal Bell operator, which is of particular interest for the case of two-qubit systems. As an application of our results, we present a closed expression for the maximal possible violation of the Clauser-Horne-Shimony-Holt (CHSH) inequality [23] given some fixed amount of entanglement or purity and a given level of measurement incompatibility. This enables us to establish a surprising link between the incompatibility of quantum measurements and the minimal entanglement needed. More precisely, we show that highly incompatible projective measurements need, in some instances, a higher amount of entanglement in order to show some fixed CHSH nonlocality than less incompatible projective measurements. In other words, a smaller resource on the measurement side does *not* require a higher resource on the state side, which is counterintuitive. An analogous result follows for the case of the two-setting linear steering inequality [24].

Preliminaries. In general, we are considering Hermitian Bell operators of the form

$$I = \sum_{a,b,x,y} c_{ab|xy} M_{a|x} \otimes M_{b|y}, \tag{1}$$

*Lucas.Tendick@hhu.de

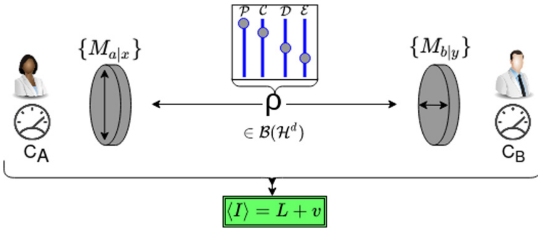

FIG. 1. Illustration of a Bell experiment. A (bounded) quantum state $\rho \in \mathcal{B}(\mathcal{H}^d)$ with adjustable resources purity $\mathcal{P}$, coherence $\mathcal{C}$, discord $\mathcal{D}$, and entanglement $\mathcal{E}$ is distributed to Alice and Bob who perform measurements $\{M_{a|x}\}$ and $\{M_{b|y}\}$ with also adjustable incompatibilities C_A and C_B. The interplay between the state resources and the measurement resources results in the observed Bell value $\langle I \rangle$. Minimal resource requirements for an observed Bell violation v beyond the local bound L are derived in the text.

where the real coefficients $c_{ab|xy}$ together with the local bound L (see below) describe the corresponding Bell inequality. The measurements are described by positive semidefinite operators $M_{a|x}, M_{b|y}$ with outcomes a, b and inputs x, y which form a positive operator-valued measure such that $\sum_a M_{a|x} = \mathbb{1}$ and $\sum_b M_{b|y} = \mathbb{1}$. A Bell inequality is given by

$$\sum_{a,b,x,y} c_{ab|xy}\, p(ab|xy)_{\text{LHV}} \leqslant L, \tag{2}$$

with the (real) local bound L for all correlations obeying a so-called local hidden-variable model (LHV). This inequality may be violated by some entangled quantum states ρ, where the probability distribution is given by $p(ab|xy) = \text{Tr}[(M_{a|x} \otimes M_{b|y})\rho]$. We call states which violate (at least) one Bell inequality nonlocal. The achieved Bell value is denoted by $\langle I \rangle = \text{Tr}(I\rho) = L + v$ where $v > 0$ is the amount by which the bound L is violated. During the course of this Letter, we will often use the spectral decomposition of a quantum state $\rho = \sum_i^d \lambda_i |\phi_i\rangle\langle\phi_i|$ with $\lambda_i \geqslant 0$ and $\sum_i^d \lambda_i = 1$ and the Bell operator $I = \sum_j^d \mu_j |\Psi_j\rangle\langle\Psi_j|$ with real eigenvalues μ_j where d is the dimension of $\rho \in \mathcal{B}(\mathcal{H}^d)$ and $\mathcal{B}(\mathcal{H}^d)$ denotes the set of bounded operators. The sets $\{|\phi_i\rangle\}$, $\{|\Psi_j\rangle\}$ form orthonormal bases. We order (without loss of generality) the eigenvalues in descending order, i.e., $\lambda_i \geqslant \lambda_s$ for $i < s$ and $\mu_j \geqslant \mu_t$ for $j < t$.

Main task. We want to quantify the minimal quantum resources of a state ρ of dimension d in order to achieve some given violation v for a given Bell operator I (i.e., the measurements are fixed). Thus, we want to minimize a general resource quantifier $R(\rho)$ such that ρ is consistent with the observed data in terms of the Bell expectation value $\langle I \rangle$, i.e., we want to find

$$R^* = \min_\rho \{R(\rho) | \langle I \rangle = \text{Tr}(\rho I) = L + v\}. \tag{3}$$

Optimizations of this form naturally occur in inference schemes based on entanglement witnesses [25–28]. The important difference to the task we consider here is that nonlocality itself is also a resource. In the context of nonlocality this problem has only been addressed for the CHSH inequality [29–34] with the main focus on entanglement. This approach based on the Bell operator makes use of the full information available and therefore allows us to study in a simple way how the required state resources depend on the chosen measurements.

Let us specify what we mean by the term quantum resource without going into detail. In any resource theory, one first defines the states which are no resource, the so-called void states (or free states), which constitute the set $\mathcal{V}$. Second, one defines the (maximal) set of operations Λ (free operations) that cannot turn a void state into a resource state. Finally, one has to find measures R which quantify the respective resource. The measures have to be faithful monotones, i.e., $R(\rho) = 0$ iff $\rho \in \mathcal{V}$ and $R[\Lambda(\rho)] \leqslant R(\rho)$ $\forall \rho$ and free operations Λ. Additional properties of many measures are normalization for the maximal resource and additivity under tensor products. For more details, see [35]. For example, in the resource theory of entanglement, the free states are the separable ones, the free operations are local operations and classical communication, and a quantifier is the relative entropy of entanglement.

Purity. Our main result is an analytical result for the minimal purity of ρ needed to achieve the Bell value $\langle I \rangle = L + v$ for a general Bell operator I of any dimension d, any number of parties n, settings k, and outcomes m. We want to emphasize that the commonly used expression $\text{Tr}(\rho^2)$ (known as linear purity) is not a proper purity measure [10] since it lacks additivity and normalization $[P(|\Psi\rangle) = \log_2(d)$ for any d-dimensional pure state $|\Psi\rangle]$ and does not vanish for the maximally mixed state. Instead one should use the Rényi 2-purity $\mathcal{P}_2(\rho) = \log_2[d\,\text{Tr}(\rho^2)]$. Here, we will employ the generalized robustness of purity, which is easier to handle mathematically. It is defined via the general robustness quantifier

$$G_R(\rho) := \min_\tau \left\{ x | x \geqslant 0, \exists \text{ a state } \tau, \frac{\rho + x\tau}{1+x} \in \mathcal{V} \right\}, \tag{4}$$

where the set $\mathcal{V}$ consists of the void states. $G_R(\rho)$ leads to the log robustness $\log_2[G_R(\rho) + 1]$, which is a proper measure of the resources considered in this Letter [16,36]. Because it is fully determined via $G_R(\rho)$, we focus the main discussions in this Letter on the generalized robustness for simplicity. Since τ can be any state, $G_R(\rho)$ can be seen as general noise robustness of ρ with respect to a void set $\mathcal{V}$ and can therefore be used to quantify a general resource G. In the case of purity, the void set $\mathcal{V}$ consists of only the maximally mixed state $\mathbb{1}/d$. It was shown in [36] that the generalized robustness of purity is given by

$$P_R(\rho) = d\lambda_1(\rho) - 1. \tag{5}$$

Thus, minimizing $P_R(\rho)$ reduces to minimizing $\lambda_1(\rho)$. In order to show our main result we first answer the (easier to solve) reverse question: Given $P_R(\rho)$, what is the maximal possible Bell value $\langle I \rangle_{\max} = L + v_{\max}$ the state ρ can achieve for a fixed Bell operator I?

Theorem 1. Given the Hermitian operator $I = \sum_{j=1}^d \mu_j |\Psi_j\rangle\langle\Psi_j|$ with $\mu_j \geqslant \mu_t$ for $j < t$ and a fixed robustness of purity $P_R(\rho)$ of a quantum state ρ. The maximal expectation value $\langle I \rangle_{\max}$ can be achieved by $\rho = \sum_{i=1}^r \lambda_i |\Psi_i\rangle\langle\Psi_i|$, where $\lambda_i \geqslant 0$, $\sum_{i=1}^r \lambda_i = 1$, $\lambda_i \geqslant \lambda_s$

for $i < s$, and is given by

$$\langle I\rangle_{\max} = \sum_{j=1}^{r} \mu_j \lambda_j, \tag{6}$$

where r is an integer such that $\frac{1}{r-1} > \lambda_1 \geqslant \frac{1}{r}$ and all eigenvalues λ_i for $i \in \{1, \dots, r-1\}$ are equal to $\lambda_1 = (1 + P_R)/d$.

Proof. The theorem follows from the generalization of Ruhe's trace inequality [37] and the fact that it is optimal to choose all eigenvalues λ_i equal to λ_1 except the lowest nonzero one, which is given by normalization. The integer r defines the rank of the optimal ρ which we construct from the $\{\lambda_i\}$ and the eigenstates of I. This choice is unique for nondegenerate eigenvalues of I. See [38] for the specifics of the proof. ■

Theorem 1 can be used reversely (see Lemma 1 in the Supplemental Material [38]), which provides our first main result. Namely, for given $\langle I\rangle_{\max}$ we can use Eq. (6) to determine the minimal $P_R(\rho)$ or $\lambda_1(\rho)$ needed to achieve the Bell value $\langle I\rangle_{\max}$. In order to determine $\lambda_1(\rho)$ one only needs to find the integer r such that Theorem 1 is valid. The usefulness of Theorem 1 lies in its simplicity. Not only does it allow one to minimize the generalized robustness of purity $P_R(\rho)$ for a fixed expectation value of the most general Bell operator via an easily accessible criterion; also, one needs to check at most d linear equations. We also proved a more involved analogon to Theorem 1 with respect to the Rényi 2-purity $\mathcal{P}_2(\rho)$. See the Supplemental Material [38] for a detailed discussion.

Equality of quantum resources for two qubits. In the following we show which effect minimizing the purity has on the other state resources. In other words, we demonstrate the power of Theorem 1 by showing that for the subset of two-qubit correlation inequalities, i.e., inequalities without single party correlation terms, the states of minimal generalized robustness of purity for a fixed violation v also minimize the respective generalized robustnesses of coherence $C_R(\rho)$, discord $D_R(\rho)$, and entanglement $E_R(\rho)$, which in fact turn out to be equal. This is of particular interest since for every quantum state the hierarchy [10]

$$\mathcal{P}(\rho) \geqslant \mathcal{C}(\rho) \geqslant \mathcal{D}(\rho) \geqslant \mathcal{E}(\rho) \tag{7}$$

holds when quantified by the same distance-based [39] measure and coherence is quantified with respect to any product basis. We will in particular choose the product basis that minimizes the coherence of the state ρ. This notion of coherence coincides with the notion of symmetric quantum discord with respect to all subsystems [10,40]. Therefore, we will only summarize the concept of coherence [12] here; for more details about discord, see [41]. Coherence in general is a basis-dependent concept and is connected to the ability of a state to be in a superposition of some (fixed) basis states. The void states δ are called incoherent states. These are diagonal with respect to a fixed basis $|i\rangle$, i.e,

$$\delta = \sum_i p_i |i\rangle\langle i|, \quad p_i \geqslant 0, \quad \sum_i p_i = 1. \tag{8}$$

Note that our notion of coherence corresponds to a minimization over all states equivalent to ρ under local unitaries.

Our result is summarized in the following theorem.

Theorem 2. Given a Bell operator of the form

$$I = \sum_{x,y} g_{x,y} A_x \otimes B_y, \tag{9}$$

with real coefficients $g_{x,y}$ and local observables $A_x = \vec{a}_x \cdot \vec{\sigma}$, $B_y = \vec{b}_y \cdot \vec{\sigma}$ where $\vec{a}_x, \vec{b}_y$ are Bloch vectors and $\vec{\sigma}$ is the vector containing the Pauli matrices. For a fixed expectation value $\langle I\rangle = L + v$, where L is the local bound and $v > 0$, there exists a two-qubit quantum state ρ_{opt} which simultaneously minimizes the generalized robustness of purity P_R, coherence with respect to all product bases C_R, and entanglement E_R.

Proof. The proof relies on the fact that the states of minimal entanglement are Bell-diagonal states (BDS), which are entangled if and only if $\lambda_1 > 1/2$. The generalized robustness of entanglement [42] reduces for two-qubit BDS to $E_R(\rho_{\text{BDS}}) = 2\lambda_1(\rho_{\text{BDS}}) - 1$. Using this fact and Lemma 1 (see the Supplemental Material [38]) the optimal state ρ_{opt} can always be chosen to be of at most rank 2. This enables us to show that the closest separable state is always incoherent in some product basis. Therefore minimizing λ_1 minimizes all state resources. We relocated the specifics of the proof to the Supplemental Material [38]. ■

Note that an equivalence between coherence and entanglement for maximally correlated states has also been shown in different contexts [16,17]. We want to highlight that there is a straightforward generalization to genuine-multipartite entanglement (GME) quantification for N-qubit Greenberger-Horne-Zeilinger (GHZ) -diagonal Bell operators (e.g., two-setting full-correlation inequalities [43]) when we ask for a violation v which requires GME [44], since the optimal states will then be diagonal in the GHZ basis and the GME of these states is completely characterized by $\lambda_1 > 1/2$ [45], analogously to two-qubit BDS.

However, in general the hierarchy (7) will not be tight. Based on numerical optimization we find that there is indeed a (nontrivial) gap between purity, coherence, and entanglement for judiciously chosen observables in the $I3322$ inequality [46], an inequality with three settings and two outcomes for both parties including single party expectation values. One reason for this is the fact that the considered Bell operator is, in contrast to those for the previous discussed correlation inequalities, not diagonal in the Bell basis. That leads to different optimal states for the respective resources. See the Supplemental Material [38] for more details.

CHSH inequality. Remarkably, our results bring insights into the well-known CHSH inequality and systems of two qubits. The CHSH operator [23] is defined as

$$I = A_1 \otimes B_1 + A_1 \otimes B_2 + A_2 \otimes B_1 - A_2 \otimes B_2, \tag{10}$$

with $|\langle I\rangle| \leqslant 2$ for local-realistic models. The general form of Eq. (6) can for the case of two-qubit states be reduced to at most rank-2 solutions

$$\lambda_1 \mu_1 + (1 - \lambda_1)\mu_2 = L + v, \tag{11}$$

which recovers the finding made in [31]. Furthermore it is well known [47] that if the observables fulfill $A_i^2 = B_j^2 = \mathbb{1}$, it holds

$$I^2 = 4\mathbb{1} \otimes \mathbb{1} - [A_1, A_2] \otimes [B_1, B_2], \tag{12}$$

where $[X, Y]$ denotes the commutator between X and Y. The observables X and Y describing projective measurements are called incompatible, i.e., they cannot be measured jointly if and only if $[X, Y] \neq 0$. This quantum effect is the central aspect of the famous Heisenberg-Robertson uncertainty relation [48]. The use of incompatible measurements is necessary but not sufficient for Bell nonlocality [6,7]. There exists a resource theory [22] which allows the quantification of measurement incompatibility of one party. Let us introduce as a quantifier C for the (global) incompatibility the product of the single party measurement incompatibilities defined by the operator norm (largest absolute eigenvalue) of the commutators. Namely, $C = C_A C_B = \|[A_1, A_2]\| \|[B_1, B_2]\|$. This is well motivated since $C = 0$ if and only if one of the parties holds compatible measurements, i.e., the CHSH inequality cannot be violated and $C = 4$ is achieved with Pauli commutation relations only. Note that the single party incompatibility C_A is directly related to incompatibility quantifiers studied in [22]. After some algebra, we obtain the eigenvalues of I as a function of C, i.e. [49,50],

$$\mu_{1/4} = \pm\sqrt{4 + C}, \quad \mu_{2/3} = \pm\sqrt{4 - C}. \tag{13}$$

This shows that the quantity C quantifies the maximal nonlocality which can possibly be revealed by the given observables. By introducing the global measurement incompatibility we can study relations between the necessary resources contained in the states and those contained in the measurements, when wanting to achieve a certain nonlocality. The maximal possible violation given in Eq. (6) reduces to

$$\langle I \rangle_{\max} = \sqrt{4 + C}\lambda_1 + \sqrt{4 - C}(1 - \lambda_1). \tag{14}$$

Note that after inserting the optimal incompatibility $C_{\max} = \frac{4(2\lambda_1 - 1)}{2\lambda_1^2 - 2\lambda_1 + 1}$ to maximize the Bell value $\langle I \rangle_{\max}$ for fixed λ_1 one easily recovers the special case [30] and notably the result [34] where a formula for the maximal CHSH value of a two-qubit state in terms of its concurrence was found.

Intuitively one would expect now for a fixed violation of the CHSH inequality, that there is a trade-off between the necessary measurement resources and the necessary state resources in the sense that more of the resource in the measurements requires less resource in the state. This, however, is not always the case. As one can see in Fig. 2 there are parameter regions where less resources on the measurement side go together with less resources on the state side. Especially for very small violations, weakly incompatible measurements require much less entanglement for the same amount of nonlocality. We want to emphasize that the behavior of the other resources with respect to the quantifier C is qualitatively the same, since these are also monotonic functions of $\lambda_1(\rho)$. We further highlight that extensions of the considered Bell operators to higher dimensions, such as those in [49], can in the case of suboptimal extensions only increase the necessary purity while keeping the quantifier C constant. Is the surprising behavior discussed above a generic feature, or does it possibly depend on the chosen quantifiers for measurement incompatibility and/or the resources in the state? We discuss other possible quantifiers for state resources in the Supplemental Material [38] and conclude that the behavior is generic, by arguing that when other quantifiers are chosen

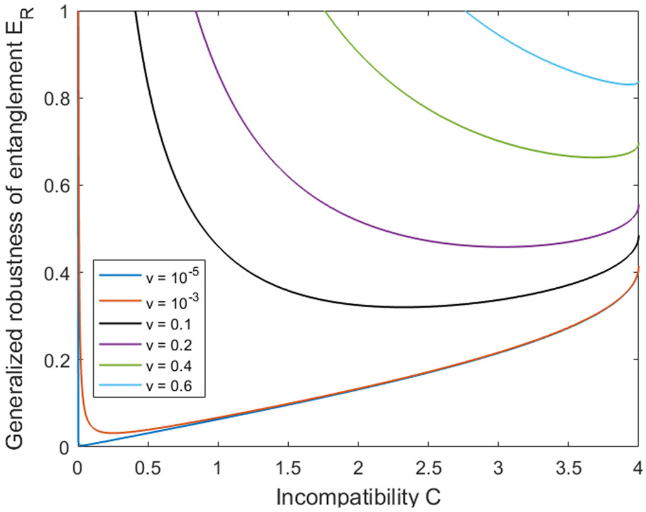

FIG. 2. The minimal generalized robustness of entanglement $E_R(\rho)$ for a given level of incompatibility C for different amounts of desired violation v. The curves diverge at some C because there is no state achieving the given violation. For low violations the effect that less entanglement for lower C is necessary becomes clearly visible, for a large regime of C.

only purity could possibly show a qualitatively different behavior. We show that this is indeed the case for the relative entropy of purity, while the Rényi 2-purity shows a similar behavior as the generalized robustness of purity as a function of C. For measurement incompatibility, we also show that the generalized robustness of incompatibility displays the same qualitative behavior. However, in general, it is still an open question whether these results are influenced by the particular choice of the incompatibility quantifier.

We strengthen this conclusion by highlighting that plots of the same qualitative behavior follow for the two-setting linear steering inequality [24] given by

$$F_2 = \left| \sum_{i=1}^{2} \langle A_i \otimes B_i \rangle \right| \leqslant \sqrt{2}, \tag{15}$$

where Bob's measurements have to be aligned orthonormally while Alice is free to choose any projective measurements. In this case, the eigenvalues of the steering operator of F_2 only depend on $C_A = \|[A_1, A_2]\|$ in an analogous way to the CHSH inequality, i.e.,

$$\tilde{\mu}_{1/4} = \pm\sqrt{2 + C_A}, \quad \tilde{\mu}_{2/3} = \pm\sqrt{2 - C_A}, \tag{16}$$

from which a behavior of the resources that is analogous to that for the CHSH inequality follows. This shows that the qualitative dependency of the state resources on the measurement incompatibility is not just due to our definition of the bipartite quantifier C, but a true physical phenomenon.

Discussion. In the present Letter we have analyzed the minimal resource requirements on the states and measurements for a given level of Bell nonlocality. We have shown that the minimal purity necessary to achieve a certain Bell value for the most general Bell operator can be found analytically via an easily accessible criterion. Since the purity of a state is its most fundamental resource which bounds *all* other resources of this state, this has major consequences for the inference of other necessary resources such as coherence and entanglement. We demonstrated this concretely by showing that the generalized robustness of all state resources can be minimized by the same state for two-qubit correlation inequalities. Finally, we have connected the nonlocality of quantum correlations, the incompatibility of quantum measurements, and the state's resources via the CHSH inequality. This revealed the counterintuitive effect, that sometimes more state resources are required to reach the same level of nonlocality, when the measurement resources are increased. While the CHSH inequality is by far the most studied Bell inequality, this behavior has, to the best of our knowledge, not been reported so far. The same effect is also prevalent for a steering inequality and thus excludes the existence of any possible conservation law for the necessary resources in states and measurements, regarding steering.

Several points are open for future research. First, one should investigate more general Bell scenarios, including the optimization over all Bell operators for a particular Bell inequality. Second, one could investigate further important resource measures. Finally, one should further investigate how the spectrum of Bell operators depends on the properties of the used measurement operators.

Acknowledgments. We thank Lennart Bittel, Federico Grasselli, Martin Kliesch, and Gláucia Murta for helpful discussions. This research was partially supported by the EU H2020 QuantERA ERA-NET Cofund in Quantum Technologies project QuICHE, and by the Federal Ministry of Education and Research (BMBF).

[1] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner, Bell nonlocality, Rev. Mod. Phys. **86**, 419 (2014).
[2] J. S. Bell, On the Einstein Podolsky Rosen paradox, Phys. Phys. Fiz. **1**, 195 (1964).
[3] J. I. de Vicente, On nonlocality as a resource theory and nonlocality measures, J. Phys. A: Math. Theor. **47**, 424017 (2014).
[4] A. Einstein, B. Podolsky, and N. Rosen, Can quantum-mechanical description of physical reality be considered complete? Phys. Rev. **47**, 777 (1935).
[5] R. F. Werner, Quantum states with Einstein-Podolsky-Rosen correlations admitting a hidden-variable model, Phys. Rev. A **40**, 4277 (1989).
[6] F. Hirsch, M. T. Quintino, and N. Brunner, Quantum measurement incompatibility does not imply Bell nonlocality, Phys. Rev. A **97**, 012129 (2018).
[7] E. Bene and T. Vértesi, Measurement incompatibility does not give rise to Bell violation in general, New J. Phys. **20**, 013021 (2018).
[8] A. A. Méthot and V. Scarani, An anomaly of non-locality, Quantum Inf. Comput. **7**, 157 (2007).
[9] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, Quantum entanglement, Rev. Mod. Phys. **81**, 865 (2009).
[10] A. Streltsov, H. Kampermann, S. Wölk, M. Gessner, and D. Bruß, Maximal coherence and the resource theory of purity, New J. Phys. **20**, 053058 (2018).
[11] J. Aberg, Quantifying superposition, arXiv:quant-ph/0612146.
[12] A. Streltsov, G. Adesso, and M. B. Plenio, *Colloquium*: Quantum coherence as a resource, Rev. Mod. Phys. **89**, 041003 (2017).
[13] H. Ollivier and W. H. Zurek, Quantum Discord: A Measure of the Quantumness of Correlations, Phys. Rev. Lett. **88**, 017901 (2001).
[14] S. Luo, Quantum discord for two-qubit systems, Phys. Rev. A **77**, 042303 (2008).
[15] K. Modi, T. Paterek, W. Son, V. Vedral, and M. Williamson, Unified View of Quantum and Classical Correlations, Phys. Rev. Lett. **104**, 080501 (2010).
[16] H. Zhu, M. Hayashi, and L. Chen, Coherence and entanglement measures based on Rényi relative entropies, J. Phys. A: Math. Theor. **50**, 475303 (2017).
[17] A. Streltsov, U. Singh, H. S. Dhar, M. N. Bera, and G. Adesso, Measuring Quantum Coherence with Entanglement, Phys. Rev. Lett. **115**, 020403 (2015).
[18] T. R. Bromley, M. Cianciaruso, and G. Adesso, Frozen Quantum Coherence, Phys. Rev. Lett. **114**, 210401 (2015).
[19] C. Spee, Certifying the purity of quantum states with temporal correlations, Phys. Rev. A **102**, 012420 (2020).
[20] M. Navascués, S. Pironio, and A. Acín, A convergent hierarchy of semidefinite programs characterizing the set of quantum correlations, New J. Phys. **10**, 073013 (2008).
[21] T. Moroder, J.-D. Bancal, Y.-C. Liang, M. Hofmann, and O. Gühne, Device-Independent Entanglement Quantification and Related Applications, Phys. Rev. Lett. **111**, 030501 (2013).
[22] T. Heinosaari, J. Kiukas, and D. Reitzner, Noise robustness of the incompatibility of quantum measurements, Phys. Rev. A **92**, 022115 (2015).
[23] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, Proposed Experiment to Test Local Hidden-Variable Theories, Phys. Rev. Lett. **23**, 880 (1969).
[24] E. G. Cavalcanti, S. J. Jones, H. M. Wiseman, and M. D. Reid, Experimental criteria for steering and the Einstein-Podolsky-Rosen paradox, Phys. Rev. A **80**, 032112 (2009).
[25] K. M. R. Audenaert and M. B. Plenio, When are correlations quantum?—Verification and quantification of entanglement by simple measurements, New J. Phys. **8**, 266 (2006).
[26] J. Eisert, F. G. S. L. Brandão, and K. M. R. Audenaert, Quantitative entanglement witnesses, New J. Phys. **9**, 46 (2007).
[27] O. Gühne, M. Reimpell, and R. F. Werner, Estimating Entanglement Measures in Experiments, Phys. Rev. Lett. **98**, 110502 (2007).
[28] D. Cavalcanti and M. O. T. Cunha, Estimating entanglement of unknown states, Appl. Phys. Lett. **89**, 084102 (2006).
[29] R. Horodecki, M. Horodecki, and P. Horodecki, Entanglement processing and statistical inference: The Jaynes principle can produce fake entanglement, Phys. Rev. A **59**, 1799 (1999).
[30] J. Batle and M. Casas, Nonlocality and entanglement in qubit systems, J. Phys. A: Math. Theor. **44**, 445304 (2011).
[31] J. Batle, M. Casas, A. R. Plastino, and A. Plastino, Inference schemes and entanglement determination, Phys. Rev. A **65**, 024304 (2002).

[32] J. Batle, M. Casas, A. R. Plastino, and A. Plastino, On the fake inferred entanglement associated with the maximum entropy inference of quantum states, J. Phys. A: Math. Gen. **34**, 6443 (2001).
[33] K. Bartkiewicz, B. Horst, K. Lemr, and A. Miranowicz, Entanglement estimation from Bell inequality violation, Phys. Rev. A **88**, 052105 (2013).
[34] F. Verstraete and M. M. Wolf, Entanglement versus Bell Violations and their Behavior under Local Filtering Operations, Phys. Rev. Lett. **89**, 170401 (2002).
[35] E. Chitambar and G. Gour, Quantum resource theories, Rev. Mod. Phys. **91**, 025001 (2019).
[36] Y. Luo and Y. Li, Robustness of purity: An operational approach to resource theory of purity, Eur. Phys. J. D **73**, 89 (2019).
[37] A. W. Marshall, I. Olkin, and B. C. Arnold, *Inequalities: Theory of Majorization and Its Applications* (Springer, New York, 2011).
[38] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevResearch.4.L012002 for more details on derivations, as well as for more details on the hierarchy of quantum resources in the $I3322$ scenario and a discussion about the incompatibility quantifier of C. This includes Refs. [51–54].
[39] Note that the hierarchy also holds for measures which are no distance in a strict mathematical sense, like the relative entropy, due to the inclusion relations of the void sets. It also holds for the generalized robustness.
[40] Y. Yao, X. Xiao, L. Ge, and C. P. Sun, Quantum coherence in multipartite systems, Phys. Rev. A **92**, 022112 (2015).
[41] A. Bera, T. Das, D. Sadhukhan, S. S. Roy, A. Sen(De), and U. Sen, Quantum discord and its allies: A review of recent progress, Rep. Prog. Phys. **81**, 024001 (2017).
[42] M. Steiner, Generalized robustness of entanglement, Phys. Rev. A **67**, 054305 (2003).
[43] V. Scarani and N. Gisin, Spectral decomposition of Bell's operators for qubits, J. Phys. A: Math. Gen. **34**, 6043 (2001).
[44] G. Svetlichny, Distinguishing three-body from two-body nonseparability by a Bell-type inequality, Phys. Rev. D **35**, 3066 (1987).
[45] M. Cianciaruso, T. R. Bromley, and G. Adesso, Accessible quantification of multiparticle entanglement, npj Quantum Inf. **2**, 16030 (2016).
[46] D. Collins and N. Gisin, A relevant two qubit Bell inequality inequivalent to the CHSH inequality, J. Phys. A: Math. Gen. **37**, 1775 (2004).
[47] L. J. Landau, On the violation of Bell's inequality in quantum theory, Phys. Lett. A **120**, 54 (1987).
[48] H. P. Robertson, The uncertainty principle, Phys. Rev. **34**, 163 (1929).
[49] S. L. Braunstein, A. Mann, and M. Revzen, Maximal Violation of Bell Inequalities for Mixed States, Phys. Rev. Lett. **68**, 3259 (1992).
[50] J. Kiukas and R. F. Werner, Maximal violation of Bell inequalities by position measurements, J. Math. Phys. **51**, 072105 (2010).
[51] S. Boyd and L. Vandenberghe, *Convex Optimization* (Cambridge University Press, Cambridge, UK, 2004).
[52] V. Vedral, M. B. Plenio, M. A. Rippin, and P. L. Knight, Quantifying Entanglement, Phys. Rev. Lett. **78**, 2275 (1997).
[53] H. Zhu, L. Chen, and M. Hayashi, Additivity and non-additivity of multipartite entanglement measures, New J. Phys. **12**, 083002 (2010).
[54] R. Uola, C. Budroni, O. Gühne, and J.-P. Pellonpää, One-to-One Mapping between Steering and Joint Measurability Problems, Phys. Rev. Lett. **115**, 230402 (2015).

Supplemental Material for "Quantifying necessary quantum resources for nonlocality"

Maximal expectation value for fixed generalized robustness of purity $P_R(\rho)$. —Here, we give more detailed derivation of Theorem 1 in the main text.

Theorem 1. *Given the Hermitian operator $I = \sum_{j=1}^{d} \mu_j |\Psi_j\rangle\langle\Psi_j|$ with $\mu_j \geq \mu_t$ for $j < t$ and a fixed robustness of purity $P_R(\rho)$ of a quantum state ρ. The maximal expectation value $\langle I\rangle_{max}$ can be achieved by $\rho = \sum_{i=1}^{r} \lambda_i |\Psi_i\rangle\langle\Psi_i|$, where $\lambda_i \geq 0$, $\sum_{i=1}^{r} \lambda_i = 1$, $\lambda_i \geq \lambda_s$ for $i < s$, and is given by*

$$\langle I\rangle_{max} = \sum_{j=1}^{r} \mu_j \lambda_j, \tag{1}$$

where r is an integer s.t. $\dfrac{1}{r-1} > \lambda_1 \geq \dfrac{1}{r}$ and all eigenvalues λ_i for $i \in \{1, \cdots, r-1\}$ are equal to $\lambda_1 = (1 + P_R)/d$.

Proof. It holds $P_R(\rho) = d\lambda_1(\rho) - 1$, i.e. the constraint of fixed purity $P_R(\rho)$ depends only on the largest eigenvalue of ρ and we have to do the optimization w.r.t. the remaining degrees of freedom. The generalization of Ruhe's trace inequality [1] states that if A, B are $d \times d$ Hermitian matrices, then

$$\sum_{i=1}^{d} \eta_i(A)\eta_{d-i+1}(B) \leq \mathrm{Tr}(AB) \leq \sum_{i=1}^{d} \eta_i(A)\eta_i(B) \tag{2}$$

where $\eta_i(A)$ denotes the i-th eigenvalue of A and the eigenvalues η_i are ordered in descending order. This simply means that in order to maximize (minimize) the expectation value $\langle I\rangle = \mathrm{Tr}(\rho I)$, ρ has to be diagonal in the same basis as I. Further, the i-th largest eigenvalue of ρ has to be multiplied with the i-th largest (smallest) eigenvalue of I.This means the ordering of the eigenstates of I and ρ are the same with respect to their eigenvalues. Since λ_1 is by assumption the largest eigenvalue of ρ, all other λ_i cannot be larger. In order to maximize the magnitude of the expectation value all λ_i should be as big as possible, which means equal to λ_1. This however is only possible for $r-1$ eigenvalues, where r (which describes the rank of ρ) is the largest integer such that the $\{\lambda_i\}$ describe a normalized quantum state. The remaining non-zero eigenvalue λ_r is given by the normalization constraint which also provides the upper bound to λ_1. The lower bound comes from the requirement that λ_1 is the largest eigenvalue, which finishes the proof. □

As an extension, we show in the following Lemma that Theorem 1 can also be used to determine the minimal λ_1 for a fixed expectation value $\langle I\rangle_{max}$.

Lemma 1. *Given the Hermitian operator $I = \sum_{j=1}^{d} \mu_j |\Psi_j\rangle\langle\Psi_j|$ with $\mu_j \geq \mu_t$ for $j < t$. The minimal robustness of purity $P_R(\rho)$ of a quantum state $\rho = \sum_{i=1}^{d} \lambda_i |\phi_i\rangle\langle\phi_i|$, where $\lambda_i \geq 0$ and $\sum_{i=1}^{d} \lambda_i = 1$, $\lambda_i \geq \lambda_s$ for $i < s$, achieving the expectation value $\langle I\rangle_{max} \geq \dfrac{1}{d}\mathrm{Tr}(I)$ is determined by the equation*

$$\langle I\rangle_{max} = \sum_{j=1}^{r} \mu_j \lambda_j, \tag{3}$$

provided that $\dfrac{1}{r-1} > \lambda_1 \geq \dfrac{1}{r}$, and all eigenvalues λ_i for $i \in \{1, \cdots, r-1\}$ are equal to λ_1.

Proof. Since $P_R(\rho) = d\lambda_1(\rho) - 1$, we need to minimize λ_1. The expectation value in general is given by

$$\langle I\rangle_{\max} = \mathrm{Tr}(I\rho) = \sum_{i,j} \mu_i \lambda_j |\langle\Psi_i|\phi_j\rangle|^2. \tag{4}$$

We achieve a minimization of λ_1 by exploiting the following two observations. First, since $\langle I\rangle_{max} \geq \dfrac{1}{d}\mathrm{Tr}(I)$ [2] we need to maximize for fixed j the term $\sum_i \mu_i \lambda_j |\langle\Psi_i|\phi_j\rangle|^2$, i.e. we need to appropriately choose the eigenbasis of ρ which according to Theorem 1 will be done by choosing ρ diagonal in the same basis as I, or more specifically $|\phi_j\rangle = |\Psi_j\rangle \;\forall\, j$. This can be seen by realizing that $\sum_i \mu_i \lambda_j |\langle\Psi_i|\phi_j\rangle|^2$ is upper bounded by $\lambda_j \mu_j$. Hence, if $|\phi_j\rangle \neq |\Psi_j\rangle$, λ_1 has to be larger than necessary because some part of the contribution towards the expectation value is lost due to a sub-optimal basis choice. Second, λ_1 will be minimal for maximal possible $\lambda_2, \lambda_3, \cdots, \lambda_r$, where r is just an index for now. More precisely, it is optimal to choose as many λ_i equal to λ_1 as possible, since by definition λ_1 is the maximal eigenvalue and for any lower value of the λ_i, λ_1 would again be larger than necessary since we did not choose the maximal contribution of the terms $\lambda_i \mu_i \ \forall\ i$ towards the expectation value. However, we still have to incorporate that ρ is a normalized quantum state, which means not all λ_i can actually be equal to λ_1. In general it possible to choose all λ_i equal to λ_1 for $i \in \{1, \cdots, r-1\}$ and to determine the smallest non-zero eigenvalue λ_r by normalization. Hence r denotes the rank of ρ. This leads to

$$\langle I \rangle_{\max} = \sum_{j=1}^{r} \mu_j \lambda_j, \tag{5}$$

where we still do not know the value of the rank r. However, we can just make an Ansatz for some $r \in \{1, \cdots, d\}$ and check whether the conditions $\frac{1}{r-1} > \lambda_1 \geq \frac{1}{r}$, which are necessary for ρ to be a normalized density matrix and for λ_1 to be the largest eigenvalue of ρ are fulfilled or not. □

Equality of quantum resources for two qubits.—Here, we give a detailed derivation of Theorem 2 in the main text.

Theorem 2. *Given a Bell operator of the form*

$$I = \sum_{x,y} g_{x,y}\ A_x \otimes B_y, \tag{6}$$

with real coefficients $g_{x,y}$ and local observables $A_x = \vec{a}_x \cdot \vec{\sigma}$, $B_y = \vec{b}_y \cdot \vec{\sigma}$ where $\vec{a}_x, \vec{b}_y$ are Bloch vectors and $\vec{\sigma}$ is the vector containing the Pauli matrices. For a fixed expectation value $\langle I \rangle = L + v$, where L is the local bound and $v > 0$, there exists a two-qubit quantum state ρ_{opt} which simultaneously minimizes the generalized robustnesses of purity P_R, coherence with respect to all product bases C_R, and entanglement E_R.

Proof. The generalized robustness is defined as

$$G_R(\rho) := \min_{\tau} \left\{ x | x \geq 0, \exists \text{ a state } \tau, \frac{\rho + x\tau}{1+x} \in \mathcal{V} \right\}, \tag{7}$$

where the set $\mathcal{V}$ consists of the respective void states and τ can be any quantum state. For more information see the main text. Note that Bell operators of the form (6) are (up to local unitaries) diagonal in the Bell basis [3]. This means $I = \sum_j \mu_j |\Psi_j\rangle\langle\Psi_j|$ where the $\{|\Psi_j\rangle\}$ are maximally entangled states. As a consequence, the states of minimal entanglement (for all measures) $\rho_{BDS} = \sum_i \lambda_i |\Psi_j\rangle\langle\Psi_j|$ are also diagonal in this Bell basis [4]. These states are entangled iff $\lambda_1(\rho_{BDS}) > 1/2$ and it was shown in [5] that the generalized robustness of entanglement for two-qubit BDS is given by $E_R(\rho_{\text{BDS}}) = 2\lambda_1(\rho_{\text{BDS}}) - 1$. This means minimizing $E_R(\rho_{\text{BDS}})$ is equivalent to minimizing the generalized robustness of purity $P_R(\rho) = d\lambda_1(\rho) - 1$, since both are monotonic functions of $\lambda_1(\rho)$. Due to Lemma (1) and the fact that $\lambda_1(\rho) > 1/2$, the state ρ_{opt} can always chosen to be of the form $\rho_{\text{opt}} = \lambda_1 |\Psi_1\rangle\langle\Psi_1| + (1-\lambda_1)|\Psi_2\rangle\langle\Psi_2|$ where λ_1 can be determined from eq. (5). It is convenient to choose $\tau = |\Psi_2\rangle\langle\Psi_2|$ as optimal noisy state in eq.(7) to minimize the generalized robustness of entanglement. This is always possible since mixing the minimal noise necessary, i.e. $x = 2\lambda_1 - 1$ of τ with ρ_{opt} will end up in a separable state

$$\frac{\rho_{\text{opt}} + (2\lambda_1 - 1)\tau}{1 + (2\lambda_1 - 1)} = \frac{1}{2}(|\Psi_1\rangle\langle\Psi_1| + |\Psi_2\rangle\langle\Psi_2|) =: \xi. \tag{8}$$

Since all considered quantifiers are invariant under local unitaries, we fix the Bell operator w.l.o.g. to be of the form

$$\begin{aligned} I =& \mu_1 |\Phi^+\rangle\langle\Phi^+| + \mu_2 |\Phi^-\rangle\langle\Phi^-| \\ &+ \mu_3 |\Psi^+\rangle\langle\Psi^+| + \mu_4 |\Psi^-\rangle\langle\Psi^-|, \end{aligned} \tag{9}$$

where $|\Phi^\pm\rangle = \frac{1}{\sqrt{2}}(|00\rangle \pm |11\rangle)$ and $|\Psi^\pm\rangle = \frac{1}{\sqrt{2}}(|01\rangle \pm |10\rangle)$ or any permutation of the eigenvalues. It follows now directly from the form of the closest separable state ξ in eq. (8) that it is also incoherent in some product basis, which means it is also the closest incoherent state to ρ_{opt}. This is because the hierarchy $P_R \geq C_R \geq D_R \geq E_R$ has to hold [6] (see also the main text), which means $C_R(\rho_{\text{opt}})$ is lower bounded by $E_R(\rho_{\text{opt}})$. It can easily be seen that ξ is incoherent for the mixtures of $\{|\Phi^+\rangle\langle\Phi^+|, |\Phi^-\rangle\langle\Phi^-|\}$ or $\{|\Psi^+\rangle\langle\Psi^+|, |\Psi^-\rangle\langle\Psi^-|\}$ and the computational basis. For the other combinations one finds as optimal bases tensor products of the eigenstates of the Pauli matrices σ_x or σ_y. The eigenstates of the the Pauli matrix σ_z are given by $\{|0\rangle, |1\rangle\}$, the eigenstates of σ_x by $|\pm\rangle = \frac{1}{\sqrt{2}}(|0\rangle \pm |1\rangle)$ and those of σ_y are given by $|R/L\rangle = \frac{1}{\sqrt{2}}(|0\rangle \pm i|1\rangle)$. Below we list for any of the possible rank-2 combinations the specific decomposition of the closest separable state ξ into these states. Namely, it can easily be verified that

$$
\begin{aligned}
&\{|\Phi^+\rangle\langle\Phi^+|, |\Phi^-\rangle\langle\Phi^-|\} \Rightarrow \xi = \frac{1}{2}(|00\rangle\langle 00| + |11\rangle\langle 11|), \\
&\{|\Phi^+\rangle\langle\Phi^+|, |\Psi^+\rangle\langle\Psi^+|\} \Rightarrow \xi = \frac{1}{2}(|++\rangle\langle ++| + |--\rangle\langle --|), \\
&\{|\Phi^+\rangle\langle\Phi^+|, |\Psi^-\rangle\langle\Psi^-|\} \Rightarrow \xi = \frac{1}{2}(|RL\rangle\langle RL| + |LR\rangle\langle LR|), \\
&\{|\Phi^-\rangle\langle\Phi^-|, |\Psi^+\rangle\langle\Psi^+|\} \Rightarrow \xi = \frac{1}{2}(|RR\rangle\langle RR| + |LL\rangle\langle LL|), \\
&\{|\Phi^-\rangle\langle\Phi^-|, |\Psi^-\rangle\langle\Psi^-|\} \Rightarrow \xi = \frac{1}{2}(|-+\rangle\langle -+| + |+-\rangle\langle +-|), \\
&\{|\Psi^+\rangle\langle\Psi^+|, |\Psi^-\rangle\langle\Psi^-|\} \Rightarrow \xi = \frac{1}{2}(|01\rangle\langle 01| + |10\rangle\langle 10|),
\end{aligned}
\tag{10}
$$

all other cases of rank-2 BDS are equivalent under local unitaries to one of the above cases. This finishes the proof. □

Maximal expectation value for fixed Rényi 2-purity.—Here, we give a detailed derivation of an analogon to Theorem 1 regarding the Rényi 2-purity $\mathcal{P}_2(\rho) = \log_2(d\mathrm{Tr}(\rho^2))$.

Theorem 3. *Given the Hermitian operator $I = \sum_{j=1}^{d} \mu_j |\Psi_j\rangle\langle\Psi_j|$ with $\mu_j \geq \mu_t$ for $j < t$ and a fixed Rényi 2-purity $\mathcal{P}_2(\rho)$ of a quantum state ρ. The maximal expectation value $\langle I\rangle_{max}$ can be achieved by $\rho = \sum_{i=1}^{r} \lambda_i |\Psi_i\rangle\langle\Psi_i|$, where $\lambda_i \geq 0$, $\sum_{i=1}^{r} \lambda_i = 1$, $\lambda_i \geq \lambda_s$ for $i < s$, and is given by*

$$
\langle I\rangle_{max} = \frac{G + \sqrt{(1 - \frac{r}{d} 2^{\mathcal{P}_2})(G^2 - Hr)}}{r},
\tag{11}
$$

where $G = \sum_i^r \mu_i$, $H = \sum_i^r \mu_i^2$, and $r \in \{1, \cdots, d\}$ is the largest integer s.t.

$$
\lambda_i = \frac{(r\langle I\rangle_{max} - G)\mu_i + H - G\langle I\rangle_{max}}{Hr - G^2} \geq 0 \ \forall \ i \leq r.
\tag{12}
$$

Proof. We first show that we can solve a different but connected optimization task which will lead to a proof of the theorem. The first simplification will be, that instead of considering the Rényi 2-purity $\mathcal{P}_2$ directly, we can just consider $P_L = \mathrm{Tr}(\rho^2)$, since the logarithm is a monotonic function of P_L. We emphasize again that, P_L is no proper measure of purity (see main text), even though it is known in the literature as linear-purity. Note that for any given P_L, the maximal achievable expectation value is s.t. $\langle I\rangle_{\max} \geq \frac{1}{d}\mathrm{Tr}(I)$. This follows from the fact that $\langle I\rangle = \frac{1}{d}\mathrm{Tr}(I)$ is the expectation value achieved by the maximally mixed state (which minimizes P_L) and for any other $P_L' > P_L(\mathbb{1}/d)$ one is able to choose a state $\Omega = x|\Psi_1\rangle\langle\Psi_1| + (1-x)\frac{\mathbb{1}}{d}$ with appropriately chosen $x \in [0,1]$ s.t. $P_L' = \mathrm{Tr}(\Omega^2)$. Since $|\Psi_1\rangle\langle\Psi_1|$ is the eigenstate corresponding to the largest eigenvalue of I it follows trivially that $\langle I\rangle_\Omega \geq \frac{1}{d}\mathrm{Tr}(I)$. This allows us to formulate an alternative optimization problem which proves the theorem. Given the Hermitian operator $I = \sum_j \mu_j |\Psi_j\rangle\langle\Psi_j|$ with fixed expectation value $\langle I'\rangle$. We want to find the minimal P_L^* of a valid quantum state that achieves the expectation value $\langle I'\rangle$. We will now show by contradiction that for the minimal P_L^* it holds $\langle I'\rangle = \langle I\rangle_{\max}$. First, it is trivial that $\langle I'\rangle > \langle I\rangle_{\max}$ leads to a contradiction since $\langle I\rangle_{\max}$ is by assumption the maximal expectation value for a given P_L^*. Second, if $\langle I'\rangle < \langle I\rangle_{\max}$ we could construct a state $\tilde{\rho} = t\rho_{\max} + (1-t)\frac{\mathbb{1}}{d}$ where $\rho_{\max}$ is a state with P_L^* achieving the expectation value $\langle I\rangle_{\max}$ and choose $t \in (0,1)$ s.t. $\langle I\rangle_{\tilde{\rho}} = \langle I'\rangle$. It follows now for $P_L(\tilde{\rho})$ that

$$
P_L(\tilde{\rho}) = P_L(t\rho_{\max} + (1-t)\frac{\mathbb{1}}{d}) < tP_L(\rho_{\max}) + (1-t)P_L(\frac{\mathbb{1}}{d}) \leq P_L^*,
\tag{13}
$$

where we used the strict convexity of P_L (which is the square of the Frobenius norm) and the fact that $P_L(\rho_{\max}) = P_L^*$. This however is a contradiction, since P_L^* is by assumption the minimum for states achieving the expectation value

$\langle I' \rangle$ and we showed that $P_L(\tilde{\rho})$ would be smaller while achieving the expectation value $\langle I' \rangle$. This allows us to solve the minimization problem for a fixed expectation value and use the optimal state of minimal P_L to solve the problem we consider in the theorem. With the same argumentation as for the generalized robustness or alternatively the proof shown in [3] (originally in the context of the CHSH inequality) we can reduce the problem of minimizing P_L under the Bell constraint s.t. the Bell operator $I = \sum_i \mu_i |\Psi_i\rangle\langle\Psi_i|$ and the optimal quantum state $\rho_{\text{opt}} = \sum_i \lambda_i |\Psi_i\rangle\langle\Psi_i|$ will be diagonal in the same basis. Note that the eigenvalues of both operators are ordered in descending order i.e., $\lambda_i \geq \lambda_s$ for $i < s$ and $\mu_j \geq \mu_t$ for $j < t$. The Lagrangian of the problem is given by

$$\mathcal{L}(\lambda_i, \alpha, \beta) = \sum_i \lambda_i^2 - \alpha(\sum_i \lambda_i - 1) - \beta(\sum_i \lambda_i \mu_i - \langle I \rangle), \tag{14}$$

where α is the multiplier according to the normalization constraint and β the multiplier for the expectation value constraint. We ignored for now the non-negativity of the eigenvalues but will come back to it later. Note that due to the eigenvalue ordering $\langle I \rangle \geq \frac{1}{d}\sum\limits_i \mu_i = \frac{1}{d}\mathrm{Tr}(I)$, since $\lambda_i \mu_i > \lambda_s \mu_s$ for $i < s$. Obviously this does not represent a loss of generality for our theorem since we are only interested in values $\langle I \rangle \geq \frac{1}{d}\sum\limits_i \mu_i = \frac{1}{d}\mathrm{Tr}(I)$ anyway, as shown above. To find an optimum, we have to take the partial derivatives of $\mathcal{L}(\lambda_i, \alpha, \beta)$ with respect to the eigenvalues and the Lagrange multiplier. In the case of the multipliers we simply retrieve the constraints, for the eigenvalues we find

$$\frac{\partial \mathcal{L}}{\partial \lambda_k} = 2\lambda_k - \alpha - \beta\mu_k \stackrel{!}{=} 0, \tag{15}$$

which results in

$$\lambda_k = \frac{1}{2}(\beta\mu_k + \alpha). \tag{16}$$

From the normalization constraint $\sum_k \lambda_k = 1$ we get

$$\alpha = \frac{2 - \beta \sum\limits_{k=1}^{r} \mu_k}{r}, \tag{17}$$

where the sum runs from $k = 1$ until the rank r, which means we make an Ansatz for a rank r solution of the problem, i.e. all eigenvalues λ_i of the state are zero $\forall\ i > r$. This method is not restrictive, in the sense that we are still able to find the true optimizer of the problem including the positivity constraints. $P_L = \mathrm{Tr}(\rho^2)$ is strictly convex, as it is the square of the Frobenius norm, which means there is unique global minimum. The projection (in the Frobenius norm) of a quasi-state η with negative eigenvalues onto the feasible set of proper quantum states ρ will be of smaller rank than η itself. Therefore, we can start by making an Ansatz for an optimal state ρ_{opt} of full rank $r = d$. If this is a proper quantum state according to (16), it is the state of minimal P_L consistent with the expectation value constraint. If the state is not a proper quantum state, we consider solutions of rank $r = d - 1$. By iteratively decreasing the rank, we will find a solution ρ_{opt} which is a proper quantum state and we do not have to consider states with an even lower rank. In the following, we will introduce the quantities $G = \sum_k^r \mu_k$ and $H = \sum_k^r \mu_k^2$. Using the Bell value constraint we find

$$\beta(Hr - G^2) = 2\langle I \rangle r - 2G, \tag{18}$$

which defines β provided that $Hr - G^2 \neq 0$. It is easy to see that $Hr - G^2 = 0$ iff $\mu_i = \mu_1\ \forall\ i \leq r$ i.e. the operator has r degenerate largest eigenvalues μ_1. This is a consequence of the Cauchy-Schwarz inequality when we consider the inner product of a vector $\vec{\mu}$ containing the r eigenvalues μ_i and a vector $\vec{\mathbb{1}}$ containing only ones. The case $Hr - G^2 = 0$ can however only lead to the optimal state ρ_{opt} iff $\langle I \rangle = \mu_1$ which follows from the requirement that the r.h.s. of eq. (18) also vanishes. In this case the problem is independent of β and we find $\lambda_k = \frac{1}{r}\ \forall\ k \leq r$. If $Hr - G^2 \neq 0$ we find the optimal state ρ_{opt} by taking the eigenvalues from eq. (16) with the corresponding eigenstates from the operator I. Reversely, this spectrum of eigenvalues can now be used to find the maximum expectation value $\langle I \rangle_{\max}$ for a fixed P_L. Multiplying (16) with λ_k and summing over k leads to $P_L = \frac{1}{2}(\beta\langle I \rangle + \alpha)$, where we identified $P_L = \mathrm{Tr}(\rho^2) = \sum_{k=1}^r \lambda_k^2$. By solving for $\langle I \rangle = \langle I \rangle_{\max} = \sum_{k=1}^r \mu_k \lambda_k$ we will find

$$\langle I \rangle_{\max} = \frac{G + \sqrt{(1 - P_L r)(G^2 - Hr)}}{r}. \tag{19}$$

From eq. (19) we see that for $P_L \geq 1/r_{\text{deg}}$ where r_{deg} is the number of degenerate largest eigenvalues of I the maximum is given by μ_1, since $G^2 - Hr_{\text{deg}} = 0$ which can trivially be checked to be correct. While for $P_L \leq 1/r_{\text{deg}}$ for only r_{deg} times degenerate eigenvalues μ_1 of I, we see that $G^2 - Hr \neq 0$ which shows that eq. (19) is valid for all operators I. The theorem follows now by rewriting the P_L in terms of the Rényi 2-purity $\mathcal{P}_2$ and the requirement that all $\lambda_k \; \forall \; k = 1, \cdots, r$ are positive semidefinite (the remaining $d - r$ eigenvalues are zero) and inserting the Lagrange multipliers into the expression (16). Note that the optimal r for a given P_L can be found by the above described feasibility check. In other words, we check whether (16) leads to a valid quantum state for the rank r, given P_L and maximal expectation value $\langle I \rangle_{\max}$. Note further that in order to derive (19), we have to solve a quadratic equation, which generally has two solutions. However, we are only interested in the maximum of those two solutions which is the solution corresponding to taking the positive root in (19). □

The hierarchy of quantum resources in the context of the I3322 inequality.—Here, we show that the hierarchy

$$P_R \geq C_R \geq D_R \geq E_R, \tag{20}$$

is not always tight for the states minimizing the respective generalized robustnesses for a given Bell operator I. The I3322 inequality is given by [7]

$$\begin{aligned} &\langle A_1 \rangle + \langle A_2 \rangle - \langle B_1 \rangle - \langle B_2 \rangle \\ &+\langle A_1 B_1 \rangle + \langle A_2 B_1 \rangle + \langle A_3 B_1 \rangle + \langle A_1 B_2 \rangle \\ &+\langle A_2 B_2 \rangle - \langle A_3 B_2 \rangle + \langle A_1 B_3 \rangle - \langle A_2 B_3 \rangle \leq 4. \end{aligned} \tag{21}$$

By generating random projective measurements $\{A_{a|x}\}$ (and similar for Bob) which lead to the observables $A_x = A_{2|x} - A_{1|x}$ we searched for Bell operators I, which will lead to a non-tight hierarchy. We found that the following measurements of Alice and Bob do lead to such a case. The measurements (rounded to four digits) are given in the computational basis $\{|0\rangle, |1\rangle\}$ by

$$\begin{aligned} A_{1|1} &= \begin{pmatrix} 0.4379 & 0.3455 + 0.3560i \\ 0.3455 - 0.3560i & 0.5621 \end{pmatrix}, \\ A_{1|2} &= \begin{pmatrix} 0.6885 & 0.3964 - 0.2394i \\ 0.3964 + 0.2394i & 0.3115 \end{pmatrix}, \\ A_{1|3} &= \begin{pmatrix} 0.9187 & -0.0737 + 0.2632i \\ -0.0737 - 0.2632i & 0.0813 \end{pmatrix}, \\ B_{1|1} &= \begin{pmatrix} 0.6973 & 0.0630 - 0.4551i \\ 0.0630 + 0.4551i & 0.3027 \end{pmatrix}, \\ B_{1|2} &= \begin{pmatrix} 0.8982 & -0.2538 + 0.1645i \\ -0.2538 - 0.1645i & 0.1018 \end{pmatrix}, \\ B_{1|3} &= \begin{pmatrix} 0.6472 & -0.0110 + 0.4777i \\ -0.0110 - 0.4777i & 0.3528 \end{pmatrix}, \end{aligned} \tag{22}$$

where the remaining POVM-elements are obtained by the completeness relation $\sum_a A_{a|x} = \mathbb{1}$ (and similar for Bob). For the generalized robustnesses of purity, coherence, and entanglement we find for these settings and a required Bell value of $\langle I \rangle = 4.001$ that,

$$P_R > C_R > E_R, \tag{23}$$

with $P_R = 2.6756$, $C_R = 0.8418$, and $E_R = 0.8291$. We calculated the purity robustness analytically with the methods described in the main text. The entanglement robustness was determined by semidefinite programming [8], and for the coherence robustness we used a combination of a simplex algorithm and semidefinite programming over all Bell operators of the form

$$\tilde{I} = (U_A \otimes U_B) I (U_A \otimes U_B)^\dagger, \tag{24}$$

where U_A, U_B are local unitaries. For the simplex algorithm we used different randomly initialized starting points which all lead to the same result, suggesting it is the true minimum. Note that the gap between purity and coherence is not just due to a trivial factor like for correlation inequalities (see main text) but due to entirely different optimal states. One reason for this is the fact that the considered Bell-operator I is, in comparison to the correlation inequality case, not diagonal in the Bell-basis. As a consequence of that, the states of minimal entanglement are no longer diagonal in the eigenbasis of I and therefore different from the states of minimal purity.

Discussion on the definition of the quantifier C.—In the main text we defined the quantifier $C = |||[A_1, A_2]||| \cdot |||[B_1, B_2]|||$ in order to judge the quality of the observables in the CHSH scenario. While the magnitude of a single party's commutator is a valid incompatibility monotone for projective measurements and directly related to the robustness of the observable with respect to white noise [9], it is not clear that C has a similar meaning for the measurements of both parties. As we argued in the main text, C is meaningful for the CHSH inequality since it determines the eigenvalues of the Bell operator and especially the maximal possible violation enabled by the observables $\{A_1, A_2, B_1, B_2\}$. In more general Bell scenarios, the ability to show nonlocality with some measurements is a distinct resource from measurement incompatibility [10, 11]. However, there exists so far no resource theory or straightforward quantification for the ability of observables to show nonlocality. In Fig. 1 we show that indeed, the general resource requirement of two-qubit states for higher Bell values increases. We can reduce the discussion to Bell-diagonal states ρ_{BDS} as they minimize the needed resources. (see Theorem 3). The curves in Fig. 1 were obtained by minimizing the largest eigenvalue $\lambda_1(\rho_{\mathrm{BDS}})$ for a given CHSH-Bell value. To do so, we used that the optimal C for fixed $\lambda_1(\rho_{\mathrm{BDS}})$ is given by $C_{\max} = \dfrac{4(2\lambda_1 - 1)}{2\lambda_1^2 - 2\lambda_1 + 1}$. We also use that the formula for the maximal expectation value $\langle I\rangle_{\max} = \sqrt{4+C}\lambda_1 + \sqrt{4-C}(1-\lambda_1)$ can also be used to fix $\langle I\rangle_{\max}$ and calculate the minimal $\lambda_1(\rho_{\mathrm{BDS}})$ needed. The generalized robustness of purity and entanglement are both determined by $\lambda_1(\rho_{\mathrm{BDS}})$. If we consider the sum of

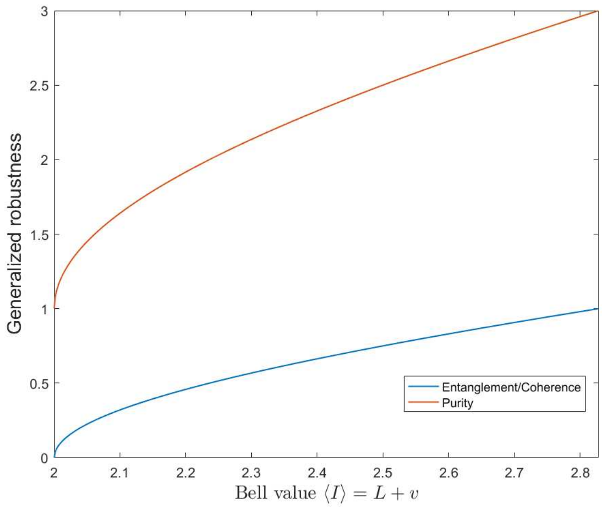

Figure 1. The minimal resources in terms of generalized robustness of purity $P_R(\rho_{\mathrm{BDS}}) = 4\lambda_1(\rho_{\mathrm{BDS}}) - 1$ and entanglement $E_R(\rho_{\mathrm{BDS}}) = 2\lambda_1(\rho_{\mathrm{BDS}}) - 1$ needed to achieve a fixed Bell value. As stated in the main text, coherence and entanglement are equivalent (and discord as well) while purity is the largest of these resources.

the single party's incompatibilities $\tilde{C} = C_A + C_B = |||[A_1, A_2]||| + |||[B_1, B_2]|||$, which due to linearity might be more intuitive to do, it would lead to eigenvalues

$$\mu_{1/4} = \pm\sqrt{4 + \frac{(C_A + C_B)^2 - C_A^2 - C_B^2}{2}},\qquad(25)$$
$$\mu_{2/3} = \pm\sqrt{4 - \frac{(C_A + C_B)^2 - C_A^2 - C_B^2}{2}}.$$

This however, is a problem since for fixed values of $\tilde{C}$ there are multiple possible eigenvalue distributions $\{\mu_i\}$ as one can easily see by fixing $\tilde{C} = 2$ and comparing the cases $C_A = 2, C_B = 0$ and $C_A = 1, C_B = 1$. This means there cannot be a well defined function of the quantum resources of the state depending $\tilde{C}$ in general. Note that in the special case $C_A = C_B$, the eigenvalues will again be well defined. Finally, we could also be interested in the functional form of the resources depending on C_A and C_B. In this case, the fundamental effect that maximal incompatible measurements for Alice and Bob will not lead to the minimal required state resources will remain. Especially for very small violations $v = 10^{-3}$, we show in Fig. 2 that the behavior is essentially the same. While it might be possible that the behaviour depends also qualitatively on the choice of the incompatibility measure for the individual parties we want to argue that this is unlikely. The commonly used measures of measurement incompatibility are all based on the robustness with respect to the set of jointly measurable POVMs and measurement incompatibility reduces to

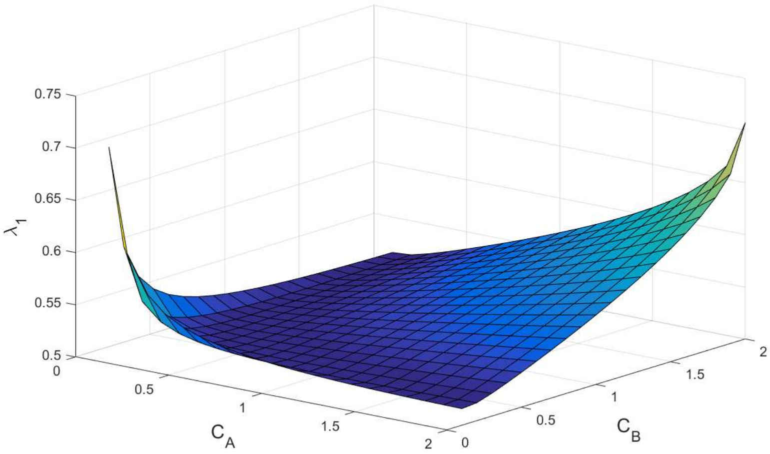

Figure 2. The necessary λ_1 vs. the single party's incompatibility quantifier C_A, C_B for a violation of $v = 10^{-3}$. The plot shows that higher incompatibility resources need higher entanglement to achieve the same violation for some parameter region of C_A, C_B.

commutation for projective measurements [9]. We therefore expect that the behaviour reported here does not depend on the specific measure. To further support this claim, we analyze the qualitative behaviour of the state resources depending on the generalized robustness of measurement incompatibility [12]. For a set of POVMs $\{M_{a|x}\}_{a,x}$, the generalized robustness of measurement incompatibility is given by

$$\mathrm{C_{GRA}}(\{M_{a|x}\}_{a,x}) = \min\{t \geq 0 | \frac{M_{a|x} + tN_{a|x}}{1+t} = O_{a|x} \ \in \mathrm{JM} \ \forall \ a, x\}, \tag{26}$$

where $\{N_{a|x}\}_{a,x}$ is any set of POVMs and JM denotes the set of POVMs $\{O_{a|x}\}_{a,x}$ that are jointly measurable (as generalization of commutativity from projective measurements to POVMs), hence classical or resource free. For more details, see [12]. We define now in correspondence to Fig. (2) in the main text the (global) incompatibility $C_{GR} = \mathrm{C_{GRA}C_{GRB}}$ as the product of the incompatibility of Alice and Bob. As one can see from Fig. (3), the qualitative dependence of the state resources is the same as in Fig. (2) of the main text. To obtain these results, we

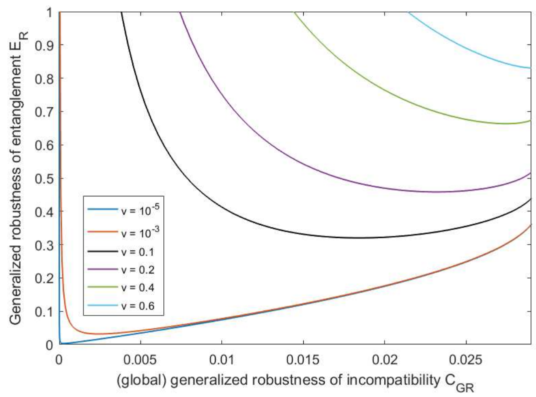

Figure 3. The minimal generalized robustness of entanglement $E_R(\rho)$ for a given level of incompatibility C_{GR} for different amounts of desired violation v. The qualitative behaviour is equivalent to that of Fig. (2) in the main text.

sampled projective measurements over the whole range of incompatibility values, which we calculated via semidefinite programming [8, 12] in tandem with Theorem 1 and Lemma 1 to calculate the minimal needed largest eigenvalue $\lambda_1(\rho)$ and therefore E_R. However, to get a complete understanding on the functional behaviour of the minimal state resources on the incompatibility, further studies of different incompatibility monotones are necessary.

Variation of state resource quantifiers. Here, we discuss the qualitative influence when we chose a quantifier for the state's resources different than the generalized robustness. It turns out that only purity can show a qualitatively different behaviour as a function of C. This is the case for the relative entropy of purity.

First, since the entanglement of Bell diagonal states

$$\rho_{BDS} = \sum_{i}^{4} \lambda_i |\Psi_j\rangle\langle\Psi_j|, \tag{27}$$

where the $|\Psi_j\rangle$ span an orthonormal basis of maximally entangled states, is completely characterized by their largest eigenvalue λ_1 and the resource measures are by definition resource monotones, it does not matter which distance-based entanglement quantifier we choose, since all of them will be monotonic functions of λ_1. By the same argumentation we can see, that for any rank-2 BDS state, the closest separable state ξ (see eq. (8)) does not change when we change the quantifier. Now the proof of Theorem 2 guarantees, that also the minimal coherence only depends on λ_1 and not on the concrete quantifier, since the closest separable state ξ is incoherent in some product basis. However, if we chose a different purity measure, it will in general depend on the whole spectrum of the state and not just the largest eigenvalue. This can lead to a potentially different functional behaviour with respect to the incompatibility quantifier C. We show that this is indeed the case by considering the relative entropy of purity. The relative entropy is defined as

$$S(\rho||\xi) = \mathrm{Tr}(\rho \log_2 \rho) - \mathrm{Tr}(\rho \log_2 \xi). \tag{28}$$

In the case of purity, the relative entropy reduces to

$$S_P(\rho||\mathbb{1}/d) = \log d - S(\rho), \tag{29}$$

where $S(\rho) = -\mathrm{Tr}(\rho \log \rho)$ denotes the von Neumann entropy. Mathematically, maximizing the von Neumann entropy under the constraint that $\langle I \rangle = \mathrm{Tr}(\rho I)$ for quantum states ρ is a standard textbook task [8]. Physically, we find that the relative entropy of purity shows a distinct qualitative behaviour with respect to the incompatibility quantifier C from all the other considered resource measures. More precisely, the relative entropy of purity is not only minimized by different states, it also monotonically decreases with increasing C as depicted in Fig. (4).

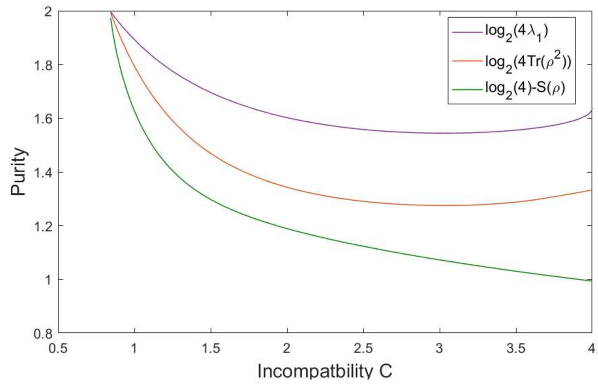

Figure 4. Comparison of the logarithm of the generalized robustness of purity, the so-called log-robustness (purple) with the Rényi 2-purity $\mathcal{P}_2$ (orange), and the relative entropy of purity (green) for an exemplary violation of $v = 0.2$. The minimal relative entropy decreases with increasing C, for all C, contrary to the other measures as described in the text.

[1] A. W. Marshall, I. Olkin, and B. C. Arnold, *Inequalities: Theory of Majorization and Its Applications* (Springer New York, 2011).

[2] As $P_R(\mathbb{1}/d) = 0$ and $\langle I \rangle_{\mathbb{1}/d} = \mathrm{Tr}(I)/d$. For expectation values $\langle I \rangle <$ I$/d$ the proof works similar by simply exchanging the order of the μ_i from decreasing to ascending.

[3] F. Verstraete and M. M. Wolf, Entanglement versus bell violations and their behavior under local filtering operations, Phys. Rev. Lett. **89**, 170401 (2002).

[4] R. Horodecki, M. Horodecki, and P. Horodecki, Entanglement processing and statistical inference: The jaynes principle can produce fake entanglement, Phys. Rev. A **59**, 1799 (1999).

[5] H. Zhu, L. Chen, and M. Hayashi, Additivity and non-additivity of multipartite entanglement measures, New Journal of Physics **12**, 083002 (2010).

[6] A. Streltsov, H. Kampermann, S. Wölk, M. Gessner, and D. Bruß, Maximal coherence and the resource theory of purity, New Journal of Physics **20**, 053058 (2018).

[7] D. Collins and N. Gisin, A relevant two qubit bell inequality inequivalent to the CHSH inequality, Journal of Physics A: Mathematical and General **37**, 1775 (2004).

[8] S. Boyd and L. Vandenberghe, *Convex Optimization* (Cambridge University Press, 2004).

[9] T. Heinosaari, J. Kiukas, and D. Reitzner, Noise robustness of the incompatibility of quantum measurements, Phys. Rev. A **92**, 022115 (2015).

[10] F. Hirsch, M. T. Quintino, and N. Brunner, Quantum measurement incompatibility does not imply bell nonlocality, Phys. Rev. A **97**, 012129 (2018).

[11] E. Bene and T. Vértesi, Measurement incompatibility does not give rise to bell violation in general, New Journal of Physics **20**, 013021 (2018).

[12] R. Uola, C. Budroni, O. Gühne, and J.-P. Pellonpää, One-to-one mapping between steering and joint measurability problems, Phys. Rev. Lett. **115**, 230402 (2015).

C Paper: Distance-based resource quantification for sets of quantum measurements

Title:	Distance-based resource quantification for sets of quantum measurements
Authors:	Lucas Tendick, Martin Kliesch, Hermann Kampermann, and Dagmar Bruß
Journal:	Quantum
Impact factor:	6.777 (2021)
Date of submission:	14 September 2022
Publication status:	Under review
Contribution by LT:	First author (input approx. 85%)

This work corresponds to the paper [65]. The summary of the results can be found in Section 4.3.

I had the initial idea to study distance-based resource quantifier for the resource of measurement incompatibility. In the need of an operational meaningful distance between two POVMs, I discussed several options with MK, who set up initial notes on the project. MK also suggested using the diamond norm by associating measure-and-prepare channels to the POVMs. Based on this discussion, I constructed distances between measurement assemblages and singled out the diamond distance as the most promising approach. After obtaining some preliminary results, I discussed, together with MK, HK, and DB, how to proceed with the project. Afterward, I saw that our work applies to other resources of quantum measurements, particularly any convex resource theory of measurement assemblages. After some more discussions, especially with HK, I derived the analytical results of the manuscript. I also obtained the numerical results and was partially assisted by HK in studying the incompatibility of MUB. After an initial draft of the version, I discussed with all co-authors how to present the results. Then, I prepared the whole manuscript based on comments from my co-authors on several draft versions. In particular, MK helped me improve the presentation style and contributed by designing Figure 2.2 (Figure 2 in the manuscript).

Distance-based resource quantification for sets of quantum measurements

Lucas Tendick,* Martin Kliesch, Hermann Kampermann, and Dagmar Bruß
Institute for Theoretical Physics, Heinrich Heine University Düsseldorf, D-40225 Düsseldorf, Germany

The advantage that quantum systems provide for certain quantum information processing tasks over their classical counterparts can be quantified within the general framework of resource theories. Certain distance functions between quantum states have successfully been used to quantify resources like entanglement and coherence. Perhaps surprisingly, such a distance-based approach has not been adopted to study resources of quantum measurements, where other geometric quantifiers are used instead. Here, we define distance functions between sets of quantum measurements and show that they naturally induce resource monotones for convex resource theories of measurements. By focusing on a distance based on the diamond norm, we establish a hierarchy of measurement resources and derive analytical bounds on the incompatibility of any set of measurements. We show that these bounds are tight for certain projective measurements based on mutually unbiased bases and identify scenarios where different measurement resources attain the same value when quantified by our resource monotone. Our results provide a general framework to compare distance-based resources for sets of measurements and allow us to obtain limitations on Bell-type experiments.

I. INTRODUCTION

It is arguably one of the most astonishing features of quantum theory that certain quantum systems exhibit behaviours without any classical analogue. While these quantum phenomena were first just regarded as a strange feature of nature which led to many philosophical questions [1–3], it has later been realized that these phenomena can actually be used as a resource in real world applications such as computation [4], sensing [5], or cryptography [6]. To understand the potential of these upcoming technologies, it is important to understand how much advantage they can provide over conventional methods and which physical phenomena enable it. To achieve this advantage, properties of both, quantum states and measurements, are relevant. Among the most important types of quantum correlations that are know to enable advantages over classical systems are entanglement [7], EPR-steering [8–10], and Bell nonlocality [11, 12].

The latter two are similar in the sense that both types of correlations can be seen as resources that require one out of several judiciously chosen quantum measurements to be performed on a resourceful quantum state in each round of an experiment. In particular, it is well-known that entangled states and incompatible measurements are necessary to witness steering or nonlocality [9] and also limit these phenomena quantitatively [13, 14]. However, states and measurements posses a variety of different resources in general [15–23].

Quantum resource theorys (QRTs) [24] allow to identify, study, and quantify quantum resources for certain quantum information processing tasks in a general framework. Moreover, this allows to identify similarities among different resources, adapt concepts and quantification methods [25–28] from one to another, and to establish relations between different resources [13–15, 29]. Any

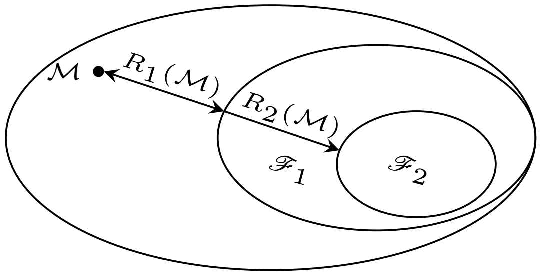

Figure 1. Distance-based resource quantification. A set of measurements $\mathcal{M}$ contains different quantum resources in general. These different quantum resources are associated with their respective sets of free measurements, here denoted by $\mathscr{F}_1$ for QRT Q_1 and $\mathscr{F}_2$ for QRT Q_2. The amount of resource in $\mathcal{M}$ associated to Q_1 and Q_2 is quantified by its distance $\mathrm{R}_1(\mathcal{M})$ to the set $\mathscr{F}_1$ and the distance $\mathrm{R}_2(\mathcal{M})$ to the set $\mathscr{F}_2$, respectively. As all free measurements $\mathcal{F} \in \mathscr{F}_2$ are also contained in $\mathscr{F}_1$ it follows that $\mathrm{R}_1(\mathcal{M}) \leq \mathrm{R}_2(\mathcal{M})$.

QRT aims to answer at least the following three questions: (i) Which objects (e.g. states or measurements) are resources for a certain task and which ones are free, i.e., do not provide any advantage? (ii) Which transformations are free, i.e., cannot create resources from free objects? (iii) How can we quantify the amount of the resource? A standard approach to quantify any quantum resource, illustrated in Figure 1, is to ask how far away a given resource is from the set of free objects, as measured by some distance-based function. Together with the class of robustness-based [22, 30–34] and weight-based quantifiers [28, 35–38] the class of distance-based [39–42] resource quantifiers form the class of so-called geometric quantifiers. One main advantage of these quantifiers is that they generally can be defined for any convex QRT (i.e., a QRT where the set of free objects is convex), which for instance allows a practical way to compare certain resources with each other. The combined insights from all three classes of geometric quantifiers usually gives a de-

* lucas.tendick@hhu.de

Resource	Monotone	Free objects	Free operations	Optimization	Type
General	• $\mathrm{R}_\diamond(\mathcal{M}_\mathbf{p})$ (12)	$\mathcal{F} \in \mathscr{F}$	$\Lambda^\dagger \in \mathbb{F}$, simulations $\xi \in \mathbb{S}$	SDP (36),(37)	Set\Average
Informativeness	• $\mathrm{IF}_\diamond(\mathcal{M}_\mathbf{p})$ (17)	$F_{a\|x} = q(a\|x)\mathbb{1}_d$ (16)	Unital maps $\Lambda^\dagger$ [a], simulations ξ [22]	SDP (F1), (F2)	Average
Coherence	• $\mathrm{C}_\diamond(\mathcal{M}_\mathbf{p})$ (19)	$F_{a\|x} = \sum_i \alpha_{i\|(a,x)} \|i\rangle\langle i\|$ (18)	SI-operations $\Lambda^\dagger_{\mathrm{SIO}}$ [b] [23], simulations ξ [c]	SDP (F4), (F5)	Average
Incompatibility	• $\mathrm{I}_\diamond(\mathcal{M}_\mathbf{p})$ (21)	$F_{a\|x} = \sum_\lambda v(a\|x,\lambda) G_\lambda$ (20)	Unital maps $\Lambda^\dagger$ [33], simulations ξ [21]	SDP (E1), (E2)	Set
Steering	$\mathrm{S}(\vec{\sigma}_\mathbf{p})$ (24), [46]	$\tau_{a\|x} = \sum_\lambda v(a\|x,\lambda)\sigma_\lambda$ (23)	(Restricted) 1W-LOCC [d] [46, 49]	SDP (C6), (C12)	Set
Nonlocality	$\mathrm{N}(\mathbf{q}_\mathbf{p})$ (26), [47]	$t(a,b\|x,y)$ (25) $= \sum_\lambda \pi(\lambda) v_A(a\|x,\lambda) v_B(b\|y,\lambda)$	WCCPI [e][50]	Linear (C3), (C5)	Set

[a] Even though it is not discussed in [22], it follows directly from the definition of unitality, that no quantum channel $\Lambda^\dagger$ can create informativeness from uninformative measurements.
[b] SIO stands for *strictly incoherent operations*.
[c] Even though it is not discussed in [23], it follows directly from the definition of the classical simulations ξ that they cannot create coherence, as linear combinations of diagonal matricies are diagonal.
[d] 1W-LOCC stands for *one-way local operations and classical communication*
[e] WCCPI stands for *wirings and classical communication prior to the inputs.*

Table I. Overview over the resources analyzed in this work. The different resources are presented in terms of the monotones we consider, the respective free objects, and the set of free operations associated to the considered QRT. The monotones we introduce in this work are marked with a bullet point •. Furthermore, we present by which kind of optimization the respective monotone can be computed and whether the resources are genuine properties of a set of objects or an average over single object properties. The free operations for steering and nonlocality are listed for completeness here and we refer to the references in the table for more details.

tailed picture of any convex QRT.

Historically, quantum states were recognized first as quantum resources and the approach of geometric quantification has been employed with great success to resources like entanglement [7] or coherence [16]. More recently, quantum measurements became the focus of QRT research. Interestingly, for quantum measurements, the analysis of resources developed in a different direction than for quantum states. While weight and robustness quantifiers for measurements are well-established, distances between different measurements have only been studied recently [43, 44] and distance-based resource quantification for sets of measurements remains widely unexplored until now.

In this work, we complete the class of geometric quantifiers for convex QRTs of sets of measurements (so-called assemblages) by introducing distance-based resource quantifiers. First, we discuss necessary properties any distance between sets of measurements has to fulfil. Then, we show that every such distance induces a convex resource monotone. We propose one particular quantifier, which is based on the diamond norm [45] between different measure-and-prepare channels and is especially tailored to Bell-type experiments, as it captures the idea that only one particular measurement out of a given collection is applied at a time in a round-by-round protocol. Based on this quantifier, we establish a hierarchy of measurement resources including recently introduced steering [46] and nonlocality monotones [47]. See Table I for an overview of the resources and the quantities we analyze in this work. We show that our quantifier can be computed efficiently by means of a semidefinite program (SDP) which we use to obtain analytical upper and lower bounds on the incompatibility (i.e., the non-joint measureability) [19] for any set of measurements. Finally, we show that these bounds are tight for special instances of projective measurements based on mutually unbiased bases (MUB) [48], which also play a special role for cases when different measurement resources attain the same value when quantified with our proposed quantifier.

II. DISTANCE-BASED RESOURCE QUANTIFICATION

Consider the canonical example of the trace distance [51]. The trace distance between two quantum states $\rho, \tau \in \mathcal{S}(\mathcal{H})$, where $\mathcal{S}(\mathcal{H})$ is the set of density matrices acting on a Hilbert space $\mathcal{H} \cong \mathbb{C}^d$ of dimension d, is given by

$$\mathrm{D}_1(\rho,\tau) = \frac{1}{2}\|\rho - \tau\|_1 \geq 0, \tag{1}$$

where $\|X\|_1 = \mathrm{Tr}[\sqrt{X^\dagger X}]$ is the trace norm of X. The trace distance is a useful tool to distinguish ρ and τ, as it fulfils all necessary properties of a metric between quantum states. Consider $\rho, \tau, \chi \in \mathcal{S}(\mathcal{H})$ and any completely positive and trace preserving (CPT) map Λ also known as quantum channel. It holds that

$$\begin{aligned} &\mathrm{D}_1(\rho,\tau)=0 \iff \rho=\tau, \\ &\mathrm{D}_1(\rho,\tau)=\mathrm{D}_1(\tau,\rho), \\ &\mathrm{D}_1(\rho,\tau)\leq \mathrm{D}_1(\rho,\chi)+\mathrm{D}_1(\chi,\tau), \\ &\mathrm{D}_1(\rho,\tau)\geq \mathrm{D}_1(\Lambda(\rho),\Lambda(\tau)), \end{aligned} \tag{2}$$

i.e., $\mathrm{D}_1(\rho,\tau)$ is a faithful and symmetric function that obeys the triangle inequality and monotonicity (i.e. it does not increase) under arbitrary CPT maps Λ. In addition to these minimal requirements, it is well known that $\mathrm{D}_1(\rho,\tau)$ has an operational interpretation in terms of the optimal probability to distinguish ρ and τ in a single-shot experiment [51]. That is, the optimal guessing probability is given by $p_{1,\text{guess}}^{(\rho,\tau)} = \frac{1}{2}(1+\mathrm{D}_1(\rho,\tau))$. These properties make the trace distance a viable tool to quantify (convex) resources.

Let us consider the prime example of a resource, the entanglement of a bipartite state $\rho\in\mathcal{S}(\mathcal{H}\otimes\mathcal{H})$. One can quantify the entanglement of ρ by its distance to the set $\mathrm{Sep}(\mathcal{H}\otimes\mathcal{H})$ of separable quantum states [39] given as

$$\mathrm{E}_1(\rho)=\min_{\rho_S\in\mathrm{Sep}(\mathcal{H}\otimes\mathcal{H})}\mathrm{D}_1(\rho,\rho_S). \tag{3}$$

It is now readily verified that $\mathrm{E}_1(\rho)$ is a non-negative, convex function with $\mathrm{E}_1(\rho)=0 \iff \rho\in\mathrm{Sep}(\mathcal{H}\otimes\mathcal{H})$ obeying the monotonicity $\mathrm{E}_1(\rho)\geq\mathrm{E}_1(\Lambda_{\mathrm{LOCC}}(\rho))$ under any local operations and classical communication (LOCC) [7] map Λ_{LOCC}. That is, $\mathrm{E}_1(\rho)$ is a faithful (i.e. $\mathrm{E}_1(\rho)=0 \iff \rho\in\mathrm{Sep}(\mathcal{H}\otimes\mathcal{H})$) convex resource monotone. Note that the monotonicity $\mathrm{E}_1(\rho)\geq\mathrm{E}_1(\Lambda_{\mathrm{LOCC}}(\rho))$ captures the fact that LOCC maps cannot create entanglement. The monotonicity of resources under these so-called free operations is sometimes also referred to as *golden rule* of QRTs [24].

We can use the insights for distances and resources of quantum states to define distance-based resource monotones for sets of quantum measurements in the following. A quantum measurement is most generally described by a positive operator valued measure (POVM) i.e., a set $\{M_a\}_a$ of effect operators $0\leq M_a\leq \mathbb{1}_d$, acting on a d-dimensional Hilbert space $\mathcal{H}$ such that $\sum_a M_a=\mathbb{1}_d$. A set of POVMs with outcomes a for different settings x is known as measurement assemblage $\mathcal{M}=\{M_{a|x}\}_{a,x}$. Note that we will omit in the following the set-indices and simply write $\mathcal{M}=\{M_{a|x}\}$ when there is no risk of confusion. If we talk about a specific element of the assemblage $\mathcal{M}$, for instance the POVM corresponding to setting x, we will write $\mathcal{M}_x=\{M_{a|x}\}_a$. Here, we consider assemblages with m measurement settings and o outcomes in each setting, i.e. $x=1,\cdots,m$ and $a=0,\cdots,o-1$. The outcome statistics of a measurement on any state ρ is given by $p(a,x)=p(x)p(a|x)=p(x)\mathrm{Tr}[M_{a|x}\rho]$, where $p(x)$ is the probability to choose the setting x.

A measurement assemblage can be converted by two different processes to another assemblage. First, as any quantum state ρ can be transformed via any CPT map Λ to another state $\Lambda(\rho)$, it follows from $\mathrm{Tr}[M_{a|x}\Lambda(\rho)]=\mathrm{Tr}[\Lambda^\dagger(M_{a|x})\rho]$ that an assemblage $\mathcal{M}$ can be transformed via the Hilbert-Schmidt adjoint (unital) map $\Lambda^\dagger$ to another assemblage $\Lambda^\dagger(\mathcal{M})$. Second, mixtures and classical post-processing maps $\mathcal{M}'=\xi(\mathcal{M})$ with $M'_{b|y}=\sum_x p(x|y)\sum_a q(b|y,x,a)M_{a|x}$ can be used to simulate [21] the assemblage $\mathcal{M}'$ from $\mathcal{M}$ via the conditional probabilities $p(x|y)$ and $q(b|y,x,a)$ for all y, respectively for all y,x,a. Note that as $p(x)=\sum_y q(y)p(x|y)$ one also obtains the probability $q(y)$ to perform setting y.

We use the probability distribution $\mathbf{p}=\{p(x)\}$ to capture the fact that typically only one quantum measurement can be performed at a time and it is also natural to assume that the likelihood of the settings x influences the capabilities of $\mathcal{M}$ in experiments. Note that we consider only the case $p(x)>0\ \forall\ x$, as measurements that are never performed can be discarded trivially. We define a distance between sets of measurements weighted with the distribution $\mathbf{p}$ as follows.

Definition 1. *Let $\mathcal{M}$ be a measurement assemblage containing m POVMs and let $\mathbf{p}$ be a probability distribution with $p(x)>0\ \forall\ x=1,\cdots,m$. We call the tuple $\mathcal{M}_{\mathbf{p}}:=(\mathcal{M},\mathbf{p})$ a* weighted measurement assemblage (WMA). *Let $\mathcal{M}_{\mathbf{p}}$, $\mathcal{N}_{\mathbf{p}}$, and $\mathcal{K}_{\mathbf{p}}$ be WMAs. Any function $\mathrm{D}(\mathcal{M}_{\mathbf{p}},\mathcal{N}_{\mathbf{p}})$ that fulfils the conditions*

$$\begin{aligned} &\mathrm{D}(\mathcal{M}_{\mathbf{p}},\mathcal{N}_{\mathbf{p}})=0 \iff \mathcal{M}=\mathcal{N}, \\ &\mathrm{D}(\mathcal{M}_{\mathbf{p}},\mathcal{N}_{\mathbf{p}})=\mathrm{D}(\mathcal{N}_{\mathbf{p}},\mathcal{M}_{\mathbf{p}}), \\ &\mathrm{D}(\mathcal{M}_{\mathbf{p}},\mathcal{N}_{\mathbf{p}})\leq\mathrm{D}(\mathcal{M}_{\mathbf{p}},\mathcal{K}_{\mathbf{p}})+\mathrm{D}(\mathcal{K}_{\mathbf{p}},\mathcal{N}_{\mathbf{p}}), \\ &\mathrm{D}(\mathcal{M}_{\mathbf{p}},\mathcal{N}_{\mathbf{p}})\geq\mathrm{D}(\Lambda^\dagger(\mathcal{M}_{\mathbf{p}}),\Lambda^\dagger(\mathcal{N}_{\mathbf{p}})), \\ &\mathrm{D}(\mathcal{M}_{\mathbf{p}},\mathcal{N}_{\mathbf{p}})\geq\mathrm{D}(\xi(\mathcal{M}_{\mathbf{p}})_{\mathbf{q}},\xi(\mathcal{N}_{\mathbf{p}})_{\mathbf{q}}). \end{aligned} \tag{4}$$

is a distance *between $\mathcal{M}_{\mathbf{p}}$ and $\mathcal{N}_{\mathbf{p}}$.*

Note that all conditions are in direct correspondence to the conditions in Eq. (2) for quantum states. Any distance that fulfills the conditions in Definition 1 can be used to define a faithful convex resource monotone for convex QRTs of measurement assemblages.

Definition 2. *Let $\mathscr{F}$ be a convex and compact set of measurement assemblages, $\mathbb{F}$ the (maximal) set of free quantum maps $\Lambda^\dagger$ such that $\Lambda^\dagger(\mathcal{F})\in\mathscr{F}$ for any $\mathcal{F}\in\mathscr{F}$, and let $\mathbb{S}$ be the set of simulations ξ such that $\xi(\mathcal{F})\in\mathscr{F}$ for any $\mathcal{F}\in\mathscr{F}$. The tuple $Q:=(\mathscr{F},\mathbb{F},\mathbb{S})$ is called a* QRT of measurement assemblages.

Definition 3. *Let $Q=(\mathscr{F},\mathbb{F},\mathbb{S})$ be a QRT of WMAs $\mathcal{M}_{\mathbf{p}}$, $\mathcal{N}_{\mathbf{p}}$. Any function $\mathrm{R}(\mathcal{M}_{\mathbf{p}})$ that fulfils*

$$\begin{aligned} &\mathrm{R}(\mathcal{M}_{\mathbf{p}})=0 \iff \mathcal{M}\in\mathscr{F}, \\ &\mathrm{R}(\mathcal{M}_{\mathbf{p}})\geq\mathrm{R}(\Lambda^\dagger(\mathcal{M})_{\mathbf{p}}),\ \forall\ \Lambda^\dagger\in\mathbb{F}, \\ &\mathrm{R}(\mathcal{M}_{\mathbf{p}})\geq\mathrm{R}(\xi(\mathcal{M}_{\mathbf{p}})_{\mathbf{q}}),\ \forall\ \xi\in\mathbb{S}, \\ &\mathrm{R}(\eta\mathcal{M}_{\mathbf{p}}+(1-\eta)\mathcal{N}_{\mathbf{p}})\leq\eta\mathrm{R}(\mathcal{M}_{\mathbf{p}})+(1-\eta)\mathrm{R}(\mathcal{N}_{\mathbf{p}}), \end{aligned} \tag{5}$$

for any $\eta\in[0,1]$ is a faithful convex resource monotone *of WMAs.*

With these definitions we obtain the following lemma, showing that every distance between measurement assemblages induces a faithful convex resource monotone.

Lemma 1. *Let $Q = (\mathscr{F}, \mathbb{F}, \mathbb{S})$ be any QRT of WMAs $\mathcal{M}_{\mathbf{p}}$ and $\mathrm{D}(\mathcal{M}_{\mathbf{p}}, \mathcal{F}_{\mathbf{p}})$ a distance function. The function*

$$\mathrm{R}(\mathcal{M}_{\mathbf{p}}) := \min_{\mathcal{F} \in \mathscr{F}} \mathrm{D}(\mathcal{M}_{\mathbf{p}}, \mathcal{F}_{\mathbf{p}}), \tag{6}$$

is a faithful convex resource monotone.

Proof. The proof relies on the conditions in Definition 1. The convexity, non-negativity, and faithfulness (i.e. $\mathrm{R}(\mathcal{M}_{\mathbf{p}}) = 0 \iff \mathcal{M} \in \mathscr{F}$) follow directly, and the monotonicity conditions follow from

$$\begin{aligned}\mathrm{R}(\mathcal{M}_{\mathbf{p}}) &= \min_{\mathcal{F} \in \mathscr{F}} \mathrm{D}(\mathcal{M}_{\mathbf{p}}, \mathcal{F}_{\mathbf{p}}) \\ &\geq \min_{\mathcal{F} \in \mathscr{F}} \mathrm{D}(\Lambda^\dagger(\mathcal{M})_{\mathbf{p}}, \Lambda^\dagger(\mathcal{F})_{\mathbf{p}}) \\ &\geq \min_{\mathcal{F}' \in \mathscr{F}} \mathrm{D}(\Lambda^\dagger(\mathcal{M})_{\mathbf{p}}, \mathcal{F}'_{\mathbf{p}}) = \mathrm{R}(\Lambda^\dagger(\mathcal{M})_{\mathbf{p}}),\end{aligned} \tag{7}$$

where we used the monotonicity of the distance and the fact that free operations $\Lambda^\dagger \in \mathbb{F}$ map free assemblages to free assemblages. An analogous calculation follows for the simulations $\xi \in \mathbb{S}$. Note that the arguments used here are similar to those for distance-based resource monotones of quantum states. □

We propose in the following a specific distance on which we focus on in the remainder of the work (see however the appendix for alternatives). More specifically, we associate to any POVM $\mathcal{M}_x = \{M_{a|x}\}_a$ a measure-and-prepare channel defined by

$$\Lambda_{\mathcal{M}_x}(\rho) = \sum_a \mathrm{Tr}[M_{a|x}\rho] |a\rangle\langle a|, \tag{8}$$

where the register states $|a\rangle$ form an orthonormal basis $\{|a\rangle\}_{0 \leq a \leq o-1}$. Note that the channel $\Lambda_{\mathcal{M}_x}$ can equivalently be described by its Choi–Jamiołkowski-matrix (see e.g. [52]). The Choi–Jamiołkowski-matrix of a quantum channel is obtained by applying a given channel to the first subsystem of the (unnormalized) maximally entangled state $|\tilde{\Phi}^+\rangle = \sum_{i=0}^{d-1} |ii\rangle$. More precisely, the Choi–Jamiołkowski-matrix of a measure-and-prepare channel as described in Eq. (8) is given by

$$J(\mathcal{M}_x) = (\Lambda_{\mathcal{M}_x} \otimes \mathbb{1})(|\tilde{\Phi}^+\rangle\langle\tilde{\Phi}^+|) = \sum_a |a\rangle\langle a| \otimes M_{a|x}^T, \tag{9}$$

where the transpose is with respect to the computational basis.

We denote the diamond distance between two quantum channels Λ_1, Λ_2 by

$$\mathrm{D}_\diamond(\Lambda_1, \Lambda_2) = \max_{\rho \in \mathcal{S}(\mathcal{H} \otimes \mathcal{H})} \frac{1}{2} \|((\Lambda_1 - \Lambda_2) \otimes \mathbb{1}_d)\rho\|_1. \tag{10}$$

Due to the connection to the trace distance, the diamond distance determines the optimal single-shot probability $p_{\diamond,\text{guess}}^{(\Lambda_1,\Lambda_2)} = \frac{1}{2}(1 + \mathrm{D}_\diamond(\Lambda_1, \Lambda_2))$ to distinguish between Λ_1 and Λ_2. Based on the diamond distance, we propose the distance $\mathrm{D}_\diamond(\mathcal{M}_{\mathbf{p}}, \mathcal{N}_{\mathbf{p}})$ between the WMAs defined as

$$\mathrm{D}_\diamond(\mathcal{M}_{\mathbf{p}}, \mathcal{N}_{\mathbf{p}}) := \sum_x p(x)\, \mathrm{D}_\diamond(\Lambda_{\mathcal{M}_x}, \Lambda_{\mathcal{N}_x}), \tag{11}$$

and its induced resource monotone

$$\mathrm{R}_\diamond(\mathcal{M}_{\mathbf{p}}) := \min_{\mathcal{F} \in \mathscr{F}} \sum_x p(x)\, \mathrm{D}_\diamond(\Lambda_{\mathcal{M}_x}, \Lambda_{\mathcal{F}_x}). \tag{12}$$

Note that the diamond distance between measure-and prepare-channels has also been introduced in the context of single POVM discrimination [43, 44]. To prove that $\mathrm{R}_\diamond(\mathcal{M}_{\mathbf{p}})$ is indeed a resource monotone, we need to show that $\mathrm{D}_\diamond(\mathcal{M}_{\mathbf{p}}, \mathcal{N}_{\mathbf{p}})$ is a distance function according to the conditions in Definition 1.

Theorem 1. *The function $\mathrm{D}_\diamond(\mathcal{M}_{\mathbf{p}}, \mathcal{N}_{\mathbf{p}})$ is a distance function between the WMAs $\mathcal{M}_{\mathbf{p}}$ and $\mathcal{N}_{\mathbf{p}}$, i.e., it fulfils all the conditions stated in Definition 1.*

Proof. The proof relies mostly on the properties of the diamond distance. It is possible to rewrite

$$\mathrm{D}_\diamond(\mathcal{M}_{\mathbf{p}}, \mathcal{N}_{\mathbf{p}}) = \frac{1}{2} \sum_x p(x) \max_\rho \sum_a \|\sigma_{a|x}(\rho) - \tau_{a|x}(\rho)\|_1, \tag{13}$$

where we have introduced $\sigma_{a|x}(\rho) = \mathrm{Tr}_1[(M_{a|x} \otimes \mathbb{1})\rho]$ and $\tau_{a|x}(\rho) = \mathrm{Tr}_1[(N_{a|x} \otimes \mathbb{1})\rho]$. Note that we omit here and in the following the Hilbert space ρ acts on. All conditions in Definition 1 can now be verified by direct computation. See Appendix A for all details. □

Note that it follows directly from its definition that $\mathrm{R}_\diamond(\mathcal{M}_{\mathbf{p}})$ is upper bounded by $\mathrm{R}_\diamond(\mathcal{M}_{\mathbf{p}}) \leq 1$, and that it fulfills the continuity condition

$$|\mathrm{R}_\diamond(\mathcal{M}_{\mathbf{p}}) - \mathrm{R}_\diamond(\mathcal{N}_{\mathbf{p}})| \leq \mathrm{D}_\diamond(\mathcal{M}_{\mathbf{p}}, \mathcal{N}_{\mathbf{p}}), \tag{14}$$

due to the triangle inequality for the diamond norm. Moreover, it can be rewritten as

$$\mathrm{R}_\diamond(\mathcal{M}_{\mathbf{p}}) = \min_{\mathcal{F} \in \mathscr{F}} 2 \sum_x p(x) p_{\diamond,\text{guess}}^{(\mathcal{M},\mathcal{F})}(x) - 1, \tag{15}$$

which is up to normalization the average optimal probability to distinguish the resources $\mathcal{M}$ from the free measurements $\mathcal{F}$ in a single-shot experiment. See also Figure 2 for an illustration of the operational meaning of $\mathrm{R}_\diamond(\mathcal{M}_{\mathbf{p}})$.

III. HIERARCHY OF MEASUREMENT RESOURCES

One main goal while studying QRTs is to obtain relations between different resources. In particular, we want

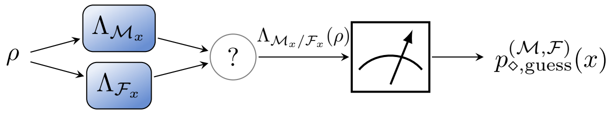

Figure 2. Illustration of the idea to use the diamond distance as resource monotone. Quantum measurements $\mathcal{M}_x, \mathcal{F}_x$ are associated with quantum channels $\Lambda_{\mathcal{M}_x}, \Lambda_{\mathcal{F}_x}$. These are distinguished by applying the channels to an optimal quantum state ρ and performing an ideal dichotomic measurement afterwards to distinguish between the output of the channels $\Lambda_{\mathcal{M}_x}$ and $\Lambda_{\mathcal{F}_x}$. The probability $p_{\diamond,\mathrm{guess}}^{(\mathcal{M},\mathcal{F})}(x)$ tells us how distinguishable the resourceful measurement $\mathcal{M}_x$ is from the free measurements $\mathcal{F}_x$.

to understand how one resource limits another quantitatively. This will show one strength of a geometric quantifier, as it can be defined for various resource theories and the discussion often reduces to an analysis of the free sets $\mathscr{F}$. In the following, we will establish a hierarchy of measurement resources based on the newly introduced quantifier $\mathrm{R}_\diamond(\mathcal{M}_\mathbf{p})$. We start by introducing the different resources.

The most basic resource of an assemblage is its informativeness [22]. The informativeness of a WMA quantifies how valuable it is to actually perform measurements compared to randomly guessing the outcomes in an experiment. An assemblage $\mathcal{M}$ is called uninformative (UI) if

$$M_{a|x} = q(a|x)\mathbb{1}_d \ \forall\ a, x, \tag{16}$$

where $\{q(a|x)\}$ are some probability distributions of a conditioned on setting x. These measurements are UI as their measurement result does not depend on the quantum state. We denote the set of UI assemblages by $\mathscr{F}_{\mathrm{UI}}$ and introduce the informativeness monotone

$$\mathrm{IF}_\diamond(\mathcal{M}_\mathbf{p}) = \min_{\mathcal{F}\in\mathscr{F}_{\mathrm{UI}}} \sum_x p(x)\,\mathrm{D}_\diamond(\Lambda_{\mathcal{M}_x}, \Lambda_{\mathcal{F}_x}). \tag{17}$$

Note that measurement informativeness was initially introduced only for a single POVM and studied in terms of the generalized robustness [22]. We have extended the notion here by considering the average informativeness of $\mathcal{M}_\mathbf{p}$.

A resource that is the foundation for the distinction between *classical* and *quantum* systems is the coherence of measurements [23]. An assemblage $\mathcal{M}$ is incoherent (in some predefined orthonormal basis $\{|i\rangle\}$) if

$$M_{a|x} = \sum_i \alpha_{i|(a,x)} |i\rangle\langle i| \ \forall\ a, x, \tag{18}$$

where $\alpha_{i|(a,x)} = \langle i|M_{a|x}|i\rangle$. These measurements cannot distinguish quantum states ρ from their fully dephased versions $\Delta(\rho) = \sum_i |i\rangle\langle i|\rho|i\rangle\langle i|$, hence they cannot detect coherence. We denote the set of incoherent assemblages by $\mathscr{F}_{\mathrm{IC}}$ and introduce the coherence monotone

$$\mathrm{C}_\diamond(\mathcal{M}_\mathbf{p}) = \min_{\mathcal{F}\in\mathscr{F}_{\mathrm{IC}}} \sum_x p(x)\,\mathrm{D}_\diamond(\Lambda_{\mathcal{M}_x}, \Lambda_{\mathcal{F}_x}). \tag{19}$$

Similarly to the informativeness, the coherence of measurements was initially introduced for a single POVM and we have extended it here by considering the average coherence of $\mathcal{M}_\mathbf{p}$. See also [26] for a different approach to coherence of measurement assemblages.

The incompatibility of measurements is probably the best-known example of a QRT for measurements and has been studied extensively in recent years [19, 33, 34, 53–55]. Contrary to classical physics, different quantum measurements may be incompatible, i.e., they cannot be performed simultaneously and one cannot access their joint measurement statistics as famously illustrated by the Heisenberg-Robertson uncertainty relation [3]. Initially interpreted as a drawback, this phenomenon lies at the heart of Bell-type experiments, as incompatibility is a necessary prerequisite to witness steering and nonlocality. An assemblage $\mathcal{M}$ is called compatible or jointly measurable (JM) if the statistics of $\mathcal{M}$ can be simulated by a single measurement via some POVM $\{G_\lambda\}$ and classical post-processing via the deterministic probability distributions $\{v(a|x,\lambda)\}$ such that

$$M_{a|x} = \sum_\lambda v(a|x,\lambda) G_\lambda \ \forall\ a, x, \tag{20}$$

and is called incompatible otherwise. Note that using deterministic post-processings $\{v(a|x,\lambda)\}$ (which represent the vertices of the corresponding probability polytope) is not a restriction as all randomness from non-deterministic distributions can be put inside G_λ. We denote the set of JM assemblages by $\mathscr{F}_{\mathrm{JM}}$ and introduce the incompatibility monotone

$$\mathrm{I}_\diamond(\mathcal{M}_\mathbf{p}) = \min_{\mathcal{F}\in\mathscr{F}_{\mathrm{JM}}} \sum_x p(x)\,\mathrm{D}_\diamond(\Lambda_{\mathcal{M}_x}, \Lambda_{\mathcal{F}_x}). \tag{21}$$

It is important to note that the incompatibility in Eq. (21) is not the average of single POVM properties, as incompatibility is always a property of sets of measurements. Therefore, the incompatibility $\mathrm{I}_\diamond(\mathcal{M}_\mathbf{p})$ is qualitatively different from the coherence or informativeness.

Similar to entanglement, incompatibility can be witnessed in a Bell-type experiment, as both are necessary resources for steering and nonlocality. Consider the WMA $\mathcal{M}_\mathbf{p}$ and any bipartite quantum state ρ shared by two-parties, Alice and Bob. By performing the measurements $\mathcal{M}_\mathbf{p}$ on her share of the state, Alice prepares the conditional states

$$\sigma_{a|x} = \mathrm{Tr}_1[(M_{a|x} \otimes \mathbb{1})\rho], \tag{22}$$

for Bob. Here $p(a|x) = \mathrm{Tr}[\sigma_{a|x}]$ is the probability to obtain $\sigma_{a|x}$. We denote the obtained state assemblage by $\vec{\sigma} = \{\sigma_{a|x}\}$ and its weighted version by $\vec{\sigma}_\mathbf{p} = (\vec{\sigma}, \mathbf{p})$.

To make sure that Alice performs incompatible measurements on an entangled state, she can prove that she can demonstrate steering. A state assemblage $\vec{\sigma}$ is said to be steerable if it cannot be obtained from a local hidden-state model (LHS) given by

$$\sigma_{a|x} = \sum_\lambda v(a|x,\lambda)\sigma_\lambda \ \forall a, x, \tag{23}$$

where the σ_λ are operators that satisfy $\sigma_\lambda \geq 0\ \forall\lambda$ and $\mathrm{Tr}[\sum_\lambda \sigma_\lambda] = 1$. Otherwise we say $\vec{\sigma}$ is unsteerable which we denote by $\vec{\sigma} \in \mathrm{LHS}$. Steering can also be quantified and we use the distance-based monotone introduced by Ku et al. [46] as

$$\mathrm{S}(\vec{\sigma}_{\mathbf{p}}) = \min_{\vec{\tau}\in\mathrm{LHS}} \frac{1}{2}\sum_{a,x} p(x)\|\sigma_{a|x} - \tau_{a|x}\|_1. \tag{24}$$

Note that originally an additional consistency constraint $\sum\limits_a \tau_{a|x} = \sum\limits_a \sigma_{a|x}$ was introduced [46], which we do not require here.

Consider now that both parties, Alice and Bob, want to prove that they perform incompatible measurements on an entangled state. Let $\mathcal{M}_{\mathbf{p}_A}$ and $\mathcal{N}_{\mathbf{p}_B}$ be the WMAs of Alice and Bob, respectively, and let ρ be their shared quantum state. Alice and Bob obtain the probability distribution $\mathbf{q} = \{q(a,b|x,y)\}$ via $q(a,b|x,y) = \mathrm{Tr}[(M_{a|x} \otimes M_{b|y})\rho]$. Note that $p(x,y) = p_A(x)p_B(y)$ is the probability to choose setting x for Alice and y for Bob and we introduce the tuple $\mathbf{q_p} = (\mathbf{q},\mathbf{p})$. To assure themselves that they share an entangled state and perform incompatible measurements, they can check whether they can demonstrate nonlocality. A probability distribution $\mathbf{q}$ is local if it can be obtained from a local hidden-variable model (LHV) given by

$$q(a,b|x,y) = \sum_\lambda \pi(\lambda) v_A(a|x,\lambda) v_B(b|y,\lambda)\ \forall a,b,x,y, \tag{25}$$

where $\pi(\lambda)$ is the probability distribution of the hidden variable λ and $\{v_A(a|x,\lambda)\}$ and $\{v_B(b|y,\lambda)\}$ are deterministic probability distributions of Alice and Bob, respectively. In this case we denote $\mathbf{q} \in \mathrm{LHV}$ and we say $\mathbf{q}$ is nonlocal otherwise. To quantify the nonlocality, we use the distance-based resource monotone for nonlocality introduced by Brito et al. [47] as

$$\mathrm{N}(\mathbf{q_p}) = \frac{1}{2}\min_{\mathbf{t}\in\mathrm{LHV}} \sum_{a,b,x,y} p(x,y)|q(a,b|x,y) - t(a,b|x,y)|. \tag{26}$$

Having introduced all these different notions of quantum resources, we can complete our goal to establish relations among them.

Theorem 2. *Let $\mathcal{M}_{\mathbf{p}_A}$, $\mathcal{N}_{\mathbf{p}_B}$ be any WMAs and ρ any bipartite quantum state of appropriate dimensions. Let $\vec{\sigma}_{\mathbf{p}_A}$ be a state assemblage obtained via $\sigma_{a|x} = \mathrm{Tr}_1[(M_{a|x}\otimes\mathbb{1})\rho]$ and let $\mathbf{q_p} = (\mathbf{q},\mathbf{p})$ be a probability distribution obtained via $q(a,b|x,y) = \mathrm{Tr}[N_{b|y}\sigma_{a|x}]$ and $p(x,y) = p_A(x)p_B(y)$. The following sequence of inequalities holds:*

$$\mathrm{IF}_\diamond(\mathcal{M}_{\mathbf{p}_A}) \geq \mathrm{C}_\diamond(\mathcal{M}_{\mathbf{p}_A}) \geq \mathrm{I}_\diamond(\mathcal{M}_{\mathbf{p}_A}) \geq \mathrm{S}(\vec{\sigma}_{\mathbf{p}_A}) \geq \mathrm{N}(\mathbf{q_p}). \tag{27}$$

Proof. The inequalities $\mathrm{IF}_\diamond(\mathcal{M}_{\mathbf{p}_A}) \geq \mathrm{C}_\diamond(\mathcal{M}_{\mathbf{p}_A}) \geq \mathrm{I}_\diamond(\mathcal{M}_{\mathbf{p}_A})$ follow from the nested structure of the sets of free assemblages. More formally, $\mathscr{F}_{\mathrm{UI}} \subset \mathscr{F}_{\mathrm{IC}} \subset \mathscr{F}_{\mathrm{JM}}$ which can be seen by realising that POVM effects that are proportional to the identity are also incoherent (in any basis) and as incoherent POVMs commute pairwise, they are jointly measurable [26]. Since we are minimizing the distance with respect to these sets, the inequalities hold. To prove that $\mathrm{I}_\diamond(\mathcal{M}_{\mathbf{p}_A}) \geq \mathrm{S}(\vec{\sigma}_{\mathbf{p}_A})$ holds, we use that incompatibility is necessary for steering. This allows us to use $\vec{\tau} = \{\tau_{a,x} = \mathrm{Tr}_1[(F^*_{a|x}\otimes\mathbb{1})\rho]\}$ as an unsteerable assemblage for any state ρ, as the closest JM measurements $\mathcal{F}^*$ (with respect to the assemblage $\mathcal{M}$) cannot lead to steerable assemblages. It follows,

$$\begin{aligned}&\mathrm{S}(\vec{\sigma}_{\mathbf{p}_A}) \\ &\leq \frac{1}{2}\sum_x p_A(x)\sum_a \|\mathrm{Tr}_1[(M_{a|x}\otimes\mathbb{1})\rho] - \mathrm{Tr}_1[(F^*_{a|x}\otimes\mathbb{1})\rho]\|_1 \\ &\leq \frac{1}{2}\sum_x p_A(x)\max_\rho \sum_a \|\mathrm{Tr}_1[((M_{a|x} - F^*_{a|x})\otimes\mathbb{1})\rho]\|_1 \\ &= \mathrm{I}_\diamond(\mathcal{M}_{\mathbf{p}_A}),\end{aligned} \tag{28}$$

where we used the representation of $\mathrm{I}_\diamond(\mathcal{M}_{\mathbf{p}_A})$ according to Eq. (13) in the last line. We employ a similar approach to show that $\mathrm{S}(\vec{\sigma}_{\mathbf{p}_A}) \geq \mathrm{N}(\mathbf{q_p})$ in Lemma 2 in Appendix B. □

Note that hierarchies related to that in Eq. (27) have also been established, at least partly, for weight- and robustness-based resource quantifiers [13, 26]. The connection between incompatibility, steering, and nonlocality has been studied by Cavalcanti et al. [13] extensively for for weight- and robustness-based quantifiers while Designolle et al. [26] discussed the relation between coherence and incompatibility and, for a single POVM, between the informativeness and the coherence in terms of the generalized robustness.

The hierarchy (27) in Theorem 2 gives insights how resources like the incompatibility limit steering and nonlocal correlations quantitatively. On the other hand, every detection of these quantum correlations gives a lower bound to the measurement resources. In particular, the violation of every appropriately normalized steering or Bell inequality, in the nonlocal game formulation [56, 57], can lower bound these measurement resources. We show in Appendix C that $\mathrm{S}(\vec{\sigma}_{\mathbf{p}_A})$ is the maximal possible steering inequality violation given by

$$\mathrm{S}(\vec{\sigma}_{\mathbf{p}_A}) = \max_{G_{a,x},\ell} \sum_{a,x} p_A(x)\mathrm{Tr}[\sigma_{a|x}G_{a|x}] - \ell, \tag{29}$$

where $\ell = \max\limits_{\vec{\tau}\in\mathrm{LHS}} \sum\limits_{a,x} p_A(x)\mathrm{Tr}[\tau_{a|x}G_{a|x}]$ is the classical bound obeyed by all unsteerable assemblages $\vec{\tau} \in \mathrm{LHS}$ and the $G_{a|x}$ are positive semidefinite matrices s.t. $\|G_{a|x}\|_\infty \leq 1$, where $\|\cdot\|_\infty$ is the spectral norm.

Moreover, the nonlocality $\mathrm{N}(\mathbf{q_p})$ can be reformulated as the violation of a Bell inequality given by

$$\mathrm{N}(\mathbf{q_p}) = \max_{C_{ab|xy},\ell} \sum_{a,b,x,y} p(x,y)C_{ab|xy}q(a,b|x,y) - \ell, \tag{30}$$

where $\ell = \max_{\mathbf{t}\in\text{LHV}} \sum_{a,b,x,y} p(x,y)C_{ab|xy}t(a,b|x,y)$ is the local bound obeyed by all local correlations $\mathbf{t} \in$ LHV and $C_{ab|xy}$ are Bell coefficients s.t. $0 \leq C_{ab|xy} \leq 1$.

It is worth to highlight that the hierarchy (27) is reminiscent of the resource hierarchy for quantum states formulated by Streltsov et al. [15]. For quantum states, it holds that

$$\mathrm{P}(\rho) \geq \mathrm{C}(\rho) \geq \mathrm{D}(\rho) \geq \mathrm{E}(\rho), \tag{31}$$

where $\mathrm{P}(\rho)$, $\mathrm{C}(\rho)$, $\mathrm{D}(\rho)$, and $\mathrm{E}(\rho)$ denote the quantum state's purity, coherence with respect to product bases, discord, and entanglement, respectively, using the same geometric quantifier. Comparing both hierarchies, it becomes clear that the informativeness of measurements is in some sense the analogue to a state's purity, as both quantify the deviation from their respective uninformative element. We also observe that coherence is an important resource for states as well as measurements, which allows for more complex phenomena such as entanglement and incompatibility. Incompatibility and entanglement both play a similar role in their respective hierarchies, as both are the smallest known resource that is necessary for steering and nonlocality. Interestingly, incompatibility and entanglement also share similarities in their respective resource breaking maps [58].

Moreover, we show in Appendix D that the entanglement $\mathrm{E}_1(\rho)$ as defined in Eq. (3) also upper bounds the steerability $\mathrm{S}(\vec{\sigma}_{\mathbf{p}_A}) \leq \mathrm{E}_1(\rho)$. This leads to the conclusion that the nonlocality $\mathrm{N}(\mathbf{q_p})$ and the steerability $\mathrm{S}(\vec{\sigma}_{\mathbf{p}_A})$ are upper bounded by the smallest of the used resources to obtain $\mathbf{q_p}$, respectively $\vec{\sigma}_{\mathbf{p}_A}$.

Corollary 1. *Let $\mathcal{M}_{\mathbf{p}_A}$, $\mathcal{N}_{\mathbf{p}_B}$ be any WMAs and ρ any bipartite quantum state of appropriate dimensions. Let $\vec{\sigma}_{\mathbf{p}_A}$ be a state assemblage obtained via $\sigma_{a|x} = \mathrm{Tr}_1[(M_{a|x} \otimes \mathbb{1})\rho]$ and let $\mathbf{q_p} = (\mathbf{q}, \mathbf{p})$ be a probability distribution obtained via $q(a,b|x,y) = \mathrm{Tr}[N_{b|y}\sigma_{a|x}]$ and $p(x,y) = p_A(x)p_B(y)$. The following inequalities hold*

$$\mathrm{N}(\mathbf{q_p}) \leq \min\{\mathrm{E}_1(\rho), \mathrm{I}_\diamond(\mathcal{M}_{\mathbf{p}_A}), \mathrm{I}_\diamond(\mathcal{N}_{\mathbf{p}_B})\}, \tag{32}$$
$$\mathrm{S}(\vec{\sigma}_{\mathbf{p}_A}) \leq \min\{\mathrm{E}_1(\rho), \mathrm{I}_\diamond(\mathcal{M}_{\mathbf{p}_A})\}. \tag{33}$$

An example illustrating the hierarchy (27) is given by considering the respective resources of the Collins-Gisin-Linden-Massar-Popescu (CGLMP) measurements [59, 60] applied to the maximally entangled state $|\Phi^+\rangle = \frac{1}{\sqrt{d}} \sum_{i=0}^{d-1} |ii\rangle$. In the CGLMP scenario, Alice and Bob perform two projective measurements in dimension d, given by $\{M_{a|x} = |a_x\rangle\langle a_x|\}, \{M_{b|y} = |b_y\rangle\langle b_y|\}$, where

$$|a_x\rangle = \frac{1}{\sqrt{d}} \sum_{q=0}^{d-1} \exp\left[\frac{2\pi i}{d} q(a - \alpha_x)\right] |q\rangle, \tag{34}$$

for Alice's measurements and

$$|b_y\rangle = \frac{1}{\sqrt{d}} \sum_{q=0}^{d-1} \exp\left[-\frac{2\pi i}{d} q(b - \beta_y)\right] |q\rangle, \tag{35}$$

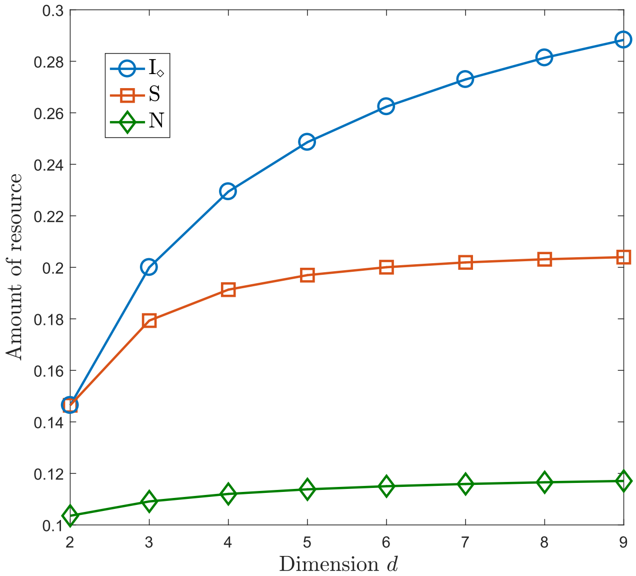

Figure 3. Comparison of the measurement resources. The incompatibility $\mathrm{I}_\diamond$, steerability S (both from Alice), and nonlocality N are shown here for the CGLMP measurements, for different dimensions d and $m = 2$ settings. The informativeness $\mathrm{IF}_\diamond$ and the coherence $\mathrm{C}_\diamond$, which are not shown here, coincide in this particular case. More specifically, $\mathrm{IF}_\diamond = \mathrm{C}_\diamond = 1 - \frac{1}{d}$ as we show in section V. The hierarchy $\mathrm{IF}_\diamond \geq \mathrm{C}_\diamond \geq \mathrm{I}_\diamond \geq \mathrm{S} \geq \mathrm{N}$ is clearly obeyed. While the nonlocality and the steerability converge quickly for growing d, the incompatibility increases further. The numerical methods used to obtain the results are explained in section IV.

for Bob's measurements, with $\alpha_x = (x - 1/2)/2$, $\beta_y = y/2$, and $a, b = 0, \cdots, d-1$ for $x, y = 1, 2$. We visualize our results in Figure 3.

IV. SDP FORMULATIONS

To study the hierarchy from Theorem 2 and the resources in more detail, an efficient method to numerically compute the respective resource quantifiers is needed. This can be done by formulating the quantifiers in terms of an SDP, which also allows us to study the quantifiers analytically by exploiting duality theory. The computation of the general quantifier $\mathrm{R}_\diamond(\mathcal{M}_\mathbf{p})$ from Eq. (12) can be stated as the following optimization problem:

$$\begin{aligned}
&\underline{\text{Primal problem (general):}} \qquad\qquad (36)\\
&\text{given}: \quad \mathcal{M}_\mathbf{p}\\
&\underset{Z_x, \mathcal{F}}{\text{minimize}} \ \sum_x p(x) \|\mathrm{Tr}_1[Z_x]\|_\infty\\
&\text{subject to:}\\
&Z_x \geq J(M_x) - J(F_x), \ Z_x \geq 0 \ \forall\, x, \ \mathcal{F} \in \mathscr{F},
\end{aligned}$$

where the Z_x are positive semidefinite matrices, $J(M_x)$ is the Choi–Jamiołkowski-matrix (see Eq. (9)) associated to setting x of the assemblage $\mathcal{M}$, and $\mathcal{F}$ are the elements of the set of free assemblages $\mathscr{F}$. The formulation of the optimization in Eq. (36) mainly relies on the SDP formulation of the diamond distance due to Watrous [61].

This compact representation of $\mathrm{R}_\diamond(\mathcal{M}_\mathbf{p})$ can be brought into an explicit SDP formulation whenever the set $\mathscr{F}$ admits an SDP formulation as we show in Appendix E for the resources considered in this work. Every SDP comes with a dual formulation which under some mild conditions (Slater's condition see e.g. [62]) returns the same optimal value as the primal problem. This condition is always satisfied for the SDP (36). Hence, $\mathrm{R}_\diamond(\mathcal{M}_\mathbf{p})$ can also be written as optimal value of the optimization problem:

Dual problem (general): (37)

$$
\begin{aligned}
&\text{given}: \quad \mathcal{M}_\mathbf{p} \\
&\underset{C_{a|x},\rho_x,\mathcal{F},O}{\text{maximize}} \ O \\
&\text{subject to:} \\
&O = \sum_{a,x} p(x)\mathrm{Tr}[M_{a|x}C_{a|x}] - \max_{\mathcal{F}\in\mathscr{F}} \sum_{a,x} p(x)\mathrm{Tr}[F_{a|x}C_{a|x}], \\
&0 \le C_{a|x} \le \rho_x \ \forall\ a,x,\ \ \rho_x \ge 0, \mathrm{Tr}[\rho_x] = 1\ \forall\ x,
\end{aligned}
$$

where the $C_{a|x}$, ρ_x are positive semidefinite matrices and $\mathcal{F}$ are the elements of the set of free assemblages $\mathscr{F}$. Note that the dual formulation in Eq. (37) is in direct correspondence to the steering and Bell inequality formulations in Eq. (29) and Eq. (30), as we maximize the difference of the resource value and the classical bound. The matrices $C_{a|x}$ describe a hyperplane in the assemblage space, while the states ρ_x fix the scale (i.e. $O \le 1$ for any $\mathcal{M}_\mathbf{p}$) of the dual program.

Since $\mathrm{R}_\diamond(\mathcal{M}_\mathbf{p})$ can be formulated as an SDP, it is efficiently computable (in the Hilbert space dimension d) and one can resort to standard toolboxes for its computation [63–66]. We want to remark that it is also possible to use a variation of the SDP (37) to obtain the optimal setting distribution $\mathbf{p}$ instead of fixing one in advance. This can be seen by introducing $C'_{a|x} = p(x)C_{a|x}$ and adjusting the constraints accordingly. See Appendix I for an example where optimizing over $\mathbf{p}$ leads to an advantage over the uniform distribution for the incompatibility $\mathrm{I}_\diamond(\mathcal{M}_\mathbf{p})$, even when only two measurement settings are considered.

Even though SDPs are mainly used for numerical optimization, the underlying structure of an SDP also offers a method to obtain analytical upper and lower bounds or even exact analytical expressions for $\mathrm{R}_\diamond(\mathcal{M}_\mathbf{p})$ depending on the complexity of the considered resource. More precisely, every feasible (but possibly sub-optimal) solution of the primal problem corresponds to an upper bound on $\mathrm{R}_\diamond(\mathcal{M}_\mathbf{p})$, while every feasible solution of the dual problem results in a lower bound. If we find feasible solutions of the primal and dual that result in the same value, we can conclude that this value is exactly $\mathrm{R}_\diamond(\mathcal{M}_\mathbf{p})$. We make use of this approach to derive bounds on the incompatibility $\mathrm{I}_\diamond(\mathcal{M}_\mathbf{p})$ for any assemblage $\mathcal{M}$ weighted with a uniform distribution $\mathbf{p}$ in Theorem 3 and to identify cases in which the hierarchy in Theorem 2 is tight in section V.

Theorem 3. *Given any WMA $\mathcal{M}_\mathbf{p}$ consisting of m POVMs in dimension d, with uniformly distributed measurement settings, i.e., $p(x) = 1/m\ \forall\ x$. The incompatibility $\mathrm{I}_\diamond(\mathcal{M}_\mathbf{p})$ is upper and lower bounded by*

$$
\begin{aligned}
\mathrm{I}_\diamond(\mathcal{M}_\mathbf{p}) \ge& \frac{1}{md}\sum_{a,x}\mathrm{Tr}[M_{a|x}^2] - \frac{1}{m}\Big(\max_{a,x}\|M_{a|x}\|_\infty + \qquad (38a)\\
&(m-1)\max_{a,a',x,x'\neq x}\|M_{a|x}^{1/2}M_{a'|x'}^{1/2}\|_\infty\Big), \\
\mathrm{I}_\diamond(\mathcal{M}_\mathbf{p}) \le& \frac{m-1}{(d+1)m^2}\sum_x \|d\mathbb{1} - \sum_a \mathrm{Tr}[M_{a|x}]M_{a|x}\|_\infty. \qquad (38b)
\end{aligned}
$$

Proof. The proof relies on finding feasible solutions of the primal (upper bound) and dual problem (lower bound) in Eq. (36) and Eq. (37) for the specific set of JM measurements $\mathscr{F}_{\mathrm{JM}}$. For the primal, we choose

$$
Z_x = (1-\eta)\sum_a |a\rangle\langle a| \otimes \frac{d - \mathrm{Tr}[M_{a|x}]}{d} M_{a|x}^T, \qquad (39)
$$

where $\eta \in [0,1]$ is the largest number such that $\mathcal{F}$ obtained from

$$
F_{a|x} = \eta M_{a|x} + (1-\eta)\mathrm{Tr}[M_{a|x}]\frac{\mathbb{1}}{d} \qquad (40)
$$

is JM. The coefficient η is known as the depolarizing robustness of the assemblage $\mathcal{M}$. Now by design, $\mathcal{F}$ is JM and $Z_x \ge 0$. The remaining constraint, $Z_x \ge \sum_a |a\rangle\langle a| \otimes (M_{a|x} - F_{a|x})^T$, can be verified by direct computation. It follows that $\mathrm{I}_\diamond(\mathcal{M}_\mathbf{p}) \le \frac{1-\eta}{md}\sum_x \|d\mathbb{1} - \sum_a \mathrm{Tr}[M_{a|x}]M_{a|x}\|_\infty$, where we used that $\mathbf{p}$ is uniformly distributed. Finally, the upper bound follows from [34], where it was found that $\eta^{\mathrm{low}} = \frac{1}{m}(1 + \frac{m-1}{d+1})$ is a lower bound to the depolarizing robustness and therefore always leads to jointly measurable measurements for general measurement assemblages $\mathcal{M}$ with m measurements of dimension d.

To obtain the lower bound from the dual problem, we rewrite $O = \sum_{a,x} p(x)\mathrm{Tr}[M_{a|x}C_{a|x}] - \mathrm{Tr}[L]$, where L is a matrix such that $L \ge \sum_{a,x} p(x)v(a|x,\lambda)C_{a|x}\ \forall\ \lambda$. Note that such an L always exists, which can be verified by multiplying both sides of the inequality with the POVM effect G_λ before summing over all λ and taking the trace. We choose as feasible solution $C_{a|x} = \frac{M_{a|x}}{d}$, $\rho_x = \sum_a C_{a|x} = \frac{\mathbb{1}}{d}$, and $L = l\mathbb{1}$ with some free parameter l. Clearly, in this way all constraints are satisfied for some appropriately chosen parameter l which still needs to be determined. We obtain that the incompatibility is lower bounded such that $\frac{1}{md}\sum_{a,x}\mathrm{Tr}[M_{a|x}^2] - dl \le \mathrm{I}_\diamond(\mathcal{M}_\mathbf{p})$. The constraint to find l is now given by

$l\mathbb{1} \geq \sum_{a,x} \frac{1}{md} v(a|x,\lambda) M_{a|x}\ \forall\ \lambda$, which means l is the spectral norm $\|\frac{1}{md} v(a|x,\lambda) M_{a|x}\|_\infty$ maximized over the deterministic distributions $\{v(a|x,\lambda)\}$. This means to find a valid l, we need to find an l such that

$$l \geq \frac{1}{md} \|\sum_{a,x} v^*(a|x,\lambda) M_{a|x}\|_\infty, \tag{41}$$

where $T := \|\sum_{a,x} v^*(a|x,\lambda) M_{a|x}\|_\infty$ and $\{v^*(a|x,\lambda)\}_{a,x}$ are the deterministic probability distributions that maximize the right-hand side of Eq (41), respectively the spectral norm in the definition of T. Using the results in [67] it follows that $l = \frac{1}{md}(\max_{a,x} \|M_{a|x}\|_\infty + (m-1) \max_{x,x' \neq x, a, a'} \|M_{a|x}^{1/2} M_{a'|x'}^{1/2}\|_\infty)$ is a valid choice, from which the lower bound on $\mathrm{I}_\diamond(\mathcal{M}_\mathbf{p})$ follows. □

While the bounds in Theorem 3 look complicated, we highlight that they become much simpler in the case of rank-1 projective measurements and especially for measurements based on MUB. Two orthonormal bases $\{|v_a\rangle\}_{0\leq a\leq d-1}$ and $\{|w_b\rangle\}_{0\leq b\leq d-1}$ are MUB if

$$|\langle v_a|w_b\rangle| = \frac{1}{\sqrt{d}}\ \forall\ a,b. \tag{42}$$

The set of projectors onto the vectors of a basis form a measurement $\mathcal{M} = \{M_a = |v_a\rangle\langle v_a|\}$. An MUB measurement assemblage is a set of measurements where the condition (42) holds for any two projections from different bases. MUB measurement assemblages find many applications in quantum information [48] and are natural candidates for highly incompatible measurements as studied in [34, 54, 68]. It is known, that in every dimension $d \geq 2$ there exist at least $m = 3$ different and at most $m = d+1$ mutually unbiased bases. While it is in general an open problem how many MUB really exist in a given dimension d, explicit constructions for $m = d+1$ MUB are known when d is a prime-power. Due to the fact that $\mathrm{Tr}[M_{a|x}] = \|M_{a|x}\|_\infty = 1$ for rank-1 projections, we obtain the following corollary.

Corollary 2. *The incompatibility $\mathrm{I}_\diamond(\mathcal{M}_\mathbf{p})$ of any measurement assemblage $\mathcal{M}$ consisting of m rank-1 projective measurements, weighted with a uniformly distributed $p(x) = \frac{1}{m}\ \forall\ x$, is bounded as*

$$1 - \frac{T}{m} \leq \mathrm{I}_\diamond(\mathcal{M}_\mathbf{p}) \leq (1-\eta)\frac{d-1}{d}, \tag{43}$$

with $T := \|\sum_{a,x} v^*(a|x,\lambda) M_{a|x}\|_\infty$ and the depolarizing robustness η defined via Eq. (40).

Using the overlap relation in Eq. (42) and the same lower bound on η as in Theorem 3, it follows for a uniformly weighted MUB measurement assemblage that

$$1 - \frac{1}{m}\big(1 + \frac{(m-1)}{\sqrt{d}}\big) \leq \mathrm{I}_\diamond(\mathcal{M}_\mathbf{p}) \leq \frac{(d-1)(m-1)}{(d+1)m}. \tag{44}$$

Some asymptotic behaviours for the incompatibility of MUB measurement assemblages can be observed. In the case of large dimensions d for a fixed number of measurements, the incompatibility approaches $\mathrm{I}_\diamond(\mathcal{M}_\mathbf{p}) \approx 1 - \frac{1}{m}$ (as the upper and lower bound collapse onto each other) i.e. it asymptotically approaches the value 1 for large d and m. To get an impression of the quality of the bounds in Eq. (44) we investigate a specific construction of MUB in prime dimensions d based on the Heisenberg-Weyl operators

$$\hat{X} = \sum_{k=0}^{d-1} |k+1\rangle\langle k|,\ \hat{Z} = \sum_{k=0}^{d-1} \omega^k |k\rangle\langle k|, \tag{45}$$

where $\{|k\rangle\}_{0\leq k\leq d-1}$ is the computational basis and $\omega = \exp\big(\frac{2\pi i}{d}\big)$ is a root of unity. In prime dimensions d, the eigenbases of the $d+1$ operators $\hat{X}, \hat{Z}, \hat{X}\hat{Z}, \hat{X}\hat{Z}^2, \cdots, \hat{X}\hat{Z}^{d-1}$ are mutually unbiased [69]. We use these eigenbases to form sets of projective POVMs. Note that it matters which subset of eigenbases we choose. For example, the set of measurements associated with the eigenbases of $\mathcal{M}^{(1)} = \{\hat{X}, \hat{Z}, \hat{X}\hat{Z}\}$ can possibly have a different incompatibility than the measurements associated with the set $\mathcal{M}^{(2)} = \{\hat{X}\hat{Z}^{d-3}, \hat{X}\hat{Z}^{d-2}, \hat{X}\hat{Z}^{d-1}\}$. This is indeed the case for MUB measurement assemblages in dimension $d = 5$ and $m = 3$ settings. We find that $\mathrm{I}_\diamond(\mathcal{M}^{(1)}_\mathbf{p}) = 0.3750$, while $\mathrm{I}_\diamond(\mathcal{M}^{(2)}_\mathbf{p}) = 0.3685$. This shows that different MUB are operationally inequivalent, which has also been demonstrated for the depolarizing robustness [54]. For the values in Table II, we used the assignment of MUB according to the WMA $\mathcal{M}^{(1)}_\mathbf{p}$, i.e., we take the first m eigenbases.

As one can see in Table II, the upper and lower bounds combined give a good idea on how incompatible this implementation of MUB is in practical scenarios. The lower bound can be tightened significantly by using the bound from Corollary 2 directly. Note that this requires an optimization over all $N_{\det} = o^m$ deterministic assignments $\{v(a|x,\lambda)\}$, where o is the number of measurement outcomes for each of the m settings. Surprisingly, the tightened lower bound coincides with the numerical values for the incompatibility $\mathrm{I}_\diamond(\mathcal{M}_\mathbf{p})$ for all m, d in Table II up to the fourth digit. While we were not able to show that the lower bound from Corollary 2 is tight for MUB measurement assemblages in general, we are able to identify important cases where this is indeed the case.

More specifically, it was shown by Designolle et al. [54] that $\eta = \frac{dT-m}{dm-m}$ is the depolarising robustness for the *standard construction* of MUB measurement assemblages in prime power dimensions given in [70] for $m = 2$, $m = d$, and $m = d+1$ measurements. It is important to highlight that the construction used above, based on the Heisenberg-Weyl operators, is an equivalent reformulation of this construction for prime dimensions [69]. From Eq. (43), it follows directly that $\mathrm{I}_\diamond(\mathcal{M}_\mathbf{p}) = 1 - \frac{T}{m}$, since

$m\backslash d$	2	3	5	7
2	0.1667 0.1464 0.1464	0.2500 0.2113 0.2113	0.3333 0.2764 0.2764	0.3750 0.3110 0.3110
3	0.2222 0.2113 0.1953	0.3333 0.2876 0.2818	0.4444 0.3750 0.3685	0.5000 0.4154 0.4147
4		0.3750 0.3455 0.3170	0.5000 0.4307 0.4146	0.5625 0.4724 0.4665
5			0.5333 0.4657 0.4422	0.6000 0.5040 0.4976
6			0.5556 0.4910 0.4607	0.6250 0.5257 0.5184
7				0.6429 0.5413 0.5332
8				0.6563 0.5728 0.5443

Table II. Incompatibility $\mathrm{I}_\diamond(\mathcal{M}_\mathbf{p})$ of MUB in prime dimensions d with m settings. In each cell, the first number is the upper bound on the incompatibility, the second number is the actual incompatibility which can be computed via the SDPs (36) and (37), marked in blue, and the third number is the lower bound on the incompatibility. The bounds are obtained from Eq. (44). Note that the lower bound is tight for $m = 2$ measurements. Furthermore, it is shown in the text, that the incompatibilities for $m = 2$, $m = d$, and $m = d + 1$ measurements can be obtained analytically.

the upper and lower bound coincide for these cases. Note that while this result holds only for this special construction, it was conjectured [54] that $\eta = \dfrac{dT-m}{dm-m}$ holds for all constructions of MUB and $m = 2$, $m = d$, and $m = d + 1$. Note further that the bounds in Eq. (44) can also be used to study cases where it is not known whether MUB exist. For instance, if there exists a set of $m = 4$ MUB in $d = 6$, the WMA needs to have an incompatibility in between $0.4438 \leq \mathrm{I}_\diamond(\mathcal{M}_\mathbf{p}) \leq 0.5357$.

To conclude this section, we want to emphasize that analogous discussions to obtain bounds on the resource quantifier $\mathrm{R}_\diamond(\mathcal{M}_\mathbf{p})$ can be made for any QRT with a free set $\mathcal{F}$ that can be described by SDP constraints. For instance, we show in Appendix F that the informativeness $\mathrm{IF}_\diamond(\mathcal{M}_\mathbf{p})$ of rank-1 projective measurements is given by $\mathrm{IF}_\diamond(\mathcal{M}_\mathbf{p}) = 1-\dfrac{1}{d}$ for any probability distribution $\mathbf{p}$. Note that since the set $\mathscr{F}_{\mathrm{UI}}$ of UI assemblages (see Eq. (16)) has a much simpler structure than the set of JM measurements, it is also easier to obtain exact expressions.

V. TIGHTNESS OF THE HIERARCHY

It is particularly interesting to study the optimal conversion of one resource to another, i.e., to study for which measurements (and states) the bounds in Eq. (27) are tight. Obviously, for UI measurements it holds $\mathrm{IF}_\diamond(\mathcal{M}_\mathbf{p}) = 0$ and all bounds are trivially tight. We study nontrivial cases of resource equivalences where $\mathrm{IF}_\diamond(\mathcal{M}_\mathbf{p}) = \mathrm{C}_\diamond(\mathcal{M}_\mathbf{p})$ and $\mathrm{I}_\diamond(\mathcal{M}_\mathbf{p}) = \mathrm{S}(\vec{\sigma}_\mathbf{p})$ holds. We start with the latter.

Incompatibility and steerability are known to be deeply connected and equivalences have been reported for robustness and weight-based quantifiers [13, 29]. We consider again the situation of uniformly distributed measurements, i.e., $p(x) = 1/m$. Let $\rho = |\Phi^+\rangle\langle\Phi^+|$ be the maximally entangled state, where $|\Phi^+\rangle = \frac{1}{\sqrt{d}}\sum_{i=0}^{d-1}|ii\rangle$. It is readily verified that

$$\sigma_{a|x} = \mathrm{Tr}_1[(M_{a|x}\otimes \mathbb{1})\rho] = \frac{M_{a|x}^T}{d}, \tag{46}$$

where the transposition is with respect to the computational basis. Using the state assemblage $\vec{\sigma} = \{\sigma_{a|x}\}$ obtained via Eq. (46) is the standard approach to map incompatibility problems to steering problems and proves also to be useful here. In section IV, we showed that for the construction of MUB in [70] and $m = 2$, $m = d$, $m = d + 1$ measurements $\mathrm{I}_\diamond(\mathcal{M}_\mathbf{p}) = 1 - \dfrac{T}{m}$ holds, where $T = \|\sum_{a,x} v^*(a|x,\lambda)M_{a|x}\|_\infty$. It follows that $1 - \dfrac{T}{m} \geq \mathrm{S}(\vec{\sigma}_\mathbf{p})$. Using the state assemblage $\vec{\sigma}$ obtained from Eq. (46), it is possible to show that this bound is indeed fulfilled. To show this, we employ the steering inequality formulation of $\mathrm{S}(\vec{\sigma}_\mathbf{p})$ as discussed in Eq. (29), which we repeat here for convenience:

$$\mathrm{S}(\vec{\sigma}_\mathbf{p}) = \max_{G_{a,x},\ell} \sum_{a,x} p(x)\mathrm{Tr}[\sigma_{a|x}G_{a|x}] - \ell, \tag{47}$$

where $\ell = \max_{\vec{\tau}\in\mathrm{LHS}} \sum_{a,x} p(x)\mathrm{Tr}[\tau_{a|x}G_{a|x}]$ is the classical bound obeyed by all unsteerable assemblages $\vec{\tau} \in \mathrm{LHS}$ and $\|G_{a|x}\|_\infty \leq 1$. We want to emphasize that ℓ can equivalently be written such that $\ell\mathbb{1} \geq \sum_{a,x} p(x)v(a|x,\lambda)G_{a|x}$ for all λ. By multiplying both sides of the inequalities with the hidden states σ_λ and taking the trace trace afterwards it follows $\ell \geq \max_{\vec{\tau}\in\mathrm{LHS}} \sum_{a,x} p(x)\mathrm{Tr}[\tau_{a|x}G_{a|x}]$ and the equality follows from the fact that we maximize over ℓ. Now, by choosing $G_{a|x} = M_{a|x}^T$ and $\ell = \dfrac{T}{m}$, clearly all constraints are fulfilled and the steerability is lower bounded by $\mathrm{S}(\vec{\sigma}_\mathbf{p}) \geq 1 - \dfrac{T}{m}$, which coincides with the upper bound. Therefore, it follows that $\mathrm{I}_\diamond(\mathcal{M}_\mathbf{p}) = \mathrm{S}(\vec{\sigma}_\mathbf{p})$. While this result is only valid for $m = 2$, $m = d$, and $m = d + 1$ and the special construction of MUB in [70] we conjecture that the equivalence between incompatibility and steerability holds for general constructions of MUB and $2 \leq m \leq d + 1$.

We searched numerically for other cases with an equality between incompatibility and steerability. However, apart from the case of generic qubit projective measurements we were not able to identify any other scenarios. Note that this finding deviates from the observations for consistent weight and robustness quantifiers studied by Cavalcanti et al. [13], where an equivalence between incompatibility and steerability was found for all assemblages. This difference is not artificial, as it remains even if we include the consistency constraint for the steerability below Eq. (24).

The second equivalence of resources we want to discuss is that between the informativeness and the coherence of assemblages. More precisely, we discuss when $\mathrm{IF}_\diamond(\mathcal{M}_\mathbf{p}) = \mathrm{C}_\diamond(\mathcal{M}_\mathbf{p})$ holds. Interestingly, this equivalence is achieved by WMAs $\mathcal{M}_\mathbf{p}$ that are mutually unbiased to the set of projective measurements onto the incoherent basis $\{|i\rangle\langle i|\}$. To see this, we note first that $\mathrm{IF}_\diamond(\mathcal{M}_\mathbf{p}) = 1-\frac{1}{d}$ holds for all rank-1 projective measurements as shown in section IV. From there it follows that the coherence of MUB is bounded by $1-\frac{1}{d} \geq \mathrm{C}_\diamond(\mathcal{M}_\mathbf{p})$. To show that this bound can be achieved, we use the dual formulation of $\mathrm{C}_\diamond(\mathcal{M}_\mathbf{p})$. More specifically, $\mathrm{C}_\diamond(\mathcal{M}_\mathbf{p})$ is given by

$$\mathrm{C}_\diamond(\mathcal{M}_\mathbf{p}) = \max_{C_{a|x},\ell_{x,i}} \sum_{a,x} p(x)\mathrm{Tr}[M_{a|x}C_{a|x}] - \sum_{x,i} \ell_{x,i}, \tag{48}$$

where the $\ell_{x,i}$ are scalars such that $\ell_{x,i} \geq p(x)\mathrm{Tr}[C_{a|x}|i\rangle\langle i|]\ \forall\ a,x,i$, and the $C_{a|x}$ are matrices such that $0 \leq C_{a|x} \leq \rho_x\ \forall\ a,x$, where the ρ_x are quantum states. The optimal solutions of the dual problem are $C_{a|x} = \frac{M_{a|x}}{d}$ and $\ell_{x,i} = \frac{p(x)}{d^2}$. Clearly, these choices are feasible, which can be verified by using that $\mathrm{Tr}[|i\rangle\langle i|M_{a|x}] = \frac{1}{d}$, due to the unbiasedness of $\mathcal{M}$ and the incoherent basis $\{|i\rangle\langle i|\}$. Further, they are optimal since they lead to $\mathrm{C}_\diamond(\mathcal{M}_\mathbf{p}) \geq 1-\frac{1}{d}$, which coincides with the upper bound.

The fact that measurements that are mutually unbiased to the incoherent basis maximize the coherence is very similar to the situation for quantum states [15]. There, for a fixed spectrum, the coherence is maximized by states that have an eigendecomposition in a mutually unbiased basis with respect to the incoherent basis. Note that the measurements within $\mathcal{M}$ do not need to be MUB measurement assemblages themselves, as long as they are mutually unbiased to the incoherent bases. Indeed, we show in Appendix F that the CGLMP measurements defined via Eq. (34) and Eq. (35) also maximize the coherence in the sense that $\mathrm{IF}_\diamond(\mathcal{M}_\mathbf{p}) = \mathrm{C}_\diamond(\mathcal{M}_\mathbf{p}) = 1-\frac{1}{d}$. Note that it is known that the maximal coherence of a single POVM in terms of the generalized robustness can be achieved by measurements in the Fourier basis of the incoherent basis [71] .

Let us briefly comment on the other two inequalities of the hierarchy in Eq. (27). The remaining two inequalities are $\mathrm{C}_\diamond(\mathcal{M}_\mathbf{p}) \geq \mathrm{I}_\diamond(\mathcal{M}_\mathbf{p})$ and $\mathrm{S}(\vec{\sigma}_{\mathbf{p}_A}) \geq \mathrm{N}(\mathbf{q}_\mathbf{p})$. While the relation between steering and nonlocality is notoriously hard to study, even for two-qubit states, the connection between coherence and incompatibility has only recently gained some attention [26, 72]. Our numerical search suggests that both bounds are true inequalities in non-trivial scenarios. However, future research is needed to come to a conclusion.

VI. CONCLUSION AND OUTLOOK

The quantification of quantum advantages plays an important role in modern quantum information theory and in particular in the framework of QRTs. In the present work, we have introduced the general notion of distance-based resource quantification for sets of measurements. We have studied which prerequisites are necessary for a function to be a proper distance between measurement assemblages and showed that every such distance induces a resource monotone. We have proposed one particular quantifier, based on the diamond norm, with a clear operational meaning in terms of the optimal single-shot distinguishability of different measurement assemblages.

On the basis of this particular quantifier, we have established a hierarchy of measurement resources in Theorem 2 and showed that recently introduced steering [46] and nonlocality quantifiers [47] fit naturally into this hierarchy. Furthermore, we have shown that our quantifier can be studied numerically and analytically in terms of SDPs. We have used this insight to establish analytical upper and lower bounds on the incompatibility of any measurement assemblage in Theorem 3. Noteworthy, by focussing on rank-1 projective measurements we have shown that the bounds on the incompatibility in Corollary 2 are tight for particular MUB measurement assemblages, which play a special role in the established measurement hierarchy. More precisely, we showed in section V that the incompatibility of MUB measurement assemblages attains the same value as the steerability of the state assemblages obtained from performing these measurements on one part of a maximally entangled state. Furthermore, we showed that measurements that are mutually unbiased to the incoherent basis maximize the coherence among all rank-1 projective measurements.

It would be interesting to see which insights can be obtained when distance-based quantifiers like the one presented here are studied for other resource theories like projective-simulability [20] or the QRT of imaginarity [18] applied to measurements. Furthermore, distance-based quantifiers should also be compared to possible entropic resource quantifiers of measurements assemblages. So far entropic quantifiers have only been considered very recently [73] for a single POVM. With the definition of a distance between measurement assemblages, it is also possible to study continuity of functions of measurement resources, which could be of independent interest for robust self-testing [74] or measurement tomography [75].

ACKNOWLEDGMENTS

We thank Lennart Bittel, Federico Grasselli, and Gláucia Murta for helpful discussions. This research was partially supported by the EU H2020 QuantERA ERA-NET Cofund in Quantum Technologies project QuICHE, and by the Federal Ministry of Education and Research (BMBF). MK's work is supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – project number 441423094.

Appendix A: Proof of Theorem 1

Here, we show that the function $\mathrm{D}_\diamond(\mathcal{M}_\mathbf{p}, \mathcal{N}_\mathbf{p})$ defined in Eq. (11) is a distance function between the two WMAs $\mathcal{M}_\mathbf{p}$ and $\mathcal{N}_\mathbf{p}$. In the following, we use that a measure-and-prepare channel (see Eq. (8)) corresponding to setting x of the assemblage $\mathcal{M}$ applied to the first subsystem of a bipartite state is given by

$$(\Lambda_{\mathcal{M}_x} \otimes \mathbb{1})(\rho) = \sum_a (|a\rangle\langle a| \otimes \mathrm{Tr}_1[(M_{a|x} \otimes \mathbb{1})\rho]). \quad \text{(A1)}$$

Furthermore, we introduce the quantities $\sigma_{a|x}(\rho) = \mathrm{Tr}_1[(M_{a|x} \otimes \mathbb{1})\rho]$ and $\tau_{a|x}(\rho) = \mathrm{Tr}_1[(N_{a|x} \otimes \mathbb{1})\rho]$.

Theorem 1. *The function $\mathrm{D}_\diamond(\mathcal{M}_\mathbf{p}, \mathcal{N}_\mathbf{p})$ is a distance function between the WMAs $\mathcal{M}_\mathbf{p}$ and $\mathcal{N}_\mathbf{p}$, i.e., it fulfils all the conditions stated in Definition 1.*

Proof. We start by writing $\mathrm{D}_\diamond(\mathcal{M}_\mathbf{p}, \mathcal{N}_\mathbf{p})$ in a more convenient form. More precisely, we use in the following that the triangle inequality for the trace norm $\|\cdot\|_1$ results in an equality due to the support on different subspaces of the terms with different a within the sum over the outcomes a. Furthermore, we used the multiplicity of the trace norm under tensor products and the fact that $\||a\rangle\langle a|\|_1 = 1 \;\forall\, a$. It follows that

$$\begin{aligned}
\mathrm{D}_\diamond(\mathcal{M}_\mathbf{p}, \mathcal{N}_\mathbf{p}) &= \frac{1}{2}\sum_x p(x) \max_\rho \|\sum_a |a\rangle\langle a| \otimes [\sigma_{a|x}(\rho) - \tau_{a|x}(\rho)]\|_1 \\
&= \frac{1}{2}\sum_x p(x) \max_\rho \sum_a \||a\rangle\langle a| \otimes [\sigma_{a|x}(\rho) - \tau_{a|x}(\rho)]\|_1 \\
&= \frac{1}{2}\sum_x p(x) \max_\rho \sum_a \|\sigma_{a|x}(\rho) - \tau_{a|x}(\rho)\|_1 .
\end{aligned} \tag{A2}$$

We use the form of $\mathrm{D}_\diamond(\mathcal{M}_\mathbf{p}, \mathcal{N}_\mathbf{p})$ according to Eq. (A2) in the following.

Clearly $\mathrm{D}_\diamond(\mathcal{M}_\mathbf{p}, \mathcal{N}_\mathbf{p})$ is a non-negative function with $\mathrm{D}_\diamond(\mathcal{M}_\mathbf{p}, \mathcal{N}_\mathbf{p}) = 0$ if and only if $\mathcal{M} = \mathcal{N}$. The triangle inequality is obeyed due to the linearity of $\mathrm{D}_\diamond(\mathcal{M}_\mathbf{p}, \mathcal{N}_\mathbf{p})$ in the trace norm. The monotonicity under quantum channel $\Lambda^\dagger$ follows from direct calculation,

$$\begin{aligned}
\mathrm{D}_\diamond(\mathcal{M}_\mathbf{p}, \mathcal{N}_\mathbf{p}) &= \frac{1}{2}\sum_x p(x) \max_\rho \sum_a \|\mathrm{Tr}_1[((M_{a|x} - N_{a|x}) \otimes \mathbb{1})\rho]\|_1 \\
&\geq \frac{1}{2}\sum_x p(x) \max_{\rho'} \sum_a \|\mathrm{Tr}_1[((M_{a|x} - N_{a|x}) \otimes \mathbb{1})(\Lambda \otimes \mathbb{1})(\rho')]\|_1 \\
&= \frac{1}{2}\sum_x p(x) \max_{\rho'} \sum_a \|\mathrm{Tr}_1[(\Lambda^\dagger(M_{a|x} - N_{a|x}) \otimes \mathbb{1})\rho']\|_1 \\
&= \mathrm{D}_\diamond(\Lambda^\dagger(\mathcal{M}_\mathbf{p}), \Lambda^\dagger(\mathcal{N}_\mathbf{p})),
\end{aligned} \tag{A3}$$

where we introduced in the second line a new quantum state ρ' (acting on a possibly different Hilbert space) and a CPT map Λ acting on the first subsystem of ρ'. The resulting state $(\Lambda \otimes \mathbb{1})(\rho')$ is clearly already included in the optimization of the first line, hence the inequality. In the third line, we used the fact that we can swap the evolution under the CPT map Λ onto the POVMs by introducing the adjoint channel $\Lambda^\dagger$. This results exactly in the definition of $\mathrm{D}_\diamond(\Lambda^\dagger(\mathcal{M}_\mathbf{p}), \Lambda^\dagger(\mathcal{N}_\mathbf{p}))$, from which the monotonicity under quantum channel $\Lambda^\dagger$ follows.

The monotonicity under measurement simulations ξ can be shown in a similar fashion,

$$
\begin{aligned}
\mathrm{D}(\xi(\mathcal{M}_{\mathbf{p}})_{\mathbf{q}}, \xi(\mathcal{N}_{\mathbf{p}})_{\mathbf{q}}) &= \frac{1}{2}\sum_y q(y) \max_\rho \sum_b \|\mathrm{Tr}_1[((M'_{b|y} - N'_{b|y}) \otimes \mathbb{1})\rho]\|_1 \\
&= \frac{1}{2}\sum_y q(y) \max_\rho \sum_b \|\sum_x p(x|y) \sum_a q(b|y,x,a) \mathrm{Tr}_1[((M_{a|x} - N_{a|x}) \otimes \mathbb{1})\rho]\|_1 \\
&\leq \frac{1}{2}\sum_y q(y) \max_\rho \sum_b \sum_x p(x|y) \sum_a q(b|y,x,a) \|\mathrm{Tr}_1[((M_{a|x} - N_{a|x}) \otimes \mathbb{1})\rho]\|_1 \\
&= \frac{1}{2}\sum_y q(y) \max_\rho \sum_x p(x|y) \sum_a \|\mathrm{Tr}_1[((M_{a|x} - N_{a|x}) \otimes \mathbb{1})\rho]\|_1 \\
&\leq \frac{1}{2}\sum_{x,y} q(y) p(x|y) \max_\rho \sum_a \|\mathrm{Tr}_1[((M_{a|x} - N_{a|x}) \otimes \mathbb{1})\rho]\|_1 \\
&= \frac{1}{2}\sum_x p(x) \max_\rho \sum_a \|\mathrm{Tr}_1[((M_{a|x} - N_{a|x}) \otimes \mathbb{1})\rho]\|_1 \\
&= \mathrm{D}_\diamond(\mathcal{M}_{\mathbf{p}}, \mathcal{N}_{\mathbf{p}}),
\end{aligned}
\tag{A4}
$$

where we used the following properties. In the first line, we used the definition of $\mathrm{D}(\xi(\mathcal{M}_{\mathbf{p}})_{\mathbf{q}}, \xi(\mathcal{N}_{\mathbf{p}})_{\mathbf{q}})$ by introducing the assemblages $\mathcal{M}'_{\mathbf{q}}$ and $\mathcal{N}'_{\mathbf{q}}$ with measurement outcomes b for the settings y, associated with the probability distribution $\mathbf{q}$. In the second line, we use that $M'_{b|y} = \sum_x p(x|y) \sum_a q(b|y,x,a) M_{a|x}$ and the analogous expression for $N'_{b|y}$. In the third line, we used the triangle inequality. In the fourth line, we performed the sum over b. In the fifth line, we interchanged the maximization with the sum over x, which leads to more degrees of freedom since we can now chose a different ρ for each x. Finally, in the sixth line, we used that $\sum_y q(y)p(x|y) = p(x)$, which leads exactly to the definition of $\mathrm{D}_\diamond(\mathcal{M}_{\mathbf{p}}, \mathcal{N}_{\mathbf{p}})$ from which the monotonicity under measurement simulations $\xi(\mathcal{M}_{\mathbf{p}})_{\mathbf{q}}$ follows. This concludes the proof. □

Appendix B: Steerability as upper bound to nonlocality

Here, we show that the steerability of a state assemblage $\vec{\sigma}_{\mathbf{p}_A}$ upper bounds the nonlocality of any probability distribution $\mathbf{q}_{\mathbf{p}}$ obtained from it. This completes the proof of Theorem 2.

Lemma 2. *Let $\vec{\sigma}_{\mathbf{p}_A} = (\vec{\sigma}, \mathbf{p}_A)$ be any state assemblage weighted with the probability distribution $\mathbf{p}_A$, $\mathcal{N}_{\mathbf{p}_B}$ any WMA of appropriate dimension and $\mathbf{q}_{\mathbf{p}} = (\mathbf{q}, \mathbf{p})$ a probability distribution obtained via $q(a,b|x,y) = \mathrm{Tr}[N_{b|y}\sigma_{a|x}]$ and $p(x,y) = p_A(x)p_B(y)$. Then, it holds that*

$$\mathrm{S}(\vec{\sigma}_{\mathbf{p}_A}) \geq \mathrm{N}(\mathbf{q}_{\mathbf{p}}). \tag{B1}$$

Proof. Let $\vec{\tau}$ be the closest LHS assemblage to $\vec{\sigma}$ with respect to the quantifier $\mathrm{S}(\vec{\sigma}_{\mathbf{p}_A})$. We use the fact that unsteerable assemblages always lead to local probability distributions. It follows that

$$
\begin{aligned}
\mathrm{N}(\mathbf{q}_{\mathbf{p}}) &\leq \frac{1}{2} \sum_{a,b,x,y} p(x,y) |\mathrm{Tr}[N_{b|y}(\sigma_{a|x} - \tau^*_{a|x})]| \\
&\leq \frac{1}{2} \sum_{a,b,x,y} p(x,y) \mathrm{Tr}[N_{b|y}|\sigma_{a|x} - \tau^*_{a|x}|] \\
&= \frac{1}{2} \sum_{a,x} p_A(x) \mathrm{Tr}[|\sigma_{a|x} - \tau^*_{a|x}|] \\
&= \frac{1}{2} \sum_{a,x} p_A(x) \|\sigma_{a|x} - \tau^*_{a|x}\|_1 = \mathrm{S}(\vec{\sigma}_{\mathbf{p}_A}),
\end{aligned}
\tag{B2}
$$

where we used in the first line the definition in Eq. (26) of the nonlocality $\mathrm{N}(\mathbf{q}_{\mathbf{p}})$ and the fact that any measurement on the closest LHS assemblage $\vec{\tau}$ (with respect to $\vec{\sigma}$) results in a local probability distribution. In the second line, we used some basic property of the absolute value and the fact that we can always decompose the difference of two Hermitian matrices like $\sigma_{a|x} - \tau^*_{a|x} = T_{a|x} - S_{a|x}$, where $T_{a|x}$ and $S_{a|x}$ are positive operators with orthogonal support. It follows that $|\sigma_{a|x} - \tau^*_{a|x}| = T_{a|x} + S_{a|x}$, where $|X| = \sqrt{X^\dagger X}$. Finally, we used in the third line that $\sum_b N_{b|y} = \mathbb{1}_d \ \forall\ y$, $\sum_y p(x,y) = p_A(x)$, and the definition of $\mathrm{S}(\vec{\sigma}_{\mathbf{p}_A})$. Therefore, it follows that the steerability is an upper bound to the nonlocality. □

Appendix C: Dual formulation of steerability and nonlocality

Here, we show that the steerability $\mathrm{S}(\vec{\sigma}_{\mathbf{p}_A})$ and the nonlocality $\mathrm{N}(\mathbf{q}_{\mathbf{p}})$ can be understood as optimal steering, respectively Bell inequality violation.

Theorem 4. *Let* $\mathrm{S}(\vec{\sigma}_{\mathbf{p}_A})$ *be the steerability of the state assemblage* $\vec{\sigma}_{\mathbf{p}_A}$*. The steerability* $\mathrm{S}(\vec{\sigma}_{\mathbf{p}_A})$ *can be reformulated as the violation of an optimized steering inequality given by*

$$\mathrm{S}(\vec{\sigma}_{\mathbf{p}_A}) = \max_{G_{a,x},\ell} \sum_{a,x} p_A(x)\mathrm{Tr}[\sigma_{a|x}G_{a|x}] - \ell, \tag{C1}$$

where $\ell = \max_{\vec{\tau}\in \mathrm{LHS}} \sum_{a,x} p_A(x)\mathrm{Tr}[\tau_{a|x}G_{a|x}]$ *is the classical bound obeyed by all unsteerable assemblages* $\vec{\tau} \in \mathrm{LHS}$ *and the* $G_{a|x}$ *are positive semidefinite matrices s.t.* $\|G_{a|x}\|_\infty \leq 1$*.*

Moreover, the nonlocality $\mathrm{N}(\mathbf{q_p})$ *of any probability distribution* $\mathbf{q_p}$ *can be reformulated as the violation of an optimized Bell inequality*

$$\mathrm{N}(\mathbf{q_p}) = \max_{C_{ab|xy},\ell'} \sum_{a,b,x,y} p(x,y)C_{ab|xy}q(a,b|x,y) - \ell', \tag{C2}$$

where $\ell' = \max_{\mathbf{t}\in \mathrm{LHV}} \sum_{a,b,x,y} p(x,y)C_{ab|xy}t(a,b|x,y)$ *is the local bound obeyed by all local correlations* $\mathbf{t} \in \mathrm{LHV}$ *and* $C_{ab|xy}$ *are Bell coefficients s.t.* $0 \leq C_{ab|xy} \leq 1$*.*

Proof. The proof relies on the dual formulations of $\mathrm{S}(\vec{\sigma}_{\mathbf{p}_A})$ which can be written in terms of an SDP and $\mathrm{N}(\mathbf{q_p})$ which can be written as a linear program. We start with the nonlocality $\mathrm{N}(\mathbf{q_p})$ by stating the optimization for an optimal Bell inequality violation given the distribution $\mathbf{q}$ and by showing that it is dual to $\mathrm{N}(\mathbf{q_p})$. Note that all of the following optimization problems require the knowledge of the deterministic probability distributions of the corresponding problem, which for the nonlocality $\mathrm{N}(\mathbf{q_p})$ are denoted by $\{v_A(a|x,\lambda)v_B(a|y,\lambda)\}$. Since these are fixed for a given problem and are trivially accessible, we will not treat them as input variables.

$$\begin{aligned}
&\underline{\text{Dual problem (nonlocality):}} \qquad\qquad \text{(C3)}\\
&\text{given}: \ \mathbf{q_p}\\
&\underset{C_{ab|xy},\ell'}{\text{maximize}} \sum_{a,b,x,y} p(x,y)C_{ab|xy}q(a,b|x,y) - \ell'\\
&\text{subject to:}\\
&\ell' \geq \sum_{a,b,x,y} p(x,y)C_{ab|xy}v_A(a|x,\lambda)v_B(b|y,\lambda) \ \forall \ \lambda,\\
&0 \leq C_{ab|xy} \leq 1 \ \forall \ a,b,x,y,
\end{aligned}$$

where $C_{ab|xy}$ are the Bell coefficients of the Bell inequality and ℓ' is the local bound. Note that $\ell' = \max_{\mathbf{t}\in\mathrm{LHV}} \sum_{a,b,x,y} p(x,y)C_{ab|xy}t(a,b|x,y)$ follows directly from the first constraint. Remember that $\mathbf{t}$ admits an LHV decomposition according to Eq. (25). The equality be seen by multiplying the constraints $\ell' \geq \sum_{a,b,x,y} p(x,y)C_{ab|xy}v_A(a|x,\lambda)v_B(b|y,\lambda) \ \forall \ \lambda$ with the probabilities $\pi(\lambda)$ before summing all the constraints together. This leads to the bound $\ell' \geq \max_{\mathbf{t}\in\mathrm{LHV}} \sum_{a,b,x,y} p(x,y)C_{ab|xy}t(a,b|x,y)$, with $t(a,b|x,y) = \sum_\lambda \pi(\lambda)v_A(a|x,\lambda)v_B(b|y,\lambda)$. The equality follows from the fact that we maximize the objective function.

Now, we show that the optimal value of the optimization in Eq. (C3) is equal to $\mathrm{N}(\mathbf{q_p})$ by deriving the primal program. Note that this generally requires dealing with inequality constraints, which can be done by generalizing the method of Lagrange multipliers to using the Karush–Kuhn–Tucker conditions (see e.g. [62]). However, since we are interested in formulating dual formulations of convex optimization problems, we can rely on simpler but less general conditions for the equivalence of the primal and the dual problem, which we come back to down below. We start by stating the Lagrangian of the problem:

$$\begin{aligned}
\mathcal{L} = &- \sum_{a,b,x,y} p(x,y)C_{ab|xy}q(a,b|x,y) + \ell'\\
&+ \sum_\lambda \pi(\lambda)\Big(\sum_{a,b,x,y} p(x,y)C_{ab|xy}v_A(a|x,\lambda)v_B(b|y,\lambda) - \ell'\Big)\\
&+ \sum_{a,b,x,y} A_{ab|xy}(C_{ab|xy} - 1) + \sum_{a,b,x,y} B_{ab|xy}(-C_{ab|xy}),
\end{aligned}$$

where we rewrote the maximization of the objective function as a minimization for convenience and introduced the Lagrange parameters $\pi(\lambda)$, $A_{ab|xy}$, and $B_{ab|xy}$, which are non-negative numbers to make the constraints explicit. Note that $\mathrm{N}(\mathbf{q_p}) \geq \mathcal{L}$ for any feasible point of the dual problem in Eq. (C3). We obtain the dual function by taking the infimum of the Lagrangian over the primal variables. More precisely, the dual function is given by

$$\begin{aligned}
&G(\{\pi(\lambda)\},\{A_{ab|xy}\},\{B_{ab|xy}\}) \qquad\qquad \text{(C4)}\\
&= \inf_{C_{ab|xy},\ell'} \Big\{ \ell'(1 - \sum_\lambda \pi(\lambda)) - \sum_{a,b,x,y} A_{ab|xy}\\
&+ \sum_{a,b,x,y} C_{ab|xy}\Big(- p(x,y)q(a,b|x,y)\\
&+ p(x,y)\sum_\lambda \pi(\lambda)v_A(a|x,\lambda)v_B(b|y,\lambda) + A_{ab|xy} - B_{ab|xy}\Big)\Big\}.
\end{aligned}$$

The dual function is unbounded from below, unless certain constraints (the primal constraints) are met. This can for instance be seen by realizing that unless $1 - \sum_\lambda \pi(\lambda) = 0$, it is always possible to make the term $\ell'(1 - \sum_\lambda \pi(\lambda))$ arbitrarily small, since ℓ' is now treated as an unconstrained variable. We obtain the primal program by maximizing the dual function under these constraints. Note that we rewrite the maximization as minimization to avoid a sign. The primal program is formally given by

$$\underline{\text{Primal problem (nonlocality):}} \tag{C5}$$
$$\text{given}: \ \mathbf{q_p}$$
$$\underset{A_{ab|xy},B_{ab|xy},\pi(\lambda),R}{\text{minimize}} \sum_{a,b,x,y} A_{ab|xy}$$
$$\text{subject to:}$$
$$A_{ab|xy} - B_{ab|xy} = R$$
$$R = p(x,y)\big(q(a,b|x,y) - \sum_\lambda \pi(\lambda) v_A(a|x,\lambda) v_B(b|y,\lambda)\big),$$
$$\sum_\lambda \pi(\lambda) = 1,\ \pi(\lambda) \geq 0\ \forall\ \lambda,\ A_{ab|xy}, B_{ab|xy} \geq 0\ \forall\ a,b,x,y.$$

Now, by definition of the ℓ_1-distance between two normalized probability distributions, the optimal value of Eq. (C5) is exactly $\mathrm{N}(\mathbf{q_p})$. Since we are dealing with a linear program, there is no duality gap between the primal and the dual formulation. Hence, $\mathrm{N}(\mathbf{q_p})$ describes the maximal Bell violation possible with the probability distribution $\mathbf{q_p}$.

Next, we need to show that $\mathrm{S}(\vec{\sigma}_{\mathbf{p}_A})$ corresponds to the optimal steering inequality violation. The procedure is the same as before for the nonlocality. However, this time we start from the primal problem.

$$\underline{\text{Primal problem (steerability):}} \tag{C6}$$
$$\text{given}: \ \vec{\sigma}_{\mathbf{p}_A}$$
$$\underset{\sigma_\lambda}{\text{minimize}}\ \frac{1}{2}\sum_{a,x} p_A(x) \|\sigma_{a|x} - \sum_\lambda v(a|x,\lambda)\sigma_\lambda\|_1$$
$$\text{subject to:}$$
$$\mathrm{Tr}[\sum_\lambda \sigma_\lambda] = 1, \sigma_\lambda \geq 0\ \forall\ \lambda.$$

First, we need to rewrite the trace norm explicitly in SDP form. We use the following formulation of the trace norm: $\|Z\|_1 = \min_Y \Big\{ \frac{\mathrm{Tr}[Y_1]}{2} + \frac{\mathrm{Tr}[Y_2]}{2} : \begin{bmatrix} Y_1 & Z \\ Z^\dagger & Y_2 \end{bmatrix} \geq 0 \Big\}$. This leads to the primal problem in explicit SDP form

$$\underline{\text{Primal problem (steerability):}} \tag{C7}$$
$$\text{given}: \ \vec{\sigma}_{\mathbf{p}_A}$$
$$\underset{\sigma_\lambda, U_{a|x}, W_{a|x}}{\text{minimize}}\ \frac{1}{4}\sum_{a,x} p_A(x)\mathrm{Tr}[U_{a|x} + W_{a|x}]$$
$$\text{subject to:}$$
$$\begin{bmatrix} U_{a|x} & \sigma_{a|x} - \sum_\lambda v(a|x,\lambda)\sigma_\lambda \\ \sigma_{a|x} - \sum_\lambda v(a|x,\lambda)\sigma_\lambda & W_{a|x} \end{bmatrix} \geq 0\ \forall\ a,x,$$
$$\mathrm{Tr}[\sum_\lambda \sigma_\lambda] = 1, \sigma_\lambda \geq 0\ \forall\ \lambda,$$

where we introduced the Hermitian matrices $U_{a|x}, W_{a|x}$.

This allows us to state the Lagrangian

$$\mathcal{L} = \frac{1}{4}\sum_{a,x} p_A(x)\mathrm{Tr}[U_{a|x} + W_{a|x}] + \ell'(1 - \mathrm{Tr}[\sum_\lambda \sigma_\lambda]) \tag{C8}$$
$$- \sum_{a,x}\Big(\mathrm{Tr}[H^{11}_{a|x} U_{a|x}] + \mathrm{Tr}[H^{22}_{a|x} W_{a|x}]$$
$$+ \mathrm{Tr}[2H^{12}_{a|x}(\sigma_{a|x} - \sum_\lambda v(a|x,\lambda)\sigma_\lambda)]\Big),$$

where ℓ' is a scalar and $H^{11}_{a|x}$, $H^{12}_{a|x}$, and $H^{22}_{a|x}$ are block-matrices such that

$$H_{a|x} = \begin{bmatrix} H^{11}_{a|x} & H^{12}_{a|x} \\ H^{12}_{a|x} & H^{22}_{a|x} \end{bmatrix} \geq 0\ \forall\ a,x. \tag{C9}$$

Note that $\mathrm{S}(\vec{\sigma}_{\mathbf{p}_A}) \geq \mathcal{L}$ for any feasible point of the dual problem in Eq. (C6). Analogous to the optimization for the nonlocality before, we can now formulate the dual function $G(\{H_{a|x}\},\ell') = \inf_{\sigma_\lambda, U_{a|x}, W_{a|x}} \mathcal{L}$. By identifying the conditions (the dual constraints) that make the dual function bounded we obtain the following dual program

$$\underline{\text{Dual problem (steerability):}} \tag{C10}$$
$$\text{given}: \ \vec{\sigma}_{\mathbf{p}_A}$$
$$\underset{H_{a|x},\ell'}{\text{maximize}}\ -\sum_{a,x}\mathrm{Tr}[2\sigma_{a|x}H^{12}_{a|x}] + \ell'$$
$$\text{subject to:}$$
$$-\ell'\mathbb{1} \geq -\sum_{a,x} v(a|x,\lambda)2H^{12}_{a|x}\ \forall\ \lambda,$$
$$\begin{bmatrix} \frac{1}{4}p_A(x)\mathbb{1} & H^{12}_{a|x} \\ H^{12}_{a|x} & \frac{1}{4}p_A(x)\mathbb{1} \end{bmatrix} \geq 0\ \forall\ a,x.$$

We can rewrite the dual problem in a more convenient form. By identifying the SDP formulation of the spectral norm $\|Z\|_\infty = \min_t \Big\{ t : \begin{bmatrix} t\mathbb{1} & Z \\ Z^\dagger & t\mathbb{1} \end{bmatrix} \geq 0 \Big\}$ and substituting $\ell' = -\tilde{\ell}$ and $H^{12}_{a|x} = -p(x)\frac{G'_{a|x}}{2}$, we arrive at

$$\underline{\text{Dual problem (steerability):}} \tag{C11}$$
$$\text{given}: \ \vec{\sigma}_{\mathbf{p}}$$
$$\underset{G'_{a|x},\tilde{\ell}}{\text{maximize}}\ \sum_{a,x} p(x)\mathrm{Tr}[G'_{a|x}\sigma_{a|x}] - \tilde{\ell}$$
$$\text{subject to:}$$
$$\tilde{\ell}\mathbb{1} \geq \sum_{a,x} p(x) v(a|x,\lambda) G'_{a|x}\ \forall\ \lambda,$$
$$-\frac{1}{2}\mathbb{1} \leq G'_{a|x} \leq \frac{1}{2}\mathbb{1}.$$

To finally arrive at the dual formulation equivalent to the statement in Eq. (C1) we shift the variables such that $G_{a|x} = G'_{a|x} + \frac{1}{2}\mathbb{1}$ and $\tilde{\ell} = \ell - \frac{1}{2}$. This leads to

$$
\begin{aligned}
&\underline{\text{Dual problem (steerability):}} \qquad\qquad (\text{C12})\\
&\text{given}: \ \vec{\sigma}_{\mathbf{p}_A}\\
&\underset{G_{a|x},\ell}{\text{maximize}} \ \sum_{a,x} p_A(x)\mathrm{Tr}[G_{a|x}\sigma_{a|x}] - \ell\\
&\text{subject to:}\\
&\ell\mathbb{1} \geq \sum_{a,x} p_A(x)v(a|x,\lambda)G_{a|x} \ \forall\ \lambda,\\
&0 \leq G_{a|x} \leq \mathbb{1}.
\end{aligned}
$$

Note that it follows again directly that the classical bound ℓ fulfills $\ell = \max_{\vec{\tau}\in\mathrm{LHS}} \sum_{a,x} p_A(x)\mathrm{Tr}[\tau_{a|x}G_{a|x}]$, where $\vec{\tau}$ admits an LHS as defined in Eq. (23).

As last step of the proof, we note that we can always find a strictly feasible point in the SDP corresponding to Eq. (C12) by choosing the $G_{a|x}$ proportional to the identity and ℓ sufficiently large. Hence there is no duality gap due to Slater's theorem (see e.g. [62]). Therefore, $\mathrm{S}(\vec{\sigma}_{\mathbf{p}_A})$ can be written as optimized steering inequality, which concludes the proof. □

Appendix D: Entanglement as upper bound for the steerability

Here, we show that the geometric entanglement $\mathrm{E}_1(\rho)$ defined in Eq. (3) as

$$
\mathrm{E}_1(\rho) = \min_{\rho_S\in\mathrm{Sep}(\mathcal{H}\otimes\mathcal{H})} \mathrm{D}_1(\rho,\rho_S), \qquad (\text{D1})
$$

where $\mathrm{Sep}(\mathcal{H}\otimes\mathcal{H})$ with $\mathcal{H}\cong\mathbb{C}^d$ is the set of separable states, upper bounds the steerability $\mathrm{S}(\vec{\sigma}_{\mathbf{p}})$ of the assemblage $\vec{\sigma}_{\mathbf{p}}$. More precisely, when $\vec{\sigma}_{\mathbf{p}}$ is obtained by performing d-dimensional measurements form any WMA $\mathcal{M}_{\mathbf{p}}$ onto any state $\rho\in\mathcal{S}(\mathcal{H}\otimes\mathcal{H})$ via $\sigma_{a|x} = \mathrm{Tr}_1[(M_{a|x}\otimes\mathbb{1})\rho]$ we show that it follows $\mathrm{S}(\vec{\sigma}_{\mathbf{p}}) \leq \mathrm{E}_1(\rho)$. Let ρ_S^* be the closest separable state with respect to the given state ρ. It holds,

$$
\begin{aligned}
&\mathrm{S}(\vec{\sigma}_{\mathbf{p}}) \qquad\qquad (\text{D2})\\
&\leq \frac{1}{2}\sum_{a,x} p(x)\|\mathrm{Tr}_1[(M_{a|x}\otimes\mathbb{1})(\rho-\rho_S^*)]\|_1\\
&= \frac{1}{2}\sum_{a,x} p(x)\max_{\|O_a\|_\infty\leq 1}|\mathrm{Tr}[O_a\mathrm{Tr}_1[(M_{a|x}\otimes\mathbb{1})(\rho-\rho_S^*)]]|\\
&= \frac{1}{2}\sum_{a,x} p(x)\max_{\|O_a\|_\infty\leq 1}|\mathrm{Tr}[(M_{a|x}\otimes O_a)(\rho-\rho_S^*)]|\\
&= \frac{1}{2}\sum_{x} p(x)\max_{\|O_a\|_\infty\leq 1}\sum_a \mathrm{Tr}[(M_{a|x}\otimes O_a)(\rho-\rho_S^*)]\\
&= \frac{1}{2}\sum_{x} p(x)\max_{\|O_a\|_\infty\leq 1}\mathrm{Tr}[\sum_a (M_{a|x}\otimes O_a)(\rho-\rho_S^*)]\\
&\leq \frac{1}{2}\sum_{x} p(x)\max_{\|O_a\|_\infty\leq 1}\|\sum_a (M_{a|x}\otimes O_{a|x})\|_\infty\|\rho-\rho_S^*\|_1\\
&\leq \frac{1}{2}\sum_{x} p(x)\|\rho-\rho_S^*\|_1 = \frac{1}{2}\|\rho-\rho_S^*\|_1 = \mathrm{E}_1(\rho),
\end{aligned}
$$

where we used in the second line the definition of the steerability $\mathrm{S}(\vec{\sigma}_{\mathbf{p}})$ and the fact that separable states ρ_S cannot lead to steering. In the third line, we used the variational characterization of the trace norm by introducing the optimization variables O_a. In the fourth line, we used some basic property of the trace and the partial trace. Next, we used in the fifth line, that we can interchange the sum over a and the maximization. Furthermore, we used that we can omit the absolute value, since we simply can change the sign of O_a if necessary. In the seventh line, we used the Hölder inequality. Finally, in the last line we used the fact that $\sum_a(M_{a|x}\otimes O_a) \leq \sum_a(M_{a|x}\otimes\mathbb{1})$ in the positive semidefinite sense. This lets us find as an upper bound $\|\sum_a M_{a|x}\otimes\mathbb{1}\|_\infty = 1$, due to the completeness relation of the POVMs. Therefore, the entanglement $\mathrm{E}_1(\rho)$ limits the steerability $\mathrm{S}(\vec{\sigma}_{\mathbf{p}}) \leq \mathrm{E}_1(\rho)$.

Appendix E: SDP formulations of incompatibility

Here, we give detailed information about the SDP formulations in Eq. (36) and Eq. (37). As an example, we explicitly derive the primal and the dual formulation of the incompatibility quantifier $\mathrm{I}_\diamond(\mathcal{M}_{\mathbf{p}})$. More specifically, we show that the incompatibility $\mathrm{I}_\diamond(\mathcal{M}_{\mathbf{p}})$ is the optimal value of the following two SDPs.

$$\begin{aligned}&\underline{\text{Primal problem (incompatibility):}}\\&\text{given}: \mathcal{M}_{\mathbf{p}}\\&\underset{a_x,Z_x,G_\lambda}{\text{minimize}}\sum_x p(x)a_x\\&\text{subject to:}\\&a_x\mathbb{1}-\mathrm{Tr}_1[Z_x]\geq 0\ \forall\ a,x,\\&Z_x\geq\sum_a|a\rangle\langle a|\otimes(M_{a|x}-F_{a|x})^T\ \forall\ x,\\&F_{a|x}=\sum_\lambda v(a|x,\lambda)G_\lambda\ \forall\ x,a,\ G_\lambda\geq 0\ \forall\ \lambda,\sum_\lambda G_\lambda=\mathbb{1},\\&Z_x\geq 0,\ a_x\geq 0\ \forall\ x,\end{aligned}\tag{E1}$$

$$\begin{aligned}&\underline{\text{Dual problem (incompatibility):}}\\&\text{given}:\ \mathcal{M}_{\mathbf{p}}\\&\underset{C_{a|x},\rho_x,L}{\text{maximize}}\quad\sum_{a,x}p(x)\mathrm{Tr}[M_{a|x}C_{a|x}]-\mathrm{Tr}[L]\\&\text{subject to:}\\&L\geq\sum_{a,x}p(x)v(a|x,\lambda)C_{a|x}\ \forall\ \lambda,\\&0\leq C_{a|x}\leq\rho_x\ \forall\ a,x,\ \rho_x\geq 0,\mathrm{Tr}[\rho_x]=1\ \forall\ x,\end{aligned}\tag{E2}$$

where $\{v(a|x,\lambda)\}$ are the deterministic probability distributions. The optimization variables of the primal problem are the positive coefficients a_x, and the positive semidefinite matrices Z_x and G_λ. The optimization variables of the dual problem are the positive semidefinite matrices $C_{a|x}$, ρ_x, and L.

The formulation of the primal problem heavily relies on the SDP formulation of the diamond norm due to Watrous [61], see also [76]. Let us recall that the Choi–Jamiołkowski-matrix of a measure-and-prepare channel (see Eq. (9)) corresponding to one POVM $\mathcal{M}_x=\{M_{a|x}\}_a$ is given by

$$J(\mathcal{M}_x)=\sum_a|a\rangle\langle a|\otimes M_{a|x}^T,\tag{E3}$$

where the transpose is with respect to the computational basis. The diamond distance between the quantum channels $\Lambda_{\mathcal{M}_x}$ and $\Lambda_{\mathcal{F}_x}$ can now be computed as

$$\begin{aligned}&\text{given}:\quad J(\mathcal{M}_x),J(\mathcal{F}_x)\\&\underset{Z_x}{\text{minimize}}\quad\|\mathrm{Tr}_1[Z_x]\|_\infty\\&\text{subject to: }Z_x\geq J(\mathcal{M}_x)-J(\mathcal{F}_x),\ Z_x\geq 0.\end{aligned}\tag{E4}$$

Using this form of the diamond norm, the primal problem in Eq. (36) follows by summing over the settings x weighted with probabilities $p(x)$ and by explicitly minimizing over the Choi–Jamiołkowski-matrices $J(\mathcal{F}_x)$, where $\mathcal{F}_x$ is the POVM corresponding to setting x of the free assemblages $\mathcal{F}$.

To arrive at the specific SDP for the incompatibility in Eq. (E1) from the general formulation in Eq. (36) we first note that the spectral norm $\|\mathrm{Tr}_1[Z_x]\|_\infty$ of a positive semidefinite matrix $\mathrm{Tr}_1[Z_x]$ can be written as the minimal value a_x such that $a_x\mathbb{1}\geq\mathrm{Tr}_1[Z_x]$ holds. Next, we write out the Choi–Jamiołkowski-matrices corresponding to the channels $\Lambda_{\mathcal{M}_x}$ and $\Lambda_{\mathcal{F}_x}$ in terms of the POVM elements $M_{a|x}$ and $F_{a|x}$. Finally, we constraint the $F_{a|x}$ explicitly to be JM i.e., that it holds $F_{a|x}=\sum_\lambda v(a|x,\lambda)G_\lambda\ \forall\ x,a$, where $\{G_\lambda\}$ is the POVM simulating $\mathcal{F}$.

To derive the dual formulation in Eq. (E2), we formulate the Lagrangian of the primal problem by incorporating the constraints explicitly. The Lagrangian is given by

$$\begin{aligned}\mathcal{L}=&\sum_x p(x)a_x+\sum_x\mathrm{Tr}[H_x(\mathrm{Tr}_1[Z_x]-a_x\mathbb{1})]\\&+\sum_x\mathrm{Tr}[C_x(\sum_a|a\rangle\langle a|\otimes[M_{a|x}\\&-\sum_\lambda v(a|x,\lambda)G_\lambda]^T-Z_x)]+\mathrm{Tr}[L(\sum_\lambda G_\lambda-\mathbb{1})],\end{aligned}\tag{E5}$$

where the H_x are d-dimensional positive semidefinite matrices, the C_x are $(o\times d)$-dimensional positive semidefinite matrices, where o is the number of measurement outcomes, and L is a d-dimensional Hermitian matrix. Note that for every feasible point in Eq. (E1), it holds $\mathrm{I}_\diamond(\mathcal{M}_{\mathbf{p}})\geq\mathcal{L}$.

By using the property $\mathrm{Tr}[\mathrm{Tr}_1(Z_x)H_x]=\mathrm{Tr}[Z_x(\mathbb{1}\otimes H_x)]$, we can formally state the dual function $G(\{H_x\},\{C_x\},L)$, which is obtained by taking the infimum of the Lagrangian $\mathcal{L}$ over the variables of the primal problem. More precisely,

$$\begin{aligned}&G(\{H_x\},\{C_x\},L)=\\&\inf_{a_x,Z_x,G_\lambda\geq 0}\Big\{\sum_x a_x(p(x)-\mathrm{Tr}[H_x])-\mathrm{Tr}[L]\\&+\sum_x\mathrm{Tr}[Z_x(\mathbb{1}\otimes H_x-C_x)]\\&+\sum_\lambda\mathrm{Tr}[G_\lambda(L-\sum_{a,x}v(a|x,\lambda)\mathrm{Tr}_1[(|a\rangle\langle a|\otimes\mathbb{1})C_x]^T)]\\&+\sum_x\mathrm{Tr}[C_x(\sum_a|a\rangle\langle a|\otimes M_{a|x}^T)]\Big\}.\end{aligned}\tag{E6}$$

It is clear that $G(\{H_x\},\{C_x\},L)$ is unbounded from below, unless certain conditions (the dual constraints) are met. For instance, if $p(x)-\mathrm{Tr}[H_x]<0$ holds for some x, the corresponding term $a_x(p(x)-\mathrm{Tr}[H_x])$ can be made arbitrarily small by increasing a_x, which is only constrained to be non-negative. The corresponding dual program is obtained by maximizing the dual function G over the dual variables $\{C_x\},\{D_x\},L$ under the dual constraints. This leads to the optimal lower bound to the primal problem.

We obtain

$$\underline{\text{Dual problem (incompatibility):}} \tag{E7}$$
$$\begin{aligned}
&\text{given}: \ \mathcal{M}_{\mathbf{p}} \\
&\underset{C_x, H_x, L}{\text{maximize}} \quad \sum_x \mathrm{Tr}[C_x(\sum_a |a\rangle\langle a| \otimes M_{a|x}^T)] - \mathrm{Tr}[L] \\
&\text{subject to:} \\
&L \geq \sum_{a,x} v(a|x,\lambda)\mathrm{Tr}_1[(|a\rangle\langle a| \otimes \mathbb{1})C_x]^T \ \forall \ \lambda, \\
&H_x \geq 0, \ C_x \geq 0, \ p(x) \geq \mathrm{Tr}[H_x] \ \forall \ x, \\
&\mathbb{1} \otimes H_x - C_x \geq 0, \ \forall \ x, \ L = L^\dagger,
\end{aligned}$$

which is formally the dual program to the primal formulation of Eq. (E1). However, we can rewrite the program in Eq. (E7) in a more useful form.

First, we can get rid of all transposes by using $\mathrm{Tr}[AB^T] = \mathrm{Tr}[A^T B]$, i.e., by swapping the transposition on the optimization variables C_x. Since the transpose C_x^T is already included in the optimization, we can simply ignore it. Second, we rewrite the first term of the objective function as $\mathrm{Tr}[C_x(\sum_a |a\rangle\langle a| \otimes M_{a|x})] = \sum_a \mathrm{Tr}[M_{a|x}(\mathrm{Tr}_1[(|a\rangle\langle a| \otimes \mathbb{1})C_x])]$. This shows that only the block-diagonal entries of C_x are important. Note that the same observation holds for the constraints C_x is involved in. It is therefore no loss of generalization to assume C_x as block diagonal. We denote $\mathrm{Tr}_1[(|a\rangle\langle a| \otimes \mathbb{1})C_x] = C'_{a|x}$. With this, we arrive at

$$\underline{\text{Dual problem (incompatibility):}} \tag{E8}$$
$$\begin{aligned}
&\text{given}: \ \mathcal{M}_{\mathbf{p}} \\
&\underset{C'_{a|x}, H_x, L}{\text{maximize}} \quad \sum_{a,x} \mathrm{Tr}[M_{a|x}C'_{a|x}] - \mathrm{Tr}[L] \\
&\text{subject to:} \\
&L \geq \sum_{a,x} v(a|x,\lambda)C'_{a|x} \ \forall \ \lambda, \\
&H_x \geq 0, p(x) \geq \mathrm{Tr}[H_x] \ \forall \ x, H_x \geq C'_{a|x} \geq 0 \ \forall \ a, x, L = L^\dagger.
\end{aligned}$$

Next, we note that it is always possible (without loss of optimality) to chose $p(x) = \mathrm{Tr}[H_x]$, since H_x constraints other variables only from above. We rewrite $H_x = p(x)\rho_x$, introducing the variables $\rho_x \geq 0$ with $\mathrm{Tr}[\rho_x] = 1$. Finally, we rewrite $C'_{a|x}$ such that $C'_{a|x} = p(x)C_{a|x}$ which leads to

$$\underline{\text{Dual problem (incompatibility):}} \tag{E9}$$
$$\begin{aligned}
&\text{given}: \ \mathcal{M}_{\mathbf{p}} \\
&\underset{C_{a|x}, \rho_x, L}{\text{maximize}} \quad \sum_{a,x} p(x)\mathrm{Tr}[M_{a|x}C_{a|x}] - \mathrm{Tr}[L] \\
&\text{subject to:} \\
&L \geq \sum_{a,x} p(x)v(a|x,\lambda)C_{a|x} \ \forall \ \lambda, \\
&0 \leq p(x)C_{a|x} \leq p(x)\rho_x \ \forall \ a, x, \ \rho_x \geq 0, \mathrm{Tr}[\rho_x] = 1 \ \forall \ x,
\end{aligned}$$

which is equivalent to the SDP (E2). Note that by virtue of the first constraint in Eq. (E10), it follows that

$$\mathrm{Tr}[L] = \max_{\mathcal{F} \in \mathscr{F}_{\mathrm{JM}}} \sum_{a,x} p(x)\mathrm{Tr}[F_{a|x}C_{a|x}]. \tag{E10}$$

The bound $\mathrm{Tr}[L] \geq \max_{\mathcal{F} \in \mathscr{F}_{\mathrm{JM}}} \sum_{a,x} p(x)\mathrm{Tr}[F_{a|x}C_{a|x}]$ can be seen by multiplying all the inequalities $L \geq \sum_{a,x} p(x)v(a|x,\lambda)C_{a|x} \ \forall \ \lambda$ with G_λ, then summing over all λ and taking the trace. The equality follows from the fact, that a strict inequality would contradict with the maximization of the objective function.

As a final step in the proof, we need to show that there is no duality gap between the primal and the dual program. This follows from Slater's theorem (see e.g. [62]) since it always possible to find a strictly feasible point in either the primal or the dual problem. This can be seen directly for the dual program in Eq. (E2), as we can chose all $C_{a|x}$ to be proportional to the identity and adjust the ρ_x and L accordingly.

Appendix F: SDP formulation of informativeness and coherence

While the calculations in Appendix E are specific to the quantifier $\mathrm{I}_\diamond(\mathcal{M}_{\mathbf{p}})$ and the QRT of incompatibility, analogous considerations can be made for any resource that has a free set $\mathscr{F}$ that admits a formulation as an SDP. In order to not repeat almost the same calculation as above, we simply state the corresponding SDP formulations for the coherence and the informativeness in the following. We start with the latter. The informativeness $\mathrm{IF}_\diamond(\mathcal{M}_{\mathbf{p}})$ is given as the optimal value of the following SDPs.

$$\underline{\text{Primal problem (informativeness):}} \tag{F1}$$
$$\begin{aligned}
&\text{given}: \ \mathcal{M}_{\mathbf{p}} \\
&\underset{a_x, Z_x, q(a|x)}{\text{minimize}} \sum_x p(x)a_x \\
&\text{subject to:} \\
&a_x\mathbb{1} - \mathrm{Tr}_1[Z_x] \geq 0 \ \forall \ a, x, \\
&Z_x \geq \sum_a |a\rangle\langle a| \otimes (M_{a|x} - F_{a|x})^T \ \forall \ x, \\
&F_{a|x} = q(a|x)\mathbb{1}, \quad q(a|x) \geq 0 \ \forall \ a, x, \sum_a q(a|x) = 1, \forall \ a, \\
&Z_x \geq 0, a_x \geq 0 \ \forall \ x,
\end{aligned}$$

$$\begin{aligned}
&\underline{\text{Dual problem (informativeness):}} && \text{(F2)}\\
&\text{given :}\quad \mathcal{M}_{\mathbf{p}}\\
&\underset{C_{a|x},\rho_x,\ell_x}{\text{maximize}}\quad \sum_{a,x} p(x)\mathrm{Tr}[M_{a|x}C_{a|x}] - \sum_x \ell_x\\
&\text{subject to:}\\
&\ell_x \geq p(x)\mathrm{Tr}[C_{a|x}]\ \forall\ a,x,\\
&0 \leq C_{a|x} \leq \rho_x\ \forall\ a,x,\quad \rho_x \geq 0, \mathrm{Tr}[\rho_x] = 1\ \forall\ x.
\end{aligned}$$

The optimization variables of the primal problem are the positive coefficients a_x, the positive semidefinite matrices Z_x and the probabilities $q(a|x)$. The optimization variables of the dual problem are the positive semidefinite matrices $C_{a|x}$, ρ_x, and the scalars ℓ_x. Note that it follows directly from the first constraint of the dual that

$$\sum_x \ell_x = \max_{\mathcal{F}\in\mathscr{F}_{\mathrm{UI}}} \sum_{a,x} p(x)\mathrm{Tr}[F_{a|x}C_{a|x}]. \tag{F3}$$

This can be seen by realizing that $\ell_x \geq p(x)\mathrm{Tr}[C_{a|x}]\ \forall a, x$ implies $\ell_x q(a|x) \geq p(x)\mathrm{Tr}[C_{a|x}q(a|x)\mathbb{1}]\ \forall a, x$, where we multiplied both sides with the conditional probabilities $q(a|x)$. We identify $q(a|x)\mathbb{1} = F_{a|x}$ due to the definition of the UI measurements in Eq. (16). Finally, we sum both sides over a and x. The equality follows from the fact that we maximize the objective function.

Like for the incompatibility, the SDP formulations of the informativeness $\mathrm{IF}_\diamond(\mathcal{M}_{\mathbf{p}})$ allow us to gain additional inside on the informativeness of WMA $\mathcal{M}_{\mathbf{p}}$. In particular, we show in the following that the informativeness $\mathrm{IF}_\diamond(\mathcal{M}_{\mathbf{p}})$ of any set of rank-1 projective measurements is given by $\mathrm{IF}_\diamond(\mathcal{M}_{\mathbf{p}}) = 1 - \frac{1}{d}$ for any probability distribution $p(x)$.

This follows by choosing $C_{a|x} = \frac{M_{a|x}}{d}$, $\rho_x = \sum_a \frac{M_{a|x}}{d} = \frac{\mathbb{1}}{d}$, and $\ell_x = \frac{p(x)}{d}$ as feasible solution for the dual problem in Eq. (F2). It follows by direct calculation that $\mathrm{IF}_\diamond(\mathcal{M}_{\mathbf{p}}) \geq 1 - \frac{1}{d}$. As feasible solution for the primal problem in Eq. (F1) we chose $q(a|x) = \frac{1}{d}$ and $Z_x = (1-\frac{1}{d})\sum_a |a\rangle\langle a| \otimes M_{a|x}$. This leads by direct computation to the upper bound $\mathrm{IF}_\diamond(\mathcal{M}_{\mathbf{p}}) \leq 1 - \frac{1}{d}$, which equals the lower bound.

State discrimination.—Note that in the particular case of rank-1 projective measurements, the informativeness $\mathrm{IF}_\diamond(\mathcal{M}_{\mathbf{p}})$ can be understood as state discrimination game. More precisely, we consider the following task. Given the (known) states $\{\rho_{a|x}\}$ that are distributed with probability $p(a|x) = \frac{1}{d}$ given that we chose the setting x with probability $p(x)$. That is, with probability $p(a,x) = p(a|x)p(x)$ we receive the state $\rho_{a|x}$. The goal of the game is now to identify the label a correctly with as high probability as possible, given that we have access to the measurement assemblage $\mathcal{M}$. We compare this to the situation where we do not perform a measurement and simply guess the label a. The term $\sum_{a,x} p(x)\mathrm{Tr}[M_{a|x}C_{a|x}] = \sum_{a,x} \frac{1}{d}p(x)\mathrm{Tr}[M_{a|x}\rho_{a|x}]$ can be seen as average probability to correctly guess the label a of the states $\rho_{a|x}$ that are send out with probability $\frac{1}{d}$ given that the setting x has been chosen with probability $p(x)$. The value $\mathrm{IF}_\diamond(\mathcal{M}_{\mathbf{p}}) = 1 - \frac{1}{d}$ describes now the difference of the optimal average probability, achieved with the assemblage $\mathcal{M}_{\mathbf{p}}$ compared to randomly guessing the label a for each setting x.

Coherence.—Finally, the coherence $\mathrm{C}_\diamond(\mathcal{M}_{\mathbf{p}})$ can also be computed by SDPs. In particular, $\mathrm{C}_\diamond(\mathcal{M}_{\mathbf{p}})$ is the optimal value of the following two SDPs.

$$\begin{aligned}
&\underline{\text{Primal problem (coherence):}} && \text{(F4)}\\
&\text{given :}\quad \mathcal{M}_{\mathbf{p}}\\
&\underset{a_x,Z_x,\alpha_{i|(a,x)}}{\text{minimize}}\ \sum_x p(x)a_x\\
&\text{subject to:}\\
&a_x\mathbb{1} - \mathrm{Tr}_1[Z_x] \geq 0\ \forall\ a,x,\\
&Z_x \geq \sum_a |a\rangle\langle a| \otimes (M_{a|x} - \sum_i \alpha_{i|(a,x)}|i\rangle\langle i|)^T\ \forall\ x,\\
&\alpha_{i|(a,x)} \geq 0\ \forall\ i,a,x, \sum_a \alpha_{i|(a,x)} = 1, \forall\ i,x,\\
&Z_x \geq 0, a_x \geq 0\ \forall\ x,\\
&\underline{\text{Dual problem (coherence):}} && \text{(F5)}\\
&\text{given :}\quad \mathcal{M}_{\mathbf{p}}\\
&\underset{C_{a|x},\rho_x,\ell_{x,i}}{\text{maximize}}\quad \sum_{a,x} p(x)\mathrm{Tr}[M_{a|x}C_{a|x}] - \sum_{x,i} \ell_{x,i}\\
&\text{subject to:}\\
&\ell_{x,i} \geq p(x)\mathrm{Tr}[C_{a|x}|i\rangle\langle i|]\ \forall\ a,x,i,\\
&0 \leq C_{a|x} \leq \rho_x\ \forall\ a,x,\quad \rho_x \geq 0, \mathrm{Tr}[\rho_x] = 1\ \forall\ x.
\end{aligned}$$

The optimization variables of the primal problem are the positive coefficients a_x, the positive semidefinite matrices Z_x and the coefficients $\alpha_{i|(a,x)}$. The optimization variables of the dual problem are the positive semidefinite matrices $C_{a|x}$, ρ_x, and the scalars $\ell_{x,i}$. With the same reasoning as with the previous resources, it can directly be seen that

$$\sum_{x,i} \ell_{x,i} = \max_{\mathcal{F}\in\mathscr{F}_{\mathrm{IC}}} \sum_{a,x} p(x)\mathrm{Tr}[F_{a|x}C_{a|x}]. \tag{F6}$$

We use these insights about the informativeness $\mathrm{IF}_\diamond(\mathcal{M}_{\mathbf{p}})$ and the coherence $\mathrm{C}_\diamond(\mathcal{M}_{\mathbf{p}})$ to identify non-trivial cases for which it holds that $\mathrm{IF}_\diamond(\mathcal{M}_{\mathbf{p}}) = \mathrm{C}_\diamond(\mathcal{M}_{\mathbf{p}})$ in the following. We start by considering assemblages $\mathcal{M}_{\mathbf{p}}$ where every POVM is a rank-1 projective measurement that is mutually unbiased to the incoherent basis.

More formally, it has to hold

$$\mathrm{Tr}[|i\rangle\langle i|M_{a|x}] = \frac{1}{d} \;\forall\; i, a, x. \tag{F7}$$

Note that this holds true for appropriately chosen assemblages $\mathcal{M}$ of MUB measurement assemblages that are also mutually unbiased to the incoherent basis. However, it is actually not necessary that the measurements within the assemblage are MUB themselves. All what is needed is that Eq. (F7) holds true for an assemblage $\mathcal{M}$ of rank-1 projections. For instance, the CGLMP (see Eq. (34) and Eq. (35)) measurements are also a valid choices. Under the condition in Eq. (F7) it is easy to see that the choices $C_{a|x} = \frac{M_{a|x}}{d}$, $\rho_x = \sum_a \frac{M_{a|x}}{d} = \frac{\mathbb{1}}{d}$, and $\ell_{x,i} = \frac{p(x)}{d^2}$ for the dual problem in Eq. (F5) lead to $\mathrm{C}_\diamond(\mathcal{M}_\mathbf{p}) \geq 1 - \frac{1}{d}$ for any probability distribution $p(x)$. Since $\mathrm{IF}_\diamond(\mathcal{M}_\mathbf{p}) = 1 - \frac{1}{d}$ and $\mathrm{IF}_\diamond(\mathcal{M}_\mathbf{p}) \geq \mathrm{C}_\diamond(\mathcal{M}_\mathbf{p})$ for any assemblage of rank-1 projections it has to hold $\mathrm{IF}_\diamond(\mathcal{M}_\mathbf{p}) = \mathrm{C}_\diamond(\mathcal{M}_\mathbf{p})$ whenever the condition in Eq. (F7) is fulfilled.

Appendix G: More distances

In the main text, we defined general distances between measurement assemblages in Definition 1. However, so far we only focused on one particular distance. Here, we introduce more examples of distances for measurements and discuss their basic properties.

We start by introducing the Schatten p−norm functions $\mathrm{D}_\mathrm{p}(\mathcal{M}_\mathbf{p}, \mathcal{N}_\mathbf{p})$ for $\mathrm{p} \in [1, \infty)$, defined as

$$\mathrm{D}_\mathrm{p}(\mathcal{M}_\mathbf{p}, \mathcal{N}_\mathbf{p}) = \sum_{a,x} p(x) \frac{1}{2} \|M_{a|x} - N_{a|x}\|_\mathrm{p}, \tag{G1}$$

where $\|X\|_\mathrm{p} = (\mathrm{Tr}[|X|^\mathrm{p}])^{1/\mathrm{p}}$ is the Schatten p−norm of X. Note that the cases $\mathrm{p} = 1$ and $\mathrm{p} = \infty$ correspond to the trace norm, respectively the spectral norm. While the functions $\mathrm{D}_\mathrm{p}(\mathcal{M}_\mathbf{p}, \mathcal{N}_\mathbf{p})$ will generally not fulfil the monotonicity under Hilbert-Schmidt adjoint channels $\Lambda^\dagger$ according to Definition 1, we will show in the following that the $\mathrm{p} = \infty$ case corresponds to a proper distance. Note that for $\mathrm{p} = 1$, the monotonicity under quantum channel $\Lambda^\dagger$ is not fulfilled, which can be seen by considering trivial extensions of the form $\Lambda^\dagger(M_{a|x}) = \mathbb{1} \otimes M_{a|x}$. Nevertheless, we also define the induced functions

$$\mathrm{R}_\mathrm{p}(\mathcal{M}_\mathbf{p}) = \min_{\mathcal{F} \in \mathscr{F}} \mathrm{D}_\mathrm{p}(\mathcal{M}_\mathbf{p}, \mathcal{F}_\mathbf{p}) \tag{G2}$$

We formulate the following theorem to show that $\mathrm{D}_\infty(\mathcal{M}_\mathbf{p}, \mathcal{N}_\mathbf{p})$ is a distance between measurement assemblages.

Theorem 5. *The function $\mathrm{D}_\infty(\mathcal{M}_\mathbf{p}, \mathcal{N}_\mathbf{p})$ is a distance function between the WMAs $\mathcal{M}_\mathbf{p}$ and $\mathcal{N}_\mathbf{p}$, i.e., it fulfils all the conditions stated in Definition 1.*

Proof. The proof can be split into several parts. Note first, that since $\mathrm{D}_\infty(\mathcal{M}_\mathbf{p}, \mathcal{N}_\mathbf{p})$ is a weighted sum of spectral norms, it follows that $\mathrm{D}_\infty(\mathcal{M}_\mathbf{p}, \mathcal{N}_\mathbf{p}) \geq 0$ with equality holding if and only if $\mathcal{M} = \mathcal{N}$. Note further that the symmetry and triangle inequality condition in Definition 1 are fulfilled trivially.

For the monotonicity under quantum channel, we consider a more general data-processing type inequality for the $\infty-$ distance between two POVM elements. Namely, for $\Lambda^\dagger(M_{a|x})$ and $\Lambda^\dagger(N_{a|x})$, where $\Lambda^\dagger$ is a unital completely positive map, it follows

$$\begin{aligned}
\|\Lambda^\dagger(M_{a|x}) - \Lambda^\dagger(N_{a|x})\|_\infty &= \max_\rho |\mathrm{Tr}[(\Lambda^\dagger(M_{a|x}) - \Lambda^\dagger(N_{a|x}))\rho]| \\
&= \max_\rho |\mathrm{Tr}[(M_{a|x} - N_{a|x})\Lambda(\rho)]| \\
&\leq \max_{\rho'} |\mathrm{Tr}[(M_{a|x} - N_{a|x})\rho']| \\
&= \|M_{a|x} - N_{a|x}\|_\infty,
\end{aligned} \tag{G3}$$

where we used the dual representation of the Schatten-∞ norm and the fact that maximum is always achieved for a density matrix ρ (more specifically the projector onto the eigenvalue of largest absolute value of $\Lambda^\dagger(M_{a|x}) - \Lambda^\dagger(N_{a|x})$). Furthermore, we used that the adjoint of the unital completely positive map $\Lambda^\dagger$ is a CPT map Λ which maps density matrices onto density matrices and therefore shrinks the state-space one optimizes over. Since $\mathrm{D}_\infty(\mathcal{M}_\mathbf{p}, \mathcal{N}_\mathbf{p})$ is a sum of norms $\|M_{a|x} - N_{a|x}\|_\infty$, it follows that $\mathrm{D}_\infty(\mathcal{M}_\mathbf{p}, \mathcal{N}_\mathbf{p}) \geq \mathrm{D}_\infty(\Lambda^\dagger(\mathcal{M}_\mathbf{p}), \Lambda^\dagger(\mathcal{N}_\mathbf{p}))$.

The monotonicity under measurement simulations $\xi(\mathcal{M}_\mathbf{p})_\mathbf{q}$ follows by direct computation,

$$\begin{aligned}\mathrm{D}_\infty(\xi(\mathcal{M}_\mathbf{p})_\mathbf{q},\xi(\mathcal{N}_\mathbf{p})_\mathbf{q}) &= \frac{1}{2}\sum_{b,y} q(y)\|\sum_{x,a} p(x|y)q(b|y,x,a)[M_{a|x}-N_{a|x}]\|_\infty \\ &\leq \frac{1}{2}\sum_{b,y} q(y)\sum_{x,a} p(x|y)q(b|y,x,a)\|M_{a|x}-N_{a|x}\|_\infty \\ &= \frac{1}{2}\sum_{y,x,a} q(y)p(x|y)\|M_{a|x}-N_{a|x}\|_\infty \\ &= \frac{1}{2}\sum_{x,a} p(x)\|M_{a|x}-N_{a|x}\|_\infty = \mathrm{D}_\infty(\mathcal{M}_\mathbf{p},\mathcal{N}_\mathbf{p}),\end{aligned} \tag{G4}$$

where we used the following properties. In the first line, we used the definition of $\mathrm{D}_\infty(\xi(\mathcal{M}_\mathbf{p})_\mathbf{q},\xi(\mathcal{N}_\mathbf{p})_\mathbf{q})$ by introducing the assemblages $\mathcal{M}'_\mathbf{q}=\xi(\mathcal{M}_\mathbf{p})_\mathbf{q}$ and $\mathcal{N}'_\mathbf{q}=\xi(\mathcal{N}_\mathbf{p})_\mathbf{q}$ where we inserted $M'_{b|y}=\sum_x p(x|y)\sum_a q(b|y,x,a)M_{a|x}$ and $N'_{b|y}=\sum_x p(x|y)\sum_a q(b|y,x,a)N_{a|x}$ directly. In the second line, we used the triangle inequality. In the third line, we performed the sum over b. Finally, in the fourth line, we used that $\sum_y q(y)p(x|y)=p(x)$, which leads exactly to the definition of $\mathrm{D}_\diamond(\mathcal{M}_\mathbf{p},\mathcal{N}_\mathbf{p})$ from which the monotonicity under measurement simulations ξ follows. Therefore, $\mathrm{D}_\infty(\mathcal{M}_\mathbf{p},\mathcal{N}_\mathbf{p})$ is a distance between measurement assemblages according to Definition 1. □

Even though they are not resources monotones generally, the functions $R_n(\mathcal{M}_\mathbf{p})$ in Eq. (G2) can be used to bound the resource quantifier $\mathrm{R}_\diamond(\mathcal{M}_\mathbf{p})$ defined in Eq. (12). More specifically, we derive in the following the bounds on the diamond distance based quantifier $\mathrm{R}_\diamond(\mathcal{M}_\mathbf{p})$ given by

$$\begin{aligned}\frac{1}{d}\mathrm{R}_\infty(\mathcal{M}_\mathbf{p}) \leq \frac{1}{d}\mathrm{R}_1(\mathcal{M}_\mathbf{p}) &\leq \mathrm{R}_\diamond(\mathcal{M}_\mathbf{p}) \\ &\leq \mathrm{R}_\infty(\mathcal{M}_\mathbf{p}) \leq \mathrm{R}_1(\mathcal{M}_\mathbf{p}),\end{aligned} \tag{G5}$$

where d is the dimension of the Hilbert space $\mathcal{H}$ the POVMs from $\mathcal{M}$ act on. Note that due to the monotonicity of Schatten norms, it holds $\|X\|_\mathrm{p} \leq \|X\|_{\mathrm{p}'}$ for $\mathrm{p}\geq\mathrm{p}'$ from which the bounds $\frac{1}{d}\mathrm{R}_\infty(\mathcal{M}_\mathbf{p}) \leq \frac{1}{d}\mathrm{R}_1(\mathcal{M}_\mathbf{p})$ and $\mathrm{R}_\infty(\mathcal{M}_\mathbf{p}) \leq \mathrm{R}_1(\mathcal{M}_\mathbf{p})$ follow directly.

The bound $\mathrm{R}_\diamond(\mathcal{M}_\mathbf{p}) \leq \mathrm{R}_\infty(\mathcal{M}_\mathbf{p})$ follows from

$$\begin{aligned}\mathrm{R}_\diamond(\mathcal{M}_\mathbf{p}) &= \min_{\mathcal{F}\in\mathscr{F}} \frac{1}{2}\sum_x p(x)\max_\rho \sum_a \|\mathrm{Tr}_1[(M_{a|x}\otimes\mathbb{1})\rho]-\mathrm{Tr}_1[(F_{a|x}\otimes\mathbb{1})\rho]\|_1 \\ &\leq \min_{\mathcal{F}\in\mathscr{F}} \frac{1}{2}\sum_x p(x)\max_\rho \sum_a \|(M_{a|x}\otimes\mathbb{1})\rho-(F_{a|x}\otimes\mathbb{1})\rho\|_1 \\ &\leq \min_{\mathcal{F}\in\mathscr{F}} \frac{1}{2}\sum_{a,x} p(x)\max_\rho \|(M_{a|x}\otimes\mathbb{1})-(F_{a|x}\otimes\mathbb{1})\|_\infty\|\rho\|_1 \\ &= \min_{\mathcal{F}\in\mathscr{F}} \frac{1}{2}\sum_{a,x} p(x)\|(M_{a|x}\otimes\mathbb{1})-(F_{a|x}\otimes\mathbb{1})\|_\infty = \mathrm{R}_\infty(\mathcal{M}_\mathbf{p}),\end{aligned} \tag{G6}$$

where we used the definition of $\mathrm{R}_\diamond(\mathcal{M}_\mathbf{p})$ in the first line and the monotonicity of the trace norm under partial trace in the second line. In the third line, we used the Hölder inequality and in the last line identified the definition of $\mathrm{R}_\infty(\mathcal{M}_\mathbf{p})$.

The last remaining bound $\frac{1}{d}\mathrm{R}_1(\mathcal{M}_\mathbf{p}) \leq \mathrm{R}_\diamond(\mathcal{M}_\mathbf{p})$ can directly be obtained by using $\rho=|\Phi^+\rangle\langle\Phi^+|$ within the optimization of the diamond norm. Here, $|\Phi^+\rangle=\frac{1}{\sqrt{d}}\sum_{i=0}^{d-1}|ii\rangle$ is the maximally entangled state, where (as before) d

is the dimension the POVMs of the measurement assemblage $\mathcal{M}$ act on. It follows

$$\begin{aligned}
\mathrm{R}_\diamond(\mathcal{M}_\mathbf{p}) &= \min_{\mathcal{F}\in\mathscr{F}} \frac{1}{2}\sum_x p(x)\max_\rho \sum_a \|\mathrm{Tr}_1[(M_{a|x}\otimes \mathbb{1})\rho] - \mathrm{Tr}_1[(F_{a|x}\otimes \mathbb{1})\rho]\|_1 \qquad (G7)\\
&\geq \min_{\mathcal{F}\in\mathscr{F}} \frac{1}{2}\sum_x p(x)\sum_a \|\mathrm{Tr}_1[(M_{a|x}\otimes \mathbb{1})|\Phi^+\rangle\langle\Phi^+|] - \mathrm{Tr}_1[(F_{a|x}\otimes \mathbb{1})|\Phi^+\rangle\langle\Phi^+|]\|_1\\
&= \min_{\mathcal{F}\in\mathscr{F}} \frac{1}{2}\sum_x p(x)\sum_a \frac{1}{d}\|(M_{a|x}-F_{a|x})^T\|_1\\
&= \min_{\mathcal{F}\in\mathscr{F}} \frac{1}{2}\sum_x p(x)\sum_a \frac{1}{d}\|M_{a|x}-F_{a|x}\|_1 = \frac{1}{d}\mathrm{R}_1(\mathcal{M}_\mathbf{p}),
\end{aligned}$$

where we used in the first line the definition of $\mathrm{R}_\diamond(\mathcal{M}_\mathbf{p})$ and in the second line that $\rho = |\Phi^+\rangle\langle\Phi^+|$ is a feasible point within the maximization over the quantum states within the diamond norm. In the third line, we used that $\mathrm{Tr}_1[(M_{a|x}\otimes \mathbb{1})|\Phi^+\rangle\langle\Phi^+|] = \frac{1}{d}M_{a|x}^T$, where the transposition is with respect to the computational basis. Finally, we can use that a transposition does not change the singular values.

The monotone $\mathrm{R}_\infty(\mathcal{M}_\mathbf{p})$ in particular is not only a valuable tool to bound the diamond distance $\mathrm{R}_\diamond(\mathcal{M}_\mathbf{p})$ but is also interesting in itself. More specifically, we show in the following that $\mathrm{R}_\infty(\mathcal{M}_\mathbf{p})$ obeys also a measurement hierarchy similar to that in Eq. (27). Let $\mathrm{IF}_\infty(\mathcal{M}_\mathbf{p}), \mathrm{C}_\infty(\mathcal{M}_\mathbf{p})$, and $\mathrm{I}_\infty(\mathcal{M}_\mathbf{p})$ be the informativeness, coherence, and incompatibility of $M_\mathbf{p}$ as measured by the distance $\mathrm{R}_\infty(\mathcal{M}_\mathbf{p})$ in Eq. (G2) with respect to the free sets $\mathscr{F}_{\mathrm{UI}}$, $\mathscr{F}_{\mathrm{IC}}$, and $\mathscr{F}_{\mathrm{JM}}$. It follows directly from $\mathscr{F}_{\mathrm{UI}} \subset \mathscr{F}_{\mathrm{IC}} \subset \mathscr{F}_{\mathrm{JM}}$ that the hierarchy

$$\mathrm{IF}_\infty(\mathcal{M}_{\mathbf{p}_A}) \geq \mathrm{C}_\infty(\mathcal{M}_{\mathbf{p}_A}) \geq \mathrm{I}_\infty(\mathcal{M}_{\mathbf{p}_A}), \qquad (G8)$$

holds. Moreover, it can be shown that $\mathrm{I}_\infty(\mathcal{M}_{\mathbf{p}_A}) \geq \mathrm{S}(\vec{\sigma}_{\mathbf{p}_A})$ (remember that we already showed that $\mathrm{S}(\vec{\sigma}_{\mathbf{p}_A}) \geq \mathrm{N}(\mathbf{q_p})$). This follows from the direct computation for any quantum state ρ of appropriate dimension and the closest JM assemblage $\mathcal{F}^*$ to $\mathcal{M}$ (with respect to the monotone $\mathrm{I}_\infty(\mathcal{M}_\mathbf{p})$):

$$\begin{aligned}
&\mathrm{S}(\vec{\sigma}_{\mathbf{p}_A}) \qquad (G9)\\
&\leq \frac{1}{2}\sum_{a,x} p_A(x)\|\mathrm{Tr}_1[((M_{a|x}-F^*_{a|x})\otimes\mathbb{1})\rho]\|_1\\
&\leq \frac{1}{2}\sum_{a,x} p_A(x)\|((M_{a|x}-F^*_{a|x})\otimes\mathbb{1})\rho\|_1\\
&\leq \frac{1}{2}\sum_{a,x} p_A(x)\|(M_{a|x}-F^*_{a|x})\otimes\mathbb{1}\|_\infty\|\rho\|_1\\
&= \frac{1}{2}\sum_{a,x} p_A(x)\|M_{a|x}-F^*_{a|x}\|_\infty = \mathrm{I}_\infty(\mathcal{M}_{\mathbf{p}_A}),
\end{aligned}$$

where we first used that JM measurements always lead to unsteerable assemblages. Second, we used that the trace norm is non-increasing under partial traces. Third, we used the Hölder inequality. It therefore follows, that the hierarchy

$$\mathrm{IF}_\infty(\mathcal{M}_{\mathbf{p}_A}) \geq \mathrm{C}_\infty(\mathcal{M}_{\mathbf{p}_A}) \geq \mathrm{I}_\infty(\mathcal{M}_{\mathbf{p}_A}) \geq \mathrm{S}(\vec{\sigma}_{\mathbf{p}_A}) \geq \mathrm{N}(\mathbf{q_p}), \qquad (G10)$$

holds.

Another distance for measurement assemblages that can be considered is based on the ℓ_1-distance between probability distributions. More specifically, the induced ℓ_1-distance between two WMAs is given by

$$\mathrm{D}_{\ell_1}(\mathcal{M}_\mathbf{p},\mathcal{N}_\mathbf{p}) = \frac{1}{2}\sum_x p(x)\max_{\rho_A}\sum_a |\mathrm{Tr}[(M_{a|x}-N_{a|x})\rho_A]|, \qquad (G11)$$

which is the ℓ_1-distance of the probability distributions $\{\mathrm{Tr}[M_{a|x}\rho_A]\}$ and $\{\mathrm{Tr}[N_{a|x}\rho_A]\}$ maximized over all quantum states ρ_A. With the same methods as for the distances $\mathrm{D}_\infty(\mathcal{M}_\mathbf{p},\mathcal{N}_\mathbf{p})$ and $\mathrm{D}_\diamond(\mathcal{M}_\mathbf{p},\mathcal{N}_\mathbf{p})$ it can be shown that $\mathrm{D}_{\ell_1}(\mathcal{M}_\mathbf{p},\mathcal{N}_\mathbf{p})$ fulfills all the conditions in Definition 4. Hence, $\mathrm{D}_{\ell_1}(\mathcal{M}_\mathbf{p},\mathcal{N}_\mathbf{p})$ is a distance function which induces the distance-based monotone

$$\mathrm{R}_{\ell_1}(\mathcal{M}_\mathbf{p}) = \min_{\mathcal{F}\in\mathscr{F}} \mathrm{D}_{\ell_1}(\mathcal{M}_\mathbf{p},\mathcal{F}_\mathbf{p}). \qquad (G12)$$

Note that while it follows directly that $\mathrm{R}_{\ell_1}(\mathcal{M}_\mathbf{p})$ will naturally induce a hierarchy between the informativeness, coherence, and the incompatibility of a WMA $\mathcal{M}_\mathbf{p}$, it is not clear whether there exist steering or nonlocality monotones that are in natural correspondence to it. Note further that in the context of coherence of single POVMs, this kind of statistical measure has also been defined by Baek et al. [23].

Even though it is not clear whether a complete hierarchy of measurement resources holds, the quantifier $\mathrm{R}_{\ell_1}(\mathcal{M}_\mathbf{p})$ is important, as it can be seen as limiting case of the quantifier $\mathrm{R}_\diamond(\mathcal{M}_\mathbf{p})$ when the maximization is performed only over product states. More formally, it holds

$$\begin{aligned}
&\mathrm{R}_\diamond(\mathcal{M}_\mathbf{p}) &\text{(G13)}\\
&\geq \min_{\mathcal{F}\in F}\frac{1}{2}\sum_x p(x)\max_{\rho=\rho_A\otimes\rho_B}\sum_a \|\mathrm{Tr}_1[((M_{a|x}-F_{a|x})\otimes\mathbb{1})\rho]\|_1\\
&=\min_{\mathcal{F}\in F}\frac{1}{2}\sum_x p(x)\max_{\rho=\rho_A\otimes\rho_B}\sum_a \|\mathrm{Tr}[(M_{a|x}-F_{a|x})\rho_A]\rho_B\|_1\\
&=\min_{\mathcal{F}\in F}\frac{1}{2}\sum_x p(x)\max_{\rho=\rho_A\otimes\rho_B}\sum_a |\mathrm{Tr}[(M_{a|x}-F_{a|x})\rho_A]|\|\rho_B\|_1\\
&=\min_{\mathcal{F}\in F}\frac{1}{2}\sum_x p(x)\max_{\rho_A}\sum_a |\mathrm{Tr}[(M_{a|x}-F_{a|x})\rho_A]|\\
&=\mathrm{R}_{\ell_1}(\mathcal{M}_\mathbf{p}),
\end{aligned}$$

where we used in the first line that we maximize only over the set of product states. In the second line we used the definition of the partial trace and finally, we used that states ρ_B are normalized in the 1−norm and identified the last line with the definition of the induced ℓ_1-distance quantifier $\mathrm{R}_{\ell_1}(\mathcal{M}_\mathbf{p})$.

Appendix H: Dichotomic measurements

Here, we show an additional property of $\mathrm{R}_\diamond(\mathcal{M}_\mathbf{p})$ that can be useful for the case where we consider measurement assemblages $\mathcal{M}$ with only two outcomes for each setting x. We show that in this special case, the diamond distance quantifier $\mathrm{R}_\diamond(\mathcal{M}_\mathbf{p})$ is equivalent to $\mathrm{R}_\infty(\mathcal{M}_\mathbf{p})$ and $\mathrm{R}_{\ell_1}(\mathcal{M}_\mathbf{p})$. Consider the WMAs $\{M_{1|x}, \mathbb{1}-M_{1|x}\}_x$ and $\{F_{1|x}, \mathbb{1}-F_{1|x}\}_x$. Remember that we already showed previously that $\mathrm{R}_\diamond(\mathcal{M}_\mathbf{p}) \leq \mathrm{R}_\infty(\mathcal{M}_\mathbf{p})$, so we only need to show that in the case of dichotomic measurements it also holds $\mathrm{R}_\diamond(\mathcal{M}_\mathbf{p}) \geq \mathrm{R}_\infty(\mathcal{M}_\mathbf{p}) = \mathrm{R}_{\ell_1}(\mathcal{M}_\mathbf{p})$. This follows directly via

$$\begin{aligned}
&\mathrm{R}_\diamond(\mathcal{M}_\mathbf{p}) \geq \mathrm{R}_{\ell_1}(\mathcal{M}_\mathbf{p})\\
&= 2\frac{1}{2}\min_{\mathcal{F}\in\mathscr{F}}\sum_x p(x)\max_{\rho_A}|\mathrm{Tr}[(M_{a|x}-F_{a|x})\rho_A]| = \mathrm{R}_\infty(\mathcal{M}_\mathbf{p}),
\end{aligned} \tag{H1}$$

where we used the bound $\mathrm{R}_\diamond(\mathcal{M}_\mathbf{p}) \geq \mathrm{R}_{\ell_1}(\mathcal{M}_\mathbf{p})$ and the fact that for dichotomic measurements both outcomes contribute equally towards $\mathrm{R}_{\ell_1}(\mathcal{M}_\mathbf{q})$. Finally, we used that this holds also true for $\mathrm{R}_\infty(\mathcal{M}_\mathbf{p})$. Alternatively, it is also enough to see that the same ρ_A is optimal for both outcomes, which leads to the conclusion that $\mathrm{R}_\infty(\mathcal{M}_\mathbf{p}) = \mathrm{R}_{\ell_1}(\mathcal{M}_\mathbf{p})$. Note that the above result shows that entanglement does not offer an advantage in distinguishing two measure-and-prepare channels for dichotomic measurements by means of the diamond norm.

Appendix I: Optimal input distribution

Here, we give an example where an optimization over the input distribution $\mathbf{p} = \{p(x)\}$ for the settings x is relevant to optimize the available resources. In particular, we show that for the incompatibility $\mathrm{I}_\diamond(\mathcal{M}_\mathbf{p})$ (see Eq. (21)) of a measurement assemblage with only two settings, the optimal incompatibility is not always achieved for a uniform distribution $p(1) = p(2) = \frac{1}{2}$. The idea is to introduce noise in only one of the measurement settings, here for $x = 2$. Let us consider an MUB measurement assemblage $\mathcal{N}$ containing $m = 2$ POVMs, constructed in the same way as the assemblages considered in Table II. From the MUB measurement assemblage $\mathcal{N}$ we obtain the measurement assemblage $\mathcal{M}$ via

$$\begin{aligned}
M_{a|1} &= N_{a|1}\ \forall\ a, \\
M_{a|2} &= \mu N_{a|2} + (1-\mu)\mathrm{Tr}[N_{a|2}]\frac{\mathbb{1}}{d}\ \forall\ a,
\end{aligned} \tag{I1}$$

where $\mu \in [0,1]$ is a depolarizing noise parameter for the second measurement. In the following, we analyze how to choose the probability distribution $\mathbf{p} = \{p(x)\}$ such that the incompatibility $\mathrm{I}_\diamond(\mathcal{M}_\mathbf{p})$ is maximized for the given assemblage $\mathcal{M}$. As mentioned in section IV, the SDP (37) can be rewritten such that it includes a maximization over the input distribution $\mathbf{p} = \{p(x)\}$. We illustrate our results in Figure 4 for the optimal setting probabilities $p(x)$ of the assemblage $\mathcal{M}$ in dimension d with noise parameter μ.

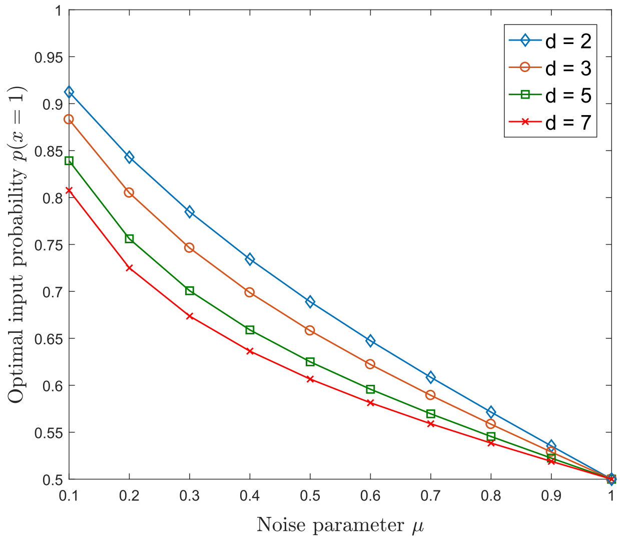

Figure 4. The optimal input probability $p(x = 1)$ for the first (noise free) measurement setting of the assemblage $\mathcal{M}$ in Eq. (I1) depending on the dimension d and the noise parameter μ. The plot shows the optimal probability $p(x = 1)$ which maximizes the incompatibility $\mathrm{I}_\diamond(\mathcal{M}_\mathbf{p})$. It can be seen that a uniform distribution is only optimal in the absence of noise (i.e. $\mu = 1$). Especially for high noise regimes (e.g. $\mu = 0.1$) a strong bias towards the noise free measurement can be seen. However, this strong bias decreases with increasing dimension d.

As one can see, even for only two measurements, strong biases towards one setting can be necessary in order to

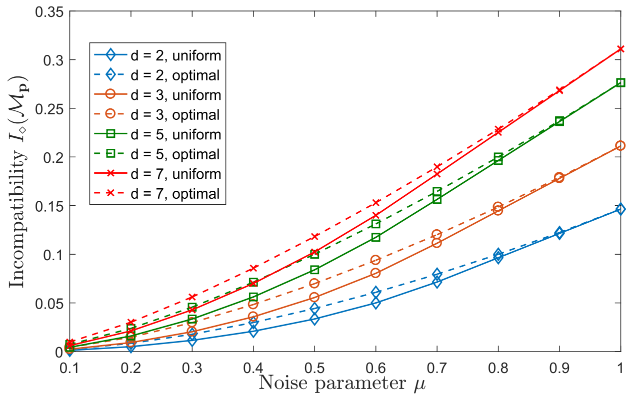

Figure 5. Comparison of the incompatibility $\mathrm{I}_\diamond(\mathcal{M}_\mathbf{p})$ between the optimal input distribution (dashed lines) and the uniform distribution (solid lines) depending on the noise parameter μ for a given dimension d. It can be seen that the optimized input distribution outperforms the uniform distribution for the measurement assemblage described in Eq. (I1). For low noise regime (μ close to 1) the solid and the dashed lines approach each other, as the uniform distribution is optimal for $\mu = 1$.

maximize the incompatibility $\mathrm{I}_\diamond(\mathcal{M}_\mathbf{p})$. We want to remark that except for the noise free case, i.e. $\mu = 1$, the optimized input distribution leads to a strictly larger incompatibility than with a uniform distribution. Note that in this particular example, the advantage is weak as can be seen in Figure 5. However, an optimization over the distribution $\mathbf{p}$ can lead to a strong increase in incompatibility for $m \geq 3$, by essentially neglecting weakly incompatibly subsets of measurements. On a qualitative basis, this effect can be explained in terms of the informativeness $\mathrm{IF}_\diamond(\mathcal{M}_\mathbf{p})$ (see Eq. (17)). For large noise (e.g. $\mu = 0.1$) the distribution is strongly biased towards the noise-free, hence more informative measurement.

ACRONYMS

CGLMP Collins-Gisin-Linden-Massar-Popescu . . . 7

CPT completely positive and trace preserving . 2

JM jointly measurable 5

LHS local hidden-state model 5

LHV local hidden-variable model 6

LOCC local operations and classical communication 3

MUB mutually unbiased bases 2

POVM positive operator valued measure 3

QRT quantum resource theory 1

SDP semidefinite program 2

UI uninformative 5

WMA weighted measurement assemblage 3

[1] A. Einstein, B. Podolsky, and N. Rosen, Can quantum-mechanical description of physical reality be considered complete?, Phys. Rev. **47**, 777 (1935).

[2] J. S. Bell, On the Einstein Podolsky Rosen paradox, Physics Physique Fizika **1**, 195 (1964).

[3] H. P. Robertson, The uncertainty principle, Phys. Rev. **34**, 163 (1929).

[4] J. Preskill, Quantum computing 40 years later (2021), arXiv:2106.10522.

[5] C. L. Degen, F. Reinhard, and P. Cappellaro, Quantum sensing, Rev. Mod. Phys. **89**, 035002 (2017).

[6] S. Pirandola, U. L. Andersen, L. Banchi, M. Berta, D. Bunandar, R. Colbeck, D. Englund, T. Gehring, C. Lupo, C. Ottaviani, J. L. Pereira, M. Razavi, J. S. Shaari, M. Tomamichel, V. C. Usenko, G. Vallone, P. Villoresi, and P. Wallden, Advances in quantum cryptography, Adv. Opt. Photon. **12**, 1012 (2020).

[7] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, Quantum entanglement, Rev. Mod. Phys. **81**, 865 (2009).

[8] R. Gallego and L. Aolita, Resource Theory of Steering, Phys. Rev. X **5**, 041008 (2015).

[9] D. Cavalcanti and P. Skrzypczyk, Quantum steering: a review with focus on semidefinite programming, Reports on Progress in Physics **80**, 024001 (2016).

[10] R. Uola, A. C. S. Costa, H. C. Nguyen, and O. Gühne, Quantum steering, Rev. Mod. Phys. **92**, 015001 (2020).

[11] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner, Bell nonlocality, Rev. Mod. Phys. **86**, 419 (2014).

[12] J. I. de Vicente, On nonlocality as a resource theory and nonlocality measures, Journal of Physics A: Mathematical and Theoretical **47**, 424017 (2014).

[13] D. Cavalcanti and P. Skrzypczyk, Quantitative relations between measurement incompatibility, quantum steering, and nonlocality, Phys. Rev. A **93**, 052112 (2016).

[14] L. Tendick, H. Kampermann, and D. Bruß, Quantifying necessary quantum resources for nonlocality, Phys. Rev. Research **4**, L012002 (2022).

[15] A. Streltsov, H. Kampermann, S. Wölk, M. Gessner, and D. Bruß, Maximal coherence and the resource theory of purity, New Journal of Physics **20**, 053058 (2018).

[16] A. Streltsov, G. Adesso, and M. B. Plenio, Colloquium: Quantum coherence as a resource, Rev. Mod. Phys. **89**, 041003 (2017).

[17] A. Bera, T. Das, D. Sadhukhan, S. S. Roy, A. Sen(De),

and U. Sen, Quantum discord and its allies: a review of recent progress, Reports on Progress in Physics **81**, 024001 (2017).

[18] K.-D. Wu, T. V. Kondra, S. Rana, C. M. Scandolo, G.-Y. Xiang, C.-F. Li, G.-C. Guo, and A. Streltsov, Operational Resource Theory of Imaginarity, Phys. Rev. Lett. **126**, 090401 (2021).

[19] O. Gühne, E. Haapasalo, T. Kraft, J.-P. Pellonpää, and R. Uola, Incompatible measurements in quantum information science (2021), arXiv:2112.06784.

[20] M. Oszmaniec, L. Guerini, P. Wittek, and A. Acín, Simulating Positive-Operator-Valued Measures with Projective Measurements, Phys. Rev. Lett. **119**, 190501 (2017).

[21] L. Guerini, J. Bavaresco, M. T. Cunha, and A. Acín, Operational framework for quantum measurement simulability, Journal of Mathematical Physics **58**, 092102 (2017).

[22] P. Skrzypczyk and N. Linden, Robustness of Measurement, Discrimination Games, and Accessible Information, Phys. Rev. Lett. **122**, 140403 (2019).

[23] K. Baek, A. Sohbi, J. Lee, J. Kim, and H. Nha, Quantifying coherence of quantum measurements, New Journal of Physics **22**, 093019 (2020).

[24] E. Chitambar and G. Gour, Quantum resource theories, Rev. Mod. Phys. **91**, 025001 (2019).

[25] R. Uola, T. Kraft, J. Shang, X.-D. Yu, and O. Gühne, Quantifying Quantum Resources with Conic Programming, Phys. Rev. Lett. **122**, 130404 (2019).

[26] S. Designolle, R. Uola, K. Luoma, and N. Brunner, Set Coherence: Basis-Independent Quantification of Quantum Coherence, Phys. Rev. Lett. **126**, 220404 (2021).

[27] R. Takagi and B. Regula, General Resource Theories in Quantum Mechanics and Beyond: Operational Characterization via Discrimination Tasks, Phys. Rev. X **9**, 031053 (2019).

[28] A. F. Ducuara and P. Skrzypczyk, Operational Interpretation of Weight-Based Resource Quantifiers in Convex Quantum Resource Theories, Phys. Rev. Lett. **125**, 110401 (2020).

[29] R. Uola, C. Budroni, O. Gühne, and J.-P. Pellonpää, One-to-One Mapping between Steering and Joint Measurability Problems, Phys. Rev. Lett. **115**, 230402 (2015).

[30] G. Vidal and R. Tarrach, Robustness of entanglement, Phys. Rev. A **59**, 141 (1999).

[31] M. Steiner, Generalized robustness of entanglement, Phys. Rev. A **67**, 054305 (2003).

[32] M. Piani and J. Watrous, Necessary and Sufficient Quantum Information Characterization of Einstein-Podolsky-Rosen Steering, Phys. Rev. Lett. **114**, 060404 (2015).

[33] T. Heinosaari, J. Kiukas, and D. Reitzner, Noise robustness of the incompatibility of quantum measurements, Phys. Rev. A **92**, 022115 (2015).

[34] S. Designolle, M. Farkas, and J. Kaniewski, Incompatibility robustness of quantum measurements: a unified framework, New Journal of Physics **21**, 113053 (2019).

[35] A. C. Elitzur, S. Popescu, and D. Rohrlich, Quantum nonlocality for each pair in an ensemble, Physics Letters A **162**, 25 (1992).

[36] M. Lewenstein and A. Sanpera, Separability and Entanglement of Composite Quantum Systems, Phys. Rev. Lett. **80**, 2261 (1998).

[37] P. Skrzypczyk, M. Navascués, and D. Cavalcanti, Quantifying Einstein-Podolsky-Rosen Steering, Phys. Rev. Lett. **112**, 180404 (2014).

[38] T. Baumgratz, M. Cramer, and M. B. Plenio, Quantifying Coherence, Phys. Rev. Lett. **113**, 140401 (2014).

[39] V. Vedral, M. B. Plenio, M. A. Rippin, and P. L. Knight, Quantifying Entanglement, Phys. Rev. Lett. **78**, 2275 (1997).

[40] T.-C. Wei and P. M. Goldbart, Geometric measure of entanglement and applications to bipartite and multipartite quantum states, Phys. Rev. A **68**, 042307 (2003).

[41] Y. Liu and X. Yuan, Operational resource theory of quantum channels, Phys. Rev. Research **2**, 012035 (2020).

[42] B. Dakić, V. Vedral, and C. Brukner, Necessary and Sufficient Condition for Nonzero Quantum Discord, Phys. Rev. Lett. **105**, 190502 (2010).

[43] Z. Puchała, L. Pawela, A. Krawiec, and R. Kukulski, Strategies for optimal single-shot discrimination of quantum measurements, Phys. Rev. A **98**, 042103 (2018).

[44] M. Sedlák and M. Ziman, Optimal single-shot strategies for discrimination of quantum measurements, Phys. Rev. A **90**, 052312 (2014).

[45] A. Kitaev, A. Shen, and M. Vyalyi, *Classical and Quantum Computation* (American Mathematical Society, 2002).

[46] H.-Y. Ku, S.-L. Chen, C. Budroni, A. Miranowicz, Y.-N. Chen, and F. Nori, Einstein-Podolsky-Rosen steering: Its geometric quantification and witness, Phys. Rev. A **97**, 022338 (2018).

[47] S. G. A. Brito, B. Amaral, and R. Chaves, Quantifying Bell nonlocality with the trace distance, Phys. Rev. A **97**, 022111 (2018).

[48] T. Durt, B. Englert, I. Bengstsson, and K. Życzkowski, On Mutually Unbiased Bases, International Journal of Quantum Information **08**, 535 (2010).

[49] E. Kaur, X. Wang, and M. M. Wilde, Conditional mutual information and quantum steering, Phys. Rev. A **96**, 022332 (2017).

[50] R. Gallego, L. E. Würflinger, A. Acín, and M. Navascués, Operational Framework for Nonlocality, Phys. Rev. Lett. **109**, 070401 (2012).

[51] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information: 10th Anniversary Edition* (Cambridge University Press, 2010).

[52] J. Watrous, *The Theory of Quantum Information* (Cambridge University Press, 2018).

[53] T. Heinosaari, T. Miyadera, and M. Ziman, An invitation to quantum incompatibility, Journal of Physics A: Mathematical and Theoretical **49**, 123001 (2016).

[54] S. Designolle, P. Skrzypczyk, F. Fröwis, and N. Brunner, Quantifying Measurement Incompatibility of Mutually Unbiased Bases, Phys. Rev. Lett. **122**, 050402 (2019).

[55] P. Skrzypczyk, I. Šupić, and D. Cavalcanti, All Sets of Incompatible Measurements give an Advantage in Quantum State Discrimination, Phys. Rev. Lett. **122**, 130403 (2019).

[56] R. Cleve, P. Hoyer, B. Toner, and J. Watrous, Consequences and limits of nonlocal strategies, in *Proceedings. 19th IEEE Annual Conference on Computational Complexity, 2004.* (IEEE, 2004).

[57] M. Araújo, F. Hirsch, and M. T. Quintino, Bell nonlocality with a single shot, Quantum **4**, 353 (2020).

[58] T. Heinosaari, J. Kiukas, D. Reitzner, and J. Schultz, Incompatibility breaking quantum channels, Journal of Physics A: Mathematical and Theoretical **48**, 435301 (2015).

[59] D. Collins, N. Gisin, N. Linden, S. Massar, and

S. Popescu, Bell Inequalities for Arbitrarily High-Dimensional Systems, Phys. Rev. Lett. **88**, 040404 (2002).

[60] J. Barrett, A. Kent, and S. Pironio, Maximally Nonlocal and Monogamous Quantum Correlations, Phys. Rev. Lett. **97**, 170409 (2006).

[61] J. Watrous, Semidefinite programs for completely bounded norms (2009), arXiv:0901.4709.

[62] S. Boyd and L. Vandenberghe, *Convex Optimization* (Cambridge University Press, 2004).

[63] M. Grant and S. Boyd, CVX: Matlab software for disciplined convex programming, version 2.1, `http://cvxr.com/cvx` (2014).

[64] M. Grant and S. Boyd, Graph implementations for nonsmooth convex programs, in *Recent Advances in Learning and Control*, Lecture Notes in Control and Information Sciences, edited by V. Blondel, S. Boyd, and H. Kimura (Springer-Verlag Limited, 2008) pp. 95–110, `http://stanford.edu/~boyd/graph_dcp.html`.

[65] K. Toh, M. Todd, and R. Tutuncu, Sdpt3 — a Matlab software package for semidefinite programming, Optimization Methods and Software (1999).

[66] M. ApS, *The MOSEK optimization toolbox for MATLAB manual. Version 9.0.* (2019).

[67] D. Popovici and Z. Sebestyén, Norm estimations for finite sums of positive operators, Journal of Operator Theory **56**, 3 (2006).

[68] J. Bavaresco, M. T. Quintino, L. Guerini, T. O. Maciel, D. Cavalcanti, and M. T. Cunha, Most incompatible measurements for robust steering tests, Phys. Rev. A **96**, 022110 (2017).

[69] S. Bandyopadhyay, P. O. Boykin, V. Roychowdhury, and F. Vatan, A new proof for the existence of mutually unbiased bases, Algorithmica **34**, 512 (2002).

[70] W. K. Wootters and B. D. Fields, Optimal state-determination by mutually unbiased measurements, Annals of Physics **191**, 363 (1989).

[71] M. Oszmaniec and T. Biswas, Operational relevance of resource theories of quantum measurements, Quantum **3**, 133 (2019).

[72] J. Kiukas, D. McNulty, and J.-P. Pellonpää, Amount of quantum coherence needed for measurement incompatibility, Phys. Rev. A **105**, 012205 (2022).

[73] H.-J. Kim and S. Lee, Relation between quantum coherence and quantum entanglement in quantum measurements (2022), arXiv:2202.12656.

[74] I. Šupić and J. Bowles, Self-testing of quantum systems: a review, Quantum **4**, 337 (2020).

[75] A. Luis and L. L. Sánchez-Soto, Complete Characterization of Arbitrary Quantum Measurement Processes, Phys. Rev. Lett. **83**, 3573 (1999).

[76] T. Theurer, D. Egloff, L. Zhang, and M. B. Plenio, Quantifying Operations with an Application to Coherence, Phys. Rev. Lett. **122**, 190405 (2019).

D

Paper: Distribution of quantum incompatibility across subsets of measurements

Title:	Distribution of quantum incompatibility across subsets of measurements
Authors:	Lucas Tendick, Hermann Kampermann, and Dagmar Bruß
Journal:	Currently on preprint server, submission to Physical Review Letters planed
Impact factor:	No Impact factor/ Physical Review Letters: 9.185 (2021)
Date of submission:	20 January 2023
Publication status:	Submitted to preprint server arXiv
Contribution by LT:	First author (input approx. 90%)

This work corresponds to the paper [66]. The summary of the results can be found in Section 4.4.

I initiated studying the incompatibility that can be gained by adding further measurements to a measurement scheme. Following our previous work [65] (Paper C), I realized that our quantifier might be well-suited to tackle this problem. After some initial calculations, I discussed the main idea of the work with HK and DB. Following the discussion, I obtained all of the analytical and numerical results. In the meantime, I further discussed with HK and started to write down the results. I then discussed these results with DB and started to write the manuscript. DB and HK proofread the manuscript and, in various steps, helped me to improve it through their comments.

Distribution of quantum incompatibility across subsets of measurements

Lucas Tendick,* Hermann Kampermann, and Dagmar Bruß
Institute for Theoretical Physics, Heinrich Heine University Düsseldorf, D-40225 Düsseldorf, Germany

Incompatible, i.e. non-jointly measurable quantum measurements are a necessary resource for many information processing tasks. It is known that increasing the number of distinct measurements usually enhances the incompatibility of a measurement scheme. However, it is generally unclear how large this enhancement is and on what it depends. Here, we show that the incompatibility which is gained via additional measurements is upper and lower bounded by certain functions of the incompatibility of subsets of the available measurements. We prove the tightness of some of our bounds by providing explicit examples based on mutually unbiased bases. Finally, we discuss the consequences of our results for the nonlocality that can be gained by enlarging the number of measurements in a Bell experiment.

arXiv:2301.08670v1 [quant-ph] 20 Jan 2023

The incompatibility of quantum measurements, i.e., the impossibility of measuring specific observable quantities simultaneously, is one of quantum physics' most prominent and striking properties. First discussed by Heisenberg [1] and Robertson [2], this counterintuitive feature was initially thought of as a puzzling curiosity that represents a drawback for potential applications. Nowadays, measurement incompatibility [3, 4] is understood as a fundamental property of nature that lies at the heart of many quantum information processing tasks, such as quantum state discrimination [5–10], quantum cryptography [11, 12], and quantum random access codes [13, 14]. Even more importantly, incompatible measurements are a crucial requirement for quantum phenomena such as quantum contextuality [15], EPR-steering [16, 17], and Bell nonlocality [18].

Its fundamental importance necessitates gaining a deep understanding of measurement incompatibility from a qualitative and quantitative perspective. By its very definition, measurement incompatibility arises when at least $m \geq 2$ measurements are considered that cannot be measured jointly by performing a single measurement instead. Generally, adding more measurements to a measurement scheme may allow for more incompatibility, hence increasing advantages in certain applications.

However, it is generally unclear how much incompatibility can be gained from adding further measurements to an existing measurement scheme and on what this potential gain depends. Similarly, it is unclear how the incompatibility of measurement pairs contributes towards the total incompatibility of the whole set. Answering these questions is crucial to understanding specific protocols' power over others, such as protocols involving different numbers of mutually unbiased bases (MUB) [19]. For example, in quantum key distribution, the six-state protocol [20] provides an advantage over the BB84 protocol [11] by using three instead of two qubit bases. While it is known [21] that the different incompatibility structures (e.g., genuine triplewise and pairwise incompatibility) arising for $m \geq 3$ measurements set different limitations on the violation of Bell inequalities, so far no systematical way to quantify the gained advantage is known. Incompatibility structures beyond two measurements have also been studied in [22–24] and measurement incompatibility was shown to be only necessary but not sufficient for nonlocality beyond the case of two dichotomic measurements [25–27].

In this work, we take a step toward answering these questions by showing how an assemblage's incompatibility depends quantitatively on its subsets' incompatibilities. More specifically, we show how the potential gain of adding measurements to an existing measurement scheme is bounded by the incompatibility of the parent positive operator valued measures (POVMs) that approximate the respective subsets of measurements by a single measurement.

Our results reveal the polygamous nature of measurement incompatibility in the sense that an assemblage of more than two measurements can only be highly incompatible if all its subsets and the respective parent POVMs of the closest jointly measurable approximation of these subsets are highly non-jointly measurable. Our considerations lead to a new notion of measurement incompatibility that accounts only for a specific measurement's incompatibility contribution. We prove the relevance of our bounds on the incompatibility that can maximally be gained by showing that they are tight for particular measurement assemblages based on MUB. Finally, we show that our results have direct consequences for steering and Bell nonlocality and discuss future applications of our results and methods.

Preliminaries.—We describe a quantum measurement most generally by a POVM, i.e., a set $\{M_a\}$ of operators $0 \leq M_a \leq \mathbb{1}$ such that $\sum_a M_a = \mathbb{1}$. Given a state ρ, the probability of obtaining outcome a is given by the Born rule $p(a) = \mathrm{Tr}[M_a \rho]$. A *measurement assemblage* is a collection of different POVMs with operators $M_{a|x}$, where x denotes the particular measurement. We write an assemblage $\mathcal{M}_{(1,2,\cdots,m)} = (\mathcal{M}_1, \mathcal{M}_2, \cdots, \mathcal{M}_m)$ of m measurements as an ordered list of POVMs, where $\mathcal{M}_x$ refers to the x-th measurement. For instance, $\mathcal{M}_{(1,2,3)} = (\mathcal{M}_1, \mathcal{M}_2, \mathcal{M}_3)$ refers to an assemblage with three (different) measurements and $\mathcal{M}_{(1,2,2)} = (\mathcal{M}_1, \mathcal{M}_2, \mathcal{M}_2)$

* lucas.tendick@hhu.de

denotes an assemblage where the second and the third POVM are equal.

An assemblage is called *jointly measurable* if it can be simulated by a single *parent POVM* $\{G_\lambda\}$ and conditional probabilities $p(a|x,\lambda)$ such that

$$M_{a|x} = \sum_\lambda p(a|x,\lambda) G_\lambda \;\forall\; a, x, \tag{1}$$

and it is called *incompatible* otherwise. Various functions can quantify measurement incompatibility [28–30]. The most suitable incompatibility quantifier for our purposes is the recently introduced *diamond distance quantifier* [31], given by

$$\mathrm{I}_\diamond(\mathcal{M}^{\mathbf{p}}) = \min_{\mathcal{F}\in \mathrm{JM}} \sum_x p(x)\, \mathrm{D}_\diamond(\Lambda_{\mathcal{M}_x}, \Lambda_{\mathcal{F}_x}), \tag{2}$$

where JM denotes the set of jointly measurable assemblages, $\Lambda_{\mathcal{M}_x} = \sum_a \mathrm{Tr}[M_{a|x}\rho]|a\rangle\langle a|$ is the *measure-and-prepare channel* associated to the measurement $\mathcal{M}_x$, and $\mathrm{D}_\diamond(\Lambda_1,\Lambda_2) = \max_{\rho\in\mathcal{S}(\mathcal{H}\otimes\mathcal{H})} \frac{1}{2}\|((\Lambda_1-\Lambda_2)\otimes \mathbb{1}_d)\rho\|_1$ is the *diamond distance* [32] between two channels Λ_1 and Λ_2, with the trace norm $\|X\|_1 = \mathrm{Tr}[\sqrt{X^\dagger X}]$. Furthermore, $\mathcal{M}^{\mathbf{p}} = (\mathcal{M},\mathbf{p})$ denotes a *weighted measurement assemblage*, where $\mathbf{p}$ contains the probabilities $p(x)$ with which measurement x is performed. Note that $\mathrm{I}_\diamond(\mathcal{M}^{\mathbf{p}})$ is induced by the general distance $\mathrm{D}_\diamond(\mathcal{M}^{\mathbf{p}},\mathcal{N}^{\mathbf{p}}) := \sum_x p(x)\,\mathrm{D}_\diamond(\Lambda_{\mathcal{M}_x},\Lambda_{\mathcal{N}_x})$ between two assemblages $\mathcal{M}^{\mathbf{p}}$ and $\mathcal{N}^{\mathbf{p}}$.

We denote by $\mathcal{M}^{\#}_{(1,2,\cdots,m)}$ the closest jointly measurable assemblage with respect to $\mathcal{M}_{(1,2,\ldots,m)}$, i.e., the *argmin* on the RHS in Eq. (2). While $\mathcal{M}^{\#}_{(1,2,\cdots,m)}$ and its underlying parent POVM are generally not unique [23, 33], all the results derived in this work hold for any valid choice, as we do not assume uniqueness. If we only approximate a subset of $n<m$ measurements of $\mathcal{M}_{(1,2,\ldots,m)}$ by jointly measurable ones, for instance the first n settings, while keeping the remaining measurements unchanged, we write $\mathcal{M}^{\#(1,2,\ldots,n)}_{(1,2,\cdots,m)}$.

The diamond distance quantifier $\mathrm{I}_\diamond(\mathcal{M}^{\mathbf{p}})$ is particularly well-suited for our purposes, as it is not only monotonous under the application of quantum channels and classical simulations but it also inherits all properties of a distance (in particular the triangle inequality of $\mathrm{D}_\diamond(\mathcal{M}^{\mathbf{p}},\mathcal{N}^{\mathbf{p}})$), and it is written in terms of a convex combination of the individual measurement's distances.

For pedagogical reasons, we focus in the main text on the scenario $2\to 3$, i.e., we consider an assemblage of $m=2$ measurements that is promoted to one with $m'=3$ settings. Furthermore, we set $p(x)$ to be uniformly distributed and simply use the symbol $\mathcal{M}$ for the weighted assemblage in this case. Moreover, we use the reoccurring example of measurements corresponding to the three Pauli observables X,Y,Z, the simplest case of measurements based on three MUB. We refer to the Supplemental Material (SM) [34] for all proofs, more background information, and generalizations to an arbitrary number of measurements and general probability distributions.

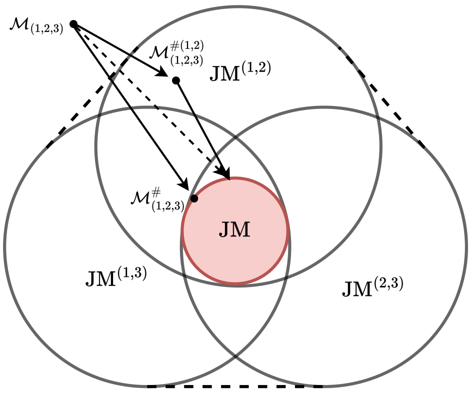

Figure 1. Different structures of incompatibility for three measurements, see also Ref. [21]. The sets $\mathrm{JM}^{(s,t)}$ contain assemblages of measurements where the pairs (s,t) are compatible. Their intersection $\mathrm{JM}^{\mathrm{pair}} := \mathrm{JM}^{(1,2)}\cap\mathrm{JM}^{(1,3)}\cap\mathrm{JM}^{(2,3)}$ contains all pairwise compatible assemblages, with the set JM of all jointly measurable assemblages as a proper subset. Assemblages not contained in the convex hull $\mathrm{JM}^{\mathrm{conv}} := \mathrm{Conv}(\mathrm{JM}^{(1,2)},\mathrm{JM}^{(1,3)},\mathrm{JM}^{(2,3)})$ of the sets $\mathrm{JM}^{(s,t)}$, i.e., those that cannot be written as a convex combination of assemblages from the sets $\mathrm{JM}^{(1,2)}$, $\mathrm{JM}^{(1,3)}$, and $\mathrm{JM}^{(2,3)}$ are genuinely triplewise incompatible. The incompatibility of $\mathcal{M}_{(1,2,3)}$ is given by the distance to its closest jointly measurable approximation $\mathcal{M}^{\#}_{(1,2,3)}$. This distance can be upper bounded using the triangle inequality via the assemblage $\mathcal{M}^{\#(1,2)}_{(1,2,3)}$.

Adding a third measurement $\mathcal{M}_3$ to the assemblage $\mathcal{M}_{(1,2)} = (\mathcal{M}_1,\mathcal{M}_2)$ is mathematically described by the concatenation of ordered lists, using the symbol $+\!\!+$, i.e., we write

$$\mathcal{M}_{(1,2,3)} = \mathcal{M}_{(1,2)} +\!\!+ \mathcal{M}_3 = (\mathcal{M}_1,\mathcal{M}_2,\mathcal{M}_3). \tag{3}$$

Using the concatenation of ordered lists, we formally define $\mathcal{M}^{\#(1,2)}_{(1,2,3)}$ such that

$$\mathcal{M}^{\#(1,2)}_{(1,2,3)} := \mathcal{M}^{\#}_{(1,2)} +\!\!+ \mathcal{M}_3. \tag{4}$$

Three measurements allow for incompatibility structures [21–24] beyond Eq. (1). We define the sets $\mathrm{JM}^{(s,t)}$ with $s\neq t\in\{1,2,3\}$ as those containing assemblages in which the measurements s and t are jointly measurable. This allows us to define *pairwise* and *genuinely triplewise incompatible* assemblages [21] as those that are *not* contained in the intersection and the convex hull of the sets $\mathrm{JM}^{(s,t)}$, respectively. See also Figure 1 for a graphical representation and more details.

Incompatibility gain.—We investigate the *incompatibility gain* obtained from adding measurements to an already available assemblage. That is, for an assemblage $\mathcal{M}_{(1,2,3)}$ defined via Eq. (3) we want to quantify the gain

$$\Delta\mathrm{I}_{(1,2)\to(1,2,3)} := \mathrm{I}_\diamond(\mathcal{M}_{(1,2,3)}) - \mathrm{I}_\diamond(\mathcal{M}_{(1,2)}). \tag{5}$$

Note that $\Delta \mathrm{I}_{(1,2)\to(1,2,3)}$ is the difference of two quantities that can be computed via semidefinite programs (SDPs) [31], however, the purely numerical value of the gained incompatibility does only provide limited physical insights by itself. While it seems generally challenging to find an exact analytical expression for the incompatibility gain, we will derive bounds on it in the following.

Our approach relies on a two-step protocol. First, we employ a *measurement splitting*, i.e., instead of considering the incompatibility of $\mathcal{M}_{(1,2,3)}$, we consider the incompatibility $\mathrm{I}_\diamond(\mathcal{M}_{(1,2,1,3,2,3)})$. That is, each measurement of $\mathcal{M}_{(1,2,3)}$ is now split up into two equivalent ones, each occurring with a probability of $\frac{1}{6}$. Furthermore, it holds $\mathrm{I}_\diamond(\mathcal{M}_{(1,2,3)}) = \mathrm{I}_\diamond(\mathcal{M}_{(1,2,1,3,2,3)})$ since the assemblages can be converted into each other by (reversible) classical post-processing [34]. The second step involves a particular instance of the triangle inequality and uses specifically that $\mathrm{I}_\diamond(\mathcal{M})$ is defined as convex combination over the individual settings. More precisely, let

$$\mathcal{N} = \mathcal{M}^{\#}_{(1,2)} \uplus \mathcal{M}^{\#}_{(1,3)} \uplus \mathcal{M}^{\#}_{(2,3)}, \tag{6}$$

be an assemblage that contains itself three assemblages (of two measurements each) that are the closest jointly measurable approximations with respect to the individual subsets of $\mathcal{M}_{(1,2,3)}$. We point out that $\mathcal{N}$ itself can be incompatible in general. Using the triangle inequality, it follows that

$$\begin{aligned}\mathrm{I}_\diamond(\mathcal{M}_{(1,2,3)}) &= \mathrm{I}_\diamond(\mathcal{M}_{(1,2,1,3,2,3)}) \\ &\leq \mathrm{D}_\diamond(\mathcal{M}_{(1,2,1,3,2,3)},\mathcal{N}) + \mathrm{I}_\diamond(\mathcal{N}).\end{aligned} \tag{7}$$

Due to our choice of $\mathcal{N}$, the term $\mathrm{D}_\diamond(\mathcal{M}_{(1,2,1,3,2,3)},\mathcal{N})$ evaluates to the average incompatibility of the subsets, as we can split the sum over all six settings into three pairs, i.e. we obtain

$$\begin{aligned}\mathrm{I}_\diamond(\mathcal{M}_{(1,2,3)}) \leq \frac{1}{3}\big[&\mathrm{I}_\diamond(\mathcal{M}_{(1,2)}) + \mathrm{I}_\diamond(\mathcal{M}_{(1,3)}) \\ &+ \mathrm{I}_\diamond(\mathcal{M}_{(2,3)})\big] + \mathrm{I}_\diamond(\mathcal{N}).\end{aligned} \tag{8}$$

That is, the incompatibility of $\mathcal{M}_{(1,2,3)}$ is upper bounded by the average incompatibility of its two-measurement subsets plus the incompatibility $\mathrm{I}_\diamond(\mathcal{N})$ that contains the information about how incompatible the respective closest jointly measurable POVMs are with each other. Notice that $\mathrm{I}_\diamond(\mathcal{N}) \leq \mathrm{I}_\diamond(\mathcal{G})$ holds, where

$$\mathcal{G} = G(\mathcal{M}^{\#}_{(1,2)}) \uplus G(\mathcal{M}^{\#}_{(1,3)}) \uplus G(\mathcal{M}^{\#}_{(2,3)}) \tag{9}$$

is the assemblage that contains the parent POVMs of the respective subsets, as $\mathcal{N}$ is a classical post-processing of $\mathcal{G}$ [34]. This shows that the incompatibility of $\mathcal{M}_{(1,2,3)}$ is limited on two different levels through its subsets. Moreover, it reveals a type of *polygamous* behavior of incompatibility. For high incompatibility of $\mathcal{M}_{(1,2,3)}$ each of the subsets, as well as the underlying parent POVMs of the respective jointly measurable approximations, have to be highly incompatible. Coming back to the incompatibility gain, we are ready to present our first main result.

Result 1. *Let $\mathrm{I}_\diamond(\mathcal{M}_{(1,2)}) \geq \max\{\mathrm{I}_\diamond(\mathcal{M}_{(1,3)}), \mathrm{I}_\diamond(\mathcal{M}_{(2,3)})\}$. It follows that the incompatibility gain as defined in Eq.(5) is bounded such that*

$$\Delta \mathrm{I}_{(1,2)\to(1,2,3)} \leq \mathrm{I}_\diamond(\mathcal{N}) \leq \mathrm{I}_\diamond(\mathcal{G}). \tag{10}$$

This means that the potential incompatibility gain is limited by the incompatibility of the assemblage $\mathcal{N}$ in Eq. (6), i.e., the concatenation of the respective closest jointly measurable approximations of the subsets. Physically more intuitive, it is limited by the incompatibility of the assemblage that contains the respective parent POVMs. The assumption $\mathrm{I}_\diamond(\mathcal{M}_{(1,2)}) \geq \max\{\mathrm{I}_\diamond(\mathcal{M}_{(1,3)}), \mathrm{I}_\diamond(\mathcal{M}_{(2,3)})\}$ represents no loss of generality for all practical purposes, as one can simply optimize over all possible two-measurement subsets.

We point out that Result 1 allows for the definition of a single *maximally incompatible additional* measurement, in the sense that it is the measurement $\mathcal{M}_3$ that maximizes the incompatibility gain $\Delta \mathrm{I}_{(1,2)\to(1,2,3)}$ for a given assemblage $\mathcal{M}_{(1,2)}$. As an illustrative example, we consider the three projective measurements $\{\Pi_{a|x}\}$ which represent the Pauli X, Y, Z observables subjected to white noise, i.e., we analyze the incompatibility of the assemblage $\mathcal{M}^\eta_{(1,2,3)} = (\mathcal{M}^\eta_1, \mathcal{M}^\eta_2, \mathcal{M}^\eta_3)$ defined via

$$M^\eta_{a|x} = \eta \Pi_{a|x} + (1-\eta)\,\mathrm{Tr}[\Pi_{a|x}]\frac{\mathbb{1}}{2}, \tag{11}$$

where $(1-\eta)$ is the noise level. It holds in this particular case that (see Figure 2):

$$\Delta \mathrm{I}_{(1,2)\to(1,2,3)}(\eta) = \mathrm{I}_\diamond(\mathcal{N}(\eta)), \tag{12}$$

which we prove analytically in the SM [34]. For the regime $\frac{1}{\sqrt{2}} \leq \eta \leq 1$ we also show that $\mathrm{I}_\diamond(\mathcal{N}(\eta)) = \mathrm{I}_\diamond(\mathcal{M}^{1/\sqrt{2}}_{(1,2,3)})$, which means that the gained incompatibility is exactly given by the incompatibility of $\mathcal{M}^\eta_{(1,2,3)}$ at the noise threshold where it becomes *pairwise compatible*.

Our methods can also be applied to obtain lower bounds. For instance, we show [34] that $\mathrm{I}_\diamond(\mathcal{M}_{(1,2,3)})$ is bounded by the average subset incompatibility:

$$\mathrm{I}_\diamond(\mathcal{M}_{(1,2,3)}) \geq \frac{1}{3}[\mathrm{I}_\diamond(\mathcal{M}_{(1,2)}) + \mathrm{I}_\diamond(\mathcal{M}_{(1,3)}) + \mathrm{I}_\diamond(\mathcal{M}_{(2,3)})]. \tag{13}$$

In general, $\mathrm{I}_\diamond(\mathcal{M}_{(1,2,3)}) < \mathrm{I}_\diamond(\mathcal{M}_{(1,2)})$ is possible, i.e., adding a measurement to an assemblage can actually decrease the incompatibility, if we do *not* optimize over the input distribution $\mathbf{p}$. While this might seem surprising, it can be explained by the fact that using measurements of little resource can be disadvantageous.

Another way to see how the incompatibility of an assemblage $\mathcal{M}_{(1,2,3)}$ can be upper bounded in terms of the incompatibility $\mathrm{I}_\diamond(\mathcal{M}_{(1,2)})$ plus the gained incompatibility due to measurement $\mathcal{M}_3$ relies solely on the structure of incompatible measurements in the metric space of all

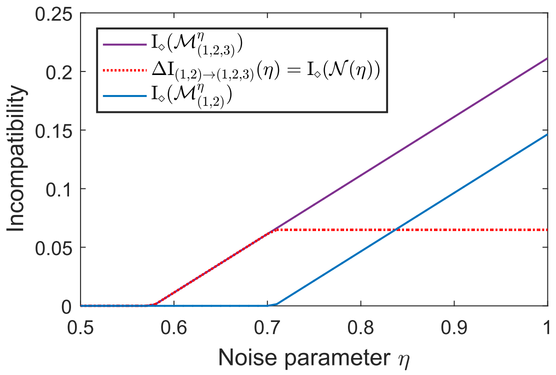

Figure 2. Incompatibility gain for adding a third Pauli measurement. The gained incompatibility is given by the red (dotted) line. In the regime where $\mathrm{I}_\diamond(\mathcal{M}_{(1,2)}) \neq 0$, the gained incompatibility remains constant. The red (dotted) curve and the blue curve add up to the violet one. The compared resources are exactly those used in the BB84 [11], respectively the six-state protocol [20], ideally with $\eta = 1$.

measurement assemblages. That means, it is sufficient to rely only on specific instances of the triangle inequality without splitting the measurements.

A new notion of incompatibility.—Consider the general assemblage $\mathcal{M}_{(1,2,3)}$ as defined in Eq. (3). Due to the triangle inequality, see also Figure 1, it holds

$$\mathrm{I}_\diamond(\mathcal{M}_{(1,2,3)}) \leq \mathrm{D}_\diamond(\mathcal{M}_{(1,2,3)}, \mathcal{N}_{(1,2,3)}) + \mathrm{I}_\diamond(\mathcal{N}_{(1,2,3)}), \tag{14}$$

for any assemblage $\mathcal{N}_{(1,2,3)}$. By choosing $\mathcal{N}_{(1,2,3)} = \mathcal{M}^{\#(1,2)}_{(1,2,3)} := \mathcal{M}^{\#}_{(1,2)} + \mathcal{M}_3$, we obtain our second main result.

Result 2. *Let $\mathcal{M}_{(1,2,3)} = \mathcal{M}_{(1,2)} + \mathcal{M}_3$ be a concatenated measurement assemblage and $\mathcal{M}^{\#}_{(1,2)}$ the closest jointly measurable approximation of $\mathcal{M}_{(1,2)}$. It holds*

$$\mathrm{I}_\diamond(\mathcal{M}_{(1,2,3)}) \leq \frac{2}{3}\,\mathrm{I}_\diamond(\mathcal{M}_{(1,2)}) + \mathrm{I}_\diamond(\mathcal{M}^{\#(1,2)}_{(1,2,3)}). \tag{15}$$

This means that the incompatibility of $\mathcal{M}_{(1,2,3)}$ is upper bounded by the incompatibility of the subset $\mathcal{M}_{(1,2)}$, weighted with the probability $p = \frac{2}{3}$, plus the incompatibility of the added measurement $\mathcal{M}_3$ with the closest jointly measurable approximation $\mathcal{M}^{\#}_{(1,2)}$ of $\mathcal{M}_{(1,2)}$. In [34] we also show that the incompatibility of $\mathcal{M}_{(1,2,3)}$ is lower bounded by

$$\mathrm{I}_\diamond(\mathcal{M}_{(1,2,3)}) \geq \tfrac{2}{3}\,\mathrm{I}_\diamond(\mathcal{M}_{(1,2)}). \tag{16}$$

The only incompatibility that contributes to $\mathrm{I}_\diamond(\mathcal{M}^{\#(1,2)}_{(1,2,3)})$ is the incompatibility of $\mathcal{M}_3$ with the assemblage $\mathcal{M}^{\#}_{(1,2)}$, which itself is jointly measurable. Therefore, this term in Eq. (15) can be understood as a new notion of incompatibility of the assemblage $\mathcal{M}_{(1,2,3)}$, where all incompatibilities apart of the contribution that comes from the presence of measurement $\mathcal{M}_3$ are omitted.

We show analytically in the SM [34] that the bound in Eq. (15) is tight for depolarized Pauli measurements (see Eq. (11)). Moreover, we show analytically that a similar bound is tight for certain measurements based on d-dimensional MUB in cases where the number of measurements m is changed such that $m = 2 \to m' = d$, $m = 2 \to m' = d+1$, and $m = d \to m' = d+1$. Namely, we prove and analyze the generalization of Eq. (15):

$$\mathrm{I}_\diamond(\mathcal{M}_{(1,2,\cdots,m)}) \leq \frac{|C|}{m}\,\mathrm{I}_\diamond(\mathcal{M}_C) + \mathrm{I}_\diamond(\mathcal{M}^{\#C}_{(1,2,\cdots,m)}), \tag{17}$$

for any assemblage $\mathcal{M}_{(1,2,\cdots,m)}$ and any subset C of measurements with cardinality $|C|$.

Incompatibility decomposition.—Looking at the results in Figure 2 leads to the question whether there exists a more general decomposition of $\mathrm{I}_\diamond(\mathcal{M}_{(1,2,3)})$ into different incompatibility structures. Indeed, since $\mathrm{I}_\diamond(\mathcal{M})$ is a *distance-based* incompatibility quantifier, our final main result follows.

Result 3. *The incompatibility of any assemblage $\mathcal{M}$ of $m = 3$ measurements is upper bounded such that*

$$\mathrm{I}_\diamond(\mathcal{M}) \leq \mathrm{I}_\diamond{}^{\mathrm{gen}}(\mathcal{M}) + \mathrm{I}_\diamond{}^{\mathrm{pair}}(\mathcal{M}) + \mathrm{I}_\diamond{}^{\mathrm{hol}}(\mathcal{M}), \tag{18}$$

where $\mathrm{I}_\diamond{}^{\mathrm{gen}}(\mathcal{M})$ is the genuine triplewise incompatibility *of $\mathcal{M}$, i.e., its minimal distance to an assemblage $\mathcal{M}^{\mathrm{conv}} \in \mathrm{JM}^{\mathrm{conv}}$. Furthermore, we define $\mathrm{I}_\diamond{}^{\mathrm{pair}}(\mathcal{M}) := \mathrm{D}_\diamond(\mathcal{M}^{\mathrm{conv}}, \mathcal{M}^{\mathrm{pair}})$ to be the* pairwise incompatibility, *where $\mathcal{M}^{\mathrm{pair}} \in \mathrm{JM}^{\mathrm{pair}}$ is the closest pairwise compatible assemblage with respect to $\mathcal{M}^{\mathrm{conv}}$. The term $\mathrm{I}_\diamond{}^{\mathrm{hol}}(\mathcal{M}) := \mathrm{I}_\diamond(\mathcal{M}^{\mathrm{pair}})$ is the* hollow incompatibility, *which implicitly depends on $\mathcal{M}$. See also Figure 1.*

We emphasize that the bound in Eq. (18) relies crucially on the distance properties of the quantifier $\mathrm{I}_\diamond(\mathcal{M})$ and *cannot* be adapted directly to robustness or weight quantifiers [28, 29]. In the SM [34] we show that the decomposition in Eq. (18) is tight for the three Pauli measurements, and give numerical indication that this is generally the case for measurements based on MUB.

Implications for steering and Bell nonlocality.—Due to the mathematical structure of our methods, they can directly be applied to quantum steering and Bell nonlocality. We describe our results regarding steering and nonlocality in more detail in the SM [34]. The analysis of the gain in nonlocal correlations in Bell experiments is particularly interesting as it seems fundamentally different from incompatibility and steering. Consider a Bell experiment where Alice performs $m_A = 3$, Bob $m_B = 2$ measurements, and we want to upper bound the nonlocality of the resulting distribution $\mathbf{q}_{(1,2,3)} = \{q(ab|xy)\}$ in terms of the nonlocality of the distributions $\mathbf{q}_{(1,2)}, \mathbf{q}_{(1,3)}$, and $\mathbf{q}_{(2,3)}$ where Alice uses only two of the measurements (analogous to Eq.(8)). For a properly chosen nonlocality distance, we obtain a corresponding bound (see SM [34]). However, this involves the average nonlocality of the distributions $\mathbf{q}_{(1,2)}, \mathbf{q}_{(1,3)}$, and $\mathbf{q}_{(2,3)}$, which, in general, is lower than the maximal obtainable nonlocality with two measurements on Alice's side. This is a consequence of the fact that there generally do not exist three different measurements for Alice, out of which any two allow for the maximal violation of a Bell inequality, while the measurements of Bob and their shared quantum state remain unchanged. We show this explicitly in the case of dichotomic measurements [34], by considering three different versions of the Clauser-Horne-Shimony-Holt (CHSH) inequality [35]. As a consequence, we observe that the nonlocality of distributions involving more than two measurements is stricter upper bounded than in the case of measurement incompatibility or steering. This observation could prove crucial in understanding why nonlocality seems to behave differently to incompatibility and steering, in the sense that using more than two measurements is not known to provide any advantages for the maximal obtainable Bell nonlocality [36, 37].

Conclusion and outlook.—In this work, we analyzed how much incompatibility can maximally be gained by adding measurements to an existing measurement scheme. We showed that this gain is upper bounded by the incompatibility of the underlying parent POVMs that approximate subsets of measurements. Our analysis shines light on a new notion of incompatibility, which decomposes the total incompatibility of an assemblage into the contributions of a single measurement that is concatenated with a jointly measurable assemblage. We proved the relevance of our bounds analytically by showing that they are tight for specific measurements based on MUB. Moreover, we showed that our methods are directly applicable to quantum steering and Bell nonlocality. For nonlocality specifically, we discovered a promising path to understand better why using more than two measurements may not provide any advantage for maximal nonlocal correlations.

Our work provides a foundation for several new directions of research. First, it would be interesting to see whether resource quantifiers such as the incompatibility robustness [29] or weight [28] can also be used to analyze how the incompatibility of an assemblage depends on its subsets. Second, our methods might prove helpful to find better bounds on the incompatibility of general assemblages, particularly assemblages based on MUB. Especially for understanding which measurements are maximally incompatible, our work provides new tools. Finally, it would be interesting to analyze the performance gain of specific cryptography [11, 20] or estimation protocols [38] with different numbers of measurements.

ACKNOWLEDGMENTS

We thank Thomas Cope, Federico Grasselli, Martin Kliesch, Nikolai Miklin, Martin Plávala, Isadora Veeren, Thomas Wagner, and Zhen-Peng Xu for helpful discussions. This research was partially supported by the EU H2020 QuantERA ERA-NET Cofund in Quantum Technologies project QuICHE, and by the Federal Ministry of Education and Research (BMBF) within the funding program "Forschung Agil - Innovative Verfahren für Quantenkommunikationsnetze" in the joint project QuKuK (grant number 16KIS1619).

[1] W. Heisenberg, Über den anschaulichen inhalt der quantentheoretischen kinematik und mechanik, Zeitschrift für Physik **43**, 172 (1927).
[2] H. P. Robertson, The uncertainty principle, Phys. Rev. **34**, 163 (1929).
[3] O. Gühne, E. Haapasalo, T. Kraft, J.-P. Pellonpää, and R. Uola, Incompatible measurements in quantum information science (2021), arXiv:2112.06784.
[4] T. Heinosaari, T. Miyadera, and M. Ziman, An invitation to quantum incompatibility, Journal of Physics A: Mathematical and Theoretical **49**, 123001 (2016).
[5] F. Buscemi, E. Chitambar, and W. Zhou, Complete resource theory of quantum incompatibility as quantum programmability, Phys. Rev. Lett. **124**, 120401 (2020).
[6] C. Carmeli, T. Heinosaari, and A. Toigo, Quantum incompatibility witnesses, Phys. Rev. Lett. **122**, 130402 (2019).
[7] R. Uola, T. Kraft, J. Shang, X.-D. Yu, and O. Gühne, Quantifying quantum resources with conic programming, Phys. Rev. Lett. **122**, 130404 (2019).
[8] M. Oszmaniec and T. Biswas, Operational relevance of resource theories of quantum measurements, Quantum **3**, 133 (2019).
[9] A. F. Ducuara and P. Skrzypczyk, Operational interpretation of weight-based resource quantifiers in convex quantum resource theories, Phys. Rev. Lett. **125**, 110401 (2020).
[10] R. Uola, T. Bullock, T. Kraft, J.-P. Pellonpää, and N. Brunner, All quantum resources provide an advantage in exclusion tasks, Phys. Rev. Lett. **125**, 110402 (2020).
[11] C. H. Bennett and G. Brassard, Quantum cryptography: Public key distribution and coin tossing, Theoretical Computer Science **560**, 7 (2014).
[12] S. Pirandola, U. L. Andersen, L. Banchi, M. Berta, D. Bunandar, R. Colbeck, D. Englund, T. Gehring, C. Lupo, C. Ottaviani, J. L. Pereira, M. Razavi, J. S. Shaari, M. Tomamichel, V. C. Usenko, G. Vallone, P. Villoresi, and P. Wallden, Advances in quantum cryptography, Adv. Opt. Photon. **12**, 1012 (2020).
[13] C. Carmeli, T. Heinosaari, and A. Toigo, Quantum random access codes and incompatibility of measurements, EPL (Europhysics Letters) **130**, 50001 (2020).
[14] H. Anwer, S. Muhammad, W. Cherifi, N. Miklin, A. Tavakoli, and M. Bourennane, Experimental charac-

terization of unsharp qubit observables and sequential measurement incompatibility via quantum random access codes, Phys. Rev. Lett. **125**, 080403 (2020).
[15] C. Budroni, A. Cabello, O. Gühne, M. Kleinmann, and J. Åke Larsson, Kochen-specker contextuality (2021), arXiv:2102.13036.
[16] R. Uola, A. C. S. Costa, H. C. Nguyen, and O. Gühne, Quantum steering, Rev. Mod. Phys. **92**, 015001 (2020).
[17] D. Cavalcanti and P. Skrzypczyk, Quantum steering: a review with focus on semidefinite programming, Reports on Progress in Physics **80**, 024001 (2016).
[18] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner, Bell nonlocality, Rev. Mod. Phys. **86**, 419 (2014).
[19] T. Durt, B. Englert, I. Bengstsson, and K. Życzkowski, On mutually unbiased bases, International Journal of Quantum Information **08**, 535 (2010).
[20] D. Bruß, Optimal eavesdropping in quantum cryptography with six states, Phys. Rev. Lett. **81**, 3018 (1998).
[21] M. T. Quintino, C. Budroni, E. Woodhead, A. Cabello, and D. Cavalcanti, Device-independent tests of structures of measurement incompatibility, Phys. Rev. Lett. **123**, 180401 (2019).
[22] R. Kunjwal, C. Heunen, and T. Fritz, Quantum realization of arbitrary joint measurability structures, Phys. Rev. A **89**, 052126 (2014).
[23] T. Heinosaari, D. Reitzner, and P. Stano, Notes on joint measurability of quantum observables, Foundations of Physics **38**, 1133 (2008).
[24] Y.-C. Liang, R. W. Spekkens, and H. M. Wiseman, Specker's parable of the overprotective seer: A road to contextuality, nonlocality and complementarity, Physics Reports **506**, 1 (2011).
[25] M. M. Wolf, D. Perez-Garcia, and C. Fernandez, Measurements incompatible in quantum theory cannot be measured jointly in any other no-signaling theory, Phys. Rev. Lett. **103**, 230402 (2009).
[26] F. Hirsch, M. T. Quintino, and N. Brunner, Quantum measurement incompatibility does not imply bell nonlocality, Phys. Rev. A **97**, 012129 (2018).
[27] E. Bene and T. Vértesi, Measurement incompatibility does not give rise to bell violation in general, New Journal of Physics **20**, 013021 (2018).
[28] M. F. Pusey, Verifying the quantumness of a channel with an untrusted device, Journal of the Optical Society of America B **32**, A56 (2015).
[29] S. Designolle, M. Farkas, and J. Kaniewski, Incompatibility robustness of quantum measurements: a unified framework, New Journal of Physics **21**, 113053 (2019).
[30] T. Cope and R. Uola, Quantifying the high-dimensionality of quantum devices (2022), arXiv:2207.05722.
[31] L. Tendick, M. Kliesch, H. Kampermann, and D. Bruß, Distance-based resource quantification for sets of quantum measurements (2022), arXiv:2205.08546.
[32] A. Kitaev, A. Shen, and M. Vyalyi, *Classical and Quantum Computation* (American Mathematical Society, 2002).
[33] L. Guerini and M. T. Cunha, Uniqueness of the joint measurement and the structure of the set of compatible quantum measurements, Journal of Mathematical Physics **59**, 042106 (2018).
[34] See Supplemental Material at [URL will be inserted by publisher] for proofs, more details, and all generalizations. This includes Refs. [39–49].
[35] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, Proposed experiment to test local hidden-variable theories, Phys. Rev. Lett. **23**, 880 (1969).
[36] M. Araújo, F. Hirsch, and M. T. Quintino, Bell nonlocality with a single shot, Quantum **4**, 353 (2020).
[37] S. G. A. Brito, B. Amaral, and R. Chaves, Quantifying Bell nonlocality with the trace distance, Phys. Rev. A **97**, 022111 (2018).
[38] D. McNulty, F. B. Maciejewski, and M. Oszmaniec, Estimating quantum hamiltonians via joint measurements of noisy non-commuting observables (2022), arXiv:2206.08912.
[39] L. Guerini, J. Bavaresco, M. T. Cunha, and A. Acín, Operational framework for quantum measurement simulability, Journal of Mathematical Physics **58**, 092102 (2017).
[40] M. Grant and S. Boyd, CVX: Matlab software for disciplined convex programming, version 2.1, `http://cvxr.com/cvx` (2014).
[41] M. Grant and S. Boyd, Graph implementations for nonsmooth convex programs, in *Recent Advances in Learning and Control*, Lecture Notes in Control and Information Sciences, edited by V. Blondel, S. Boyd, and H. Kimura (Springer-Verlag Limited, 2008) pp. 95–110, `http://stanford.edu/~boyd/graph_dcp.html`.
[42] K. Toh, M. Todd, and R. Tutuncu, Sdpt3 — a Matlab software package for semidefinite programming, Optimization Methods and Software (1999).
[43] M. ApS, *The MOSEK optimization toolbox for MATLAB manual. Version 9.0.* (2019).
[44] S. Boyd and L. Vandenberghe, *Convex Optimization* (Cambridge University Press, 2004).
[45] B. S. Cirel'son, Quantum generalizations of Bell's inequality, Letters in Mathematical Physics **4**, 93 (1980).
[46] W. K. Wootters and B. D. Fields, Optimal state-determination by mutually unbiased measurements, Annals of Physics **191**, 363 (1989).
[47] S. Designolle, P. Skrzypczyk, F. Fröwis, and N. Brunner, Quantifying measurement incompatibility of mutually unbiased bases, Phys. Rev. Lett. **122**, 050402 (2019).
[48] S. Bandyopadhyay, P. O. Boykin, V. Roychowdhury, and F. Vatan, A new proof for the existence of mutually unbiased bases, Algorithmica **34**, 512 (2002).
[49] H.-Y. Ku, S.-L. Chen, C. Budroni, A. Miranowicz, Y.-N. Chen, and F. Nori, Einstein-Podolsky-Rosen steering: Its geometric quantification and witness, Phys. Rev. A **97**, 022338 (2018).

SUPPLEMENTAL MATERIAL

In this Supplemental Material, we give detailed background information on measurement incompatibility, provide proofs for the results and statements in the main text, and discuss how to apply our results to steering and nonlocality. Furthermore, we show how to generalize our results to general sets of m measurements and weighted measurement assemblages $\mathcal{M}^{\mathbf{p}} = (\mathcal{M}, \mathbf{p})$ with general probability distributions $\mathbf{p}$.

I. BACKGROUND INFORMATION ON INCOMPATIBILITY

Here, we give detailed background information on the important properties of the diamond distance quantifier $\mathrm{I}_\diamond(\mathcal{M}^{\mathbf{p}})$ defined in Eq. (2). To provide a relatively self contained overview in this Supplemental Material, we also repeat the relevant definitions from the main text. An assemblage $\mathcal{M}_{(1,2,\cdots,m)} = (\mathcal{M}_1, \mathcal{M}_2, \cdots, \mathcal{M}_m)$ of m measurements with outcomes a and settings x is called *jointly measurable* if it can be simulated by a single *parent POVM* $\{G_\lambda\}$ and conditional probabilities $p(a|x,\lambda)$ such that

$$M_{a|x} = \sum_\lambda p(a|x,\lambda) G_\lambda \;\forall\; a, x, \tag{19}$$

and it is called *incompatible* otherwise. Note that the probabilities $p(a|x,\lambda)$ can always be identified with deterministic response functions $v(a|x,\lambda)$ since the randomness in $p(a|x,\lambda)$ can be shifted to the parent POVM by appropriately redefining the G_λ. Let us denote by JM the set of all jointly measurable assemblages. For more than two measurements, there exist different sub-structures of incompatibility. Focusing on the case of three measurements, we define the sets $\mathrm{JM}^{(s,t)}$ with $s,t \in \{1,2,3\}$ such that $s \neq t$ as the sets containing assemblages in which the measurement s and t are jointly measurable. Their intersection $\mathrm{JM}^{\mathrm{pair}} := \mathrm{JM}^{(1,2)} \cap \mathrm{JM}^{(1,3)} \cap \mathrm{JM}^{(2,3)}$ contains all assemblages in which any pair of two measurements are compatible, the so-called *pairwise* compatible assemblages. On the other hand, the set $\mathrm{JM}^{\mathrm{conv}} := \mathrm{Conv}(\mathrm{JM}^{(1,2)}, \mathrm{JM}^{(1,3)}, \mathrm{JM}^{(2,3)})$ describes the convex hull of the sets $\mathrm{JM}^{(1,2)}$, $\mathrm{JM}^{(1,3)}$, and $\mathrm{JM}^{(2,3)}$, i.e., it contains all assemblage that can be written as a convex combination of assemblages where one pair of measurements is compatible. More formally, it contains all assemblages of the form

$$\mathcal{M}_{(1,2,3)} = p_{(1,2)}\mathcal{J}^{(1,2)}_{(1,2,3)} + p_{(1,3)}\mathcal{J}^{(1,3)}_{(1,2,3)} + p_{(2,3)}\mathcal{J}^{(2,3)}_{(1,2,3)}, \tag{20}$$

where $\mathcal{J}^{(s,t)}_{(1,2,3)} \in \mathrm{JM}^{(s,t)}$ and the convex combination is to be understood on the level of the individual POVM effects. Finally an assemblage $\mathcal{M}_{(1,2,3)} \notin \mathrm{JM}^{\mathrm{conv}}$ is said to be *genuinely triplewise incompatible*. Note, these notions can straightforwardly be generalized to more than three measurements. See also [21] and for a graphical representation Figure 1 in the main text.

To quantify the incompatibility as a resource, we use the *diamond distance quantifier* [31] given by

$$\mathrm{I}_\diamond(\mathcal{M}^{\mathbf{p}}) = \min_{\mathcal{F}\in \mathrm{JM}} \sum_x p(x)\, \mathrm{D}_\diamond(\Lambda_{\mathcal{M}_x}, \Lambda_{\mathcal{F}_x}), \tag{21}$$

where $\Lambda_{\mathcal{M}_x} = \sum_a \mathrm{Tr}[M_{a|x}\rho]|a\rangle\langle a|$ is the *measure-and-prepare channel* associated to the measurement $\mathcal{M}_x$, and $\mathrm{D}_\diamond(\Lambda_1, \Lambda_2) = \max_{\rho\in\mathcal{S}(\mathcal{H}\otimes\mathcal{H})} \frac{1}{2}\|((\Lambda_1 - \Lambda_2)\otimes \mathbb{1}_d)\rho\|_1$ is the *diamond distance* [32] between two channels Λ_1, and Λ_2, with the trace norm$\|X\|_1 = \mathrm{Tr}[\sqrt{X^\dagger X}]$. Technically speaking, $\mathrm{I}_\diamond(\mathcal{M}^{\mathbf{p}})$ quantifies the incompatibility of a *weighted assemblage* $\mathcal{M}^{\mathbf{p}} = (\mathcal{M}, \mathbf{p})$ which contains the information about the probabilities $p(x)$ with which the measurement x is performed. The distance between two assemblages $\mathcal{M}^{\mathbf{p}}$ and $\mathcal{N}^{\mathbf{p}}$ that induces the quantifier $\mathrm{I}_\diamond(\mathcal{M}^{\mathbf{p}})$ is given by

$$\mathrm{D}_\diamond(\mathcal{M}^{\mathbf{p}}, \mathcal{N}^{\mathbf{p}}) := \sum_x p(x)\, \mathrm{D}_\diamond(\Lambda_{\mathcal{M}_x}, \Lambda_{\mathcal{N}_x}). \tag{22}$$

Like in the main text, we will simply write $\mathrm{I}_\diamond(\mathcal{M})$ to imply the case where $p(x) = \frac{1}{m}\forall x$. We denote by $\mathcal{M}^{\#}_{(1,2,\cdots,m)}$ the closest jointly measurable assemblage to $\mathcal{M}_{(1,2,\ldots,m)}$, i.e., the *arg-min* on the RHS in Eq. (21). Therefore, $\mathcal{M}^{\#}_{(1,2,\cdots,m)}$ can be seen as the closest jointly measurable approximation of the assemblage $\mathcal{M}_{(1,2,\ldots,m)}$. If we only approximate a subset of $n < m$ measurements of $\mathcal{M}_{(1,2,\ldots,m)}$ by jointly measurable measurements, for instance the first n settings, while keeping the remaining measurements unchanged, we write $\mathcal{M}^{\#(1,2,\ldots,n)}_{(1,2,\cdots,m)}$. Adding measurements

$\mathcal{M}'_{(m+1,m+2,\ldots,m+n)} = (\mathcal{M}'_{m+1}, \mathcal{M}'_{m+2}, \cdots, \mathcal{M}'_{m+n})$ to the assemblage $\mathcal{M}_{(1,2,\cdots,m)} = (\mathcal{M}_1, \mathcal{M}_2, \cdots, \mathcal{M}_m)$ is mathematically described by the concatenation of ordered list, using the symbol $+\!\!\!+$, i.e., we write

$$\mathcal{M}_{(1,2,\cdots,n+m)} = \mathcal{M}_{(1,2,\cdots,m)} +\!\!\!+ \mathcal{M}'_{(m+1,m+2,\cdots,m+n)} = (\mathcal{M}_1, \mathcal{M}_2, \cdots, \mathcal{M}_m, \mathcal{M}'_{m+1}, \cdots, \mathcal{M}'_{m+n}). \tag{23}$$

Using the notion of concatenation of ordered lists, we formally define

$$\mathcal{M}^{\#(1,2,\ldots,n)}_{(1,2,\cdots,m)} := \mathcal{M}^{\#}_{(1,2,\cdots,n)} +\!\!\!+ \mathcal{M}_{n+1} +\!\!\!+ \cdots +\!\!\!+ \mathcal{M}_m. \tag{24}$$

The diamond distance quantifier in Eq. (21) is a *faithful* resource quantifier, i.e., it holds that

$$\mathrm{I}_\diamond(\mathcal{M}^{\mathbf{p}}) = 0 \iff \mathcal{M} = \mathcal{M}^{\#} \in \mathrm{JM}. \tag{25}$$

For the above statement to be true, we assume that $p(x) \neq 0\ \forall x$, which is no restriction, since measurements that are never performed can be excluded from the assemblage before calculating the incompatibility.

Furthermore, $\mathrm{I}_\diamond(\mathcal{M}^{\mathbf{p}})$ is a monotone under any unital quantum channel $\Lambda^\dagger$ (these are exactly those channels that map POVMs to POVMs), i.e.,

$$\mathrm{I}_\diamond(\mathcal{M}^{\mathbf{p}}) \geq \mathrm{I}_\diamond(\Lambda^\dagger(\mathcal{M})^{\mathbf{p}}), \tag{26}$$

which follows from the fact that the trace distance is contractive under the application of completely positive and trace preserving (CPTP) maps. Note that in the resource theory of incompatibility, all unital quantum channels $\Lambda^\dagger$ are free. Indeed, it is straight forward to see that $\{\Lambda^\dagger(G_\lambda)\}$ is a parent POVM for the assemblage $\Lambda^\dagger(\mathcal{M})$ whenever $\{G_\lambda\}$ is a parent POVM for $\mathcal{M}$. That is, it holds

$$\Lambda^\dagger(M_{a|x}) = \sum_\lambda p(a|x,\lambda)\Lambda^\dagger(G_\lambda). \tag{27}$$

Additionally, $\mathrm{I}_\diamond(\mathcal{M}^{\mathbf{p}})$ is non-increasing under classical simulations $\mathcal{M}' = \xi(\mathcal{M})$ with

$$M'_{b|y} = \sum_x p(x|y) \sum_a q(b|y,x,a) M_{a|x}\ \forall\, b, y, \tag{28}$$

where $\mathcal{M}$ can be used to simulate [39] the assemblage $\mathcal{M}'$ via the conditional probabilities $p(x|y)$ and $q(b|y,x,a)$ for all y, respectively for all y, x, a. Using the classical simulations, one also obtains the possible probabilities $q(y)$ to perform setting y via $p(x) = \sum_y q(y)p(x|y)$. That means, it holds [31]:

$$\mathrm{I}_\diamond(\mathcal{M}^{\mathbf{p}}) \geq \mathrm{I}_\diamond(\xi(\mathcal{M}^{\mathbf{p}})^{\mathbf{q}}), \tag{29}$$

for all measurement simulations ξ. Eq. (29) follows ultimately from the fact that $\mathrm{I}_\diamond(\mathcal{M}^{\mathbf{p}})$ is based on a norm and that it is written as a convex combination over the settings.

Finally, since $\mathrm{I}_\diamond(\mathcal{M}^{\mathbf{p}})$ is based on the diamond distance $\mathrm{D}_\diamond(\mathcal{M}^{\mathbf{p}}, \mathcal{N}^{\mathbf{p}}) := \sum_x p(x)\, \mathrm{D}_\diamond(\Lambda_{\mathcal{M}_x}, \Lambda_{\mathcal{N}_x})$ between two weighted assemblages and the set JM of jointly measurable assemblage is convex, it is a convex function. Even more the distance $\mathrm{D}_\diamond(\mathcal{M}^{\mathbf{p}}, \mathcal{N}^{\mathbf{p}})$ fulfills the triangle inequality, i.e.,

$$\mathrm{D}_\diamond(\mathcal{M}^{\mathbf{p}}, \mathcal{N}^{\mathbf{p}}) \leq \mathrm{D}_\diamond(\mathcal{M}^{\mathbf{p}}, \mathcal{L}^{\mathbf{p}}) + \mathrm{D}_\diamond(\mathcal{L}^{\mathbf{p}}, \mathcal{N}^{\mathbf{p}}), \tag{30}$$

for any weighted measurement assemblages $\mathcal{M}^{\mathbf{p}}, \mathcal{L}^{\mathbf{p}}$, and $\mathcal{N}^{\mathbf{p}}$. It therefore follows that

$$\mathrm{I}_\diamond(\mathcal{M}^{\mathbf{p}}) \leq \mathrm{D}_\diamond(\mathcal{M}^{\mathbf{p}}, \mathcal{N}^{\#,\mathbf{p}}) \leq \mathrm{D}_\diamond(\mathcal{M}^{\mathbf{p}}, \mathcal{N}^{\mathbf{p}}) + \mathrm{I}_\diamond(\mathcal{N}^{\mathbf{p}}), \tag{31}$$

for any assemblages $\mathcal{M}$ and $\mathcal{N}$, where $\mathcal{N}^{\#} \in \mathrm{JM}$ is the closest jointly measurable assemblage with respect to $\mathcal{N}$. Note that the first inequality follows from the fact that $\mathcal{N}^{\#}$ is jointly measurable but not necessarily the closest jointly measurable assemblage to $\mathcal{M}$, i.e., $\mathcal{N}^{\#} \neq \mathcal{M}^{\#}$.

To prove the tightness of our bounds in the main text, we rely on the SDP formulation of $\mathrm{I}_\diamond(\mathcal{M}^{\mathbf{p}})$, which besides its numerical uses allows us, in some instances, to determine the incompatibility of an assemblage analytically. In [31] it was shown that that $\mathrm{I}_\diamond(\mathcal{M}^{\mathbf{p}})$ is equivalent to the optimal value of the SDP:

$$\begin{aligned}
&\underline{\text{Primal problem (incompatibility):}}\\
&\text{given}: \mathcal{M}^{\mathbf{p}}\\
&\underset{a_x, Z_x, G_\lambda}{\text{minimize}} \sum_x p(x) a_x\\
&\text{subject to:}\\
&a_x \mathbb{1} - \mathrm{Tr}_1[Z_x] \geq 0 \;\forall\; a, x,\\
&Z_x \geq \sum_a |a\rangle\langle a| \otimes (M_{a|x} - F_{a|x})^T \;\forall\; x,\\
&F_{a|x} = \sum_\lambda v(a|x,\lambda) G_\lambda \;\forall\; x, a,\;\; G_\lambda \geq 0 \;\forall\; \lambda, \sum_\lambda G_\lambda = \mathbb{1},\\
&Z_x \geq 0,\;\; a_x \geq 0 \;\forall\; x,
\end{aligned} \tag{32}$$

where the a_x are non-negative coefficients, the Z_x are positive semidefinite matrices and the G_λ are the POVM effects of the parent POVM. SDPs represent a special instance of convex optimization for which there exist off-the-shelf software [40–43] to efficiently solve them. Importantly, every SDP comes with a *dual formulation* that yields the same optimal value under some mild assumptions (see e.g., [44]). This is indeed the case here [31], i.e., $\mathrm{I}_\diamond(\mathcal{M}^{\mathbf{p}})$ can also be understood as the optimal value of the SDP:

$$\begin{aligned}
&\underline{\text{Dual problem (incompatibility):}}\\
&\text{given}:\; \mathcal{M}^{\mathbf{p}}\\
&\underset{C_{a|x}, \rho_x, L}{\text{maximize}} \quad \sum_{a,x} p(x) \mathrm{Tr}[M_{a|x} C_{a|x}] - \mathrm{Tr}[L]\\
&\text{subject to:}\\
&L \geq \sum_{a,x} p(x) v(a|x,\lambda) C_{a|x} \;\forall\; \lambda,\\
&0 \leq C_{a|x} \leq \rho_x \;\forall\; a, x,\;\; \rho_x \geq 0, \mathrm{Tr}[\rho_x] = 1 \;\forall\; x,
\end{aligned} \tag{33}$$

where the $C_{a|x}$, ρ_x, and L are positive semidefinite matricies. Since the primal problem in Eq. (32) corresponds to a minimization, every feasible point (i.e., any set of variables that fulfills all constraints) leads to an upper bound on $\mathrm{I}_\diamond(\mathcal{M}^{\mathbf{p}})$. Similarly, every feasible solution of the dual in Eq. (33) leads to a lower bound.

II. MEASUREMENT SPLITTING

In the main text, we argue that it is equivalent to consider the incompatibility of the assemblage $\mathcal{M}_{(1,2,1,3,2,3)}$ instead of $\mathcal{M}_{(1,2,3)}$, i.e., we use that $\mathrm{I}_\diamond(\mathcal{M}_{(1,2,3)}) = \mathrm{I}_\diamond(\mathcal{M}_{(1,2,1,3,2,3)})$ in order to derive the bound on the incompatibility gain in Eq. (10). Note that $\mathcal{M}_{(1,2,1,3,2,3)}$ is an assemblage in which each of the measurements $\mathcal{M}_1, \mathcal{M}_2$, and $\mathcal{M}_3$ occurs twice with probability $\frac{1}{6}$ each. On the other hand in $\mathcal{M}_{(1,2,3)}$ each of the measurements is used with a probability of $\frac{1}{3}$. To show the equivalence $\mathrm{I}_\diamond(\mathcal{M}_{(1,2,3)}) = \mathrm{I}_\diamond(\mathcal{M}_{(1,2,1,3,2,3)})$, we actually show that $\mathrm{I}_\diamond(\mathcal{M}_{(1,2,3)}) = \mathrm{I}_\diamond(\mathcal{M}'_{(1,1,2,2,3,3)})$ and finally use that the set JM of jointly measurable measurements is closed under relabeling. We first show that $\mathcal{M}'_{(1,1,2,2,3,3)} = \xi(\mathcal{M}_{(1,2,3)})$ for a measurement simulation (see also Eq. (28)) of the form

$$M'_{b|y} = \sum_x p(x|y) \sum_a q(b|y,x,a) M_{a|x} \;\forall\; b, y, \tag{34}$$

where we set $q(b|y,x,a) = \delta_{ba}$ for all b, y, x, a with δ_{ba} being the Kronecker delta. Furthermore, we use mixing probabilities $p(x|y)$ such that $p(x=1|y=1) = p(x=1|y=2) = 1$, $p(x=2|y=3) = p(x=2|y=4) = 1$, and $p(x=3|y=5) = p(x=3|y=6) = 1$ with all other probabilities set to zero. This is clearly a valid measurement simulation of $\mathcal{M}'_{(1,1,2,2,3,3)}$ using the measurements $\mathcal{M}_{(1,2,3)}$. Finally, notice that due to $p(x) = \sum_y q(y) p(x|y)$, it holds

$$\frac{1}{3} = p(x=i) = q(y=2i-1) + q(y=2i), \;\text{ for } i = 1,2,3, \tag{35}$$

which is clearly fulfilled for $q(y) = \frac{1}{6}\ \forall y$.The above equation actually shows a more general statement, i.e., any probabilities $p(y=1)+p(y=2)$ that sum to $\frac{1}{3}$ are allowed. The same holds for the other instances.

To show the other direction, i.e., $\mathcal{M}_{(1,2,3)} = \xi(\mathcal{M}'_{(1,1,2,2,3,3)})$ we use again $q(b|y,x,a) = \delta_{ba}$ for all b, y, x, a. For the mixing probabilities, we set $p(x=1|y=1) = p(x=2|y=1) = \frac{1}{2}$, $p(x=3|y=2) = p(x=4|y=2) = \frac{1}{2}$, and $p(x=5|y=2) = p(x=6|y=3) = \frac{1}{2}$ with all other probabilities set to zero. Again, it straightforward to check that this a valid measurement simulation. From the equivalence

$$p(x=1) = \frac{1}{6} = \sum_y q(y)p(1|y) = q(y=1)\frac{1}{2}, \tag{36}$$

it follows directly that $q(y=1) = \frac{1}{3}$ and similarly for the other cases. Now, since $\mathcal{M}'_{(1,1,2,2,3,3)} = \xi(\mathcal{M}_{(1,2,3)})$ and $\mathcal{M}_{(1,2,3)} = \xi(\mathcal{M}'_{(1,1,2,2,3,3)})$, it holds that $\mathrm{I}_\diamond(\mathcal{M}_{(1,2,3)}) = \mathrm{I}_\diamond(\mathcal{M}'_{(1,1,2,2,3,3)})$. Analogously follows the measurement splitting with more measurements.

Let us note here, that measurement simulations can also be used to show that the incompatibility of the parent POVMs of different subsets of jointly measurable assemblages is an upper bound on the incompatibility of these assemblages. More formally, let $\mathcal{M}_{(1,2,3)}$ be an assemblage and let $\mathcal{N} = \mathcal{M}^{\#}_{(1,2)} + \mathcal{M}^{\#}_{(1,3)} + \mathcal{M}^{\#}_{(2,3)}$ be the assemblage that contains the closest jointly measurable assemblages for the three subsets. Furthermore, let $\mathcal{G} = G(\mathcal{M}^{\#}_{(1,2)}) + G(\mathcal{M}^{\#}_{(1,3)}) + G(\mathcal{M}^{\#}_{(2,3)})$ be the assemblage that contains the parent POVMs of the respective subsets. With the above methods (and by the definition of the parent POVM in Eq. (19)) it can be seen that there exists a measurement simulation ξ such that $\xi(\mathcal{G}) = \mathcal{N}$, which directly implies that $\mathrm{I}_\diamond(\mathcal{N}) \leq \mathrm{I}_\diamond(\mathcal{G})$ holds.

III. LOWER BOUNDS

Here, we prove the lower bounds stated in Eq. (13) and Eq. (16). Remember, we consider the case in which $p(x) = \frac{1}{3}$, i.e., the input probabilities are uniformly distributed. Let us start by showing that

$$\mathrm{I}_\diamond(\mathcal{M}_{(1,2,3)}) \geq \frac{1}{3}[\mathrm{I}_\diamond(\mathcal{M}_{(1,2)}) + \mathrm{I}_\diamond(\mathcal{M}_{(1,3)}) + \mathrm{I}_\diamond(\mathcal{M}_{(2,3)})], \tag{37}$$

holds. We start by using that $\mathrm{I}_\diamond(\mathcal{M}_{(1,2,3)}) = \mathrm{I}_\diamond(\mathcal{M}_{(1,2,1,3,2,3)})$. Now, the closest jointly measurable assemblage $\mathcal{M}^{\#}_{(1,2,1,3,2,3)}$ with respect to $\mathcal{M}_{(1,2,1,3,2,3)}$ allows us to rewrite $\mathrm{I}_\diamond(\mathcal{M}_{(1,2,1,3,2,3)})$ such that

$$\mathrm{I}_\diamond(\mathcal{M}_{(1,2,1,3,2,3)}) = \mathrm{D}_\diamond(\mathcal{M}_{(1,2,1,3,2,3)}, \mathcal{M}^{\#}_{(1,2,1,3,2,3)}). \tag{38}$$

Now, concerning the measurement pairs $(1,2), (1,3)$, and $(2,3)$ the subsets of $\mathcal{M}^{\#}_{(1,2,1,3,2,3)}$ are jointly measurable by definition but not necessarily optimal for the respective subsets of $\mathcal{M}_{(1,2,1,3,2,3)}$. Using that the distance $\mathrm{D}_\diamond(\mathcal{M}_{(1,2,1,3,2,3)}, \mathcal{M}^{\#}_{(1,2,1,3,2,3)})$ is a convex combination over the individual settings, it follows that

$$\mathrm{I}_\diamond(\mathcal{M}_{(1,2,3)}) = \mathrm{I}_\diamond(\mathcal{M}_{(1,2,1,3,2,3)}) \geq \frac{1}{3}[\mathrm{I}_\diamond(\mathcal{M}_{(1,2)}) + \mathrm{I}_\diamond(\mathcal{M}_{(1,3)}) + \mathrm{I}_\diamond(\mathcal{M}_{(2,3)})]. \tag{39}$$

To show the second lower bound, i.e.,

$$\mathrm{I}_\diamond(\mathcal{M}_{(1,2,3)}) \geq \tfrac{2}{3}\,\mathrm{I}_\diamond(\mathcal{M}_{(1,2)}), \tag{40}$$

it is enough to notice that leaving out the contribution of the setting $x = 3$ can only lead to lower values than $\mathrm{I}_\diamond(\mathcal{M}_{(1,2,3)})$. Finally, we use again that the remaining measurements (for the settings $x = 1, 2$) from the closest jointly measurable assemblage $\mathcal{M}^{\#}_{(1,2,3)}$ do not need to be optimal.

IV. STEERING AND NONLOCALITY

Here, we show that our methods can directly be applied to quantum steering and Bell nonlocality. We start by considering steering. Let $\vec{\sigma}_{(1,2,\cdots,m)} = (\sigma_1, \sigma_2, \cdots, \sigma_m)$ with $\sigma_x = \{\sigma_{a|x}\}_a$ be the *steering assemblage* that Alice prepares for Bob by performing the measurements from a measurement assemblage $\mathcal{M}_{(1,2,\cdots,m)}$ on a shared state ρ such that $\sigma_{a|x} = \mathrm{Tr}_A[(M_{a|x} \otimes \mathbb{1})\rho]$. The *consistent steering distance* [49] given by

$$\mathrm{S}(\vec{\sigma}) = \min_{\vec{\tau} \in \mathrm{CLHS}} \frac{1}{2} \sum_{a,x} \frac{1}{m} \|\sigma_{a|x} - \tau_{a|x}\|_1, \tag{41}$$

can be used to quantify the steerability of any steering assemblage $\mathrm{S}(\vec{\sigma})$. Here, $\vec{\tau} \in \mathrm{CLHS}$ denotes an assemblage that admits a local hidden-state model (LHS) and fulfills the consistency condition $\sum_a \tau_{a|x} = \sum_a \sigma_{a|x} = \rho_B = \mathrm{Tr}_A[\rho]\ \forall x$. A LHS for $\vec{\tau}$ is given by

$$\tau_{a|x} = \sum_\lambda p(a|x,\lambda)\sigma_\lambda, \tag{42}$$

where the σ_λ are sub-normalized states and the $p(a|x,\lambda)$ resemble a classical post-processing, similarly to that in Eq. (19) in the definition of jointly measurable assemblages. Note that we directly used here that the choice of the settings is uniformly distributed, i.e., $p(x) = \frac{1}{m}$. However, generally, we can use any distribution with $p(x) \neq 0\ \forall x$, just like in the case for incompatibility. Note further that our following arguments are independent, as it was also the case for the incompatibility, of the number of outcomes a in the steering assemblage $\vec{\sigma}$.

Now, since $\mathrm{S}(\vec{\sigma})$ is based on a distance (the trace distance) we can directly derive the steering analog to the incompatibility bounds in the main text. In fact, our method relies only on the metric properties of the respective quantifiers, the fact they are written as a convex combination over the individual settings, and the general idea that a measurement can be split in two separate copies of itself. We make the following correspondence statements to our definitions for the incompatibility case:

$$\vec{\sigma}_{(1,2,\cdots,m)} \longleftrightarrow \mathcal{M}_{(1,2,\cdots,m)}, \tag{43a}$$
$$\mathrm{S}(\vec{\sigma}) \longleftrightarrow \mathrm{I}_\diamond(\mathcal{M}), \tag{43b}$$
$$\vec{\sigma}^{\#}_{(1,2,\cdots,m)} \longleftrightarrow \mathcal{M}^{\#}_{(1,2,\cdots,m)}, \tag{43c}$$
$$\vec{\sigma}^{\#(1,2,\cdots,n)}_{(1,2,\cdots,m)} \longleftrightarrow \mathcal{M}^{\#(1,2,\cdots,n)}_{(1,2,\cdots,m)}. \tag{43d}$$

That is, $\vec{\sigma}^{\#}_{(1,2,\cdots,m)}$ is the closest assemblage in the set CLHS to $\vec{\sigma}_{(1,2,\cdots,m)}$ with respect to the distance

$$\mathrm{D}_A(\vec{\sigma}_{(1,2,\cdots,m)}, \vec{\sigma'}_{(1,2,\cdots,m)}) := \sum_{a,x} \frac{1}{m} \|\sigma_{a|x} - \sigma'_{a|x}\|_1, \tag{44}$$

which induces the steering distance in Eq. (41). Furthermore, it holds

$$\vec{\sigma}^{\#(1,2,\cdots,n)}_{(1,2,\cdots,m)} := \vec{\sigma}^{\#}_{(1,2,\cdots,n)} \mathbin{+\!\!+} \sigma_{n+1} \mathbin{+\!\!+} \cdots \mathbin{+\!\!+} \sigma_m. \tag{45}$$

This implies, it holds that

$$\mathrm{S}(\vec{\sigma}_{(1,2,3)}) \leq \frac{1}{3}[\mathrm{S}(\vec{\sigma}_{(1,2)}) + \mathrm{S}(\vec{\sigma}_{(1,3)}) + \mathrm{S}(\vec{\sigma}_{(2,3)})] + \mathrm{S}(\vec{\tau}), \tag{46}$$

where $\vec{\tau} = \vec{\sigma}^{\#}_{(1,2)} \mathbin{+\!\!+} \vec{\sigma}^{\#}_{(1,3)} \mathbin{+\!\!+} \vec{\sigma}^{\#}_{(2,3)}$ is a state assemblage (with $m = 6$ settings) that contains itself three assemblages (of two settings each) that are the closest consistent unsteerable assemblages to the respective subsets. Note that $\vec{\tau}$ can be steerable in general. Note further that it is crucial to use a *consistent* steering quantifier here, in order to avoid *signaling* in the assemblage $\vec{\tau}$. All the other bounds follow from here on directly. That is, it follows that

$$\begin{aligned} \mathrm{S}(\vec{\sigma}_{(1,2,3)}) &\geq \frac{1}{3}[\mathrm{S}(\vec{\sigma}_{(1,2)}) + \mathrm{S}(\vec{\sigma}_{(1,3)}) + \mathrm{S}(\vec{\sigma}_{(2,3)})], \\ \mathrm{S}(\vec{\sigma}_{(1,2,3)}) &\geq \frac{2}{3}\mathrm{S}(\vec{\sigma}_{(1,2)}). \end{aligned} \tag{47}$$

Moreover, using the assemblage $\vec{\sigma}^{\#(1,2)}_{(1,2,3)}$ it holds that

$$\mathrm{S}(\vec{\sigma}_{(1,2,3)}) \leq \frac{2}{3}\mathrm{S}(\vec{\sigma}_{(1,2)}) + \mathrm{S}(\vec{\sigma}^{\#(1,2)}_{(1,2,3)}). \tag{48}$$

For nonlocality, very similar arguments can be made. However, we will see that additional constraints arise that distinguish nonlocality from steering and incompatibility. Let $\mathbf{q} = \{q(ab|xy)\}$ be a general probability distribution between two distant parties Alice and Bob. We consider the case where both, Alice and Bob, have two different measurement settings already available and Alice upgrades her measurement scheme with an additional third measurement. We denote the resulting distribution by $\mathbf{q}_{(1,2,3)}$. The nonlocality of a general distribution $\mathbf{q}$ can be quantified via the *consistent* version of the classical trace distance quantifier introduced in [37], which is given by

$$\mathrm{N}(\mathbf{q}) = \frac{1}{2} \min_{\mathbf{t} \in \mathrm{CLHV}} \sum_{a,b,x,y} \frac{1}{m_A m_B} |q(a,b|x,y) - t(a,b|x,y)|. \tag{49}$$

Here, we denote by CLHV the set of consistent local hidden-variable models (LHVs), i.e., the set of those local distributions $\mathbf{t} \in$ LHV that fulfill $\sum_a t(a,b|x,y) = t(b|y) = q(b|y) = \sum_a q(a,b|x,y)\ \forall\ b,y,x$ and similarly $\sum_b t(a,b|x,y) = t(a|x) = q(a|x) = \sum_b q(a,b|x,y)\ \forall\ a,y,x$. The (Bell) locality condition is expressed in terms of the LHV:

$$t(a,b|x,y) = \sum_\lambda \pi(\lambda) p_A(a|x,\lambda) p_B(b|y,\lambda)\ \forall a,b,x,y, \tag{50}$$

for the distribution $\mathbf{t}$. Finally, we denote by m_A the number of measurement settings of Alice and by m_B those of Bob, which we set to $m_B = 2$ here. Once again, we restrict our discussion to the case where the input probabilities $p(x,y) = p(x)p(y) = \frac{1}{m_A}\frac{1}{m_B}$ are uniformly distributed.

Since $\mathrm{N}(\mathbf{q})$ relies on a distance that is written as a convex combination over the individual settings, we can use the triangle inequality together with the measurement splitting method. We make the following correspondence statements to our definitions in the incompatibility case:

$$\mathbf{q}_{(1,2,\cdots,m_A)} \longleftrightarrow \mathcal{M}_{(1,2,\cdots,m)}, \tag{51a}$$
$$\mathrm{N}(\mathbf{q}) \longleftrightarrow \mathrm{I}_\diamond(\mathcal{M}), \tag{51b}$$
$$\mathbf{q}^{\#}_{(1,2,\cdots,m_A)} \longleftrightarrow \mathcal{M}^{\#}_{(1,2,\cdots,m)}, \tag{51c}$$
$$\mathbf{q}^{\#(1,2,\cdots,n_A)}_{(1,2,\cdots,m_A)} \longleftrightarrow \mathcal{M}^{\#(1,2,\cdots,n)}_{(1,2,\cdots,m)}. \tag{51d}$$

That is, $\mathbf{q}^{\#}_{(1,2,\cdots,m_A)}$ is the closest consistent and local distribution to $\mathbf{q}_{(1,2,\cdots,m_A)}$ with respect the the classical trace distance (ℓ_1 distance) that induces the nonlocality distance in Eq. (49). Furthermore, $\mathbf{q}^{\#(1,2,\cdots,n_A)}_{(1,2,\cdots,m_A)} = \mathbf{q}^{\#}_{(1,2,\cdots,n_A)} + \mathbf{q}_{n_A+1} + \cdots + \mathbf{q}_{m_A}$, where we treat the probability vector that describes a distribution $\mathbf{q}_{(1,2,\cdots,m_A)}$ as ordered list. We would like to emphasize that the indices $(1,2,\cdots,m_A)$ refer to the measurements of Alice, and Bob's number of measurements remains fixed here.

These correspondence relations imply that it is possible to obtain the bounds

$$\frac{2}{3}\mathrm{N}(\mathbf{q}_{(1,2)}) \leq \mathrm{N}(\mathbf{q}_{(1,2,3)}) \leq \frac{2}{3}\mathrm{N}(\mathbf{q}_{(1,2)}) + \mathrm{N}(\mathbf{q}^{\#(1,2)}_{(1,2,3)}), \tag{52}$$

Furthermore, we obtain the bounds

$$\frac{1}{3}[\mathrm{N}(\mathbf{q}_{(1,2)}) + \mathrm{N}(\mathbf{q}_{(1,3)}) + \mathrm{N}(\mathbf{q}_{(2,3)})] \leq \mathrm{N}(\mathbf{q}_{(1,2,3)}) \leq \frac{1}{3}[\mathrm{N}(\mathbf{q}_{(1,2)}) + \mathrm{N}(\mathbf{q}_{(1,3)}) + \mathrm{N}(\mathbf{q}_{(2,3)})] + \mathrm{N}(\mathbf{t}), \tag{53}$$

where $\mathbf{t} = \mathbf{q}^{\#}_{(1,2)} + \mathbf{q}^{\#}_{(1,3)} + \mathbf{q}^{\#}_{(2,3)}$ is a distribution (with $m_A = 6$ settings for Alice) which contains the closest local distributions with respect to the corresponding two-measurement subsets of Alice's measurement settings.

Interestingly, the term $\frac{1}{3}[\mathrm{N}(\mathbf{q}_{(1,2)}) + \mathrm{N}(\mathbf{q}_{(1,3)}) + \mathrm{N}(\mathbf{q}_{(2,3)})]$ behaves differently from its steering and incompatibility counterpart. Namely, it is limited by the fact that $\mathrm{N}(\mathbf{q}_{(1,2)})$, $\mathrm{N}(\mathbf{q}_{(1,3)})$, and $\mathrm{N}(\mathbf{q}_{(2,3)})$ cannot, in general, be maximal simultaneously. That is, contrary to incompatibility or steering, where all of the subset resources can be maximal at the same time.

The reason for this is that there are not enough degrees of freedom for Alice to violate a given Bell inequality with different measurements, given that Bob keeps his settings fixed (besides the state that is also fixed). To exemplify this, we consider the scenario where both parties have two outcomes for each setting. In that case, the nonlocality of $\mathrm{N}(\mathbf{q}_{(1,2)})$, $\mathrm{N}(\mathbf{q}_{(1,3)})$, and $\mathrm{N}(\mathbf{q}_{(2,3)})$ is directly linked to the amount of violation of the CHSH inequality [35], as it was shown in [37]. However, the CHSH inequality requires very specific combinations measurements to get maximal violation. Indeed, consider the three corresponding versions of the CHSH inequality:

$$\begin{aligned}\mathrm{CHSH}_{(1,2)} &\coloneqq \langle A_1\otimes B_1\rangle + \langle A_1\otimes B_2\rangle + \langle A_2\otimes B_1\rangle - \langle A_2\otimes B_2\rangle \le 2,\\ \mathrm{CHSH}_{(1,3)} &\coloneqq \langle A_1\otimes B_1\rangle + \langle A_1\otimes B_2\rangle + \langle A_3\otimes B_1\rangle - \langle A_3\otimes B_2\rangle \le 2,\\ \mathrm{CHSH}_{(2,3)} &\coloneqq \langle A_2\otimes B_1\rangle + \langle A_2\otimes B_2\rangle + \langle A_3\otimes B_1\rangle - \langle A_3\otimes B_2\rangle \le 2,\end{aligned} \tag{54}$$

and their sum

$$\mathrm{CHSH}_{(1,2,3)} \coloneqq \mathrm{CHSH}_{(1,2)} + \mathrm{CHSH}_{(1,3)} + \mathrm{CHSH}_{(2,3)} \le 6. \tag{55}$$

The inequality $\mathrm{CHSH}_{(1,2,3)} \le 6$ can also be rewritten as

$$2[\langle A_1\otimes B_1\rangle + \langle A_1\otimes B_2\rangle + \langle A_3\otimes B_1\rangle - \langle A_3\otimes B_2\rangle] + 2\langle A_2\otimes B_1\rangle \le 6, \tag{56}$$

which directly implies that the Tsirelson bound [45], i.e., the quantum bound of $\mathrm{CHSH}_{(1,2,3)}$ is given by $Q = 4\sqrt{2}+2 < 6\sqrt{2}$, i.e., quantum mechanics cannot reach the value $6\sqrt{2}$ that would correspond to all three contributions of Alice to be maximal simultaneously. The same is true for no-signaling theories, where the bound is given by NS $= 10$. Further, one needs to consider all possible combinations of different versions of the CHSH inequalities in Eq. (54). That is, one needs to consider all 8 symmetries of the CHSH inequality, corresponding to the 8 CHSH facets of the local polytope. However, going through all the combinations shows that there is no combination which allows for a higher combined CHSH value than $\mathrm{CHSH}_{(1,2,3)}$ in Eq. (55).

We want to emphasize again that such additional restrictions are not prevalent for the incompatibility and steering quantifiers analog of Eq. (53), which shows a clear separation of nonlocality to the other resources. Since the term $\frac{1}{3}[\mathrm{N}(\mathbf{q}_{(1,2)})+\mathrm{N}(\mathbf{q}_{(1,3)})+\mathrm{N}(\mathbf{q}_{(2,3)})]$ is also used in upper bounding the nonlocality $\mathrm{N}(\mathbf{q}_{(1,2,3)})$, this could be a promising path to understanding why additional settings do not seem to increase the resource of nonlocality [36, 37], in strict contrast to the resources of incompatibility and steerability. We expect that the same is true for more than two outcomes, however, more research in this direction is necessary.

V. GENERALIZATIONS

In this section, we will generalize our framework from the main text in two directions. First, we discuss the scenario for more measurements i.e, $m > 3$. Then, we will discuss the case in which the assemblage $\mathcal{M}^{\mathbf{p}} = (\mathcal{M}, \mathbf{p})$ is weighted by a general probability distribution $\mathbf{p}$, instead of a uniform one.

Using the methods from the main text and from Section II, general bounds can be derived. We demonstrate this in the following for the assemblage $\mathcal{M}_{(1,2,3,4)}$ of $m = 4$ uniformly distributed measurements. Further generalizations follow straightforwardly then. Let $\mathcal{M}^{\#(1,2,3)}_{(1,2,3,4)}$ be the closest assemblage with respect to the first three measurements of $\mathcal{M}_{(1,2,3,4)}$. Using the triangle inequality we get

$$\mathrm{I}_\diamond(\mathcal{M}_{(1,2,3,4)}) \le \mathrm{D}_\diamond(\mathcal{M}_{(1,2,3,4)}, \mathcal{M}^{\#(1,2,3)}_{(1,2,3,4)}) + \mathrm{I}_\diamond(\mathcal{M}^{\#(1,2,3)}_{(1,2,3,4)}) = \frac{3}{4}\,\mathrm{I}_\diamond(\mathcal{M}_{(1,2,3)}) + \mathrm{I}_\diamond(\mathcal{M}^{\#(1,2,3)}_{(1,2,3,4)}), \tag{57}$$

as a direct generalization of Eq. (15).

In general, let $C_0 = \{1, 2, \cdots, m\}$ be the set of of all possible measurements from an assemblage $\mathcal{M}_{(1,2,\cdots,m)}$. Furthermore, let $C \in C_0$ be any non-empty subset of C_0 with cardinality $|C|$. It follows that

$$\mathrm{I}_\diamond(\mathcal{M}_{(1,2,\cdots,m)}) \le \frac{|C|}{m}\,\mathrm{I}_\diamond(\mathcal{M}_C) + \mathrm{I}_\diamond(\mathcal{M}^{\#C}_{(1,2,\cdots,m)}), \tag{58}$$

where $|C|$ is the number of measurements contained in the subset $C \in C_0$. Since Eq. (58) holds for any subset C, we can conclude that

$$\mathrm{I}_\diamond(\mathcal{M}_{(1,2,\cdots,m)}) \le \min_{C\in C_0}\Big[\frac{|C|}{m}\,\mathrm{I}_\diamond(\mathcal{M}_C) + \mathrm{I}_\diamond(\mathcal{M}^{\#C}_{(1,2,\cdots,m)})\Big], \tag{59}$$

which in particular includes the optimization over all n measurement subsets. Note the upper bound trivially results in an equality in the case that $|C| = 1$, i.e., for practical purposes, one might exclude these cases from the minimization.

However, we can generalize our framework even more. Denote by $\{C_i\}$ a set of disjoint subsets of C_0 such that $\cup_i C_i = C_0$. It can directly be concluded that

$$\mathrm{I}_\diamond(\mathcal{M}_{(1,2,\cdots,m)}) \leq \sum_i \frac{|C_i|}{m} \mathrm{I}_\diamond(\mathcal{M}_{C_i}) + \mathrm{I}_\diamond(\mathcal{M}^\#_{C_1} \mathbin{+\!\!\!+} \mathcal{M}^\#_{C_2} \mathbin{+\!\!\!+} \cdots \mathbin{+\!\!\!+} \mathcal{M}^\#_{C_n}), \tag{60}$$

where $\mathcal{M}^\#_{C_1} \mathbin{+\!\!\!+} \mathcal{M}^\#_{C_2} \mathbin{+\!\!\!+} \cdots \mathbin{+\!\!\!+} \mathcal{M}^\#_{C_n}$ is an assemblage that contains itself the closest jointly measurable assemblages for the n respective subsets C_i. Again, it is possible to minimize Eq. (60) over a particular choice of different subsets and, in particular, over all non-trivial sets of subsets.

Besides the generalization of our bounds based solely on particular instances of the triangle inequality, we can also use the measurement splitting method in a more general setup. For instance, by splitting each measurement from $\mathcal{M}_{(1,2,3,4)}$ three times, we obtain the assemblage $\mathcal{M}_{(1,2,3,1,2,4,1,3,4,2,3,4)}$. This lets us conclude that it holds

$$\mathrm{I}_\diamond(\mathcal{M}_{(1,2,3,4)}) \leq \frac{1}{4}[\mathrm{I}_\diamond(\mathcal{M}_{(1,2,3)}) + \mathrm{I}_\diamond(\mathcal{M}_{(1,2,4)}) + \mathrm{I}_\diamond(\mathcal{M}_{(1,3,4)}) + \mathrm{I}_\diamond(\mathcal{M}_{(2,3,4)})] + \mathrm{I}_\diamond(\mathcal{N}), \tag{61}$$

with $\mathcal{N} = \mathcal{M}^\#_{(1,2,3)} \mathbin{+\!\!\!+} \mathcal{M}^\#_{(1,2,4)} \mathbin{+\!\!\!+} \mathcal{M}^\#_{(1,3,4)} \mathbin{+\!\!\!+} \mathcal{M}^\#_{(2,3,4)}$ as a direct generalization of Eq. (8). This leads for the generalization of the incompatibility gain in Eq. (10) to

$$\Delta\mathrm{I}_{(1,2,3)\to(1,2,3,4)} \leq \mathrm{I}_\diamond(\mathcal{N}) \leq \mathrm{I}_\diamond(\mathcal{G}), \tag{62}$$

where we assume $\mathrm{I}_\diamond(\mathcal{M}_{(1,2,3)}) \geq \max\{\mathrm{I}_\diamond(\mathcal{M}_{(1,2,4)}), \mathrm{I}_\diamond(\mathcal{M}_{(1,3,4)}), \mathrm{I}_\diamond(\mathcal{M}_{(2,3,4)})\}$ analogous to the condition stated in Result 1 and $\mathcal{G}$ is the assemblage containing the parent POVMs of the corresponding subsets. Again, further generalizations of (62) for other scenarios can be derived by applying our methods.

We show, in the following, that our results can be applied to any probability distribution $\mathbf{p}$ with which an assemblage $\mathcal{M}$ is weighted. Here, we focus on assemblages with $m = 3$ measurements. Further generalizations follow directly from the above discussion. Let $\mathcal{M}^{\mathbf{p}} = (\mathcal{M}, \mathbf{p})$ be a general weighted measurement assemblage. Using the triangle inequality, it holds

$$\mathrm{I}_\diamond(\mathcal{M}^{\mathbf{p}}_{(1,2,3)}) \leq \mathrm{D}_\diamond(\mathcal{M}^{\mathbf{p}}_{(1,2,3)}, \mathcal{N}^{\mathbf{p}}_{(1,2,3)}) + \mathrm{I}_\diamond(\mathcal{N}^{\mathbf{p}}_{(1,2,3)}), \tag{63}$$

for any assemblage $\mathcal{N}_{(1,2,3)}$. By setting $\mathcal{N} = \mathcal{M}^{\#(1,2)}_{(1,2,3)} \coloneqq \mathcal{M}^\#_{(1,2)} \mathbin{+\!\!\!+} \mathcal{M}_3$, it follows that

$$\mathrm{D}_\diamond(\mathcal{M}^{\mathbf{p}}_{(1,2,3)}, \mathcal{N}^{\mathbf{p}}_{(1,2,3)}) = [p(1) + p(2)]\, \mathrm{I}_\diamond(\mathcal{M}^{\mathbf{q}}_{(1,2)}), \tag{64}$$

where $\mathbf{q} = (\frac{p(1)}{p(1)+p(2)}, \frac{p(2)}{p(1)+p(2)})$ is the probability distribution weighting the assemblage $\mathcal{M}_{(1,2)}$. It is important to note here, that $\mathcal{M}^\#_{(1,2)}$ refers specifically to the closest assemblage to $\mathcal{M}_{(1,2)}$ with respect to the distribution $\mathbf{q}$. Note further that the particular instance of a uniform distribution can straightforwardly be recovered from here. This shows that $\mathrm{I}_\diamond(\mathcal{M}^{\mathbf{p}}_{(1,2,3)})$ is upper bounded by the incompatibility of its subset $\mathcal{M}_{(1,2)}$ weighted by the likelihood of choosing a measurement from that subset, plus the incompatibility $\mathrm{I}_\diamond(\mathcal{M}^{\#(1,2),\mathbf{p}}_{(1,2,3)})$. Similarly, if we want to use the measurement splitting method, we can chose any initial distribution $\mathbf{p}$ and proceed as usual to obtain bounds. As we noted in Section II the method is not limited to split a measurement into two equally likely versions of itself. The only conditions that have to be satisfied are the conditions in the second equality of Eq. (35).

VI. PROOFS REGARDING THE INCOMPATIBILITY OF MUTUALLY UNBIASED BASES

In this section, we present the proofs related to statements in the main text regarding the incompatibility of measurements based on MUB [19]. Two orthonormal bases $\{|v_a\rangle\}_{0\leq a\leq d-1}$ and $\{|w_b\rangle\}_{0\leq b\leq d-1}$ are said to be MUB if it holds that

$$|\langle v_a|w_b\rangle| = \frac{1}{\sqrt{d}}\ \forall\ a, b. \tag{65}$$

The set of projectors onto the vectors of a basis form the measurement $\mathcal{M} = \{M_a = |v_a\rangle\langle v_a|\}$. Now, an *MUB measurement assemblage* [31] is a set of measurements where the condition (65) holds for any two projections from different bases. While it is generally unknown how many MUB exist in a dimension d, it is known that for every $d \geq 2$, there exist at least $m = 3$ and at most $m = d + 1$ MUB. In the case where d is a prime-power there exist explicit constructions of MUB [46], which are known to be operationally inequivalent [31, 47]. The possibly most simple construction of a complete set of MUB, i.e., $m = d+1$ bases, can be used whenever d is a prime. In this case, we can use the Heisenberg-Weyl operators

$$\hat{X} = \sum_{k=0}^{d-1} |k+1\rangle\langle k|,\ \ \hat{Z} = \sum_{k=0}^{d-1} \omega^k |k\rangle\langle k|, \tag{66}$$

for a specific construction. Here, $\{|k\rangle\}_{0\leq k\leq d-1}$ is the computational basis and $\omega = \exp\left(\frac{2\pi i}{d}\right)$ is a root of unity. In prime dimensions d, the eigenbases of the $d+1$ operators $\hat{X}, \hat{Z}, \hat{X}\hat{Z}, \hat{X}\hat{Z}^2, \cdots, \hat{X}\hat{Z}^{d-1}$ are mutually unbiased [48]. Most notably, for $m = 2$, $m = d$, and $m = d+1$ there exists an analytical expression for the incompatibility of the MUB measurement assemblages obtained via this construction [31, 47]. For $d = 2$, our MUB measurement assemblage reduces to the projective measurements defined by the Pauli operators.

A. Tightness proof for Eq. (10)

We start by proving that Eq. (10) is tight for a noisy MUB measurement assemblage based on Pauli measurements, i.e., we show that

$$\Delta \mathrm{I}_{(1,2)\to(1,2,3)}(\eta) := \mathrm{I}_\diamond(\mathcal{M}^\eta_{(1,2,3)}) - \mathrm{I}_\diamond(\mathcal{M}^\eta_{(1,2)}) = \mathrm{I}_\diamond(\mathcal{N}(\eta)), \tag{67}$$

with $\mathcal{N}(\eta) = \mathcal{M}^{\#\eta}_{(1,2)} + \mathcal{M}^{\#\eta}_{(1,3)} + \mathcal{M}^{\#\eta}_{(2,3)}$ holds true for measurements of the form

$$M^\eta_{a|x} = \eta \Pi_{a|x} + (1-\eta)\,\mathrm{Tr}[\Pi_{a|x}]\frac{\mathbb{1}}{2}, \tag{68}$$

where the $\Pi_{a|x} = M^{\eta=1}_{a|x}$ are projectors defined via the eigenvectors of Pauli operators and η defines the amount of noise in the measurements.

We divide our proof into three different parameter regimes. Let η_2^* and η_3^* be the *white-noise robustness* of $\mathcal{M}_{(1,2)}$, respectively $\mathcal{M}_{(1,2,3)}$, i.e., the maximal η where the noisy assemblages are still jointly measurable. We consider the regimes $1): \ \eta \leq \eta_3^* \leq \eta_2^*,\ \ 2): \ \eta_3^* < \eta \leq \eta_2^*$, and $3): \ \eta_3^* \leq \eta_2^* < \eta$ corresponding to the three regimes in Figure 2.

Note that regime $1): \ \eta \leq \eta_3^*$ leads trivially to

$$\Delta \mathrm{I}_{(1,2)\to(1,2,3)}(\eta) = \mathrm{I}_\diamond(\mathcal{N}(\eta)) = 0. \tag{69}$$

For the second regime, i.e., $\eta_3^* < \eta \leq \eta_2^*$ it follows directly that

$$\Delta \mathrm{I}_{(1,2)\to(1,2,3)}(\eta) = \mathrm{I}_\diamond(\mathcal{M}^\eta_{(1,2,3)}) = \mathrm{I}_\diamond(\mathcal{N}(\eta)). \tag{70}$$

The first equality follows from the fact that $\mathrm{I}_\diamond(\mathcal{M}^\eta_{(1,2)}) = 0$ by definition. The second equality follows from the fact that $\mathcal{M}^{\#\eta}_{(s,t)} = \mathcal{M}^\eta_{(s,t)}$ for any $s,t \in \{1,2,3\}$ such that $s \neq t$, since the subset $\mathcal{M}^\eta_{(s,t)}$ is jointly measurable by definition. Now, due to the reverse direction of the measurement splitting method outlined in Section II, it holds $\mathrm{I}_\diamond(\mathcal{M}^\eta_{(1,2,3)}) = \mathrm{I}_\diamond(\mathcal{N}(\eta))$. That means, the only non-trivial case is regime $3): \ \eta_3^* \leq \eta_2^* < \eta$.

Our proof for this regime relies on solving the SDPs in Eq. (32) and Eq. (33) analytically. Starting from the dual:

$$\begin{aligned}
&\underline{\text{Dual problem (incompatibility):}} \\
&\text{given}: \ \mathcal{M}^\eta, \mathbf{p} \\
&\underset{C_{a|x},\rho_x,L}{\text{maximize}} \quad \sum_{a,x} p(x)\mathrm{Tr}[M^\eta_{a|x} C_{a|x}] - \mathrm{Tr}[L] \\
&\text{subject to:} \\
&L \geq \sum_{a,x} p(x) v(a|x,\lambda) C_{a|x} \ \forall\ \lambda, \\
&0 \leq C_{a|x} \leq \rho_x \ \forall\ a,x,\ \ \rho_x \geq 0, \mathrm{Tr}[\rho_x] = 1 \ \forall\ x,
\end{aligned} \tag{71}$$

we choose the specific instance where $C_{a|x} = \frac{\Pi_{a|x}}{2}$, $L = l\mathbb{1}$, and $\rho_x = \sum_a C_{a|x} = \frac{\mathbb{1}}{2}$ for some appropriately chosen scalar-variable l. For a qubit assemblage $\mathcal{M}$ with POVM effects of the form

$$M^{\eta}_{a|x} = \eta\Pi_{a|x} + (1-\eta)\,\mathrm{Tr}[\Pi_{a|x}]\frac{\mathbb{1}}{2} = \eta\Pi_{a|x} + (1-\eta)\frac{\mathbb{1}}{2}, \tag{72}$$

this evaluates to the lower bound $\mathrm{I}_\diamond(\mathcal{M}^\eta) \geq \eta + \frac{(1-\eta)}{2} - \frac{T}{m}$, where $T \coloneqq \|\sum_{a,x} v^*(a|x,\lambda)M_{a|x}\|_\infty$ and $\{v^*(a|x,\lambda)\}_{a,x}$ is the deterministic strategy maximizing the norm. Note that this bound results from choosing $l = \frac{T}{2m}$, which can be shown to be always a valid choice [31].

For (noise-free, i.e., $\eta = 1$) MUB measurement assemblages it was proven in [47] that whenever $m = 2$, $m = d$, or $m = d+1$, it holds that

$$\eta^*_m = \frac{dT - m}{dm - m}. \tag{73}$$

This lets us conclude (for the qubit case, i.e., d = 2) that

$$\begin{aligned} \mathrm{I}_\diamond(\mathcal{M}^\eta) &\geq \eta + \frac{(1-\eta)}{2} - \frac{\eta^*_m + 1}{2} \\ &= \frac{1}{2}(\eta - \eta^*_m). \end{aligned} \tag{74}$$

For the upper bound of $\mathrm{I}_\diamond(\mathcal{M}^\eta)$ we invoke the primal SDP:

$$\begin{aligned} &\underline{\text{Primal problem (incompatibility):}} \\ &\text{given}: \mathcal{M}^\eta, \mathbf{p} \\ &\underset{Z_x, G_\lambda}{\text{minimize}} \sum_x p(x)\|\mathrm{Tr}_1[Z_x]\|_\infty \\ &\text{subject to:} \\ &Z_x \geq \sum_a |a\rangle\langle a| \otimes (M^\eta_{a|x} - F_{a|x})^T \;\forall\, x, \\ &F_{a|x} = \sum_\lambda v(a|x,\lambda)G_\lambda \;\forall\, x,a,\;\; G_\lambda \geq 0 \;\forall\, \lambda, \sum_\lambda G_\lambda = \mathbb{1}, \\ &Z_x \geq 0,\;\forall\, x, \end{aligned} \tag{75}$$

where we have explicitly replaced the constraints involving the variables a_x in the SDP in Eq. (32) by using the spectral norm (largest singular value). By choosing $F_{a|x} = \eta^*_m\Pi_{a|x} + (1-\eta^*_m)\frac{\mathbb{1}}{2}$ and $Z_x = \frac{1}{2}(\eta - \eta^*_m)\sum_a |a\rangle\langle a| \otimes \Pi^T_{a|x}$ for $\eta \geq \eta^*_m$ all constraints can directly be verified to hold. Therefore, we obtain the upper bound

$$\mathrm{I}_\diamond(\mathcal{M}^\eta) \leq \frac{1}{2}(\eta - \eta^*_m). \tag{76}$$

That implies $\mathrm{I}_\diamond(\mathcal{M}^\eta) = \dfrac{1}{2}(\eta - \eta^*_m)$ for any assemblage involving $m = 2$ or $m = 3$ noisy MUB measurement assemblages with $\eta \geq \eta^*_m$ in $d = 2$. Therefore, the incompatibility gain $\Delta\mathrm{I}_{(1,2)\to(1,2,3)}(\eta) \coloneqq \mathrm{I}_\diamond(\mathcal{M}^\eta_{(1,2,3)}) - \mathrm{I}_\diamond(\mathcal{M}^\eta_{(1,2)})$ evaluates to

$$\begin{aligned} \Delta\mathrm{I}_{(1,2)\to(1,2,3)}(\eta) &= \frac{1}{2}[(\eta - \eta) + (\eta^*_2 - \eta^*_3)]. \\ &= \frac{1}{2}[(\eta^*_2 - \eta^*_3)]. \end{aligned} \tag{77}$$

Note that the gain is constant in this regime, as it is also evident from Figure 2. Now, to finish the proof, we have to show that $\mathrm{I}_\diamond(\mathcal{N}(\eta))$ has the same incompatibility. However, this follows almost directly, since $\mathcal{N}(\eta) = \mathcal{M}^{\#\eta}_{(1,2)} \mathbin{+\!\!\!+} \mathcal{M}^{\#\eta}_{(1,3)} \mathbin{+\!\!\!+} \mathcal{M}^{\#\eta}_{(2,3)}$ contains the closest jointly measurable assemblages with respect to the subsets. As it is known from [47] and [31] (and we confirmed it with the above calculation) all of these subsets are again just noisy versions of MUB measurement assemblages, with the same noise contained in every subset. From the reverse direction of the measurement splitting method, it follows that

$$\mathrm{I}_\diamond(\mathcal{N}(\eta)) = \mathrm{I}_\diamond(\mathcal{M}^\eta_{(1,2,3)}) \tag{78}$$

for $\eta = \eta^*_2 = \frac{1}{\sqrt{2}}$. Therefore, it follows that $\mathrm{I}_\diamond(\mathcal{N}(\eta)) = \dfrac{1}{2}[(\eta^*_2 - \eta^*_3)]$ for $\eta \geq \eta^*_2$ which concludes the proof.

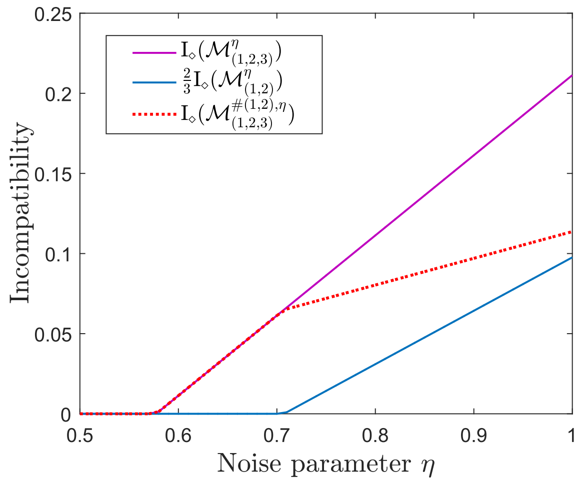

Figure 3. Incompatibility bound from Eq. (15) for measurements corresponding to the three Pauli measurements. The different contributions (depicted by the dashed red and the blue line) result together in the incompatibility $\mathrm{I}_\diamond(\mathcal{M}^\eta_{(1,2,3)})$. The two discontinuities of $\mathrm{I}_\diamond(\mathcal{M}^{\#(1,2),\eta}_{(1,2,3)})$ (dashed red line) indicate the points where $\mathcal{M}^\eta_{(1,2,3)}$ becomes compatible, respectively pairwise compatible.

B. Tightness proof for Eq. (15)

Here, we show that Eq. (15) is tight for the case of noisy Pauli measurements (see Eq. (68)). That is, we show that

$$\mathrm{I}_\diamond(\mathcal{M}^\eta_{(1,2,3)}) = \frac{2}{3}\,\mathrm{I}_\diamond(\mathcal{M}^\eta_{(1,2)}) + \mathrm{I}_\diamond(\mathcal{M}^{\#(1,2),\eta}_{(1,2,3)}), \tag{79}$$

holds for the assemblage $\mathcal{M}^\eta_{(1,2,3)}$ that contains noisy Pauli measurements. To give a better overview, we also plot the respective incompatibility contributions of $\mathrm{I}_\diamond(\mathcal{M}^\eta_{(1,2,3)})$ in Figure 3. The proof reduces to show the equality for the case $\eta > \eta_2^*$, as the other cases follow trivially from the discussions made in Section VI A. We already evaluated the values of $\mathrm{I}_\diamond(\mathcal{M}^\eta_{(1,2,3)})$ and $\mathrm{I}_\diamond(\mathcal{M}^\eta_{(1,2)})$, i.e., we only have to show that

$$\begin{aligned}\mathrm{I}_\diamond(\mathcal{M}^{\#(1,2),\eta}_{(1,2,3)}) = \mathrm{I}_\diamond(\mathcal{M}^\eta_{(1,2,3)}) - \frac{2}{3}\,\mathrm{I}_\diamond(\mathcal{M}^\eta_{(1,2)}) &= \frac{1}{2}(\eta-\eta_3^*) - \frac{1}{3}(\eta-\eta_2^*) \\ &= \frac{1}{6}\eta + \frac{1}{3}\eta_2^* - \frac{1}{2}\eta_3^*.\end{aligned} \tag{80}$$

Since we already know that $\frac{1}{6}\eta + \frac{1}{3}\eta_2^* - \frac{1}{2}\eta_3^* \leq \mathrm{I}_\diamond(\mathcal{M}^{\#(1,2),\eta}_{(1,2,3)})$, due to the general bound in Eq. (58), it is enough to show that $\mathrm{I}_\diamond(\mathcal{M}^{\#(1,2),\eta}_{(1,2,3)}) \leq \frac{1}{6}\eta + \frac{1}{3}\eta_2^* - \frac{1}{2}\eta_3^*$ also holds true.

We rely again on the primal problem in Eq. (75) using the feasible point where $F_{a|x} = \eta_3^*\Pi_{a|x} + (1-\eta_3^*)\frac{\mathbb{1}}{2}$ and $Z_x = \frac{1}{2}(\eta_2^*-\eta_3^*)\sum_a |a\rangle\langle a| \otimes \Pi_{a|x}^T$ for $x=1,2$ and $Z_x = \frac{1}{2}(\eta-\eta_3^*)\sum_a |a\rangle\langle a| \otimes \Pi_{a|x}^T$ for $x=3$. It can be checked again directly that this point is indeed feasible. Moreover, we obtain a primal objective value of

$$\sum_x \frac{1}{3}\|\mathrm{Tr}_1[Z_x]\|_\infty = 2\frac{1}{3}\cdot\frac{1}{2}(\eta_2^*-\eta_3^*) + \frac{1}{6}(\eta-\eta_3^*) = \frac{1}{6}\eta + \frac{1}{3}\eta_2^* - \frac{1}{2}\eta_3^*, \tag{81}$$

which concludes the proof.

C. Tightness proof for generalizations of Eq. (15)

Here, we show that in the scenarios $m=2 \to m'=d$, $m=2 \to m'=d+1$, and $m=d \to m'=d+1$ there exists analog bounds to Eq. (15) that are tight for d-dimensional noisy MUB measurement assemblages. As before, we only consider the non-trivial case here and in the following, i.e., the noisy regime where none of the incompatibilities vanish.

Also, we refer to the noise-free measurements, i.e., the projectors on the MUB by $\Pi_{a|x} = M_{a|x}^{\eta=1}$. The corresponding bound (see Eq. (58)) for the instance $2 \to d$ reads

$$\mathrm{I}_\diamond(\mathcal{M}^\eta_{(1,2,\cdots,d)}) \leq \frac{2}{d}\,\mathrm{I}_\diamond(\mathcal{M}^\eta_{(1,2)}) + \mathrm{I}_\diamond(\mathcal{M}^{\#(1,2),\eta}_{(1,2,\cdots,d)}), \tag{82}$$

where we know that $\mathrm{I}_\diamond(\mathcal{M}^\eta_{(1,2)}) = (\frac{d-1}{d})(\eta-\eta_2^*)$ by generalizing the previous qubit result. Indeed, carefully checking the calculation for the $d=2$ case in the Section VI A, reveals the general (dimension dependant) prefactor for the incompatibility of two noisy MUB measurements.

Furthermore, using essentially the same feasible points as before (simply extended to the case of $m=d$ instead of $m=2$ measurements) we obtain that $\mathrm{I}_\diamond(\mathcal{M}^\eta_{(1,2,\cdots,d)}) = (\frac{d-1}{d})(\eta-\eta_d^*)$. With that, we know that

$$\mathrm{I}_\diamond(\mathcal{M}^{\#(1,2),\eta}_{(1,2,\cdots,d)}) \geq \mathrm{I}_\diamond(\mathcal{M}^\eta_{(1,2,\cdots,d)}) - \frac{2}{d}\,\mathrm{I}_\diamond(\mathcal{M}^\eta_{(1,2)}) = \Big(\frac{d-1}{d}\Big)(\eta-\eta_d^*) - \frac{2}{d}\Big(\frac{d-1}{d}\Big)(\eta-\eta_2^*), \tag{83}$$

which means, it remains to show that $\mathrm{I}_\diamond(\mathcal{M}^{\#(1,2),\eta}_{(1,2,\cdots,d)}) \leq (\frac{d-1}{d})(\eta-\eta_d^*) - \frac{2}{d}(\frac{d-1}{d})(\eta-\eta_2^*)$ also holds. This can directly be verified by using the feasible point $F_{a|x} = \eta_d^*\Pi_{a|x} + (1-\eta_d^*)\frac{\mathbb{1}}{d}$ and $Z_x = (\frac{d-1}{d})(\eta_2^*-\eta_d^*)\sum_a |a\rangle\langle a| \otimes \Pi_{a|x}^T$ for $x=1,2$, and $Z_x = (\frac{d-1}{d})(\eta-\eta_d^*)\sum_a |a\rangle\langle a|\otimes \Pi^T_{a|x}$ for $x = 3,\cdots,d$. This concludes the proof.

The corresponding bound for the $2 \to d+1$ scenario (see Eq. (58)) reads

$$\mathrm{I}_\diamond(\mathcal{M}^\eta_{(1,2,\cdots,d+1)}) \leq \frac{2}{d+1}\,\mathrm{I}_\diamond(\mathcal{M}^\eta_{(1,2)}) + \mathrm{I}_\diamond(\mathcal{M}^{\#(1,2),\eta}_{(1,2,\cdots,d+1)}), \tag{84}$$

with $\mathrm{I}_\diamond(\mathcal{M}^\eta_{(1,2)}) = (\frac{d-1}{d})(\eta-\eta_2^*)$. Using the same feasible points as for the $m=2$ and $m=d$ case, it also follows that $\mathrm{I}_\diamond(\mathcal{M}^\eta_{(1,2,\cdots,d+1)}) = (\frac{d-1}{d})(\eta-\eta_{d+1}^*)$, i.e, to prove tightness, we have to show that

$$\mathrm{I}_\diamond(\mathcal{M}^{\#(1,2),\eta}_{(1,2,\cdots,d+1)}) \leq \Big(\frac{d-1}{d}\Big)(\eta-\eta_{d+1}^*) - \Big(\frac{2}{d+1}\Big)\Big(\frac{d-1}{d}\Big)(\eta-\eta_2^*), \tag{85}$$

holds true. Using the same construction as before, this can be checked directly. Namely, using the feasible point $F_{a|x} = \eta_{d+1}^*\Pi_{a|x}+(1-\eta_{d+1}^*)\frac{\mathbb{1}}{d}$ and $Z_x = (\frac{d-1}{d})(\eta_2^*-\eta_{d+1}^*)\sum_a|a\rangle\langle a|\otimes\Pi^T_{a|x}$ for $x=1,2$, and $Z_x = (\frac{d-1}{d})(\eta-\eta_{d+1}^*)\sum_a|a\rangle\langle a|\otimes\Pi^T_{a|x}$ for $x=3,\cdots,d+1$ it follows directly that Eq. (85) is indeed true, which concludes the proof.

In the case $d \to d+1$, the corresponding bound reads

$$\mathrm{I}_\diamond(\mathcal{M}^\eta_{(1,2,\cdots,d+1)}) \leq \frac{d}{d+1}\,\mathrm{I}_\diamond(\mathcal{M}^\eta_{(1,2,\cdots,d)}) + \mathrm{I}_\diamond(\mathcal{M}^{\#(1,2,\cdots,d),\eta}_{(1,2,\cdots,d+1)}), \tag{86}$$

with $\mathrm{I}_\diamond(\mathcal{M}^\eta_{(1,2,\cdots,d+1)}) = (\frac{d-1}{d})(\eta-\eta_{d+1}^*)$ and $\mathrm{I}_\diamond(\mathcal{M}^\eta_{(1,2,\cdots,d)}) = (\frac{d-1}{d})(\eta-\eta_d^*)$. That means we have to check that

$$\mathrm{I}_\diamond(\mathcal{M}^{\#(1,2,\cdots,d),\eta}_{(1,2,\cdots,d+1)}) \leq \Big(\frac{d-1}{d}\Big)(\eta-\eta_{d+1}^*) - \Big(\frac{d}{d+1}\Big)\Big(\frac{d-1}{d}\Big)(\eta-\eta_d^*), \tag{87}$$

is true. Using $F_{a|x} = \eta_{d+1}^*\Pi_{a|x} + (1-\eta_{d+1}^*)\frac{\mathbb{1}}{d}$ and $Z_x = \frac{d-1}{d}(\eta_d^*-\eta_{d+1}^*)\sum_a|a\rangle\langle a|\otimes\Pi^T_{a|x}$ for $x=1,2,\cdots,d$, and $Z_x = \frac{d-1}{d}(\eta-\eta_{d+1}^*)\sum_a|a\rangle\langle a|\otimes\Pi^T_{a|x}$ for $x=d+1$ this can be verified, just as in the above cases.

D. Additional insights on Eq. (18)

In this subsection, we give additional insights to Eq. (18) from the main text. That is, we analyse the incompatibility decomposition

$$\mathrm{I}_\diamond(\mathcal{M}_{(1,2,3)}) \leq \mathrm{I}_\diamond{}^{\mathrm{gen}}(\mathcal{M}_{(1,2,3)}) + \mathrm{I}_\diamond{}^{\mathrm{pair}}(\mathcal{M}_{(1,2,3)}) + \mathrm{I}_\diamond{}^{\mathrm{hol}}(\mathcal{M}_{(1,2,3)}), \tag{88}$$

for an arbitrary assemblage $\mathcal{M}_{(1,2,3)}$. Note that we defined here $\mathrm{I}_\diamond{}^{\mathrm{gen}}(\mathcal{M}) := \mathrm{D}_\diamond(\mathcal{M}_{(1,2,3)}, \mathcal{M}^{\mathrm{conv}})$ to be the *genuine triplewise incompatibility* of $\mathcal{M}_{(1,2,3)}$, i.e., its distance to the closest assemblage $\mathcal{M}^{\mathrm{conv}} \in \mathrm{JM}^{\mathrm{conv}} := \mathrm{Conv}(\mathrm{JM}^{(1,2)}, \mathrm{JM}^{(1,3)}, \mathrm{JM}^{(2,3)})$. Furthermore, $\mathrm{I}_\diamond{}^{\mathrm{pair}}(\mathcal{M}) := \mathrm{D}_\diamond(\mathcal{M}^{\mathrm{conv}}, \mathcal{M}^{\mathrm{pair}})$, is the *pairwise incompatibility*, where $\mathcal{M}^{\mathrm{pair}} \in \mathrm{JM}^{pair} := \mathrm{JM}^{(1,2)} \cap \mathrm{JM}^{(1,3)} \cap \mathrm{JM}^{(2,3)}$ is the closest assemblage in which all measurements are pairwise-compatible and $\mathrm{I}_\diamond{}^{\mathrm{hol}}(\mathcal{M}) := \mathrm{I}_\diamond(\mathcal{M}^{\mathrm{pair}})$ is the *hollow incompatibility* of $\mathcal{M}_{(1,2,3)}$. Note that the pairwise and hollow incompatibility depend implicitly on $\mathcal{M}_{(1,2,3)}$. See also Figure 1 for the different incompatibility structures. Indeed the incompatibilities defined here, are nothing else but the distances to the next corresponding compatibility structure in Figure 1 in the main text.

We now show that the bound in Eq. (88) is tight for the three Pauli measurements. For simplicity, we focus on the noise-free scenario in the following. From the previous discussions, we know that $\mathrm{I}_\diamond(\mathcal{M}_{(1,2,3)}) = \frac{1}{2}(1-\eta_3^*)$.

For the contribution $\mathrm{I}_\diamond{}^{\mathrm{gen}}(\mathcal{M}_{(1,2,3)})$ we can use $\mathcal{M}_{(1,2,3)}^{\#(1,2)}$ as (possibly sub-optimal) point in $\mathrm{JM}^{\mathrm{conv}}$. Therefore, we obtain the bound

$$\mathrm{I}_\diamond{}^{\mathrm{gen}}(\mathcal{M}_{(1,2,3)}) \leq \mathrm{D}_\diamond(\mathcal{M}_{(1,2,3)}, \mathcal{M}_{(1,2,3)}^{\#(1,2)}) = \tfrac{2}{3}\,\mathrm{I}_\diamond(\mathcal{M}_{(1,2)}) = \tfrac{1}{3}(1-\eta_2^*). \tag{89}$$

For the contribution $\mathrm{I}_\diamond{}^{\mathrm{pair}}(\mathcal{M}_{(1,2,3)})$ we use as a guess for $\mathcal{M}^{\mathrm{pair}}$ the depolarized version of $\mathcal{M}_{(1,2,3)}$ where it becomes pairwise compatible. That is, $\mathcal{M}^{\mathrm{pair}}$ is of the form

$$M_{a|x}^{\mathrm{pair}} = \eta_3^{\mathrm{pair}}\Pi_{a|x} + (1-\eta_3^{\mathrm{pair}})\frac{\mathbb{1}}{d}. \tag{90}$$

Using the results from previous discussions, we therefore obtain

$$\mathrm{I}_\diamond{}^{\mathrm{pair}}(\mathcal{M}_{(1,2,3)}) \leq \frac{1}{3}(\eta_2^* - \eta_3^{\mathrm{pair}}) + \frac{1}{6}(1-\eta_3^{\mathrm{pair}}), \tag{91}$$

by bounding the distance $\mathrm{D}_\diamond(\mathcal{M}_{(1,2,3)}^{\#(1,2)}, \mathcal{M}^{\mathrm{pair}})$ through the SDP for the diamond norm, i.e., we examine the SDP in Eq. (32) with $\mathcal{F} = \mathcal{M}^{\mathrm{pair}}$.

Finally, for the contribution $\mathrm{I}_\diamond{}^{\mathrm{hol}}(\mathcal{M}_{(1,2,3)})$ we use as (possibly sub-optimal) candidate for the closest jointly measurable assemblage simply the appropriately depolarized version of $\mathcal{M}_{(1,2,3)}$, i.e., we obtain the bound $\mathrm{I}_\diamond{}^{\mathrm{hol}}(\mathcal{M}_{(1,2,3)}) \leq \frac{1}{2}(\eta_3^{\mathrm{pair}} - \eta_3^*)$. Summing all these bounds up, we obtain that

$$\begin{aligned}\mathrm{I}_\diamond{}^{\mathrm{gen}}(\mathcal{M}_{(1,2,3)}) + \mathrm{I}_\diamond{}^{\mathrm{pair}}(\mathcal{M}_{(1,2,3)}) + \mathrm{I}_\diamond{}^{\mathrm{hol}}(\mathcal{M}_{(1,2,3)}) &\leq \frac{1}{3}(1-\eta_2^*) + \frac{1}{3}(\eta_2^* - \eta_3^{\mathrm{pair}}) + \frac{1}{6}(1-\eta_3^{\mathrm{pair}}) + \frac{1}{2}(\eta_3^{\mathrm{pair}} - \eta_3^*) \\ &= \frac{1}{2}(1-\eta_3^*),\end{aligned} \tag{92}$$

which equals the value for $\mathrm{I}_\diamond(\mathcal{M}_{(1,2,3)})$ for the noise free Pauli measurements, as calculated in section VI A. Therefore Eq. (88) is tight. Note that the proof crucially relies on knowing the incompatibility of $\mathrm{I}_\diamond(\mathcal{M}_{(1,2,3)})$, i.e., without having an analytical expression for this term in higher dimensions $d > 2$, this way of proving equality will not work generally. However, we can check numerically, whether the bound in Eq. (88) is tight for higher dimensional MUB. Indeed, our numerics suggest for up to $d = 7$ that Eq. (88) is tight for MUB with a deviation of the order 10^{-9}.