The power of tests for signal detection in high-dimensional data

Inaugural-Dissertation

zur Erlangung des Doktorgrades
der Mathematisch-Naturwissenschaftlichen Fakultät
der Heinrich-Heine-Universität Düsseldorf

vorgelegt von

Marc Ditzhaus
aus Duisburg

Düsseldorf, November 2016

Aus dem Mathematischen Institut
der Heinrich-Heine Universität Düsseldorf

Gedruckt mit der Genehmigung der
Mathematisch-Naturwissenschaftlichen Fakultät der
Heinrich-Heine-Universität Düsseldorf

Referent:	Prof. Dr. Arnold Janssen
Koreferenten:	Prof. Dr. Markus Pauly
	Prof. Dr. Markus Reiß

Tag der mündlichen Prüfung: 28.06.2017

Abstract

In this thesis we are interested in the testing problem, whether there are rare and weak signals (alternative) or no signals (null) within white noise background. To be more specific, we study the asymptotic behaviour of the log-likelihood ratio test (LLRT) and Tukey's higher criticism test (HC) modified by Donoho and Jin [20] when the number of observations n tends to infinity. First results were shown by Ingster [32], who studied the asymptotic behaviour of LLRT in great detail under the assumption of normal distributions. In the same context Jin [41] and Donoho and Jin [20] used the term detection boundary, which divides the plane, that represents the parametrisation of signal strength and the probability of a signal, into two areas. By doing this they illustrate their results and the ones of Ingster [32]: Underneath the boundary LLRT yields no better results than flipping a coin (as $n \to \infty$). Above the boundary LLRT can completely separate the null and the alternative (as $n \to \infty$). Moreover, Donoho and Jin [20] showed that the latter is also valid for HC. In contrast to LLRT HC does not depend on the unknown signal strength and probability of a signal. Thus, it is applicable in practice. Similar results concerning HC were also shown for other distributions, see [10, 12, 20].

The first chapter serves as an introduction for this thesis. In the second chapter we present an extension of the model which was studied in the literature. The main difference between these models is that the signal strength and the probability for a signal can differ in each observation. Moreover, we do not restrict our model to normal or other specific distributions. In the following first main part of this thesis we discuss the asymptotic behaviour of LLRT. We are especially interested in the limit distribution of the test statistic on the detection boundary. There are already results in the literature concerning this, see [10, 12, 20, 32], which we can extend to our general model. In the second main part of this thesis we show that the detection boundaries of HC and LLRT coincide for some different assumptions concerning the distributions. Moreover, we discuss the asymptotic behaviour of HC on the detection boundary for these assumptions. We want to emphasise that the last issue was an open problem until recently.

Zusammenfassung

In dieser Arbeit beschäftigen wir uns mit dem Testproblem, ob innerhalb von weißem Rauschen einige wenige Signale (Alternative) oder keine Signale (Nullhypothese) vorliegen. Hierzu studieren wir das asymptotische Verhalten des Log-Likelihood-Quotienten Tests (LLQ) und des von Donoho und Jin [20] modifizierten *Higher Criticism* Tests (HC), wenn die Anzahl n der Beobachtungen gegen unendlich strebt. Erste Resultate wurden von Ingster [32] erzielt, der unter Normalverteilungsannahmen das asymptotische Verhalten von LLQ studierte. In diesem Kontext führten Jin [41] und Donoho und Jin [20] den Begriff der Erkennungsgrenze ein, welche die Ebene, die die Parametrisierung der Signalstärke und -wahrscheinlichkeit darstellt, in zwei Bereiche teilt. Auf diese Weise visualisierten sie ihre Ergebnisse sowie diejenigen von Ingster [32]: Unterhalb dieser Grenze erzielt LLQ keine besseren Ergebnisse als ein Münzwurf (für $n \to \infty$). Oberhalb dieser Grenze kann LLQ zwischen Nullhypothese und Alternative (für $n \to \infty$) ohne Fehler unterscheiden. Weiterhin zeigten Donoho and Jin [20], dass HC Letzteres ebenfalls kann. Im Gegensatz zu LLQ hängt HC nicht von der Signalwahrscheinlichkeit und -stärke ab und ist somit in der Praxis anwendbar. Das zuvor erwähnte, asymptotisch optimale Verhalten von HC wurde auch für andere Verteilungsannahmen nachgewiesen, siehe [10, 12, 20].

Nach einem einleitenden ersten Kapitel stellen wir eine Erweiterung des bisher betrachteten Modells im zweiten Kapitel vor, indem wir zulassen, dass die Signalwahrscheinlichkeit und -stärke für verschiedene Beobachtungen unterschiedlich sein kann. Zudem schränken wir das Modell nicht auf bestimmte Verteilungsannahmen, z.B. Normalverteilung, ein. Im ersten Hauptteil der Arbeit beschäftigen wir uns mit dem asymptotischen Verhalten von LLQ. Insbesondere sind wir an dem Konvergenzverhalten der Teststatistik auf der Erkennungsgrenze interessiert. Hierbei lassen sich die bisherigen Ergebnisse bezüglich Normalverteilungsannahmen, siehe [10, 12, 20, 32], auf unser allgemeineres Modell erweitern. Im zweiten Hauptteil der Arbeit widmen wir uns HC. Wir zeigen, dass die Erkennungsgrenzen von HC und LLQ unter noch nicht betrachteten Verteilungsannahmen übereinstimmen. Zudem präsentieren wir erste Ergebnisse zum Verhalten von HC auf der Grenze, welches in der Literatur bisher noch nicht studiert wurde.

Acknowledgements

First and foremost, I want to express my sincere appreciation to my advisor Prof. Dr. Arnold Janssen for his invaluable support, his guidance, supervision of this thesis and many fruitful discussions. I am indebted to my mentor Prof. Dr. Markus Pauly for reviewing this thesis. Both, Prof. Dr. Arnold Janssen and Prof. Dr. Markus Pauly, had a major impact on my education during my bachelor and master studies.

I want to thank Prof. Dr. Markus Reiß for writing the third review.

Moreover, I wish to thank my former and current colleagues at the chair for mathematical statistics and probability theory for a very warm working atmosphere.

I am also grateful for the financial support of the Deutsche Forschungsgemeinschaft (DFG).

Finally, I thank my family, my friends and in particular Susanne Unterlaß for their love, patience and support.

Contents

1. Introduction

We are confronted with a lot of data nowadays. Thus, there is a lot of work to do for the statistician. One task, which is the goal of this thesis, is to detect signals within white noise. At the beginning of genomics, e.g., scientists were hopeful that the so called *common disease-common variant hypothesis* was true, see [23]. This hypothesis says that some known genes are differentially expressed in patients affected by a common disease. Practice, however, showed us that this assumption is incorrect. In recent research the assumption, that numerous unknown genes of an affected patient are differentially but only slightly expressed, is often used, see [16, 24, 33]. Note that the number of the differentially expressed genes is quite huge but relative to the observed genes it is small. That is why the model, which is used in this case, is called the *rare and weak model* because the number of signals is relatively small (*rare signals*) and so is the effect size (*weak signals*). Consequently, it is very difficult to decide if there are signals or not. This model is also applied in cosmology and astronomy, see [13, 40], and in local anomaly detection, especially in disease outbreak detection, see [50, 55]. Donoho and Jin [20] modified Tukey's *higher criticism*, see [59–61], for the purpose of detecting (heterogeneous) normal mixtures, i.e., for the testing problem $\mathcal{H}_{0,n}$ versus $\mathcal{H}_{1,n}$, where

$$\begin{aligned} &\mathcal{H}_{0,n}:\ Y_{n,1},\dots,Y_{n,n} \text{ are i.i.d. with distribution } P_0 := N(0,1) \\ &\mathcal{H}_{1,n}: Y_{n,1},\dots,Y_{n,n} \text{ are i.i.d. with distribution } Q := (1-\varepsilon_n)N(0,1)+\varepsilon_n N(\vartheta_n,1) \end{aligned} \tag{1.1}$$

for some $\varepsilon_n > 0$, $\vartheta_n \in \mathbb{R}$, $\vartheta_n \neq 0$ and $n \in \mathbb{N}$. The null $\mathcal{H}_{0,n}$ can be interpreted as white noise. Furthermore, the alternative $\mathcal{H}_{1,n}$ can be interpreted as white noise, where a random number of observations contains an additional signal ϑ_n. We will later see that under $\mathcal{H}_{1,n}$ each observation is additionally shifted by ϑ_n with probability ε_n, see our Lemma 2.3. In other words, the parameter ε_n is the probability for a signal and ϑ_n is the signal's strength. At first glance, this testing problem is quite easy because the null and the alternative have one element each. Consequently, it is well known that the *log-likelihood ratio test*, in short *LLRT*, achieves the best power among all tests, see, e.g., [58].

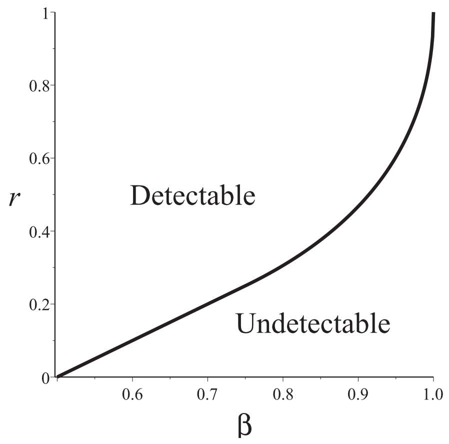

Figure 1.1.: Phase diagram for (sparse heterogeneous) normal mixtures. The solid curve represents the detection boundary given by (1.3), which splits the plane into two areas, the detectable and undetectable area.

Ingster [32] as well as Donoho and Jin [20] considered the following parametrisation of ε_n and ϑ_n:

$$\varepsilon_n := n^{-\beta} \text{ and } \vartheta_n := \sqrt{2r\log n} \text{ for some } \beta \in \left(\frac{1}{2}, 1\right),\ r \in (0,1). \tag{1.2}$$

Their results can be explained and visualised by using the so called *detection boundary* and by plotting it in a *phase diagram*. This was first done by Donoho and Jin [20]. For the testing problem (1.1) and the parametrisation (1.2) the detection boundary ρ^* is given by

$$\rho^*(\beta) := \begin{cases} \beta - \frac{1}{2} & \text{if } \frac{1}{2} < \beta \leq \frac{3}{4}. \\ \left(1 - \sqrt{1-\beta}\right)^2 & \text{if } \frac{3}{4} < \beta < 1. \end{cases} \tag{1.3}$$

It divides the r-β-plane, see Figure 1.1, into the *detectable* and the *undetectable* area. If $r > \rho^*(\beta)$ then (r, β) belongs to the detectable area, i.e., LLRT can separate the null and the alternative completely (asymptotically) and so the sum of type I and type II error probabilities tends to 0. Conversely, if $r < \rho^*(\beta)$ then (r, β) belongs to the undetectable area, i.e., the null and the alternative are asymptotically indistinguishable and so the sum of type I and type II error probabilities converges to 1 for all tests. Donoho and Jin [20] discussed also other mixture models, e.g., a Chi-squared model. Cai et al. [10] discussed heteroscedastic normal mixtures, i.e., the variance under the alternative can differ from 1. Cai and Wu [12] discussed a great class of exponential families. For all these models a phase diagram can be drawn and a detection boundary can be calculated. Recently the case of multivariate Gaussian distributed observations $Y_{n,i}$ was discussed in the literature, see [47, 62]. Beside all these continuous models, Arias-Castro and Wang [5] discussed a discrete mixture model, to be more specific a sparse Poisson mixture model. They determined the detection boundary for this model by drawing a parallel to the normal mixture model mentioned above. We also want to mention the paper of Jin [41], who discussed the detection problem for the *multiple looks* model. In this model the sample is divided into m groups of the same size, where every group contains exactly one signal under the alternative, and so the number of signals is not random. At least for normal distributions, Jin [41] showed that the results for the mixture model (1.1) can be transferred to the multiple looks model.

In practice LLRT cannot be used because it depends on the unknown parameters ε and ϑ. But the *higher criticism test* of Donoho and Jin [20], in short HC, does not depend on these parameters. The detectable areas of HC and LLRT coincide for the testing problem (1.1). This is also valid for the other above-mentioned models, see [5, 10, 12, 20].

The testing problem (1.1) can also be interpreted as a multiple hypothesis testing problem for the global null. In this context Gontscharuk et al. [25], see also [26], discussed the asymptotic behaviour of HC by using *local levels*. The well known *false discovery rate* (FDR) controlling procedure of Benjamini and Hochberg [6] can also be applied for this problem. But Donoho and Jin [20] showed that this and other known procedures for multiple testing problems do not achieve the same asymptotic optimality for the heterogeneous normal mixture model, i.e., the corresponding detectable regions are smaller. Jager and Wellner [35] studied a new class of goodness-of-fit tests based on Phi-divergences, which includes HC and which are also independent of the unknown parameters. They showed that the detectable area of each member of this class coincides with one of HC and LLRT in the normal heterogeneous mixture model. Moreover, Stepanova and Pavlenko [57] suggested another class of goodness-of-fit tests based on the ideas of Csörgő et al. [14] and proved that all these tests behave asymptotically as good as HC and so as LLRT in the heterogeneous normal mixture model. Similarly to the HC test statistic their statistics are also sup-functionals of the empirical process weighted by a certain class of functions. For all previous mentioned tests the null distribution need to be known. Arias-Castro and Wang [4] considered that the null distribution is unknown but symmetric. They suggested a variant of HC for symmetry. For the sparse generalised Gaussian mixture model they proved that their test has the same asymptotically optimal behaviour as LLRT and HC. To sum up, there are other adaptive tests that can compete with HC.

We have assumed that the observations are independent until now. But the correlated case is also of great interest, e.g., in genetics and genomics, see [16]. Hall and Jin [30] proved that a modification of HC is even optimal for the correlated case under certain assumptions for the correlation structure. But there are also unsatisfying results about HC for strong dependence, see [29]. The model considered in [29, 30] can be seen as the special case, that $n = p$, for the following linear regression model:

$$Y = X\beta + \sigma Z,$$

where $\sigma > 0$, Z is an n-dimensional multivariate normal distributed noise vector with the identity as covariance matrix, $X \in \mathbb{R}^{n\times p}$ is the design matrix and $\beta \in \mathbb{R}^p$ is the vector of regression coefficients. The optimality of tests, among other variants of HC, for the null $\beta = 0$ are discussed in [3, 31], where the non-zero entries of β are supposed to be sparse under the alternative. Beside the detection of signals, HC can be used to construct estimates for ε_n, see [11], and for feature selection, see [17, 18, 63]. For more details about HC and possible application fields we refer the reader to the survey paper of Donoho and Jin [19] and the huge number of references therein for more information.

The aim of this thesis is to extend the results mentioned above for HC and LLRT in the context of signal detection to a more general model, where, e.g., the signal probability $\varepsilon_{n,i}$ can differ in each observation. We are especially interested in the asymptotic power of both tests on the detection boundary. Some results about the asymptotic behaviour of LLRT on the boundary are already known, e.g., for the heterogeneous and heteroscedastic normal mixture model, see [10, 20, 32]. But the power of HC on the boundary was an open question until now.

This introduction is followed by seven chapters, which finally conclude with the answer to the very same question.

In Chapter 2 we introduce our more general mixture model. The new idea is to allow the distribution of the observations $Y_{n,i}$ to depend on i, so that the observations do not need to have the same distribution. We also present some examples for this model from which we want to emphasise three types:

- heteroscedastic normal mixtures (light-tailed distribution).
- exponential families (including light- and heavy-tailed distributions).
- h-model (structure model for chimeric alternatives).

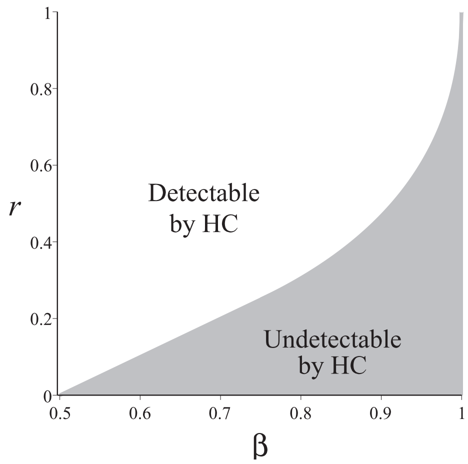

Figure 1.2.: The detectable and the undetectable area of HC are visualised for (sparse heterogeneous) normal mixtures. The boundary, which splits the plane into these two areas, belongs to the undetectable area.

For all these examples we determine the detection boundary and the asymptotic behaviour of HC and LLRT on the detection boundary in the upcoming chapters, respectively.

In Chapters 3 to 5, the first main part of this thesis, we discuss the asymptotic behaviour of LLRT. The LLRT and its test statistic are introduced in Chapter 3. At the beginning of Chapter 4 we explain the useful connection between the asymptotic behaviour of LLRT and weak convergence of binary experiments. By this it is sufficient for our purpose to determine the accumulation points of certain binary experiments. We distinguish between trivial and non-trivial accumulation points. There are two trivial cases: $\mathcal{H}_{0,n}$ and $\mathcal{H}_{1,n}$ can be completely separated by LLRT (asymptotically) or $\mathcal{H}_{0,n}$ and $\mathcal{H}_{1,n}$ merge (asymptotically). In the non-trivial case LLRT can successfully, but not completely, separate $\mathcal{H}_{0,n}$ and $\mathcal{H}_{1,n}$ (asymptotically). The non-trivial accumulation points correspond to the behaviour of LLRT on the detection boundary. In Sections 4.2, 4.3 and 4.5 we present necessary and sufficient conditions for trivial and non-trivial accumulation points of these binary experiments. In Chapter 5 we make use of them for the examples introduced in Chapter 2. By this we determine, among others, the detection boundary and the asymptotic behaviour of LLRT on it for models which are not discussed in the literature until recently, e.g., the h-model.

In the last three chapters, the second main part, we focus on HC. The structure of this part is the same as the one of the previous part. In Chapter 6 we introduce the test and the test statistic. The theoretical results are presented in Chapter 7. In Chapter 8 these results are applied to our examples introduced in Chapter 2. To determine the detectable area of HC we extend the ideas in [10, 12, 20] to our general model. By using this extension we show for all our examples that

the detectable and the undetectable areas of HC and LLRT coincide.

We also discuss the model suggested by Cai and Wu [12]. We prove that under the assumptions of Theorem 3 in [12] not only LLRT but also HC can completely separate the null and the alternative. This was an unsolved problem in [12]. Similarly to the first part we are in particular interested in the following question.

How does HC behave on the detection boundary asymptotically?

As far as we know, this question was unanswered in the literature until recently. We verify for all our examples that HC cannot successfully separate the null and the alternative, i.e., the sum of type I and type II error probabilities converges to 1. In other words,

HC has no power on the detection boundary asymptotically.

Consequently, LLRT yields better results than HC asymptotically, at least on the boundary. The results concerning HC are visualised in Figure 1.2 for the normal mixture model introduced at the beginning of this chapter, see (1.1) and (1.2).

For a better understanding of this thesis, the appendix contains additional information about infinitely divisible distributions, binary experiments and distances between probability measures for the readers, who are not familiar with these topics. Beside that, some technical results are presented.

2. The model

2.1. Introduction

At first we present the general model, which we consider throughout the whole thesis. In Sections 2.2 to 2.4 we introduce some specific and some quite general examples for this model, among others the normal mixtures mentioned in Chapter 1. These examples are referred to in the third chapter of each main part in order to apply our results to them.

Assumption 2.1. *(i) Let $(\varepsilon_{n,i})_{1\leq i\leq n\in\mathbb{N}}$ be a triangular array of real numbers in $[0,1]$ and $(\Omega,\mathcal{A})$ be a measurable space. Moreover, let $\mu_{n,i}$ and $P_{n,i}$ be two different probability measures on $(\Omega,\mathcal{A})$ for all $1\leq i\leq n\in\mathbb{N}$. Suppose that*

$$\varepsilon_{n:n} := \max_{i=1,\ldots,n}\{\varepsilon_{n,i}\} \to 0 \quad \textit{as } n\to\infty \tag{2.1}$$

$$\textit{and } \; \mu_{n,i} \ll P_{n,i} \quad \textit{for all } 1\leq i\leq n\in\mathbb{N}. \tag{2.2}$$

Let $f_{n,i}$ be the $P_{n,i}$-density of $\mu_{n,i}$, which is, here and subsequently, a short notation for the Radon-Nikodym density of $\mu_{n,i}$ with respect of $P_{n,i}$. Define for all $1\leq i\leq n\in\mathbb{N}$

$$Q_{n,i} := (1-\varepsilon_{n,i})P_{n,i} + \varepsilon_{n,i}\mu_{n,i} \tag{2.3}$$

and for all $n\in\mathbb{N}$ the product measures

$$Q_{(n)} := \bigotimes_{i=1}^{n} Q_{n,i} \;\; \textit{and} \;\; P_{(n)} := \bigotimes_{i=1}^{n} P_{n,i}.$$

For all $1\leq i\leq n\in\mathbb{N}$ let $g_{n,i}$ be the $P_{n,i}$-density of $Q_{n,i}$, i.e.,

$$g_{n,i} \;=\; \frac{\mathrm{d}([1-\varepsilon_{n,i}]P_{n,i}+\varepsilon_{n,i}\mu_{n,i})}{\mathrm{d}P_{n,i}} \;=\; 1+\varepsilon_{n,i}\left(f_{n,i}-1\right). \tag{2.4}$$

If $(\Omega, \mathcal{A}) = (\mathbb{R}, \mathcal{R})$ *we denote the distribution function of* $P_{n,i}$ *by* $F_{n,i}$ *and the left-continuous quantile function of* $P_{n,i}$ *by* $F_{n,i}^{-1}$ *for all* $1 \leq i \leq n$*, i.e.,*

$$F_{n,i}^{-1}(u) := \inf\{t \in \mathbb{R} : F_{n,i}(t) \geq u\} \quad \textit{for all} \;\; u \in (0,1). \tag{2.5}$$

(ii) Suppose that (i) *holds, where* $P_{n,i}$ *and* $\mu_{n,i}$ *do not depend on* i*, i.e.,*

$$P_{n,i} = P_{n,1} \textit{ and } \mu_{n,i} = \mu_{n,1} \textit{ for all } i \in \{1, \ldots, n\}.$$

Set for all $n \in \mathbb{N}$

$$P_n := P_{n,1},\, \mu_n := \mu_{n,1},\, Q_n := Q_{n,1},$$
$$f_n := f_{n,1},\, g_n := g_{n,1},\, F_n := F_{n,1},\, F_n^{-1} := F_{n,1}^{-1}.$$

(iii) Suppose that (ii) *holds, where* $\varepsilon_{n,i}$ *does not depend on* i *as well. Set* $\varepsilon_n := \varepsilon_{n,1}$ *for all* $n \in \mathbb{N}$*.*

Remark 2.2. (i) In the following chapters we introduce some random variables and further probability measures. For simplicity of the notation, they should also "live" on $(\Omega, \mathcal{A})$. Hence, we suppose that the measurable space $(\Omega, \mathcal{A})$ is rich enough.

(ii) Under Assumption 2.1(i) we have for all $\delta > 0$

$$\begin{aligned}
\max_{i=1,\ldots,n} \{P_{n,i}(\varepsilon_{n,i} f_{n,i} \geq \delta)\} &= \max_{i=1,\ldots,n} \left\{ \int \mathbf{1}\{\varepsilon_{n,i} f_{n,i} \geq \delta\} \; \mathrm{d}P_{n,i} \right\} \\
&\leq \max_{i=1,\ldots,n} \left\{ \int \mathbf{1}\{\varepsilon_{n,i} f_{n,i} \geq \delta\} \; \frac{\varepsilon_{n,i}\, f_{n,i}}{\delta} \; \mathrm{d}P_{n,i} \right\} \\
&= \delta^{-1} \max_{i=1,\ldots,n} \left\{ \varepsilon_{n,i} \int \mathbf{1}\{\varepsilon_{n,i} f_{n,i} \geq \delta\} \; \mathrm{d}\mu_{n,i} \right\} \\
&\leq \delta^{-1}\, \varepsilon_{n:n} \;\longrightarrow\; 0 \quad \text{as } n \to \infty.
\end{aligned}$$

□

Consequently, the generalisation of the testing problem (1.1) is given by

$$\mathcal{H}_{0,n} : \; P_{(n)} \quad \text{versus} \;\; \mathcal{H}_{1,n} : Q_{(n)}. \tag{2.6}$$

We already mentioned in Chapter 1 how the null and the alternative can be interpreted. The null represents some white noise. For the interpretation of the alternative we need the following elementary lemma.

Lemma 2.3. *Let X, Y, Z be random variables on the same probability space $(\Omega, \mathcal{A}, \mathcal{P})$, where $Z \sim B(1, \varepsilon)$ for some $\varepsilon \in [0,1]$ and X, Y take values on some vector space V over the field $\mathbb{R}$. Denote by P and μ the distribution of X and Y, respectively, i.e., $\mathcal{P}^X = P$ and $\mathcal{P}^Y = \mu$. Furthermore, let X and Z be independent as well as Y and Z. Then*

$$(1-Z)X + ZY \sim (1-\varepsilon)P + \varepsilon\mu.$$

Proof. Suppose that $X, Y : (\Omega, \mathcal{A}) \to (V, \mathcal{A}_V)$. Then for all $A \in \mathcal{A}_V$

$$\mathcal{P}\left((1-Z)X + ZY \in A\right) = \mathcal{P}\left(X \in A\right)\mathcal{P}\left(Z = 0\right) + \mathcal{P}\left(Y \in A\right)\mathcal{P}\left(Z = 1\right). \qquad \blacksquare$$

Remark 2.4. There are no assumptions on the dependence structure of X and Y. Consequently, we obtain immediately:
Let $U \sim \mathcal{U}_{(0,1)}$ and $Z \sim B(1,\varepsilon)$ for some $\varepsilon \in [0,1]$ be two random variables on the same probability space $(\Omega, \mathcal{A}, \mathcal{P})$. Moreover, let P and μ be probability measures on $(\mathbb{R}, \mathcal{B})$. Denote the distribution functions of P, μ by F, G and the corresponding left-continuous quantile functions by F^{-1}, G^{-1}, compare to (2.5). Then

$$(1-Z)F^{-1}(U) + ZG^{-1}(U) \sim (1-\varepsilon)P + \varepsilon\mu. \qquad \square$$

Thus, the alternative can be interpreted as a two-stage experiment. In the first step, it is determined if the i^{th} observation contains a signal or not, where the probability for a signal is equal to $\varepsilon_{n,i}$ (in our general model). If the i^{th} observation contains a signal then it is a realisation of the distribution $\mu_{n,i}$. Otherwise, it is a realisation of the distribution $P_{n,i}$. Hence, the number of signals is random under the alternative. Note that models with a fixed number of signals under the alternative are also discussed in the literature, e.g., the multiple looks model, see [41].

In the introduction we only mentioned the parametrisation in (1.2). But for the same model Cai et al. [10] also examined the case that $\vartheta_n \searrow 0$ and $\sqrt{n}\varepsilon_n \to \infty$. They described this case as the *dense* case and the other case, where $\vartheta_n \to \infty$ and $\sqrt{n}\varepsilon_n \to 0$, is described as the *sparse* case. The main idea of these cases is to distinguish between relatively strong but rare signals and very weak but many signals. Note that in the literature the signal strength is called weak even for the sparse case because μ_n tends to infinity very slowly. Other authors used the notation *moderately sparse* and *very sparse* case. But we prefer the one, which was used in [10], and extend it to our more general model by using the variational distance, defined in Definition and Lemma A.12(i).

Notation 2.5. Suppose that Assumption 2.1(i) holds.

(i) We denote by the *sparse case* the case in which

$$\liminf_{n\to\infty} \min_{1\le i\le n} ||P_{n,i} - \mu_{n,i}|| > 0 \quad \text{and} \quad \lim_{n\to\infty} \sum_{i=1}^{n} \varepsilon_{n,i}^2 = 0.$$

(ii) We denote by the *dense case* the case in which

$$\lim_{n\to\infty} \max_{1\le i\le n} ||P_{n,i} - \mu_{n,i}|| = 0 \quad \text{and} \quad \liminf_{n\to\infty} \sum_{i=1}^{n} \varepsilon_{n,i}^2 > 0. \qquad \square$$

2.2. First examples

At the beginning we present the heteroscedastic normal mixture model, which we already mentioned in Chapter 1.

Example 2.6 (Heteroscedastic normal mixture model). Suppose that Assumption 2.1(iii) holds, where $(\Omega, \mathcal{A}) = (\mathbb{R}, \mathcal{B})$, $P_n = N(0,1) =: P_0$ and $\mu_n = N(\vartheta_n, \tau^2)$ for some $\tau \in (0, \infty)$, $\vartheta_n \in \mathbb{R}$ and all $n \in \mathbb{N}$. Then for all $x \in \mathbb{R}$

$$f_n(x) \;=\; \frac{1}{\tau} \exp\left(\left[1 - \frac{1}{\tau^2}\right]\frac{x^2}{2} + \frac{\vartheta_n}{\tau^2}x - \frac{\vartheta_n^2}{2\tau^2}\right). \qquad \square$$

Remark 2.7. If $\tau = 1$ we get the heterogeneous normal mixture model discussed in Chapter 1, see (1.1). In this case the variances of P_n and μ_n are equal. $\square$

In the following we introduce examples corresponding to the three max-stable distributions, namely Gumbel, Fréchet and Exponential, see, e.g., [27], of which the latter one represents the Weibull distributions. These are of special interest because there is a connection between extreme value theory and determining the asymptotic behaviour of the log-likelihood ratio test, see Section 4.4.

Example 2.8 (Gumbel distribution). Suppose that Assumption 2.1(iii) holds, where $(\Omega, \mathcal{A}) = (\mathbb{R}, \mathcal{B})$ and $P_0 := P_1 = P_n$ is the standard Gumbel distribution for all $n \in \mathbb{N}$. We denote by Λ the distribution function of P_0, i.e., for all $x \in \mathbb{R}$

$$\Lambda(x) = \exp\left(-e^{-x}\right) \quad \text{and} \quad \frac{\mathrm{d}P_0}{\mathrm{d}\lambda\!\!\lambda}(x) = \exp\left(-x - e^{-x}\right). \tag{2.7}$$

Moreover, suppose that μ_n is equal to the convolution of P_0 and the dirac measure centred at $\vartheta_n \in \mathbb{R}$, i.e., the measure is uniquely determined by the shift ϑ_n:

$$\mu_n(-\infty, x] := \Lambda(x - \vartheta_n) \quad \text{for all } x \in \mathbb{R}.$$

Then

$$f_n(x) = \frac{\mathrm{d}P_0}{\mathrm{d}\lambda\!\!\lambda}(x - \vartheta_n)\left(\frac{\mathrm{d}P_0}{\mathrm{d}\lambda\!\!\lambda}(x)\right)^{-1} = e^{\vartheta_n} \exp\left(-e^{-x}\left[e^{\vartheta_n} - 1\right]\right). \qquad \square$$

Beside the previous two (location) families, the following two examples correspond to scale families.

Example 2.9 (Fréchet distribution). Suppose that Assumption 2.1(iii) holds, where $(\Omega, \mathcal{A}) = (\mathbb{R}, \mathcal{B})$ and $P_0 := P_1 = P_n$ is a Fréchet distribution on $(0, \infty)$ with parameter $\alpha > 0$ for all $n \in \mathbb{N}$, i.e., for all $x > 0$

$$P_0(0, x] = \exp\left(-\frac{1}{x^\alpha}\right) \quad \text{and} \quad \frac{\mathrm{d}P_0}{\mathrm{d}\lambda\!\!\lambda}(x) = \alpha x^{-\alpha-1} \exp\left(-\frac{1}{x^\alpha}\right).$$

Moreover, suppose μ_n is the distribution, which is uniquely determined by

$$\mu_n(0, x] := P_0\left(0, \frac{x}{\vartheta_n}\right] \quad \text{for all } x > 0 \text{ and some } \vartheta_n > 0.$$

Then for all $x > 0$

$$f_n(x) = \vartheta_n^\alpha \exp\left(-\frac{1}{x^\alpha}\left(\vartheta_n^\alpha - 1\right)\right). \qquad \square$$

Example 2.10 (Exponential distribution). Denote by $Exp(\lambda)$ the exponential distribution with parameter $\lambda > 0$. Suppose that Assumption 2.1(iii) holds, where $(\Omega, \mathcal{A}) = (\mathbb{R}, \mathcal{B})$, $P_n = Exp(1) =: P_0$ and $\mu_n = Exp(1 + \vartheta_n)$ for some $\vartheta_n > -1$ and all $n \in \mathbb{N}$. Then for every $x > 0$

$$f_n(x) = (\vartheta_n + 1)\exp(-\vartheta_n x). \qquad \square$$

2.3. Exponential families

2.3.1. Introduction

In this section we focus on exponential families $(\mathcal{Q}_\vartheta)_{\vartheta\in\Theta}$, $0 \in \Theta \subset \mathbb{R}$, of the shape

$$\frac{\mathrm{d}\mathcal{Q}_\vartheta}{\mathrm{d}\mathcal{Q}_0}(x) = C(\vartheta)\,\exp\left(\vartheta\, h(x)\right),\ x \in \Omega.$$

Given such a family $(\mathcal{Q}_\vartheta)_{\vartheta\in\Theta}$, we are interested in the model which arises from Assumption 2.1(ii), where we set $P_n := \mathcal{Q}_0$ and $\mu_n := \mathcal{Q}_{\vartheta_n}$ for all $n \in \mathbb{N}$ and some sequence $(\vartheta_n)_{n\in\mathbb{N}}$ in Θ.

Assumption 2.11. *Let $(\mathcal{Q}_\vartheta)_{\vartheta\in\Theta}$, $0 \in \Theta \subset \mathbb{R}$, be a family of probability measures on some measurable space $(\Omega, \mathcal{A})$ with $\mathcal{Q}_\vartheta \ll \mathcal{Q}_0$ for all $\vartheta \in \Theta$. Furthermore, suppose that the $\mathcal{Q}_0$-density of $\mathcal{Q}_\vartheta$, $\vartheta \in \Theta$, is given by*

$$\frac{\mathrm{d}\mathcal{Q}_\vartheta}{\mathrm{d}\mathcal{Q}_0} \;=\; C(\vartheta)\ \exp\left(-\vartheta h\right), \tag{2.8}$$

where $h : \Omega \to \mathbb{R}$ is measurable and $C(\vartheta) < \infty$ for all $\vartheta \in \Theta$. Let $(\vartheta_n)_{n\in\mathbb{N}}$ be a sequence in Θ. Suppose that Assumption 2.1(ii) *holds with $\mu_n = \mathcal{Q}_{\vartheta_n}$ and $P_n = \mathcal{Q}_0$ for all $n \in \mathbb{N}$. Note that for all $n \in \mathbb{N}$*

$$f_n \;=\; \frac{\mathrm{d}\mathcal{Q}_{\vartheta_n}}{\mathrm{d}\mathcal{Q}_0}.$$

Remark 2.12. (i) (*Examples*) Clearly, the heterogeneous normal mixture model, i.e., the model given in Example 2.6 with $\tau = 1$, and the models given in Examples 2.8 to 2.10 fulfil Assumption 2.11. Consequently, light-tailed distributions, e.g., the normal and the exponential distribution, as well as heavy-tailed distributions, e.g., Fréchet distribution, belong to the class of exponential families introduced in Assumption 2.11.

(ii) It is easy to see that

$$\omega(\vartheta) := C(\vartheta)^{-1} \;=\; \int \exp\left(-\vartheta h\right)\ \mathrm{d}\mathcal{Q}_0 \ \text{ for all } \vartheta \in \Theta. \tag{2.9}$$

The function $\omega : \Theta \to \mathbb{R}$ is called the Laplace transform of $\mathcal{Q}_0$ with respect to h. In the following subsections we discuss some useful properties of this transform. For a deeper discussion of it we refer the reader to [22, 64]. In both references only the case $h(x) = x$, $x \in \Omega$, is treated and so a Laplace transform in their sense is one with respect to the identity function in our sense. But there is no loss of generality in assuming $h(x) = x$, $x \in \Omega$, because by the transformation formula we have for all $\vartheta \in \Theta$ and every $x \in \Omega$

$$\frac{\mathrm{d}\mathcal{Q}_\vartheta^h}{\mathrm{d}\mathcal{Q}_0^h}(x) = C(\vartheta) \exp\left(-\vartheta x\right).$$

Hence, the Laplace transform of $\mathcal{Q}_0$ with respect to h and the one of $\mathcal{Q}_0^h$ with respect to the identity are equal.

(iii) Suppose, in contrast to (2.8), that the $\mathcal{Q}_0$-density of $\mathcal{Q}_\vartheta$, $\vartheta \in \Theta$, is given by

$$\frac{\mathrm{d}\mathcal{Q}_\vartheta}{\mathrm{d}\mathcal{Q}_0}(x) \;=\; \widetilde{C}(\vartheta)\; \exp\Big(-q(\vartheta)h(x)\Big) \quad \text{for all } x \in \Omega, \qquad \square$$

where $q : \Theta \to q(\Theta) =: \widetilde{\Theta}$. If q is invertible, e.g., if q is strictly monotone, and $0 \in \widetilde{\Theta}$ then (2.8) holds for the family $(\widetilde{\mathcal{Q}}_\theta)_{\theta\in\widetilde{\Theta}}$ given by $\widetilde{\mathcal{Q}}_\theta := \mathcal{Q}_{q^{-1}(\theta)}$ for all $\theta \in \widetilde{\Theta}$. In this case we would analyse the family $(\widetilde{\mathcal{Q}})_{\theta\in\widetilde{\Theta}}$ first and transmit the corresponding results to the original family $(\mathcal{Q}_\vartheta)_{\vartheta\in\Theta}$ afterwards.

As we will see later, the asymptotic behaviour of $C(\vartheta_n)$ and $\omega(\vartheta_n)$ plays a crucial role. We distinguish between $\vartheta_n \to 0$ and $\vartheta_n \to \infty$, which correspond to the distinction between the dense and the sparse case, see Notation 2.5. For both cases we need some more specific assumptions, which we introduce in the corresponding subsections.

2.3.2. Sparse case: Abelian and Tauberian theorem

At the beginning of this section we present the definition and some properties of slowly varying functions. The terms *regularly* and *slowly varying functions* as used nowadays were initiated by Karamata [42]. For a deeper discussion of these functions and possible applications we refer the reader to Bingham et al. [8]. Important applications are the Abelian and Tauberian theorems. These theorems deal with the convergence of the Laplace transform ω, see (2.9). We present such a theorem of Feller [22] and some corollaries of it, which we apply in Sections 5.2.1 and 8.4.2. Moreover, we present the assumptions for the sparse case. We want to mention that the heterogeneous normal mixture model does not fulfil these assumptions. But the exponential distribution mixture model does. We will explain later, see, e.g., Remark 5.9, that the results concerning the log-likelihood ratio test and the higher criticism test can be transferred from the exponential distribution to the Fréchet and the Gumbel distribution. Consequently, by using our assumptions for the sparse case we can get results for light- and heavy-tailed distributions.

Definition 2.13. *A measurable function $L : (0,\infty) \to (0,\infty)$ varies slowly at infinity if and only if for every fixed $x > 0$*

$$\frac{L(xt)}{L(t)} \to 1 \text{ as } t \to \infty. \tag{2.10}$$

Karamata [42] studied slowly varying functions under the assumption that L is continuous. He proved the following uniform convergence theorem and the following property of L in Lemma 2.15 under this additional assumption. Korevaar et al. [46] showed these results for the general case of measurable functions L.

Theorem 2.14 (Uniform convergence theorem, 1.2.1 in [8]). *Let L be a slowly varying function at infinity. Then the convergence in* (2.10) *holds uniformly in x lying in a compact subinterval $[a,b]$ of $(0,\infty)$. That means in particular that for any function $\psi : (0,\infty) \to (0,\infty)$ satisfying*

$$\lim_{t\to\infty} \frac{\psi(t)}{t} = x \text{ for some } x \in (0,\infty)$$

we have

$$\frac{L(\psi(t))}{L(t)} \to 1 \text{ as } t \to \infty.$$

Lemma 2.15 (1.3.6(v) in [8]). *Let L be a slowly varying function at infinity. Then*

$$L(t)t^{\delta} \to \infty \text{ and } L(t)t^{-\delta} \to 0 \text{ as } t \to \infty$$

for every $\delta > 0$. Hence,

$$\frac{\log(L(t))}{\log(t)} \to 0 \text{ and } L(t) = t^{o(1)} \text{ as } t \to \infty.$$

Assumption 2.16 (Exponential family, sparse case). *Let Assumption 2.11 and the following three conditions* (i)-(iii) *be fulfilled.*

(i) *We have $\Theta = [0,\infty)$, $(\Omega,\mathcal{A}) = ([a,\infty),\mathcal{B}([a,\infty)))$ and $\mathcal{Q}_0(\{a\}) = 0$ for some $a \in \mathbb{R}$.*

(ii) *We have $\vartheta_n \to \infty$ as $n \to \infty$.*

(iii) *There exists some $\delta > 0$ such that h is strictly increasing and continuous in $[a, a+\delta]$,*

$$h(a) = 0 \text{ and } h(t) \geq h(a+\delta) \text{ for all } t \geq a+\delta.$$

Remark 2.17. (i) We can always assume without loss of generality that $h(a) = 0$, since otherwise we work with $\widetilde{h}$ given by $\widetilde{h}(t) := h(t) - h(a)$ for all $t \geq a$.

(ii) Under Assumption 2.16 $\mathcal{Q}_0^h$ is a measure on $[0, \infty)$, where $\mathcal{Q}_0^h(\{0\}) = 0$. Thus, $\mathcal{Q}_0^h$ can be treated as a measure on $((0, \infty), \mathcal{B}((0, \infty)))$ and, consequently, the theorems in [22] can be applied to it. □

Notation 2.18. Let $t^* \in \bar{\mathbb{R}}$.

(i) We write $t \searrow t^*$ and $t \nearrow t^*$ if $t^* < t \to t^*$ and $t^* > t \to t^*$, respectively.

(ii) We call $f, g : \mathbb{R} \to \mathbb{R}$ asymptotically equivalent as $t \to t^*$ (respectively, $t \searrow t^*$ and $t \nearrow t^*$), in symbols $f(t) \sim_{\text{asy}} g(t)$, if as $t \to t^*$ (respectively, $t \searrow t^*$ and $t \nearrow t^*$)

$$f(t) \;=\; g(t)\,(1 + o(1))\,.$$

□

Using the terminology of slowly varying functions we can formulate the Abelian and Tauberian theorem of Feller [22].

Theorem 2.19. *Let L be a slowly varying function at infinity and $p \in [0, \infty)$. Suppose that Assumption 2.16 holds. Let $\Gamma : (0, \infty) \to (0, \infty)$ be the gamma function, i.e.,*

$$\Gamma(s) \;=\; \int_0^\infty x^{s-1} e^x \,\mathrm{d}x \quad \textit{for all } s > 0. \tag{2.11}$$

(i) The following two statements (2.12) and (2.13) are equivalent:

$$\omega(\vartheta) \;\sim_{asy}\; \vartheta^{-p}\, L(\vartheta) \qquad \textit{as } \vartheta \to \infty \tag{2.12}$$

$$\textit{and} \;\; \mathcal{Q}_0^h[0, t] \;\sim_{asy}\; \frac{1}{\Gamma(p+1)}\, t^p\, L\left(\frac{1}{t}\right) \qquad \textit{as } t \searrow 0. \tag{2.13}$$

(ii) Suppose that for some $c, \nu > 0$

$$h(t) \sim_{asy} c(t-a)^\nu \tag{2.14}$$

$$\textit{and} \;\; \mathcal{Q}_0[a, t] \sim_{asy} (t-a)^{\nu p} L\left(\frac{1}{t-a}\right) \quad \textit{as } t \searrow a. \tag{2.15}$$

Then

$$\omega(\vartheta) \;\sim_{asy}\; c^{-p}\,\Gamma(p+1)\vartheta^{-p}\, L\left(\vartheta^{\frac{1}{\nu}}\right) \quad \textit{as } \vartheta \to \infty.$$

Remark 2.20. (i) The necessity of (2.12) for (2.13) is called an Abelian theorem. The sufficiency is called a Tauberian theorem.

(ii) Because $\mathcal{Q}_0^h$ and $\mathcal{Q}_0$ are measures we have

$$\mathcal{Q}_0^h[0,t] = o(1) \text{ and } \mathcal{Q}_0[a,a+t] = o(1) \text{ as } t \searrow 0.$$

Hence, if (2.13) or (2.15) holds for $p = 0$ then $L(t) = o(1)$ as $t \to \infty$. Combining this and Lemma 2.15 yields

$$\omega(\vartheta) \to 0 \;\text{ as } \vartheta \to \infty$$

under (2.13), as well as under (2.14) and (2.15). □

Proof of Theorem 2.19. Due to Remark 2.12(ii) and Remark 2.17(ii) the statement in (i) is an immediate consequence of Theorems XIII.5.2 and XIII.5.3 in [22]. Suppose that (2.14) and (2.15) are fulfilled for some $c, \nu > 0$. By (iii) of Assumption 2.16 the mapping $h_{|[a,a+\delta]}$ is invertible. Clearly, its inverse $h^{-1} : [0, h(a+\delta)] \to [a, a+\delta]$ is continuous and

$$h^{-1}(u) \;\sim_{\text{asy}}\; \left(\frac{u}{c}\right)^{\frac{1}{\nu}} + a \text{ as } u \searrow 0.$$

Combining this, Theorem 2.14 and (2.15) yields

$$\mathcal{Q}_0^h[0,t] = \mathcal{Q}_0\left[a, h^{-1}(t)\right] \;\sim_{\text{asy}}\; (h^{-1}(t)-a)^{\nu p} L\left(\frac{1}{h^{-1}(t)-a}\right) \;\sim_{\text{asy}}\; \left(\frac{t}{c}\right)^{p} L\left(t^{-\frac{1}{\nu}}\right)$$

as $t \searrow 0$. Finally, (ii) follows from (i). ■

The following lemma is needed in Section 5.2.1.

Lemma 2.21. *Let L be a slowly varying function at infinity and $p, M \in (0,\infty)$. Suppose that Assumption 2.16 and (2.13) hold. Let $(\psi_n)_{n\in\mathbb{N}\cup\{0\}}$ be a sequence of functions with domain $[0,\infty)$ taking values in $[-M, M]$ such that*

$$\begin{aligned} &\psi_n(x) \;=\; 0 \textit{ for all } x \geq M,\; n \in \mathbb{N}\cup\{0\} \;\textit{ and }\; \lambda\!\!\lambda(E) \;=\; 0, \\ &\textit{where } E := \{x \in [0,\infty) : \textit{ There exists a sequence } (x_n)_{n\in\mathbb{N}} \textit{ in } [0,\infty) \textit{ such that} \\ &\qquad\qquad \lim_{n\to\infty} x_n = x \textit{ but } \lim_{n\to\infty} \psi_n(x_n) \neq \psi_0(x)\}. \end{aligned}$$

Moreover, assume that

$$\int \psi_0(x)x^{p-1}\,\mathrm{d}x \;\neq\; 0.$$

Then

$$\int \psi_n\left(\vartheta_n h(x)\right)\,\mathrm{d}\mathcal{Q}_0(x) \;\sim_{asy}\; \frac{1}{\Gamma(p)}\,\vartheta_n^{-p} L\left(\vartheta_n\right)\int \psi_0(x)x^{p-1}\,\mathrm{d}x \text{ as } n \to \infty.$$

Remark 2.22. Note that $E \in \mathcal{B}([0,\infty))$, see, e.g., p. 226 in [7]. □

Proof. By (2.13) we have for all $x \in (0, M]$

$$\frac{\mathcal{Q}_0^h[0,\vartheta_n^{-1}x]}{\mathcal{Q}_0^h[0,\vartheta_n^{-1}M]} \to \frac{x^p}{M^p} \text{ as } n \to \infty.$$

For all sufficiently large $n \in \mathbb{N}$ let ν_n and ν be the (uniquely determined) probability measures on $([0,M],\mathcal{B}([0,M]))$ that for all $x \in [0,M]$

$$\nu_n[0,x] \;=\; \frac{\mathcal{Q}_0^h[0,\vartheta_n^{-1}x]}{\mathcal{Q}_0^h[0,\vartheta_n^{-1}M]} \quad\text{and}\quad \nu[0,x] \;=\; \frac{x^p}{M^p}.$$

Obviously,

$$\nu_n \xrightarrow{\text{w}} \nu \text{ as } n \to \infty.$$

By Theorem 5.5 of Billingsley [7], an extension of the continuous mapping theorem,

$$\nu_n^{\psi_n} \xrightarrow{\text{w}} \nu^{\psi_0} \text{ as } n \to \infty.$$

Consequently, for all continuous functions $g : \mathbb{R} \to \mathbb{R}$

$$\int g\,\mathrm{d}\nu_n^{\psi_n} \;=\; \int_{-M}^{M} g\,\mathrm{d}\nu_n^{\psi_n} \;\to\; \int_{-M}^{M} g\,\mathrm{d}\nu^{\psi_0} \;=\; \int g\,\mathrm{d}\nu^{\psi_0} \text{ as } n \to \infty.$$

Finally, combining this and (2.13) completes the proof:

$$\begin{aligned}
\int \psi_n\left(\vartheta_n h(x)\right)\,\mathrm{d}\mathcal{Q}_0(x) \;&=\; \mathcal{Q}_0^h\left[0, M\vartheta_n^{-1}\right]\int x\,\mathrm{d}\nu_n^{\psi_n}(x)\\
&\sim_{\text{asy}} \frac{M^p\vartheta_n^{-p}}{\Gamma(p+1)} L\left(\vartheta_n\right)\int x\,\mathrm{d}\nu^{\psi_0}(x)\\
&=\; \frac{p}{\Gamma(p+1)}\vartheta_n^{-p} L\left(\vartheta_n\right)\int \psi_0(x)x^{p-1}\,\mathrm{d}x \text{ as } n \to \infty.
\end{aligned}$$

■

2.3.3. Dense Case

In this section we discuss the dense case briefly, i.e., $\vartheta_n \to 0$. Note that the Laplace transform ω is analytic in a neighbourhood around 0 under certain conditions, see, e.g., [22, 64]. Thus, we know the asymptotic behaviour of $\omega(t)$ as $t \to 0$. Using Remark 2.12(ii) we can obviously extend the results in [22, 64] to our more general case that h is not necessarily equal to the identity function. We first introduce the assumptions for the dense case and then present the result about the Laplace transform mentioned before. We want to emphasise that the heterogeneous normal mixture model and the models given in Examples 2.8 to 2.10 fulfil our assumption for the dense case.

Assumption 2.23 (Exponential family, dense case). *Let Assumption 2.11 be fulfilled and suppose that*

$$(\Omega, \mathcal{A}) = (\mathbb{R}, \mathcal{B}), \ \ (-\varepsilon, \varepsilon) \subset \Theta \textit{ for some } \varepsilon > 0 \textit{ and } \vartheta_n \to 0 \quad \textit{as } n \to \infty.$$

Lemma 2.24. *Let Assumption 2.23 be fulfilled. The Laplace transform $\omega : \Theta \to (0, \infty)$ given by* (2.9) *is analytic in* $(-\varepsilon, \varepsilon)$ *with*

$$\omega^{(k)}(\vartheta) = \int (-h(x))^k \exp\left(-\vartheta h(x)\right) \, \mathrm{d}\mathcal{Q}_0(x) \quad \textit{for all } k \in \mathbb{N} \cup \{0\}.$$

Remark 2.25. We are particularly interested in the case $\vartheta = 0$. For this case we have

$$(-1)^k \omega^{(k)}(0) = \mathbb{E}_{\mathcal{Q}_0}\left(h^k\right) \text{ for all } k \in \mathbb{N} \cup \{0\}.$$

That means in particular that all moments of h are finite under $\mathcal{Q}_0$. □

2.4. The h-model

As mentioned in Chapter 1 our testing problem can also be interpreted as a multiple hypothesis testing problem for the global null. A great amount of multiple testing procedures are based on p-values including the famous one of Benjamini and Hochberg [6], and so does the higher criticism test, which we discuss elaborately in Part II. In this section we introduce a model for p-values, which is quite similar to the *chimeric alternatives* of Khmaladze [43]. We call this (new) model *h-model.*

The basic idea of our h-model was analogously used in extreme value theory and can be explained as follows. If extreme values are of interest, e.g., the flood of the ocean, then the information of interest is located in the tails and the specific shape of the distribution excluding the tails can be neglected. An appropriate model for this is, e.g., the *extreme value tangent model* of Janssen and Marohn [38]. In the context of multiple testing problems the interesting information is near 0. To be more specific, small p-values indicate that the alternative is true or, in our situation, a signal is present. A possible way to model p-values containing a signal is to consider the support of the corresponding distribution being on $(0, \delta)$ for small $\delta > 0$. Generalising this idea leads to our h-model, which we present in the following.

Assumption 2.26 (h-model). *Let $h : (0,1) \to [0, \infty)$ be a measurable function with*

$$\int_0^1 h^m \, \mathrm{d}\lambda\!\!\lambda = c_m \in (0, \infty) \quad \text{for } m \in \{1, 2\}, \quad \text{where } c_1 \leq 1. \tag{2.16}$$

Let $(\varepsilon_{n,i})_{1 \leq i \leq n \in \mathbb{N}}$, $(\tau_{n,i})_{1 \leq i \leq n \in \mathbb{N}}$ be triangular arrays of real numbers in $[0,1]$ such that

$$\begin{aligned} &\tau_{n,i} > 0 \text{ for all } 1 \leq i \leq n \in \mathbb{N} \\ \text{and} \quad &\max_{1 \leq i \leq n} \{\tau_{n,i} + \varepsilon_{n,i}\} \to 0 \quad \text{as } n \to \infty. \end{aligned} \tag{2.17}$$

For $1 \leq i \leq n \in \mathbb{N}$ define $h_{n,i} : (0,1) \to \mathbb{R}$ by

$$h_{n,i}(u) = -c_1 \mathbf{1}_{[\tau_{n,i}, 1)}(u) + \frac{1 - \tau_{n,i}}{\tau_{n,i}} h\left(\frac{u}{\tau_{n,i}}\right) \mathbf{1}_{(0, \tau_{n,i})}(u), \; u \in (0,1). \tag{2.18}$$

For $1 \leq i \leq n \in \mathbb{N}$ set $P_{n,i} := P_0 := \mathcal{U}(0,1)$ and define the probability measure $\mu_{n,i}$ by its $P_{n,i}$-density

$$\frac{\mathrm{d}\mu_{n,i}}{\mathrm{d}P_{n,i}} := f_{n,i} := 1 + h_{n,i}. \tag{2.19}$$

Remark 2.27. (i) By substitution it is easy to show that

$$\int_0^1 h_{n,i}(u) \, \mathrm{d}u = 0.$$

Combining this and $c_1 \leq 1$ yields that $\mu_{n,i}$ is indeed a probability measure with $\mu_{n,i} \ll P_{n,i}$. Obviously, Assumption 2.1(i) holds under Assumption 2.26.

(ii) If $c_1 = 1$ then the support of $\mu_{n,i}$ is a subset of or equal to $(0, \tau_{n,i})$ and becomes progressively smaller as $n \to \infty$. This corresponds to what we suggested in the introduction of this section. □

Part I.

Power of the log-likelihood ratio test

3. Introduction and motivation

Suppose that Assumption 2.1(i) holds. For all $\alpha \in [0,1]$ and every $n \in \mathbb{N}$ denote by $\Phi_{\alpha,n}$ the set containing all tests of nominal level α for the null $\mathcal{H}_{0,n}$, i.e.,

$$\Phi_{\alpha,n} := \left\{\varphi : (\Omega^n, \mathcal{A}^n) \to ([0,1], \mathcal{B}([0,1])) : \varphi \text{ is measurable and } \mathbb{E}_{P_{(n)}}(\varphi) \leq \alpha\right\}.$$

Let $\alpha \in [0,1]$ and $n \in \mathbb{N}$ be fixed. By the Neyman-Pearson Lemma it is known, see, e.g., Section 2.8 of [58], that

$$\sup_{\varphi \in \Phi_{\alpha,n}} \mathbb{E}_{Q_{(n)}}(\varphi) = \mathbb{E}_{Q_{(n)}}\left(\varphi^*_{n,\alpha}\right), \tag{3.1}$$

where $\varphi^*_{n,\alpha} : (\Omega^n, \mathcal{A}^n) \to ([0,1], \mathcal{B}([0,1]))$ is a log-likelihood ratio test, in short LLRT, of nominal level α. That means that

$$\begin{aligned} &\mathbb{E}_{P_{(n)}}(\varphi^*_{n,\alpha}) \leq \alpha \\ \text{and } \varphi^*_{n,\alpha} &= \left\{\begin{array}{l} 1 \\ 0 \end{array}\right. \log\left(\frac{\mathrm{d}Q_{(n)}}{\mathrm{d}P_{(n)}}\right) \begin{array}{l} > \\ < \end{array} c_{n,\alpha} \text{ for some } c_{n,\alpha} \in [-\infty, \infty], \end{aligned} \tag{3.2}$$

where the logarithm is canonically and continuously extended to $[0,\infty]$ and $\frac{\mathrm{d}Q_{(n)}}{\mathrm{d}P_{(n)}}$ is defined as in (A.18). In other words, $\varphi^*_{n,\alpha}$ is the best test for our testing problem (2.6). Thus, in this first main part of the thesis we discuss the asymptotic power of LLRT. That is why we are interested in necessary and sufficient conditions for the case that the test statistic LLR_n of LLRT, given by

$$LLR_n := \log\left(\frac{\mathrm{d}Q_{(n)}}{\mathrm{d}P_{(n)}}\right), \tag{3.3}$$

converges in distribution under the null $\mathcal{H}_{0,n}$ and under the alternative $\mathcal{H}_{1,n}$, respectively. At the beginning of Chapter 4 we explain that there is a connection between our purpose and weak convergence of the binary experiments $\{P_{(n)}, Q_{(n)}\}$. In Sections 4.2 and 4.3 we present how accumulation points of LLR_n (in the sense of convergence in distribution) can be determined under $\mathcal{H}_{0,n}$ and under $\mathcal{H}_{1,n}$, respectively. We distinguish between trivial and non-trivial accumulation points, both terms are explained in detail in Section 4.1. There are two trivial cases: $\mathcal{H}_{0,n}$ and $\mathcal{H}_{1,n}$ can be completely separated by LLRT (asymptotically) or $\mathcal{H}_{0,n}$ and $\mathcal{H}_{1,n}$ merge (asymptotically). In the non-trivial case LLRT can successfully, but not completely, separate $\mathcal{H}_{0,n}$ and $\mathcal{H}_{1,n}$ (asymptotically). In Section 4.4 we explain, as already mentioned in Section 2.2, that there is a connection between extreme value theory and our problem. In the last section of this chapter we present some preliminary results of a current research project, which completes in some sense the theory developed in Sections 4.2 and 4.3.

In Chapter 5 we apply the results to the examples introduced in Sections 2.2 to 2.4. We give the proofs for the dense heteroscedastic normal mixture model, which are omitted in [10]. We also apply our results to the great class of exponential families suggested by Cai and Wu [12]. Doing this we slightly extend their results, see Section 5.4. Finally, we discuss the exponential family model and the h-model, which are not discussed in this context in the literature until recently. Note that all our examples correspond to the univariate case, i.e., $(\Omega, \mathcal{A}) = (\mathbb{R}, \mathcal{B})$. In Chapter 1 we mentioned briefly that the multivariate case is also of interest and that there are first results for it, see [47, 62]. We want to emphasise that almost all results in this part are applicable for general measurable spaces $(\Omega, \mathcal{A})$ and in particular for the multivariate case.

4. Theoretical results

4.1. Connection to binary experiments

First, we explain that our problem to determine the asymptotic power of LLRT can be reformulated in terms of weak convergence of binary experiments. In Appendix A.3 we collect some important definitions and results about binary experiments for readers who are not familiar with them. Second, we present a tool based on Hellinger distances, see Definition and Lemma A.12, to distinguish between the non-trivial and trivial cases. Both terms are introduced more detailed in Remark 4.2. Using this tool we get a first impression which choices of $\varepsilon_{n,i}$, $P_{n,i}$ and $\mu_{n,i}$ (may) lead to trivial or non-trivial accumulation points.

Remark 4.1. Suppose Assumption 2.1(i) holds. Let $P_{(0)}$ and $Q_{(0)}$ be two further probability measures on $(\Omega, \mathcal{A})$. Moreover, let $\{k_n : n \in \mathbb{N}\}$ be a subsequence of $\mathbb{N}$, i.e.,

$$k_n \in \mathbb{N} \text{ and } k_n < k_{n+1} \text{ for all } n \in \mathbb{N}. \tag{4.1}$$

Let $\varphi^*_{n,\alpha} : (\Omega^n, \mathcal{A}^n) \to ([0,1], \mathcal{B}([0,1]))$ and $\varphi^*_{0,\alpha} : (\Omega, \mathcal{A}) \to ([0,1], \mathcal{B}([0,1]))$ be a log-likelihood ratio test for every $n \in \mathbb{N}$, $\alpha \in [0,1]$ such that (3.1) is fulfilled for it.

(i) By Theorem 16.10 in [58] $\{P_{(k_n)}, Q_{(k_n)}\}$ converges weakly to $\{P_{(0)}, Q_{(0)}\}$ as $n \to \infty$ if and only if

$$\lim_{n\to\infty} \mathbb{E}_{Q_{(k_n)}}\left(\varphi^*_{k_n,\alpha}\right) = \mathbb{E}_{Q_{(0)}}\left(\varphi^*_{0,\alpha}\right) \text{ for every } \alpha \in [0,1].$$

(ii) Suppose that $\{P_{(k_n)}, Q_{(k_n)}\}$ converges weakly to $\{P_{(0)}, Q_{(0)}\}$ as $n \to \infty$. Moreover, let $(\alpha_n)_{n\in\mathbb{N}}$ be a sequence of nominal levels such that

$$\mathbb{E}_{P_{(k_n)}}(\varphi^*_{k_n,\alpha_n}) = \alpha_n \to \alpha \in [0,1] \quad \text{as } n \to \infty.$$

Then by Corollary 15.11 in [58]

$$\lim_{n\to\infty} \mathbb{E}_{Q_{(k_n)}}\left(\varphi^*_{k_n,\alpha_n}\right) = \mathbb{E}_{Q_{(0)}}\left(\varphi^*_{0,\alpha}\right) \text{ for every } \alpha \in [0,1].$$

(iii) Note that $1 - \varphi^*_{n,\alpha}$ is a log-likelihood ratio test for the testing problem

$$\widetilde{\mathcal{H}}_{0,n} : Q_{(n)} \quad \text{versus} \quad \widetilde{\mathcal{H}}_{1,n} : P_{(n)}$$

for all $n \in \mathbb{N}$. Thus, if $\{P_{(k_n)}, Q_{(k_n)}\}$ converges weakly to $\{P_{(0)}, Q_{(0)}\}$ as $n \to \infty$ then combining (i) and (ii) yields

$$\lim_{n\to\infty} \mathbb{E}_{P_{(k_n)}} \left(\varphi^*_{k_n,\alpha}\right) = \mathbb{E}_{P_{(0)}} \left(\varphi^*_{0,\alpha}\right) \text{ for every } \alpha \in [0,1]. \qquad \square$$

Here and subsequently, we mainly distinguish between three cases: two trivial and one non-trivial case, which are introduced in the following remark.

Remark 4.2. Suppose that the assumptions of Remark 4.1 hold and that $\{P_{(k_n)}, Q_{(k_n)}\}$ converges weakly to $\{P_{(0)}, Q_{(0)}\}$ as $n \to \infty$.

(i) If $\{P_{(0)}, Q_{(0)}\}$ is equivalent to the so called uninformative experiment $\{\epsilon_0, \epsilon_0\}$ then by Remark 4.1 (ii) and by using subsequence arguments the sum of type I and type II error probabilities of $\varphi^*_{k_n,\alpha_n}$ tends to 1 for every $\alpha \in [0,1]$, i.e.,

$$\lim_{n\to\infty} \mathbb{E}_{P_{(k_n)}} \left(\varphi^*_{k_n,\alpha_n}\right) + \mathbb{E}_{Q_{(k_n)}} \left(1 - \varphi^*_{k_n,\alpha_n}\right) = 1.$$

Thus, LLRT yields no better results than the test $\varphi \equiv \alpha$ (asymptotically) and so it is asymptotically useless.

(ii) Let $c \in \mathbb{R}$. Moreover, for all $n \in \mathbb{N}$ let $\varphi^*_{k_n,\alpha_n}$ be a log-likelihood ratio test with nominal level $\alpha_n \in [0,1]$ such that the critical value c_{k_n,α_n} equals c, see (3.2). Assume that $\{P_{(0)}, Q_{(0)}\}$ is equivalent to the so called full informative experiment $\{\epsilon_{-\infty}, \epsilon_\infty\}$, i.e., they have the same standard form. Then it is known, see Definition A.19 and Corollary A.20, that

$$\mathcal{L}\left(LLR_n | P_{(n)}\right) \xrightarrow{w} \mathcal{L}\left(\log\left(\frac{\mathrm{d}P_{(0)}}{\mathrm{d}Q_{(0)}}\right) \Big| P_{(0)}\right) = \mathcal{L}\left(\log\left(\frac{\mathrm{d}\epsilon_{-\infty}}{\mathrm{d}\epsilon_\infty}\right) \Big| \epsilon_{-\infty}\right) = \epsilon_{-\infty}$$

and $\mathcal{L}\left(LLR_n | Q_{(n)}\right) \xrightarrow{w} \epsilon_\infty$ as $n \to \infty$.

Consequently, the sum of type I and type II error probabilities of $\varphi^*_{k_n,\alpha_n}$ tends to 0. Thus, $\mathcal{H}_{0,n}$ and $\mathcal{H}_{1,n}$ can be completely separated by using LLRT (asymptotically).

(iii) Suppose that $\{P_{(0)}, Q_{(0)}\}$ is not equivalent to $\{\epsilon_{-\infty}, \epsilon_\infty\}$ and not equivalent to $\{\epsilon_0, \epsilon_0\}$. Then $P_{(0)}$ and $Q_{(0)}$ are neither singular to each other nor equal. Hence, we can conclude from Lemmas A.13 and A.22 that for every sequence $(\alpha_n)_{n\in\mathbb{N}}$ of nominal levels

$$\limsup_{n\to\infty} \mathbb{E}_{P_{(k_n)}}\left(\varphi^*_{k_n,\alpha_n}\right) + \mathbb{E}_{Q_{(k_n)}}\left(1-\varphi^*_{k_n,\alpha_n}\right) \geq 1 - ||Q_{(0)} - P_{(0)}|| > 0.$$

Furthermore, by Remark 4.1(i) the sum of type I and type II error probabilities of $\varphi^*_{k_n,\alpha}$ tends to some $C \in (0,1)$ for some $\alpha \in [0,1]$. Consequently, LLRT can successfully but not completely separate $\mathcal{H}_{0,n}$ and $\mathcal{H}_{1,n}$ (asymptotically). □

A first tool to distinguish between the three cases is based on Hellinger distances. Some definitions and results of different distances for probability measures, among others of the Hellinger distance d, are collected in Appendix A.2. Under Assumption 2.1(i) we define

$$D_n := D_n\left(P_{(n)}, Q_{(n)}\right) := \sum_{i=1}^{n} d^2\left(P_{n,i}, Q_{n,i}\right). \tag{4.2}$$

Lemma 4.3. *Let $\{k_n : n \in \mathbb{N}\}$ be a subsequence of $\mathbb{N}$ and Assumption 2.1*(i) *be fulfilled.*

(i) The binary experiment $\left\{P_{(k_n)}, Q_{(k_n)}\right\}$ converges weakly to the uninformative experiment $\{\epsilon_0, \epsilon_0\}$ if and only if $\lim_{n\to\infty} D_{k_n} = 0$.

(ii) The binary experiment $\left\{P_{(k_n)}, Q_{(k_n)}\right\}$ converges weakly to the full informative experiment $\{\epsilon_{-\infty}, \epsilon_\infty\}$ if and only if $\lim_{n\to\infty} D_{k_n} = \infty$.

(iii) If $\left\{P_{(k_n)}, Q_{(k_n)}\right\}$ converges weakly to some binary experiment $\left\{\widetilde{P}, \widetilde{Q}\right\}$, which is neither full informative nor uninformative, then $\lim_{n\to\infty} D_{k_n} \in (0,\infty)$.

(iv) All accumulation points of $\{P_{(n)}, Q_{(n)}\}$ are neither full informative nor uninformative if and only if $0 < \liminf_{n\to\infty} D_n \leq \limsup_{n\to\infty} D_n < \infty$.

Remark 4.4. By Lemma A.23 every sequence of binary experiments has at least one accumulation point. □

Proof. Set $h_{k_n,i} := d^2\left(P_{k_n,i}, Q_{k_n,i}\right)$ for all $1 \leq i \leq k_n$. Then

$$D_{k_n} = \sum_{i=1}^{k_n} h_{k_n,i} \quad \text{and} \quad \prod_{i=1}^{k_n}(1-h_{k_n,i}) = 1 - d^2(P_{(k_n)}, Q_{(k_n)})$$

for all $n \in \mathbb{N}$. Moreover, by Lemma A.14

$$\max_{1\leq i\leq k_n} h_{k_n,i} \leq \max_{1\leq i\leq k_n} ||P_{k_n,i} - Q_{k_n,i}|| \leq \max_{1\leq i\leq k_n} \varepsilon_{k_n,i}\, ||P_{k_n,i} - \mu_{k_n,i}|| \leq \varepsilon_{k_n:k_n} \to 0$$

as $n \to \infty$. Combining this, some basic calculations, Definition and Lemma A.12(iii), Lemma A.21 and Lemma A.28 yields (i)-(iii). Clearly, (iv) follows from (i) and (ii). ■

In general, it is difficult to calculate D_n exactly. In most cases, it is only possible to get lower and upper bounds for D_n. Consequently, it is not a satisfying tool for our purpose, in particular when we want to determine non-trivial accumulation points. Nevertheless, we can use it to get a first idea which choices for $\varepsilon_{n,i}$, $P_{n,i}$, $\mu_{n,i}$ (may) lead to (non-)trivial accumulation points of $\{P_{(n)}, Q_{(n)}\}$.

Remark 4.5. Suppose that Assumption 2.1(i) holds. By Lemma A.14

$$\frac{1}{2}\sum_{i=1}^{n} \varepsilon_{n,i}^2 ||P_{n,i} - \mu_{n,i}||^2 \ \leq D_n \leq \ \sum_{i=1}^{n} \varepsilon_{n,i} ||P_{n,i} - \mu_{n,i}|| \text{ for all } n \in \mathbb{N}. \tag{4.3}$$

Let $\{k_n : n \in \mathbb{N}\}$ be a subsequence of $\mathbb{N}$. Combining (4.3) and Lemma 4.3 yields:

(i) If

$$\lim_{n\to\infty} \sum_{i=1}^{n} \varepsilon_{n,i} = 0$$

then $\left\{P_{(n)}, Q_{(n)}\right\}$ converges weakly to the uninformative experiment. If

$$\limsup_{n\to\infty} \sum_{i=1}^{n} \varepsilon_{n,i} < \infty$$

then no accumulation point of $\{P_{(n)}, Q_{(n)}\}$ is full informative. Both implications hold independently of how the measures $P_{n,i}$ and $\mu_{n,i}$ are chosen.

(ii) Let us assume that one of the two following conditions, which are equivalent according to Lemma A.14, holds:

$$\liminf_{n\to\infty} \min_{i=1,\dots,k_n} ||P_{k_n,i} - \mu_{k_n,i}|| > 0$$
$$\text{or } \liminf_{n\to\infty} \min_{i=1,\dots,n} d(P_{k_n,i}, \mu_{k_n,i}) > 0.$$

This situation corresponds to the sparse case introduced in Notation 2.5.

(a) If $\{P_{(k_n)}, Q_{(k_n)}\}$ converges weakly to $\{\epsilon_0, \epsilon_0\}$ then

$$\lim_{n\to\infty} \sum_{i=1}^{k_n} \varepsilon_{k_n,i}^2 = 0.$$

(b) If $\{P_{(k_n)}, Q_{(k_n)}\}$ converges weakly to $\{\epsilon_{-\infty}, \epsilon_\infty\}$ then

$$\lim_{n\to\infty} \sum_{i=1}^{k_n} \varepsilon_{k_n,i} = \infty.$$

(c) If

$$\lim_{n\to\infty} k_n^{-\frac{1}{2}} \sum_{i=1}^{k_n} \varepsilon_{k_n,i} = \infty$$

then by Lemma A.30

$$\lim_{n\to\infty} \sum_{i=1}^{k_n} \varepsilon_{k_n,i}^2 = \infty$$

and, hence, $\{P_{(k_n)}, Q_{(k_n)}\}$ converges weakly to $\{\epsilon_{-\infty}, \epsilon_\infty\}$.

We can conclude among others from the above: if $\{P_{(n)}, Q_{(n)}\}$ has a non-trivial accumulation point $\{P, Q\}$ we can conclude from (i), (iic) and Lemma A.30 that

$$\limsup_{n\to\infty} \sum_{i=1}^{n} \varepsilon_{n,i} > 0, \quad \limsup_{n\to\infty} n \sum_{i=1}^{n} \varepsilon_{n,i}^2 > 0,$$
$$\liminf_{n\to\infty} n^{-\frac{1}{2}} \sum_{i=1}^{n} \varepsilon_{n,i} < \infty \quad \text{and} \quad \liminf_{n\to\infty} \sum_{i=1}^{n} \varepsilon_{n,i}^2 < \infty. \tag{4.4}$$

For the simplest case, that $\varepsilon_{n,i} = \varepsilon_n$ for all $1 \le i \le n$ and $\varepsilon_n \le \varepsilon_{n-1}$ for all sufficiently large $n \in \mathbb{N}$, we can simplify (4.4): there exist some C_1, $C_2 > 0$ such that

$$C_1\, n \le \varepsilon_n \le C_2\, n^{-\frac{1}{2}} \quad \text{for sufficiently large } n \in \mathbb{N}. \qquad \square$$

4.2. Non-trivial accumulation points

Due to Section 4.1 we want to find necessary and sufficient conditions for the case that $\{P_{(n)}, Q_{(n)}\}$ or at least $\{P_{(k_n)}, Q_{(k_n)}\}$ converges weakly to a non-trivial binary experiment. Note that Lemma 4.3 gives us only a necessary condition for that issue. The convergence to trivial binary experiments is the topic of the upcoming section. For the purpose mentioned above we need to examine the asymptotic behaviour of LLR_n in the sense of convergence in distribution under $\mathcal{H}_{0,n}$ and under $\mathcal{H}_{1,n}$, respectively. Here, we are only interested in the case that the limiting random variables of LLR_n, or in general of LLR_{k_n}, are real-valued. If the limiting random variables are real-valued then we show that the random variables have infinitely divisible distributions. All results and definitions in the context of infinitely divisible distributions, which are needed here, are collected in Appendix A.1. For a deeper discussion we refer the reader to the book of Petrov [51], on which Appendix A.1 is based. Finally, we want to mention that in general the limiting random variables of LLR_n take values in $\bar{\mathbb{R}}$. Preliminary results for this general case are presented and discussed in Section 4.5.

Here and subsequently, $X_{n,j} : \Omega^n \to \Omega$ denotes the projection to coordinate j for all $1 \le j \le n \in \mathbb{N}$, i.e.,

$$X_{n,j}(\omega) \;=\; \omega_j \;\text{ for all }\; \omega = (\omega_1, \dots, \omega_n) \in \Omega^n. \tag{4.5}$$

The following Condition (**A**), in short (**A**), describes the case on which we are interested.

Condition (A). *Suppose Assumption 2.1*(i). *Let $\mathcal{P}$ be a probability measure on $(\Omega, \mathcal{A})$. Furthermore, let $\{k_n : n \in \mathbb{N}\}$ be a subsequence of $\mathbb{N}$ and ξ_1, ξ_2 be real-valued random variables on $(\Omega, \mathcal{A}, \mathcal{P})$ such that*

$$L_{k_n} := \sum_{i=1}^{k_n} \log\left(g_{k_n,i}\left(X_{k_n,i}\right)\right) \xrightarrow[\text{under } P_{(k_n)}]{D} \xi_1 \quad \text{as } n \to \infty \tag{\textbf{A1}}$$

$$\text{and } \; L_{k_n} := \sum_{i=1}^{k_n} \log\left(g_{k_n,i}\left(X_{k_n,i}\right)\right) \xrightarrow[\text{under } Q_{(k_n)}]{D} \xi_2 \quad \text{as } n \to \infty. \tag{\textbf{A2}}$$

Remark 4.6. By Definition A.19 and Corollary A.20 the condition (**A2**) holds for some random variable $\xi_2 : (\Omega, \mathcal{A}, \mathcal{P}) \to (\bar{\mathbb{R}}, \bar{\mathcal{B}})$ if and only if (**A1**) does so for some random variable $\xi_1 : (\Omega, \mathcal{A}, \mathcal{P}) \to (\bar{\mathbb{R}}, \bar{\mathcal{B}})$. Moreover, if $\{P_{(k_n)}, Q_{(k_n)}\}$ converges weakly to $\{P, Q\}$ then by A.18-A.20 the conditions (**A1**) and (**A2**) hold for

$$\xi_1 \sim \mathcal{L}\Big(\log\left(\frac{\mathrm{d}Q}{\mathrm{d}P}\right) \Big| P\Big) =: \nu_1, \quad \xi_2 \sim \mathcal{L}\Big(\log\left(\frac{\mathrm{d}Q}{\mathrm{d}P}\right) \Big| Q\Big) =: \nu_2 \quad \text{and} \quad \frac{\mathrm{d}\nu_2}{\mathrm{d}\nu_1} \;=\; \exp.$$

Note that in this case ξ_1 and ξ_2 are not necessarily real-valued. In this section we focus on the case that ξ_1 and ξ_2 are real-valued, whereas preliminary results for the general case are presented in Section 4.5. □

Due to Remark 4.6 it is sufficient to show that (**A1**) holds for a real-valued ξ_1 and to ensure that ξ_2 is also real. For the latter we use the first lemma of Le Cam, see Lemma A.26. But first we show that the condition of infinitely smallness, see Definition A.5, is fulfilled for the triangular array $(Y_{n,i})_{1\le i\le n\in\mathbb{N}}$ given by $Y_{n,i} := \log(g_{n,i}(X_{n,i}))$ for all $1 \le i \le n$ under the null and under the alternative.

Lemma 4.7. *Under Assumption 2.1*(i) *we have for every* $\delta > 0$

$$\max_{1\le i\le n}\Big\{P_{n,i}\Big(|\log(g_{n,i})| \ge \delta\Big)\Big\} + \max_{1\le i\le n}\Big\{Q_{n,i}\Big(|\log(g_{n,i})| \ge \delta\Big)\Big\} \to 0 \quad \textit{as } n \to \infty.$$

Proof. Note that

$$g_{n,i} = (1-\varepsilon_{n,i}) + \varepsilon_{n,i} f_{n,i} \ge (1-\varepsilon_{n:n}) = 1 + o(1) \text{ as } n \to \infty. \tag{4.6}$$

Hence, from Remark 2.2

$$\begin{aligned}
&\max_{1\le i\le n}\Big\{P_{n,i}\Big(|\log(g_{n,i})| \ge \delta\Big)\Big\}\\
&\le \max_{1\le i\le n}\Big\{P_{n,i}\Big(g_{n,i} \ge \exp(\delta)\Big) + P_{n,i}\Big((1-\varepsilon_{n:n}) \le \exp(-\delta)\Big)\Big\}\\
&\le \max_{1\le i\le n}\Big\{P_{(n)}\Big(\varepsilon_{n,i} f_{n,i} \ge \exp(\delta) - 1\Big)\Big\} + o(1) \;=\; o(1) \quad \text{as } n \to \infty.
\end{aligned}$$

Moreover, note that for every triangular array $(B_{n,i})_{i\le n}$ of sets in $\mathcal{A}$ we have as $n \to \infty$

$$\max_{1\le i\le n}\{Q_{n,i}(B_{n,i})\} \le \max_{1\le i\le n}\{P_{n,i}(B_{n,i}) + \varepsilon_{n:n}\} = \max_{1\le i\le n}\{P_{n,i}(B_{n,i})\} + o(1). \qquad \blacksquare$$

Thus, we can conclude from Theorem A.6:

Corollary 4.8. *Suppose that (**A**) holds for a subsequence* $\{k_n : n \in \mathbb{N}\}$ *of* $\mathbb{N}$ *and real-valued random variables* ξ_1 *and* ξ_2*. Then the distribution of* ξ_j *is infinitely divisible and, thus, it is uniquely determined by its Lévy characteristic* $(\gamma_j, \sigma_j^2, \eta_j)$ *for* $j = 1, 2$*. Moreover, the characteristic function of* ξ_j *is given by* (A.2) *for* $j = 1, 2$.

A first quite simple observation is that $(-\infty, 0)$ is a null set of η_1 and η_2 under (**A**).

Lemma 4.9. *Suppose that Assumption 2.1*(i) *holds. Then for every* $y < 0$

$$\bigcup_{i=1}^{n}\{x \in \Omega : \log(g_{n,i}(x)) \le y\} \;=\; \emptyset \quad \textit{if } n \in \mathbb{N} \textit{ is sufficiently large.}$$

*If (**A**) holds and* η_1*,* η_2 *are the corresponding Lévy measures, see Corollary 4.8, then*

$$\eta_1(-\infty, 0) \;=\; 0 \;=\; \eta_2(-\infty, 0).$$

Proof. From (4.6) we obtain for every fixed $y < 0$ and all sufficiently large $n \in \mathbb{N}$

$$\bigcup_{i=1}^{n} \{x \in \Omega : \log(g_{n,i}(x)) \leq y\} \subseteq \{x \in \Omega : \log(1 - \varepsilon_{n:n}) \leq y\} = \emptyset. \tag{4.7}$$

Suppose that (**A**) is fulfilled. Let η_1 and η_2 be the corresponding Lévy measures, see Corollary 4.8. Combining Lemma 4.7 and (A.4) of Theorem A.7 yields

$$\eta_j(-\infty, y] \ = \ 0 \text{ for all } y < 0 \text{ and so } \ \eta_j(-\infty, 0) = 0.$$ ■

Janssen et al. [39] discussed binary and more general statistical experiments, which have infinitely divisible limiting experiments. The following conclusion of Discussion (8.3), Remark (8.4) and Lemma (8.7)(c) from [39] can be used to determine further properties of the Lévy characteristics under (**A**).

Lemma 4.10. *Let $C^2_{loc}(\mathbb{R})$ denote the set of all continuous and bounded functions $f : \mathbb{R} \to \mathbb{R}$ which are twice continuously differentiable in an open subset $U \ni 0$ of $\mathbb{R}$. Let the probability measure ν on $(\mathbb{R}, \mathcal{B})$ be infinitely divisible with Lévy characteristic (γ, σ^2, η).*

(i) The Lévy characteristic of ν is uniquely determined by the generating functional $A : C^2_{loc}(\mathbb{R}) \to \mathbb{R}$, which admits

$$A(f) \ = \ \gamma f^{(1)}(0) + \frac{\sigma^2}{2} f^{(2)}(0) + \int_{\mathbb{R}\setminus\{0\}} \left(f(x) - f(0) - \frac{f^{(1)}(0)x}{1+x^2} \right) \mathrm{d}\eta(x) \tag{4.8}$$

for all $f \in C^2_{loc}(\mathbb{R})$. Moreover, if (4.8) holds for some constants $\widetilde{\gamma} \in \mathbb{R}$, $\widetilde{\sigma}^2 \geq 0$ and some Lévy measure η then $(\gamma, \sigma^2, \eta) = (\widetilde{\gamma}, \widetilde{\sigma}^2, \widetilde{\eta})$.

(ii) Assume

$$\int \exp \, \mathrm{d}\nu = 1. \tag{4.9}$$

Then the probability measure $\widetilde{\nu}$ given by its ν-density

$$\frac{\mathrm{d}\widetilde{\nu}}{\mathrm{d}\nu} \ = \ \exp$$

is infinitely divisible with Lévy characteristic $(\widetilde{\gamma}, \widetilde{\sigma}^2, \widetilde{\eta})$, where $\widetilde{\sigma}^2 = \sigma^2$. Moreover, the Lévy measures η and $\widetilde{\eta}$ are mutually continuous with

$$\frac{\mathrm{d}\widetilde{\eta}}{\mathrm{d}\eta} \ = \ \exp$$

and we have

$$\widetilde{\gamma} \;=\; \gamma + \sigma^2 + \int_{\mathbb{R}\setminus\{0\}} (e^x - 1)\, \frac{x}{1+x^2}\, \mathrm{d}\eta(x) \in \mathbb{R}. \tag{4.10}$$

(iii) The equation in (4.9) *holds if and only if*

$$\int_{(-\infty,-1)\cup(1,\infty)} \exp\; \mathrm{d}\eta \;<\; \infty \tag{4.11}$$

$$\textit{and} \;\; \gamma + \frac{\sigma^2}{2} + \int_{\mathbb{R}\setminus\{0\}} \left(e^x - 1 - \frac{x}{x^2+1}\right) \mathrm{d}\eta(x) \;=\; 0. \tag{4.12}$$

Remark 4.11. Let $f \in C^2_{loc}(\mathbb{R})$ and $\varepsilon > 0$ such that $[-\varepsilon, \varepsilon] \subset U$. By Taylor's formula there exists some $\kappa_x \in [-\varepsilon, \varepsilon]$ for every $x \in [-\varepsilon, \varepsilon]$ such that

$$f(x) - f(0) \;=\; f^{(1)}(0)x + f^{(2)}(\kappa_x)\frac{x^2}{2}.$$

Because $f^{(2)}$ is continuous on $[-\varepsilon, \varepsilon]$ there exists $C > 0$ such that for every $x \in [-\varepsilon, \varepsilon]$

$$\left| f(x) - f(0) - \frac{f^{(1)}(0)x}{x^2+1} \right| \;=\; \left| f^{(2)}(\kappa_x)\frac{x^2}{2} + f^{(1)}(0)\frac{x + x^3 - x}{1+x^2} \right| \leq C\, x^2. \tag{4.13}$$

Since f is bounded, combining (4.13), the basic property (A.1) of Lévy measures and (4.13) shows that the integral from (4.8) is indeed finite. By similar arguments it can be shown that the integral from (4.12) is finite if (4.11) holds. Furthermore, if (4.9) holds and $\widetilde{\eta}$ is the Lévy measure from Lemma 4.10(ii) then

$$\begin{aligned}
0 &\leq \int_{\mathbb{R}\setminus\{0\}} |e^x - 1|\, \frac{|x|}{1+x^2}\, \mathrm{d}\eta(x) \\
&\leq \int_{[-1,1]\setminus\{0\}} \left|\frac{e^x - 1}{x}\right| x^2\, \mathrm{d}\eta(x) + \int_{(-\infty,-1)} 1\, \mathrm{d}\eta(x) + \int_{(1,\infty)} e^x\, \mathrm{d}\eta(x) \\
&\leq 2\int_{[-1,1]\setminus\{0\}} x^2\, \mathrm{d}\eta(x) + \eta(-\infty,-1) + \widetilde{\eta}(1,\infty) < \infty.
\end{aligned}$$

Hence, the integral from (4.10) is also finite if (4.9) holds. □

Proof. Combining Discussion (8.3) and Remark (8.4) of [39] proves (i). Suppose that (4.9) holds. By Lemma (8.7)(b) of [39]

$$0 \;=\; A(\exp) \;=\; \gamma + \frac{\sigma^2}{2} + \int_{\mathbb{R}\setminus\{0\}} \left(e^x - 1 - \frac{x}{x^2+1}\right) \mathrm{d}\eta(x), \tag{4.14}$$

where A can be extended canonically so that it can be evaluated at $f = \exp$, see Remark

(8.6) of [39]. Moreover, by (4.14) and Lemma (8.7)(c) of [39] the generating functional $\widetilde{A}$ of $\widetilde{\nu}$ is given for every $f \in C^2_{loc}(\mathbb{R})$ by

$$\begin{aligned}
\widetilde{A}(f) &= A(f \cdot \exp) \\
&= \gamma \left[f^{(1)}(0) + f(0) \right] + \frac{\sigma^2}{2} \left[f^{(2)}(0) + 2 f^{(1)}(0) + f(0) \right] \\
&\quad + \int_{\mathbb{R}\setminus\{0\}} \left(f(x) e^x - f(0) - \frac{\left[f^{(1)}(0) + f(0) \right] x}{1+x^2} \right) \, \mathrm{d}\eta(x) \\
&= \gamma f^{(1)}(0) + \frac{\sigma^2}{2} \left[f^{(2)}(0) + 2 f^{(1)}(0) \right] + \int_{\mathbb{R}\setminus\{0\}} \left[(f(x) - f(0)) e^x - \frac{f^{(1)}(0) x}{1+x^2} \right] \mathrm{d}\eta(x) \\
&= \widetilde{\gamma} f^{(1)}(0) + \frac{\widetilde{\sigma}^2}{2} f^{(2)}(0) + \int_{\mathbb{R}\setminus\{0\}} \left[f(x) - f(0) - \frac{f^{(1)}(0) x}{1+x^2} \right] \mathrm{d}\widetilde{\eta}(x).
\end{aligned}$$

Thus, (ii) follows. Respecting (4.14) we conclude (iii) from Lemma (8.7)(b) of [39]. ■

Applying, among others, the first lemma of Le Cam, see Lemma A.26, we get the following corollary.

Corollary 4.12. *Suppose (**A**) and let $(\gamma_j, \sigma_j^2, \eta_j)$ be the Lévy characteristic of ξ_j for $j = 1, 2$. Then the following statements* (i)-(iii) *hold.*

(i) $(P_{(k_n)})_{n\in\mathbb{N}}$ *and* $(Q_{(k_n)})_{n\in\mathbb{N}}$ *are mutually contiguous, in symbols* $P_{(k_n)} \triangleleft \triangleright Q_{(k_n)}$ *(see Definition A.24).*

(ii) The measures η_1 *and* η_2 *are mutually absolutely continuous. That means in particular that their continuity points coincide, in symbols* $C(\eta_1) = C(\eta_2)$*. Furthermore,*

$$\frac{\mathrm{d}\eta_2}{\mathrm{d}\eta_1} = \exp\,. \tag{4.15}$$

(iii) We have

$$\sigma_1^2 \;=\; \sigma_2^2 \;=:\; \sigma^2.$$

Moreover,

$$\begin{aligned}
\gamma_1 &= -\frac{\sigma^2}{2} + \int_{(0,\infty)} \left(1 - e^x + \frac{x}{1+x^2} \right) \mathrm{d}\eta_1(x) \\
\text{and} \quad \gamma_2 &= \frac{\sigma^2}{2} + \int_{(0,\infty)} \left(1 - e^x + \frac{x}{1+x^2} e^x \right) \mathrm{d}\eta_1(x).
\end{aligned} \tag{4.16}$$

Proof. Let ν_j be the distribution of ξ_j for $j = 1, 2$. Obviously, we have

$$\nu_j(\mathbb{R}) = 1 \text{ and so } \nu_j(\{-\infty, \infty\}) = 0 \text{ for } j = 1, 2. \tag{4.17}$$

Therefore, (i) and

$$\int_{\mathbb{R}} \exp \, d\nu_1 = 1 \tag{4.18}$$

follow from Lemma A.18 and the first lemma of Le Cam, see Lemma A.26. Finally, combining (4.18), Lemma 4.10(ii) and (iii) yields (ii) and (iii). ■

We want to remind the reader that due to Remark 4.6 it is sufficient to verify (**A1**) for some real-valued ξ_1 and to ensure that ξ_2 is also real. By the first lemma of Le Cam the latter is fulfilled if and only if

$$\int_{\mathbb{R}} \exp \, d\nu_1 = 1,$$

where $\xi_1 \sim \nu_1$. Finally, combining this, Remark 4.6, Lemma 4.9 and Lemma 4.10(iii) yields immediately:

Corollary 4.13. *Suppose Assumption 2.1*(i). *(**A**) holds for real-valued random variables ξ_1 and ξ_2 if and only if (**A1**) holds for real-valued ξ_1 with Lévy characteristic* $(\gamma_1, \sigma^2, \eta_1)$ *such that*

$$\int_{(1,\infty)} \exp \, \mathrm{d}\eta_1 < \infty \tag{4.19}$$

$$\text{and} \quad \gamma_1 + \frac{\sigma^2}{2} + \int_{(0,\infty)} \left(e^x - 1 - \frac{x}{x^2+1} \right) \mathrm{d}\eta_1(x) = 0. \tag{4.20}$$

Now we are interested in sufficient and necessary conditions for the case that (**A1**), (4.19) and (4.20) hold for some real-valued random variable ξ_1 with Lévy characteristic $(\gamma_1, \sigma^2, \eta_1)$. By applying Theorem A.7 and simplifying the conditions corresponding to (a)-(c) we will show that the following Condition (**B**), in short (**B**), is a possible option for this purpose.

Condition (B). *Let Assumption 2.1*(i) *be fulfilled. Let* $\sigma^2 \geq 0$, $\{k_n : n \in \mathbb{N}\}$ *be a subsequence of* $\mathbb{N}$ *and* η_1, η_2 *be measures on* $(\mathbb{R} \setminus \{0\}, \mathcal{B}(\mathbb{R} \setminus \{0\}))$ *such that*

$$\eta_1(-\infty, 0) \;=\; 0 \;=\; \eta_2(-\infty, 0) \text{ and } \eta_1 \ll \eta_2 \ll \eta_1 \text{ with } \frac{\mathrm{d}\eta_2}{\mathrm{d}\eta_1} \;=\; \exp\,. \tag{4.21}$$

*Moreover, suppose that the following conditions (**B1**), (**B2**) and (**B3**) hold.*

*(**B1**): One of the equations* (**B1a**) *and* (**B1b**) *holds for all* $x \in C(\eta_1) \cap (0, \infty)$*:*

$$\lim_{n\to\infty} P_{(k_n)}\left(\max_{1\leq i\leq k_n}\left\{\varepsilon_{k_n,i} f_{k_n,i}\left(X_{k_n,i}\right)\right\} \leq e^x - 1\right) = \exp\Big(-\eta_1\left(x,\infty\right)\Big) \in (0,1], \tag{B1a}$$

$$\lim_{n\to\infty} \sum_{i=1}^{k_n} P_{k_n,i}(\varepsilon_{k_n,i} f_{k_n,i} > e^x - 1) = \eta_1\left(x,\infty\right) \in [0,\infty). \tag{B1b}$$

*(**B2**): One of the equations* (**B2a**) *and* (**B2b**) *holds for all* $x \in C(\eta_2) \cap (0, \infty)$*:*

$$\lim_{n\to\infty} Q_{(k_n)}\left(\max_{1\leq i\leq k_n}\left\{\varepsilon_{k_n,i} f_{k_n,i}\left(X_{k_n,i}\right)\right\} \leq e^x - 1\right) = \exp\Big(-\eta_2\left(x,\infty\right)\Big) \in (0,1], \tag{B2a}$$

$$\lim_{n\to\infty} \sum_{i=1}^{k_n} Q_{k_n,i}(\varepsilon_{k_n,i} f_{k_n,i} > e^x - 1) = \eta_2\left(x,\infty\right) \in [0,\infty). \tag{B2b}$$

*(**B3**): We have*

$$\lim_{\varepsilon\searrow 0} \ \mathop{\substack{\limsup \\ \liminf}}_{n\to\infty} \sum_{i=1}^{k_n} \varepsilon_{k_n,i}^2 \, \mathbb{E}_{P_{k_n,i}}\left(f_{k_n,i}^2 \mathbf{1}_{\{\varepsilon_{k_n,i} f_{k_n,i} \leq \varepsilon\}} - 1\right) \;=\; \sigma^2,$$

where the notation $\substack{\limsup \\ \liminf}$ *is a short way to say that the equation holds for the limes superior and limes inferior.*

Remark 4.14. We do not assume in (**B**) that η_1 and η_2 are Lévy measures. □

Theorem 4.15. *(**A**) holds for a subsequence* $\{k_n : n \in \mathbb{N}\}$ *of* $\mathbb{N}$*, random variables* ξ_1*,* ξ_2 *with Lévy characteristics* $(\gamma_1, \sigma^2, \eta_1)$ *and* $(\gamma_2, \sigma^2, \eta_2)$ *if and only if (**B**) holds for* $\{k_n : n \in \mathbb{N}\}$*,* $\sigma^2 \geq 0$ *and measures* η_1, η_2*. Moreover, if (**A**) holds then* (i)-(iii) *of Corollary 4.12 is fulfilled.*

The proof of Theorem 4.15 can be found in Section 4.2.2. In the following lemma we show that the two conditions from (**B1**) as well as the ones from (**B2**) are equivalent.

Lemma 4.16. *Suppose Assumption 2.1*(i)*. Let $x > 0$ be fixed and $\{k_n : n \in \mathbb{N}\}$ be a subsequence of $\mathbb{N}$. Then*

$$\lim_{n\to\infty} \sum_{i=1}^{k_n} P_{k_n,i}\left(\varepsilon_{k_n,i} f_{k_n,i} > x\right) = b \in [0,\infty] \tag{4.22}$$

if and only if

$$\lim_{n\to\infty} P_{(k_n)}\left(\max_{1\le i\le k_n} \{\varepsilon_{k_n,i} f_{k_n,i}\left(X_{k_n,i}\right)\} \le x\right) = \exp\left(-b\right) \in [0,1], \tag{4.23}$$

where we use the convention $\exp(-\infty) = 0$*. The equivalence of* (4.22) *and* (4.23) *also holds if we replace $P_{k_n,i}$ in* (4.22) *by $Q_{k_n,i}$ and $P_{(k_n)}$ in* (4.23) *by $Q_{(k_n)}$.*
Hence, the equation in (**B1a**) *holds for a given $x \in (0,\infty)$ if and only if* (**B1b**) *does so, and* (**B2a**) *holds for a given $x \in (0,\infty)$ if and only if* (**B2b**) *does so.*

Proof. Fix $x \in (0,\infty)$. Let $h_{n,i}$ be defined by

$$h_{n,i} := P_{k_n,i}\left(\varepsilon_{k_n,i} f_{k_n,i} > x\right) \text{ for all } n \in \mathbb{N} \text{ and every } 1 \le i \le k_n.$$

Note that for all $n \in \mathbb{N}$

$$\sum_{i=1}^{k_n} h_{n,i} = \sum_{i=1}^{k_n} P_{k_n,i}\left(\varepsilon_{k_n,i} f_{k_n,i} > x\right) \quad \text{and}$$

$$\prod_{i=1}^{k_n} (1 - h_{n,i}) = P_{(k_n)}\left(\max_{1\le i\le k_n} \{\varepsilon_{k_n,i} f_{k_n,i}\left(X_{k_n,i}\right)\} \le x\right).$$

By Remark 2.2

$$\max_{1\le i\le k_n} \{h_{n,i}\} \to 0 \quad \text{as } n \to \infty.$$

Hence, the equivalence follows from Lemma A.28. Obviously, the proof is still valid if we replace $P_{k_n,i}$ by $Q_{k_n,i}$ and $P_{(k_n)}$ by $Q_{(k_n)}$. ∎

Remark 4.17. By the second binomial formula we have for every $A \in \mathcal{A}$

$$\begin{aligned}&\sum_{i=1}^{k_n} \varepsilon_{k_n,i}^2 \, \mathbb{E}_{P_{k_n,i}}\left(\left(f_{k_n,i} - 1\right)^2 \mathbf{1}_A\right) \\ = \;& \sum_{i=1}^{k_n} \varepsilon_{k_n,i}^2 \, \mathbb{E}_{P_{k_n,i}}\left(f_{k_n,i}^2 \mathbf{1}_A - 1\right) + 2\sum_{i=1}^{k_n} \varepsilon_{k_n,i}^2 \mu_{k_n,i}(A^c) - \sum_{i=1}^{k_n} \varepsilon_{k_n,i}^2 P_{k_n,i}(A^c).\end{aligned}$$

Thus, if for some $\varepsilon > 0$

$$\limsup_{n\to\infty} \sum_{i=1}^{k_n} \varepsilon_{k_n,i}\mu_{k_n,i}\left(\varepsilon_{k_n,i}f_{k_n,i} > \varepsilon\right) + \sum_{i=1}^{k_n} \varepsilon_{k_n,i}P_{k_n,i}\left(\varepsilon_{k_n,i}f_{k_n,i} > \varepsilon\right) \ < \ \infty \tag{4.24}$$

then due to $\varepsilon_{k_n:k_n} = o(1)$ as $n \to \infty$ we have

$$\begin{aligned}
&\sum_{i=1}^{k_n} \varepsilon_{k_n,i}^2 \, \mathbb{E}_{P_{k_n,i}}\Big(f_{k_n,i}^2 \mathbf{1}_{\{\varepsilon_{k_n,i}f_{k_n,i}\le\varepsilon\}} - 1\Big) \\
&= \sum_{i=1}^{k_n} \varepsilon_{k_n,i}^2 \, \mathbb{E}_{P_{k_n,i}}\Big(\Big(f_{k_n,i} - 1\Big)^2 \mathbf{1}_{\{\varepsilon_{k_n,i}f_{k_n,i}\le\varepsilon\}}\Big) + o(1) \qquad (4.25) \\
&\ge \ o(1) \quad \text{as } n \to \infty.
\end{aligned}$$

It is easy to see that by Lemma 4.16 that condition (4.24) holds for all $\varepsilon > 0$ from a dense subset of $(0, \infty)$ if (**B2a**) does so. By Lemma 4.16 the latter holds if (**B2**) is fulfilled. □

Because of (4.15) and (4.16) the Lévy characteristics and so the distributions of ξ_1, ξ_2 only depend on η_2 and σ^2. Hence, a logical consequence is to ask whether we can spare (**B1**). The answer is yes, as we show in the following. Condition (**B**) can be replaced by a new condition denoted by (**B'**).

Condition (B'). *Let Assumption 2.1*(i) *be fulfilled. Let* $\sigma^2 \ge 0$, $\{k_n : n \in \mathbb{N}\}$ *be a subsequence of* $\mathbb{N}$ *and* η_1, η_2 *be measures on* $(\mathbb{R} \setminus \{0\}, \mathcal{B}(\mathbb{R} \setminus \{0\}))$ *such that* (4.21), ***(B3)*** *and the following condition **(B'2)** hold.*

***(B'2)**: Condition* (**B2a**) *or* (**B2b**) *or the following* (**B2c**) *holds for all* $x \in C(\eta_1)$*:*

$$\lim_{n\to\infty} \sum_{i=1}^{k_n} \varepsilon_{k_n,i}\mu_{k_n,i}\left(\varepsilon_{k_n,i}f_{k_n,i} > e^x - 1\right) = (\eta_2 - \eta_1)(x, \infty) \in [0, \infty). \tag{\textbf{B2c}}$$

Remark 4.18. We do not assume in (**B'**) that η_1 and η_2 are Lévy measures. □

We show that (**B'**) is also sufficient and necessary for (**A**).

Theorem 4.19. *Condition **(A)** holds for a subsequence* $\{k_n : n \in \mathbb{N}\}$ *of* $\mathbb{N}$*, random variables* ξ_1, ξ_2 *with Lévy characteristics* $(\gamma_1, \sigma^2, \eta_1)$ *and* $(\gamma_2, \sigma^2, \eta_2)$ *if and only if Condition **(B')** holds for* $\{k_n : n \in \mathbb{N}\}$, σ^2 *and measures* η_1, η_2*. Moreover, if **(A)** holds then* (i)-(iii) *of Corollary 4.12 hold.*

Before we can prove Theorem 4.19 we need the following lemma.

Lemma 4.20. *Suppose that Assumption 2.1*(i) *holds. Then we have for all $y > 0$, every subsequence $\{k_n : n \in \mathbb{N}\}$ of $\mathbb{N}$ and all $1 \le j \le n \in \mathbb{N}$:*

$$y\,P_{n,j}\Big(\varepsilon_{n,j} f_{n,j} > y\Big) \quad \le \quad \varepsilon_{n,j}\mu_{n,j}\Big(\varepsilon_{n,j} f_{n,j} > y\Big), \tag{4.26}$$

$$\lim_{n\to\infty}\sum_{i=1}^{k_n} Q_{k_n,i}\Big(\varepsilon_{k_n,i} f_{k_n,i} > y\Big) < \infty \Leftrightarrow \lim_{n\to\infty}\sum_{i=1}^{k_n} \varepsilon_{k_n,i}\mu_{k_n,i}\Big(\varepsilon_{k_n,i} f_{k_n,i} > y\Big) < \infty, \tag{4.27}$$

$$\lim_{n\to\infty}\sum_{i=1}^{k_n} Q_{k_n,i}\Big(\varepsilon_{k_n,i} f_{k_n,i} > y\Big) = 0 \Leftrightarrow \lim_{n\to\infty}\sum_{i=1}^{k_n} \varepsilon_{k_n,i}\mu_{k_n,i}\Big(\varepsilon_{k_n,i} f_{k_n,i} > y\Big) = 0, \tag{4.28}$$

$$\lim_{n\to\infty}\sum_{i=1}^{k_n} P_{k_n,i}\Big(\varepsilon_{k_n,i} f_{k_n,i} > y\Big) = \infty \Rightarrow \lim_{n\to\infty}\sum_{i=1}^{k_n} \varepsilon_{k_n,i}\mu_{k_n,i}\Big(\varepsilon_{k_n,i} f_{k_n,i} > y\Big) = \infty. \tag{4.29}$$

Proof. Note that (4.27)-(4.29) follow immediately from (4.26) and the definition of $Q_{n,i}$, see (2.3). Fix $y > 0$ and $1 \le j \le n \in \mathbb{N}$. We conclude (4.26) from the following calculation:

$$y\,P_{n,j}\Big(\varepsilon_{n,j} f_{n,j} > y\Big) \;=\; \int_{\{\varepsilon_{n,j} f_{n,j} > y\}} y \;\mathrm{d}P_{n,j} \;\le\; \int_{\{\varepsilon_{n,j} f_{n,j} > y\}} \varepsilon_{n,j} f_{n,j} \;\mathrm{d}P_{n,j}. \qquad \blacksquare$$

Proof of Theorem 4.19. By Theorem 4.15 it remains to show that (**B'**) is sufficient for (**B**). Suppose that (**B'**) is fulfilled for the subsequence $\{k_n : n \in \mathbb{N}\}$, the constant σ^2 and the measures η_1, η_2. For this purpose we use a typical subsequence argument. We show that for every subsequence $\{k_{n,1} : n \in \mathbb{N}\}$ of $\mathbb{N}$ there exists a further subsequence $\{k_n^{(2)} : n \in \mathbb{N}\}$ of $\{k_n^{(1)} : n \in \mathbb{N}\}$ such that (**B**) holds for $\{k_{n,2} : n \in \mathbb{N}\}$, σ^2 and the measures η_1, η_2. Then we can conclude that (**B**) holds for $\{k_n : n \in \mathbb{N}\}$, σ^2 and the measures η_1, η_2.
First, assume (**B2c**) and (**B3**) are fulfilled for all $x \in C(\eta_1)\cap(0,\infty)$. Thus, by Lemma 4.20

$$(\mathrm{e}^x - 1)\sum_{i=1}^{k_n} P_{k_n,i}\left(\varepsilon_{k_n,i} f_{k_n,i} > \mathrm{e}^x - 1\right) \le \sum_{i=1}^{k_n} \varepsilon_{k_n,i}\mu_{k_n,i}\left(\varepsilon_{k_n,i} f_{k_n,i} > \mathrm{e}^x - 1\right) \to (\eta_2 - \eta_1)(x,\infty)$$

as $n \to \infty$ for every $x \in C(\eta_1) \cap (0,\infty)$. Let $\{k_n^{(1)} : n \in \mathbb{N}\}$ be an arbitrary subsequence of $\{k_n : n \in \mathbb{N}\}$ and D be a countable dense subset of $C(\eta_1) \cap (0,\infty)$. By a well known diagonalisation procedure for subsequences there exists a subsequence $\{k_{n,2} : n \in \mathbb{N}\}$ of $\{k_{n,1} : n \in \mathbb{N}\}$ and a non-increasing function $h : (0,\infty) \to (0,\infty)$ such that

$$\lim_{n\to\infty}\sum_{i=1}^{k_{n,2}} P_{k_{n,2},i}\left(\varepsilon_{k_{n,2},i} f_{k_{n,2},i} > e^x - 1\right) \;=\; h(x) \;\text{ for all } x \in D, \tag{4.30}$$

$$\text{where } \; 0 \;\le\; h(x) \;\le\; (\mathrm{e}^x - 1)^{-1}\,(\eta_2 - \eta_1)(x,\infty) \;\text{ for all } x \in D. \tag{4.31}$$

Due to monotonicity it is easy to see that (4.30) and (4.31) hold even for all $x \in C(h)$.

Let $\widetilde{\eta}_1$ and $\widetilde{\eta}_2$ be measures on $(\mathbb{R}\setminus\{0\},\mathcal{B}(\mathbb{R}\setminus\{0\}))$ such that for all $x\in C(h)$

$$\widetilde{\eta}_1(x,\infty) \;=\; h(x)\,,\; \widetilde{\eta}_2(x,\infty) \;=\; h(x) + (\eta_2-\eta_1)(x,\infty) \tag{4.32}$$

$$\text{and }\; \widetilde{\eta}_1(-\infty,0) = 0 = \widetilde{\eta}_2(-\infty,0). \tag{4.33}$$

Since $C(h)$ is a dense subset of $(0,\infty)$ we conclude from (4.7), (4.32) and (4.33) that

$$(\widetilde{\eta}_2-\widetilde{\eta}_1) \;=\; (\eta_2-\eta_1). \tag{4.34}$$

Note that by (2.1), (4.30), (4.32) and (**B2c**)

$$\begin{aligned}
&\sum_{i=1}^{k_{n,2}} Q_{k_{n,2},i}\left(\varepsilon_{k_{n,2},i} f_{k_{n,2},i} > e^x-1\right)\\
&= \sum_{i=1}^{k_{n,2}}(1-\varepsilon_{k_{n,2},i})P_{k_{n,2},i}\left(\varepsilon_{k_{n,2},i} f_{k_{n,2},i} > e^x-1\right) + \sum_{i=1}^{k_{n,2}}\varepsilon_{k_{n,2},i}\mu_{k_{n,2},i}\left(\varepsilon_{k_{n,2},i} f_{k_{n,2},i} > e^x-1\right)\\
&\to\; h(x) + (\eta_2-\eta_1)(x,\infty) \;=\; \widetilde{\eta}_2(x,\infty) \quad \text{as } n\to\infty
\end{aligned}$$

for all $x\in C(h)\cap C(\eta_1)$ and due to monotonicity also for all $x\in C(\widetilde{\eta}_2)\cap(0,\infty)$. Hence, we can conclude that (**B**) holds for the subsequence $\{k_{n,2}: n\in\mathbb{N}\}$, the constant σ^2 and the measures $\widetilde{\eta}_1,\widetilde{\eta}_2$. By Corollary 4.12 and Theorem 4.15

$$\frac{\mathrm{d}\widetilde{\eta}_2}{\mathrm{d}\widetilde{\eta}_1} \;=\; \exp\,.$$

Combining this, (4.21) and (4.34) yields

$$\eta_1 \;=\; \widetilde{\eta}_1 \;\text{ and }\; \eta_2 \;=\; \widetilde{\eta}_2.$$

Note that by Lemma 4.16 conditions (**B2b**) and (**B2c**) are equivalent. Thus, it remains to consider the case that (**B2b**) and (**B3**) hold. Due to the definition of $Q_{k_n,i}$ we have

$$Q_{k_n,i}(A) \geq \varepsilon_{k_n,i}\mu_{k_n,i}(A) \;\text{ for all } A\in\mathcal{A}.$$

Finally, using the same diagonalisation procedure as before completes the proof. ∎

Lemma 4.16 says that (**B2a**) and (**B2b**) are equivalent. If (**B3**) holds then a similar statement can be made for (**B2a**), (**B2b**) and (**B2c**).

Corollary 4.21. *If (B') holds then* (**B2a**)-(**B2c**) *hold for all* $x\in C(\eta_1)\cap(0,\infty)$.

Proof. Suppose (**B'**). By Theorem 4.19 (**A**) holds and so does (**B**) by Theorem 4.15. This means in particular that (**B2**) is fulfilled. We conclude from Lemma 4.16 that

(**B2a**) and (**B2b**) are fulfilled for all $x \in C(\eta_2) \cap (0, \infty)$. Note that $C(\eta_1) = C(\eta_2)$ by Corollary 4.12(ii). Furthermore, we deduce from Lemma 4.16 that (**B1b**) holds for all $x \in C(\eta_1) \cap (0, \infty)$. Combining (**B1b**) and (**B2b**) verifies (**B2c**) for all $x \in C(\eta_1) \cap (0, \infty)$. ■

We are in particular interested in the case that $E_n := \{P_{(k_n)}, Q_{(k_n)}\}$ converges weakly to a Gaussian experiment. That is why we reformulate Theorem 4.19 for this special case. Recall that the random variables ξ_1 and ξ_2 are normally distributed if and only if $\eta_1 \equiv \eta_2 \equiv 0$. From this and Theorem 4.15 we obtain the following corollary.

Corollary 4.22. *If (**A**) is fulfilled with $\xi_j \sim N(m_j, \sigma_j^2)$, $j \in \{1, 2\}$, then for some $\sigma^2 \geq 0$*

$$\sigma_1^2 \ = \ \sigma_2^2 = \sigma^2 \ \text{ and } \ m_j \ = \ (-1)^j \frac{\sigma^2}{2}. \tag{4.35}$$

In the following we formulate the new conditions belonging to the Gaussian case.

Condition (A normal). *Suppose that (**A**) is fulfilled for $\xi_j \sim N(m_j, \sigma_j^2)$, $j = 1, 2$, where* (4.35) *holds for some $\sigma^2 \geq 0$.*

Remark 4.23. By Corollary 4.22 the distributions of ξ_1 and ξ_2 are uniquely determined by one parameter, namely σ^2, under (**A normal**). □

Condition (B normal). *Suppose that Assumption 2.1*(i) *holds. Moreover, assume that there exist a subsequence $\{k_n : n \in \mathbb{N}\}$ of $\mathbb{N}$ and constants $\sigma^2 \geq 0$, $y_0 > 0$ such that the following two conditions (**B2 normal**) and (**B3 normal**) are fulfilled.*

*(**B2 normal**): One of the following conditions* (**B2a normal**)-(**B2c normal**) *holds for every $y > 0$:*

$$\lim_{n\to\infty} Q_{(k_n)}\Big(\max_{1\leq i\leq k_n} \{\varepsilon_{k_n,i} f_{k_n,i}(X_{k_n,i})\} \leq y\Big) = 1, \tag{B2a normal}$$

$$\lim_{n\to\infty} \sum_{i=1}^{k_n} Q_{k_n,i}\Big(\varepsilon_{k_n,i} f_{k_n,i} > y\Big) = 0, \tag{B2b normal}$$

$$\lim_{n\to\infty} \sum_{i=1}^{k_n} \varepsilon_{k_n,i} \mu_{k_n,i}\Big(\varepsilon_{k_n,i} f_{k_n,i} > y\Big) = 0. \tag{B2c normal}$$

*(**B3 normal**): We have*

$$\sigma^2 = \lim_{n\to\infty} \sum_{i=1}^{k_n} \varepsilon_{k_n,i}^2 \, \mathbb{E}_{P_{k_n,i}}\left(f_{k_n,i}^2 \mathbf{1}_{\{\varepsilon_{k_n,i} f_{k_n,i} \leq y_0\}} - 1\right).$$

Remark 4.24. Combining Lemma 4.16 and (4.28) of Lemma 4.20 shows that (**B2a normal**)-(**B2c normal**) are equivalent. □

Now we can reformulate Theorem 4.19 for the Gaussian case. Note that the conditions of Theorem 4.19 can be slightly weakened for this case, as we prove in the following.

Corollary 4.25. *Let $y_0 > 0$. (**A normal**) holds for a subsequence $\{k_n : n \in \mathbb{N}\}$ of $\mathbb{N}$ and a constant $\sigma^2 \geq 0$ if and only if (**B normal**) holds for $\{k_n : n \in \mathbb{N}\}$, σ^2 and y_0.*

Remark 4.26. (i) Assume (**B2 normal**). Because of the above-mentioned equivalence (**B3 normal**) is fulfilled for all $y_0 > 0$ if it is fulfilled for one.

(ii) Suppose (**A normal**). It can be easily shown that

$$\mathcal{L}\left(\log\left(\frac{\mathrm{d}N(\sigma,1)}{\mathrm{d}N(0,1)}\right) \Big| N(0,1)\right) = N\left(-\frac{\sigma^2}{2}, \sigma^2\right).$$

Hence, $\{P_{(k_n)}, Q_{(k_n)}\}$ converges weakly to the Gaussian shift model $\{N(0,1), N(\sigma,1)\}$ or equivalently to $\left\{N(-\frac{\sigma^2}{2}, \sigma^2), N(\frac{\sigma^2}{2}, \sigma^2)\right\}$.

(iii) We use the convention $N(a, 0) = \epsilon_a$ for $a \in \mathbb{R}$. Therefore, if (**B2 normal**) and (**B3 normal**) are fulfilled for $\sigma^2 = 0$ then (**A**) holds for $\xi_1 = \xi_2 = 0$ ($\mathcal{P}$-a.s.). In this case $\{P_{(k_n)}, Q_{(k_n)}\}$ converges weakly to the uninformative experiment $\{\epsilon_0, \epsilon_0\}$. Note that in this case (**B normal**) can be weakened, see Corollary 4.35. □

Proof. By Theorem 4.19 and Remark 4.24 it is sufficient to show that (**B3 normal**) and (**B3**) are equivalent if (**B2c normal**) holds for all $y > 0$. Suppose that (**B2c normal**) holds for all $y > 0$. Then for all $\varepsilon \in (0, y_0]$

$$\begin{aligned} 0 &\leq \sum_{i=1}^{k_n} \mathbb{E}_{P_{k_n,i}}\left(\varepsilon_{k_n,i}^2 f_{k_n,i}^2 \mathbf{1}_{\{\varepsilon < \varepsilon_{k_n,i} f_{k_n,i} \leq y_0\}}\right) \\ &\leq y_0 \sum_{i=1}^{k_n} \varepsilon_{k_n,i} \mu_{k_n,i}\Big(\varepsilon_{k_n,i} f_{k_n,i} > \varepsilon\Big) \ = \ o(1) \text{ as } n \to \infty. \end{aligned}$$

Hence, for all $\varepsilon \in (0, y_0]$

$$\sum_{i=1}^{k_n} \mathbb{E}_{P_{k_n,i}}\left(\varepsilon_{k_n,i}^2 f_{k_n,i}^2 \mathbf{1}_{\{\varepsilon_{k_n,i} f_{k_n,i} \leq y_0\}}\right) = \sum_{i=1}^{k_n} \mathbb{E}_{P_{k_n,i}}\left(\varepsilon_{k_n,i}^2 f_{k_n,i}^2 \mathbf{1}_{\{\varepsilon_{k_n,i} f_{k_n,i} \leq \varepsilon\}}\right) + o(1)$$

as $n \to \infty$. Thus, it is easy to see that (**B3 normal**) and (**B3**) are equivalent. ■

4.2.1. Some technical lemmas

In this section we discuss important properties of $\log(g_{n,i})$, $f_{n,i}$ and

$$W_{n,i} \;:=\; \sqrt{g_{n,i}} - 1 \text{ for all } i = \{1, \ldots, n\}, \tag{4.36}$$

which we need in Section 4.2.2 to prove Theorem 4.15. By Lemma 4.3, (4.2) and (A.17)

$$\sum_{i=1}^{n} \mathbb{E}_{P_{n,i}}\left(W_{n,i}^2\right) \;=\; \sum_{i=1}^{n} \mathbb{E}_{P_{n,i}}\left((\sqrt{g_{n,i}} - 1)^2\right) \;=\; 2\sum_{i=1}^{n} d^2\left(P_{n,i}, Q_{n,i}\right) \;=\; 2D_n \tag{4.37}$$

for all $n \in \mathbb{N}$. Consequently, we see that $W_{n,i}$ plays an important role for our problem.

Lemma 4.27. *Suppose Assumption 2.1*(i)*. Define $W_{n,i}$ as in* (4.36) *and set*

$$\widetilde{A}_{n,i,\varepsilon} := \{|\log(g_{n,i})| \leq \varepsilon\} \text{ for all } 1 \leq i \leq n \text{ and } \varepsilon > 0. \tag{4.38}$$

Let $\varepsilon > 0$ be fixed. If n is sufficiently large then for all $1 \leq i \leq n$

$$\widetilde{A}_{n,i,\varepsilon} = \left\{e^{-\varepsilon} \leq g_{n,i} \leq e^{\varepsilon}\right\} = \left\{g_{n,i} \leq e^{\varepsilon}\right\} = \left\{\varepsilon_{n,i} f_{n,i} \leq e^{\varepsilon} - 1 + \varepsilon_{n,i}\right\}, \tag{4.39}$$

$$\widetilde{A}^{c}_{n,i,\varepsilon} = \left\{g_{n,i} > e^{\varepsilon}\right\} = \left\{\varepsilon_{n,i} f_{n,i} > e^{\varepsilon} - 1 + \varepsilon_{n,i}\right\}, \tag{4.40}$$

$$\left\{\varepsilon_{n,i} f_{n,i} \leq e^{\varepsilon} - 1\right\} \subseteq \widetilde{A}_{n,i,\varepsilon} \subseteq \left\{\varepsilon_{n,i} f_{n,i} \leq e^{2\varepsilon} - 1\right\}, \tag{4.41}$$

$$-1 < e^{-\frac{\varepsilon}{2}} - 1 \leq W_{n,i}\, \mathbf{1}_{\widetilde{A}_{n,i,\varepsilon}} \leq e^{\frac{\varepsilon}{2}} - 1, \tag{4.42}$$

$$\frac{\varepsilon_{n,i}\, |f_{n,i} - 1|\, \mathbf{1}_{\widetilde{A}_{n,i,\varepsilon}}}{e^{\frac{\varepsilon}{2}} + 1} \leq |W_{n,i}|\, \mathbf{1}_{\widetilde{A}_{n,i,\varepsilon}} \leq \frac{\varepsilon_{n,i}\, |f_{n,i} - 1|\, \mathbf{1}_{\widetilde{A}_{n,i,\varepsilon}}}{e^{-\frac{\varepsilon}{2}} + 1}, \tag{4.43}$$

$$W_{n,i}^2 = g_{n,i} - 1 - 2W_{n,i}. \tag{4.44}$$

Furthermore, for all $i \in \mathbb{N}$

$$W_{n,i} = \Big((1 - \varepsilon_{n,i}) + \varepsilon_{n,i} f_{n,i}\Big)^{1/2} - 1 \geq (1 - \varepsilon_{n:n})^{1/2} - 1 =: a_n \to 0 \text{ as } n \to \infty, \tag{4.45}$$

$$\max_{1 \leq j \leq n} \left|\mathbb{E}_{P_{n,j}}\left(W_{n,j} \mathbf{1}_{\widetilde{A}_{n,j,\varepsilon}}\right)\right| \to 0 \text{ as } n \to \infty, \tag{4.46}$$

$$\max_{1 \leq j \leq n} \left|\mathbb{E}_{P_{n,j}}\left(W_{n,j}^2 \mathbf{1}_{\widetilde{A}_{n,j,\varepsilon}}\right)\right| \to 0 \text{ as } n \to \infty. \tag{4.47}$$

Proof. By (2.1), the definition (2.4) of $g_{n,i}$ and $f_{n,i} \geq 0$ it can easily be seen that (4.39), (4.40) and (4.41) hold. We leave the details to the reader. (4.42) follows directly from

(4.39) and the definition (4.36) of $W_{n,i}$. By the third binomial formula

$$|W_{n,i}| = \frac{|g_{n,i} - 1|}{\sqrt{g_{n,i}} + 1} = \frac{\varepsilon_{n,i}\,|f_{n,i} - 1|}{\sqrt{g_{n,i}} + 1}.$$

Hence, (4.43) follows immediately from (4.39). Moreover,

$$W_{n,i}^2 = (\sqrt{g_{n,i}} - 1)^2 = g_{n,i} - 2\sqrt{g_{n,i}} + 1 = g_{n,i} - 1 - 2W_{n,i},$$

which proves (4.44). Clearly, (4.45) holds. By Jensen's inequality, Lemma 4.7 and (4.45)

$$\begin{aligned} 0 \leftarrow a_n \le \max_{1\le i\le n} \left\{ \mathbb{E}_{P_{n,i}} \left(W_{n,i} \mathbf{1}_{\widetilde{A}_{n,i,\varepsilon}} \right) \right\} &= \max_{1\le j\le n} \left\{ \mathbb{E}_{P_{n,i}} \left(\sqrt{g_{n,i}}\, \mathbf{1}_{\widetilde{A}_{n,i,\varepsilon}} \right) - P_{n,i}(\widetilde{A}_{n,i,\varepsilon}) \right\} \\ &\le \max_{1\le i\le n} \left\{ \sqrt{\mathbb{E}_{P_{n,i}}\left(g_{n,i}\right)} - P_{n,i}(\widetilde{A}_{n,i,\varepsilon}) \right\} \\ &= \max_{1\le i\le n} \left\{ 1 - P_{n,i}(\widetilde{A}_{n,i,\varepsilon}) \right\} \\ &\le \max_{1\le i\le n} \left\{ P_{n,i}\left(\widetilde{A}^c_{n,i,\varepsilon} \right) \right\} \to 0 \quad \text{as } n \to \infty. \end{aligned}$$

Thus, (4.46) follows. By Lemma 4.7, (4.44) and (4.46)

$$\begin{aligned} \max_{1\le i\le n} \left\{ \mathbb{E}_{P_{n,i}} \left(W_{n,i}^2 \mathbf{1}_{\widetilde{A}_{n,i,\varepsilon}} \right) \right\} &= \max_{1\le i\le n} \left\{ \mathbb{E}_{P_{n,i}} \left((g_{n,i} - 1)\, \mathbf{1}_{\widetilde{A}_{n,i,\varepsilon}} \right) - 2\, \mathbb{E}_{P_{n,i}} \left(W_{n,i} \mathbf{1}_{\widetilde{A}_{n,i,\varepsilon}} \right) \right\} \\ &= \max_{1\le i\le n} \left\{ \mathbb{E}_{P_{n,i}} \left((1 - g_{n,i})\, \mathbf{1}_{\widetilde{A}^c_{n,i,\varepsilon}} \right) - 2\, \mathbb{E}_{P_{n,i}} \left(W_{n,i} \mathbf{1}_{\widetilde{A}_{n,i,\varepsilon}} \right) \right\} \\ &\le \max_{1\le i\le n} \left\{ P_{n,i}\left(\widetilde{A}^c_{n,i,\varepsilon} \right) \right\} + 2 \max_{1\le i\le n} \left\{ \left| \mathbb{E}_{P_{n,i}} \left(W_{n,i} \mathbf{1}_{\widetilde{A}_{n,i,\varepsilon}} \right) \right| \right\} \to 0 \end{aligned}$$

as $n \to \infty$. Consequently, we obtain (4.47). ■

In the following we extend a result of Witting and Nölle [65] for our purpose.

Lemma 4.28. *Suppose Assumption 2.1*(i). *Define* $W_{n,i}$ *and* $\widetilde{A}_{n,i,\varepsilon}$ *as in* (4.36) *and* (4.38) *for every* $\varepsilon > 0$ *and* $1 \le i \le n \in \mathbb{N}$. *Then there exist some real-valued random variable* $Z_{n,i,\varepsilon}$ *and some constants* $c_{1,\varepsilon} < 0 < c_{2,\varepsilon}$ *for all* $\varepsilon > 0$ *and* $1 \le i \le n \in \mathbb{N}$ *such that*

$$\log\left(g_{n,i}\right) \mathbf{1}_{\widetilde{A}_{n,i,\varepsilon}} = \left(2W_{n,i} - W_{n,i}^2 - Z_{n,i,\varepsilon} W_{n,i}^2 \right) \mathbf{1}_{\widetilde{A}_{n,i,\varepsilon}} \tag{4.48}$$

$$\text{and} \quad c_{1,\varepsilon} \le \max_{1\le i\le n} \left\{ Z_{n,i,\varepsilon} \right\} \le c_{2,\varepsilon} \tag{4.49}$$

if $n \in \mathbb{N}$ *is sufficiently large. Moreover,*

$$c_\varepsilon := \max\left\{ -c_{1,\varepsilon}, c_{2,\varepsilon} \right\} \to 0 \quad \text{as } \varepsilon \searrow 0. \tag{4.50}$$

Proof. Let $\varepsilon > 0$ and $1 \le i \le n \in \mathbb{N}$ be fixed for the first part of the proof. Witting and Nölle already showed, see p. 65 in [65], that

$$\Big(2W_{n,i} - \log\left(g_{n,i}\right)\Big)\mathbf{1}_{\widetilde{A}_{n,i,\varepsilon}} = W_{n,i}^2\,\mathbf{1}_{\widetilde{A}_{n,i,\varepsilon}} \int_0^1 \frac{2(1-z)}{(1+zW_{n,i})^2}\,\mathrm{d}z. \tag{4.51}$$

Note that their $W_{n,i}$ differs from ours by a the factor of 2 and that the integral from (4.51) is well defined because of (4.42) of Lemma 4.27. Define

$$\begin{aligned}
Z_{n,i,\varepsilon} &:= \left(\frac{-\log\left(g_{n,i}\right) + 2W_{n,i} - W_{n,i}^2}{W_{n,i}^2}\right)\mathbf{1}_{\widetilde{A}_{n,i,\varepsilon}\cap\{W_{n,i}\neq 0\}}\,,\\
c_{1,\varepsilon} &:= \int_0^1 \frac{2(1-z)}{\left(1+z\left(e^{\frac{\varepsilon}{2}}-1\right)\right)^2}\,\mathrm{d}z - 1\,,\\
c_{2,\varepsilon} &:= \int_0^1 \frac{2(1-z)}{\left(1+z\left(e^{-\frac{\varepsilon}{2}}-1\right)\right)^2}\,\mathrm{d}z - 1\,.
\end{aligned}$$

Then (4.48) holds obviously. Furthermore, (4.49) follows from (4.42) and (4.51). Finally, we can conclude (4.50) from Lebesgue's dominated convergence theorem. ∎

Lemma 4.29. *Suppose that (**A**) or (**B**) is fulfilled for a subsequence* $\{k_n : n \in \mathbb{N}\}$ *of* $\mathbb{N}$. *Define* $W_{n,i}$ *and* $\widetilde{A}_{n,i,\varepsilon}$ *as in* (4.36) *and* (4.38) *for every* $\varepsilon > 0$ *and all* $1 \le i \le n \in \mathbb{N}$. *Then there exists some constant* $C \in (0,\infty)$ *such that*

$$\limsup_{n\to\infty} \sum_{i=1}^{k_n} \mathbb{E}_{P_{k_n,i}}\left(W_{k_n,i}^2 \mathbf{1}_{\widetilde{A}_{k_n,i,\varepsilon}}\right) \le C \quad \textit{for all sufficiently small } \varepsilon > 0. \tag{4.52}$$

Proof. First, assume (**A**). We deduce from (4.37) that for all $\varepsilon > 0$, $n \in \mathbb{N}$

$$\sum_{i=1}^{k_n} \mathbb{E}_{P_{k_n,i}}\left(W_{k_n,i}^2 \mathbf{1}_{\widetilde{A}_{k_n,i,\varepsilon}}\right) \;\le\; \sum_{i=1}^{k_n} \mathbb{E}_{P_{k_n,i}}\left(W_{k_n,i}^2\right) \;=\; 2\,D_{k_n}.$$

Thus, (4.52) follows from Lemma 4.3(ii).
Now let (**B**) be fulfilled. Since (**B3**) holds there exists some $\widetilde{\varepsilon} \in C(\eta_2) \cap (0,\infty)$ such that

$$\limsup_{n\to\infty} \sum_{i=1}^{k_n} \varepsilon_{k_n,i}^2 \left(\mathbb{E}_{P_{k_n,i}}\left(f_{k_n,i}^2 \mathbf{1}_{\{\varepsilon_{k_n,i} f_{k_n,i} \le \varepsilon\}}\right) - 1\right) \;\le\; \sigma^2 + 1 =: C \tag{4.53}$$

for all $\varepsilon \in (0, \widetilde{\varepsilon})$, where by Lemma 4.16 we can assume without loss of generality that (**B2a**) holds for $\widetilde{\varepsilon}$ and so does (4.24). Finally, we conclude from (4.53), Remark 4.17, (4.41) and (4.43) of Lemma 4.27 that for all $\varepsilon \in (0, \frac{1}{2}\log(\widetilde{\varepsilon}+1))$

$$\begin{aligned}\sum_{i=1}^{k_n} \mathbb{E}_{P_{k_n,i}}\left(W_{k_n,i}^2 \mathbf{1}_{\widetilde{A}_{k_n,i,\varepsilon}}\right) &\leq \sum_{i=1}^{k_n} \varepsilon_{k_n,i}^2 \mathbb{E}_{P_{k_n,i}}\left((f_{k_n,i}-1)^2 \mathbf{1}_{\left\{\varepsilon_{k_n,i} f_{k_n,i} \leq e^{2\varepsilon}-1\right\}}\right)\\ &\leq \sum_{i=1}^{k_n} \varepsilon_{k_n,i}^2 \mathbb{E}_{P_{k_n,i}}\left((f_{k_n,i}-1)^2 \mathbf{1}_{\left\{\varepsilon_{k_n,i} f_{k_n,i} \leq \widetilde{\varepsilon}\right\}}\right)\\ &= \sum_{i=1}^{k_n} \varepsilon_{k_n,i}^2 \left(\mathbb{E}_{P_{k_n,i}}\left(f_{k_n,i}^2 \mathbf{1}_{\left\{\varepsilon_{k_n,i} f_{k_n,i} \leq \widetilde{\varepsilon}\right\}}\right)-1\right)+o(1)\\ &\leq C+o(1) \text{ as } n\to\infty. \end{aligned}$$

■

4.2.2. Proof of Theorem 4.15

In this section we give the proof of Theorem 4.15. First, we introduce a new condition denoted by (**C**). We show that (**A**) and (**C**) are equivalent under certain additional assumptions. After having done this we prove that (**C**) can be replaced by (**B**). Finally, we show that the mentioned additional assumptions are fulfilled under (**A**) as well as under (**B**). Consequently, we deduce that (**A**) and (**B**) are equivalent.

Condition (C). *Suppose Assumption 2.1*(i). *Let* $\{k_n : n \in \mathbb{N}\}$ *be a subsequence of* $\mathbb{N}$, $\sigma^2 \geq 0$ *and* η_1, η_2 *be measures on* $(\mathbb{R}\setminus\{0\}, \mathcal{B}(\mathbb{R}\setminus\{0\}))$ *such that* (4.21) *and the following conditions (**C1**)-(**C4**) are fulfilled.*

*(**C1**): For all* $x \in C(\eta_1)\cap(0,\infty)$

$$\lim_{n\to\infty}\sum_{i=1}^{k_n} P_{k_n,i}\Big(\log(g_{k_n,i})>x\Big)=\eta_1(x,\infty)\in[0,\infty). \qquad \textbf{(C1a)}$$

*(**C2**): We have*

$$\sigma^2=\lim_{\varepsilon\searrow 0}\ \substack{\limsup\\ \liminf\\ n\to\infty}\ \sum_{i=1}^{k_n} \mathbb{E}_{P_{k_n,i}}\left((\log(g_{k_n,i}))^2 \mathbf{1}_{\{|\log(g_{k_n,i})|\leq\varepsilon\}}\right).$$

***(C3)**: We have*

$$\lim_{C(\eta_1)\ni\tau\searrow 0}\left(\genfrac{}{}{0pt}{}{\limsup}{\liminf}_{n\to\infty}\left(\sum_{i=1}^{k_n}\mathbb{E}_{P_{k_n,i}}\Big(\log\left(g_{k_n,i}\right)\mathbf{1}_{\{|\log(g_{k_n,i})|\leq\tau\}}\Big)\right) + \int_{(\tau,\infty)}\frac{x}{1+x^2}\,\mathrm{d}\eta_1(x)\right)$$
$$= -\frac{\sigma^2}{2}+\int_{(0,\infty)}\left(1-e^x+\frac{x}{1+x^2}\right)\,\mathrm{d}\eta_1(x).$$

***(C4)**: For all $x\in C(\eta_2)\cap(0,\infty)$*

$$\lim_{n\to\infty}\sum_{i=1}^{k_n}Q_{k_n,i}\Big(\log\left(g_{k_n,i}\right)>x\Big)=\eta_2(x,\infty)\in[0,\infty). \tag{\textbf{C4a}}$$

Theorem 4.30. *Suppose Assumption 2.1*(i) *and let*

$$\lim_{n\to\infty}\sum_{i=1}^{k_n}\mathbb{E}_{P_{k_n,i}}\left(\log\left(g_{k_n,i}\right)\mathbf{1}_{\{|\log(g_{k_n,i})|<\varepsilon\}}\right)^2 \;=\; 0 \tag{4.54}$$

*for all sufficiently small $\varepsilon>0$. Then **(C)** holds for a subsequence $\{k_n : n\in\mathbb{N}\}$ of $\mathbb{N}$, a constant $\sigma^2\geq 0$ and measures η_1,η_2 if and only if **(A)** holds for $\{k_n : n\in\mathbb{N}\}$ and real-valued, infinitely divisible random variables ξ_1,ξ_2 with Lévy characteristics $(\gamma_1,\sigma^2,\eta_1)$ and $(\gamma_2,\sigma^2,\eta_1)$, where $\gamma_1,\gamma_2\in\mathbb{R}$ are given by* (4.16).

Proof. Note that due to Lemma 4.7, Lemma 4.9 and (4.54) we can apply Theorem A.10 for $Y_{k_n,i}:=\log\left(g_{k_n,i}\left(X_{k_n,i}\right)\right)$, $1\leq i\leq k_n$, under $P_{(k_n)}$.
Suppose (**C**). By Theorem A.10 (**A1**) holds for a real-valued random variable ξ_1 with Lévy characteristics $(\gamma_1,\sigma^2,\eta_1)$, where γ_1 is given by (4.16). Moreover, by (4.21) and (**C4**) we have for some $\tau\in C(\eta_2)\cap(0,1)$

$$\int_{(1,\infty)}\exp\,\mathrm{d}\eta_1 \;=\; \eta_2(1,\infty)\;\leq\eta_2(\tau,\infty)\;<\infty.$$

Consequently, (**A**) follows from Corollary 4.13.
Assume (**A**). By Lemma 4.9 and Corollary 4.12 the constants γ_1,γ_2 are given by (4.16) and, moreover, (4.21) is fulfilled. Applying Theorem A.10 yields (**C1**)-(**C3**). Note that we can replace the relation sign $\geq$ by $>$ in (A.5) since we consider continuity points x of a measure η. Hence, applying Theorem A.7 for $Y_{k_n,i}$ under $Q_{(k_n)}$ while considering Lemma 4.7 proves (**C4**). ■

In the first step we show that (**C1**) and (**C4**) are equivalent to (**B1**) and (**B2**), respectively.

Lemma 4.31. *Suppose Assumption 2.1*(i). (**C1a**) *holds for a subsequence* $\{k_n : n \in \mathbb{N}\}$ *of* $\mathbb{N}$*, a measure* η_1 *on* $(\mathbb{R} \setminus \{0\}, \mathcal{B}(\mathbb{R} \setminus \{0\}))$ *and a fixed* $x \in C(\eta_1) \cap (0, \infty)$ *if and only if* (**B1a**) *does. Moreover,* (**C4a**) *holds for a subsequence* $\{k_n : n \in \mathbb{N}\}$ *of* $\mathbb{N}$*, a measure* η_2 *on* $(\mathbb{R} \setminus \{0\}, \mathcal{B}(\mathbb{R} \setminus \{0\}))$ *and a fixed* $x \in C(\eta_2) \cap (0, \infty)$ *if and only if* (**B2a**) *does.*

Proof. Note that by (4.39) of Lemma 4.27 we have for all $x \in C(\eta_1) \cap (0, \infty)$

$$\sum_{i=1}^{k_n} P_{k_n,i} \left(\log\left(g_{k_n,i}\right) \le x\right) \;=\; \sum_{i=1}^{k_n} P_{k_n,i} \left(\varepsilon_{k_n,i} f_{k_n,i} \le e^x - 1 + \varepsilon_{k_n,i}\right) \tag{4.55}$$

if $n \in \mathbb{N}$ is sufficiently large. Due to the fact that x is a point of continuity, combining (2.1), (4.55) and basic calculations yields the equivalence of (**B1a**) and (**C1a**). Analogously, the equivalence of (**B2a**) and (**C4a**) follows. ∎

In the second step we verify the equivalence of (**B3**) and (**C2**) under (**A**) and under (**B**), respectively.

Lemma 4.32. *Suppose (**A**) or (**B**). Then (**B3**) holds if and only if*

$$\lim_{\varepsilon \searrow 0} \begin{matrix} \limsup \\ \liminf \\ n\to\infty \end{matrix} \sum_{i=1}^{k_n} \mathbb{E}_{P_{k_n,i}} \left(W_{k_n,i}^2 \mathbf{1}_{\widetilde{A}_{k_n,i,\varepsilon}} \right) \;=\; \frac{\sigma^2}{4}.$$

*Moreover, (**B3**) holds if and only if (**C2**) does.*

Proof. Note that by Lemma 4.31 (**C1**) and (**C4**) hold under (**B**). If (**A**) holds then applying Lemma 4.7 and Theorem A.7 with $Y_{k_n,i} := \log\left(g_{k_n,i}\left(X_{k_n,i}\right)\right)$ under $P_{(k_n)}$ and under $Q_{(k_n)}$ yields (**C1**) and (**C4**), respectively. To sum up, (**C1**) and (**C4**) are fulfilled. Let $W_{k_n,i}$ and $Z_{k_n,i,\varepsilon}$ be defined as in (4.36) and Lemma 4.28 for all $1 \le i \le n$ and $\varepsilon > 0$. We deduce from Lemma 4.28 that for all $\varepsilon > 0$ and every sufficiently large $n \in \mathbb{N}$

$$\begin{aligned}
&\sum_{i=1}^{k_n} \mathbb{E}_{P_{k_n,i}} \left(\log\left(g_{k_n,i}\right)^2 \mathbf{1}_{\widetilde{A}_{k_n,i,\varepsilon}} \right) \\
&= \sum_{i=1}^{k_n} \mathbb{E}_{P_{k_n,i}} \left(\left(2\, W_{k_n,i} - \left(1 + Z_{k_n,i,\varepsilon}\right) W_{k_n,i}^2\right)^2 \mathbf{1}_{\widetilde{A}_{k_n,i,\varepsilon}} \right) \\
&= 4 \sum_{i=1}^{k_n} \mathbb{E}_{P_{k_n,i}} \left(W_{k_n,i}^2 \mathbf{1}_{\widetilde{A}_{k_n,i,\varepsilon}} \right) &(4.56) \\
&\quad + \sum_{i=1}^{k_n} \mathbb{E}_{P_{k_n,i}} \left(\left(-4\left(1 + Z_{k_n,i,\varepsilon}\right) W_{k_n,i} + \left(1 + Z_{k_n,i,\varepsilon}\right)^2 W_{k_n,i}^2 \right) W_{k_n,i}^2 \mathbf{1}_{\widetilde{A}_{k_n,i,\varepsilon}} \right). &(4.57)
\end{aligned}$$

Note that by Lemma 4.28 and (4.42) of Lemma 4.27 we have for all sufficiently large $n \in \mathbb{N}$:

$$\max_{1\le i\le k_n} \{|Z_{k_n,i,\varepsilon}|\} \le c_\varepsilon \to 0 \quad \text{and} \quad \max_{1\le i\le k_n} \left\{|W_{k_n,i}|\,\mathbf{1}_{\widetilde{A}_{k_n,i,\varepsilon}}\right\} \le e^{\frac{\varepsilon}{2}} - 1 \to 0 \text{ as } \varepsilon \searrow 0. \quad (4.58)$$

Define the sum in (4.57) by $R_{n,\varepsilon}$. Combining (4.58) and Lemma 4.29 shows that

$$\lim_{\varepsilon\searrow 0} \limsup_{n\to\infty} |R_{n,\varepsilon}| = 0.$$

Consequently, the first mentioned equivalence follows. From this equivalence, (4.25) of Remark 4.17 and (4.43) of Lemma 4.27 we obtain the equivalence of (**B3**) and (**C2**). ∎

In the third and last step we prove that (**C3**) already holds under (**B**).

Lemma 4.33. *Let (**B**) be fulfilled for some subsequence $\{k_n : n \in \mathbb{N}\}$ of $\mathbb{N}$, $\sigma^2 \ge 0$ and η_1, η_2, where we assume that the latter ones are Lévy measures. Then (**C3**) holds for $\{k_n : n \in \mathbb{N}\}$, η_1 and σ^2.*

Proof. First, we deduce (**C1**) and (**C4**) from Lemma 4.31. Let $\widetilde{A}_{n,i,\varepsilon}$, $W_{n,i}$, $Z_{n,i,\varepsilon}$ be defined as in (4.38), (4.36) and Lemma 4.28 for all $1 \le i \le n \in \mathbb{N}$ and $\varepsilon > 0$. From Lemma 4.28 and (4.44) of Lemma 4.27 we obtain

$$\begin{aligned}
&\sum_{i=1}^{k_n} \mathbb{E}_{P_{k_n,i}} \left(\log(g_{k_n,i})\,\mathbf{1}_{\widetilde{A}_{k_n,i,\varepsilon}}\right)\\
&= \sum_{i=1}^{k_n} \mathbb{E}_{P_{k_n,i}} \left(\left(g_{k_n,i} - 1 - (2 + Z_{k_n,i,\varepsilon})W_{k_n,i}^2\right)\mathbf{1}_{\widetilde{A}_{k_n,i,\varepsilon}}\right)\\
&= \sum_{i=1}^{k_n} \mathbb{E}_{P_{k_n,i}} \left((1 - g_{k_n,i})\,\mathbf{1}_{\widetilde{A}^c_{k_n,i,\varepsilon}}\right) - \sum_{i=1}^{k_n} \mathbb{E}_{P_{k_n,i}} \left((2 + Z_{k_n,i,\varepsilon})W_{k_n,i}^2\mathbf{1}_{\widetilde{A}_{k_n,i,\varepsilon}}\right)\\
&= \sum_{i=1}^{k_n} P_{k_n,i}(\log(g_{k_n,i}) > \varepsilon) - \sum_{i=1}^{k_n} Q_{k_n,i}(\log(g_{k_n,i}) > \varepsilon) - \sum_{i=1}^{k_n} \mathbb{E}_{P_{k_n,i}} \left((2 + Z_{k_n,i,\varepsilon})W_{k_n,i}^2\mathbf{1}_{\widetilde{A}_{k_n,i,\varepsilon}}\right)
\end{aligned}$$

for all sufficiently large $n \in \mathbb{N}$ and $\varepsilon > 0$. Note that by Lemma 4.28

$$\max_{1\le i\le k_n} |Z_{k_n,i,\varepsilon}| \le c_\varepsilon \to 0 \quad \text{as} \quad \varepsilon \searrow 0$$

for all sufficiently large $n \in \mathbb{N}$. Combining this, (**C2**) and Lemma 4.32 yields

$$\lim_{\varepsilon\searrow 0} \begin{matrix}\limsup\\ \liminf\\ n\to\infty\end{matrix} \sum_{i=1}^{k_n} \mathbb{E}_{P_{k_n,i}} \left((2 + Z_{k_n,i,\varepsilon})W_{k_n,i}^2\mathbf{1}_{\widetilde{A}_{k_n,i,\varepsilon}}\right) = \frac{\sigma^2}{2}.$$

Note that by Taylor's theorem there are constants $C_1, C_2, C_3 \in (0,\infty)$ such that

$$\begin{aligned}\int_{(0,\infty)} \left|1 - e^x + \frac{x}{1+x^2}\right| \,\mathrm{d}\eta_1(x) &\le \int_{(0,1)} \left(\left|-x + \frac{x}{x^2+1}\right| + C_1 x^2\right) \mathrm{d}\eta_1(x) + \int_{[1,\infty)} C_2 e^x \,\mathrm{d}\eta_1(x)\\ &\le \int_{(0,1)} C_3\, x^2 \,\mathrm{d}\eta_1(x) + C_2\, \eta_2[1,\infty) < \infty.\end{aligned}$$

Finally, we conclude from the above, Lebesgue's dominated limit theorem, (4.43) of Lemma 4.27, (**C1**) and (**C4**) that

$$\begin{aligned}&\lim_{C(\eta_1)\ni\varepsilon\searrow 0} \left(\begin{matrix}\limsup\\ \liminf\\ n\to\infty\end{matrix} \sum_{i=1}^{k_n} \mathbb{E}_{P_{k_n,i}} \left(\log\left(g_{k_n,i}\right) \mathbf{1}_{\widetilde{A}_{k_n,i,\varepsilon}}\right) + \int_{(\varepsilon,\infty)} \frac{x}{1+x^2}\,\mathrm{d}\eta_1(x)\right)\\ &= \lim_{C(\eta_1)\ni\varepsilon\searrow 0} \left((\eta_1 - \eta_2)(\varepsilon,\infty) - \frac{\sigma^2}{2} + \int_{(\varepsilon,\infty)} \frac{x}{1+x^2}\,\mathrm{d}\eta_1(x)\right)\\ &= -\frac{\sigma^2}{2} + \lim_{C(\eta_1)\ni\varepsilon\searrow 0} \int_{(\varepsilon,\infty)} \left(1 - e^x + \frac{x}{1+x^2}\right) \mathrm{d}\eta_1(x)\\ &= -\frac{\sigma^2}{2} + \int_{(0,\infty)} \left(1 - e^x + \frac{x}{1+x^2}\right) \mathrm{d}\eta_1(x).\end{aligned}$$ ■

To conclude Theorem 4.15 from Lemmas 4.31 to 4.33 and Theorem 4.30 it remains to show (4.54) for all sufficiently small $\varepsilon > 0$ under (**A**) as well as under (**B**).

Lemma 4.34. *Let (**A**) or (**B**) be fulfilled for a subsequence $\{k_n : n \in \mathbb{N}\}$ of $\mathbb{N}$ as well as measures η_1 and η_2. Then* (4.54) *holds for all sufficiently small $\varepsilon > 0$.*

Proof. First, note that as explained at the beginning of the proof of Lemma 4.32 (**C1**) and (**C4**) hold. Let $W_{n,i}$ and $Z_{n,i,\varepsilon}$ be defined as in (4.36) and Lemma 4.28 for all $1 \le i \le n$ and every $\varepsilon > 0$. Since $(a+b)^2 \le 4a^2 + 4b^2$ for all $a, b \in \mathbb{R}$ we deduce from Lemma 4.28

$$\begin{aligned}0 &\le \sum_{i=1}^{k_n} \mathbb{E}_{P_{k_n,i}} \left(\log\left(g_{k_n,i}\right) \mathbf{1}_{\{|\log(g_{k_n,i})|<\varepsilon\}}\right)^2\\ &= \sum_{i=1}^{k_n} \mathbb{E}_{P_{k_n,i}} \left(\left(2W_{k_n,i} - \left(1 + Z_{k_n,i,\varepsilon}\right) W_{k_n,i}^2\right) \mathbf{1}_{\{|\log(g_{k_n,i})|<\varepsilon\}}\right)^2\\ &\le 4\left[\sum_{i=1}^{k_n} \mathbb{E}_{P_{k_n,i}} \left(2W_{k_n,i} \mathbf{1}_{\{|\log(g_{k_n,i})|<\varepsilon\}}\right)^2 + 2\sum_{i=1}^{k_n} \mathbb{E}_{P_{k_n,i}} \left(W_{k_n,i}^2 \mathbf{1}_{\{|\log(g_{k_n,i})|<\varepsilon\}}\right)^2\right]\end{aligned}$$

for all sufficiently small $\varepsilon > 0$ and every sufficiently large $n \in \mathbb{N}$. It remains to show that both sums in the last line of the previous inequality converge to 0 as $n \to \infty$ for all sufficiently small $\varepsilon > 0$. Let $y_\varepsilon \in C(\eta_1) \cap (0,\varepsilon)$ for every $\varepsilon > 0$. From Lemma 4.27, Lemma 4.29, (**C1**), (**C4**) and (2.1) we deduce that for all sufficiently small $\varepsilon > 0$

$$\begin{aligned}
&\sum_{i=1}^{k_n} \mathbb{E}_{P_{k_n,i}} \left(W_{k_n,i} \mathbf{1}_{\{|\log(g_{k_n,i})| < \varepsilon\}} \right)^2 \\
&= \frac{1}{4} \sum_{i=1}^{k_n} \mathbb{E}_{P_{k_n,i}} \left(\left(g_{k_n,i} - 1 - W_{k_n,i}^2 \right) \mathbf{1}_{\{|\log(g_{k_n,i})| < \varepsilon\}} \right)^2 \\
&\le \sum_{i=1}^{k_n} \mathbb{E}_{P_{k_n,i}} \left((g_{k_n,i} - 1) \mathbf{1}_{\{|\log(g_{k_n,i})| < \varepsilon\}} \right)^2 + \sum_{i=1}^{k_n} \mathbb{E}_{P_{k_n,i}} \left(W_{k_n,i}^2 \mathbf{1}_{\{|\log(g_{k_n,i})| < \varepsilon\}} \right)^2 \\
&\le \sum_{i=1}^{k_n} \varepsilon_{k_n,i}^2 \mathbb{E}_{P_{k_n,i}} \left((1 - f_{k_n,i}) \mathbf{1}_{\{|\log(g_{k_n,i})| \ge \varepsilon\}} \right)^2 + \sum_{i=1}^{k_n} \mathbb{E}_{P_{k_n,i}} \left(W_{k_n,i}^2 \mathbf{1}_{\{|\log(g_{k_n,i})| < \varepsilon\}} \right)^2 \\
&\le 4 \sum_{i=1}^{k_n} \varepsilon_{k_n,i}^2 \left[\mu_{k_n,i} \left(\log(g_{k_n,i}) > y_\varepsilon \right)^2 + P_{k_n,i} \left(\log(g_{k_n,i}) > y_\varepsilon \right)^2 \right] \\
&\quad + \max_{1 \le i \le k_n} \left\{ \mathbb{E}_{P_{k_n,i}} \left(W_{k_n,i}^2 \mathbf{1}_{\{\log(g_{k_n,i}) \le 2\varepsilon\}} \right) \right\} \sum_{i=1}^{k_n} \mathbb{E}_{P_{k_n,i}} \left(W_{k_n,i}^2 \mathbf{1}_{\{\log(g_{k_n,i}) \le 2\varepsilon\}} \right) \\
&\le 4\, \varepsilon_{k_n:k_n} \sum_{i=1}^{k_n} Q_{k_n,i} \left(\log(g_{k_n,i}) > y_\varepsilon \right) + 4\, \varepsilon_{k_n:k_n} \sum_{i=1}^{k_n} P_{k_n,i} \left(\log(g_{k_n,i}) > y_\varepsilon \right) + o(1) \\
&= o(1) \quad \text{as } n \to \infty.
\end{aligned}$$

Furthermore, we conclude from (4.47) of Lemma 4.27, Lemma 4.28 and Lemma 4.29 that for all sufficiently large $n \in \mathbb{N}$ and every sufficiently small $\varepsilon > 0$

$$\begin{aligned}
&\sum_{i=1}^{k_n} \mathbb{E}_{P_{k_n,i}} \left((1 + Z_{k_n,i,\varepsilon}) W_{k_n,i}^2 \mathbf{1}_{\{|\log(g_{k_n,i})| < \varepsilon\}} \right)^2 \\
&\le 2 \max_{1 \le i \le k_n} \left\{ \mathbb{E}_{P_{k_n,i}} \left(W_{k_n,i}^2 \mathbf{1}_{\{|\log(g_{k_n,i})| \le 2\varepsilon\}} \right) \right\} \sum_{i=1}^{k_n} \mathbb{E}_{P_{k_n,i}} \left(W_{k_n,i}^2 \mathbf{1}_{\{|\log(g_{k_n,i})| \le 2\varepsilon\}} \right) \\
&= o(1) \quad \text{as } n \to \infty. \qquad \blacksquare
\end{aligned}$$

4.3. Trivial accumulation points

In this section we present sufficient and necessary conditions for the convergence of $\{P_{(k_n)}, Q_{(k_n)}\}$ to one of the trivial binary experiments $\{\epsilon_0, \epsilon_0\}$ or $\{\epsilon_{-\infty}, \epsilon_\infty\}$. We want to emphasise that the conditions are quite similar to the ones for non-trivial limits. This reduces the effort in calculations if we are interested in the non-trivial and trivial cases for a given model. We also present some results under more restrictive conditions and first results for the exponential families introduced in Section 2.3.1.

As mentioned in Remark 4.26(iii) (**A**) and (**A normal**) include the convergence to the uninformative experiment. Thus, we can verify this convergence by showing that (**B**) or (**B'**) or (**B normal**) holds for trivial measures $\eta_1 \equiv 0 \equiv \eta_2$ and $\sigma^2 = 0$. But we can weaken these conditions more.

Corollary 4.35. *Let $y_0 > 0$. Under Assumption 2.1*(i) *the following conditions* (i)-(iii) *are equivalent for a subsequence $\{k_n : n \in \mathbb{N}\} \subseteq \mathbb{N}$.*

(i) The binary experiment $\{P_{(k_n)}, Q_{(k_n)}\}$ converges weakly to the uninformative experiment $\{\epsilon_0, \epsilon_0\}$ as $n \to \infty$.

(ii) The sum D_{k_n} given by (4.2) *converges to* 0 *as $n \to \infty$.*

(iii) One of the conditions (**B2a normal**)-(**B2c normal**) *is fulfilled for $y := y_0$. Furthermore, (**B3 normal**) is fulfilled for y_0 and $\sigma^2 = 0$.*

Remark 4.36. (i) Because y_0 can be chosen arbitrarily, condition (iii) is fulfilled for all $y_0 > 0$ if it is for one.

(ii) Suppose one of the conditions (**B2a normal**)-(**B2c normal**) for some $y_0 > 0$. By Lemma 4.16, Remark 4.17 and Lemma 4.20 (**B3 normal**) is fulfilled for y_0 and $\sigma^2 = 0$ if and only if

$$\sum_{i=1}^{k_n} \varepsilon_{k_n,i}^2 \, \mathbb{E}_{P_{k_n,i}} \left(f_{k_n,i}^2 \mathbf{1}_{\{\varepsilon_{k_n,i} f_{k_n,i} \le y_0\}} - 1 \right) \; \le \; o(1) \quad \text{as } n \to \infty.$$

(iii) To prove the sufficiency of (iii) for (ii), we slightly modify the methods of Cai et al. [10]. The following proof shows that this implication still holds if we replace $\{\varepsilon_{k_n,i} f_{k_n,i} \le y_0\}$ and $\{\varepsilon_{k_n,i} f_{k_n,i} > y_0\}$ by arbitrary sets $A_{k_n,i} \in \mathcal{A}$ and $A_{k_n,i}^c$ in (**B2 normal**) and in (**B3 normal**), respectively. The advantage of our special sets is that (iii) is even necessary. □

In the proof of Corollary 4.35 we apply the following lemma.

Lemma 4.37. *Suppose Assumption 2.1*(i). *Let D_n be given by* (4.2) *for all $n \in \mathbb{N}$. Then for every triangular array $(A_{n,i})_{i \le n}$ of sets $A_{n,i} \in \mathcal{A}$ and every $n \in \mathbb{N}$*

$$D_n \le \left[\sum_{i=1}^n \left(\frac{\varepsilon_{n,i}}{2} + 2\varepsilon_{n,i}^2 \right) \mu_{n,i}\left(A_{n,i}^c \right) \right] + \sum_{i=1}^n \varepsilon_{n,i}^2 \mathbb{E}_{P_{n,i}} \left(f_{n,i}^2 \mathbf{1}_{A_{n,i}} - 1 \right).$$

The inequality mentioned above holds, e.g., for the sets $A_{n,i}$ given by

$$A_{n,i} := \{\varepsilon_{n,i} f_{n,i} \le y_0\} \text{ for every } 1 \le i \le n \in \mathbb{N} \text{ and some } y_0 > 0. \tag{4.59}$$

Proof. Let $n \in \mathbb{N}$ be fixed. We have

$$\begin{aligned} D_n &= \sum_{i=1}^n 1 - \mathbb{E}_{P_{n,i}}(\sqrt{g_{n,i}}) \\ &\le \sum_{i=1}^n 1 - \mathbb{E}_{P_{n,i}} \left(\sqrt{1 - \varepsilon_{n,i} + \varepsilon_{n,i} f_{n,i} \mathbf{1}_{A_{n,i}}} \right). \end{aligned} \tag{4.60}$$

From the third binomial formula we deduce that for all $u \ge 0$

$$(u^2 - 1)^2 = (u-1)^2(u+1)^2 \ge \frac{1}{2}(u-1)^2.$$

Substituting $x = u^2 - 1$ yields

$$x^2 \ge \frac{1}{2}\left(\sqrt{1+x} - 1\right)^2 = \frac{x}{2} + 1 - \sqrt{1+x} \text{ for all } x \ge -1.$$

Hence,

$$1 - \sqrt{1+x} \le -\frac{x}{2} + x^2 \quad \text{for all } x \ge -1.$$

Applying this for $x = \varepsilon_{n,i}(f_{n,i}\mathbf{1}_{A_{n,i}} - 1)$ (pointwisely) we conclude from (4.60) that

$$\begin{aligned} D_n &\le \sum_{i=1}^n \left[-\frac{\varepsilon_{n,i}}{2} \mathbb{E}_{P_{n,i}} \left(f_{n,i}\mathbf{1}_{A_{n,i}} - 1 \right) + \varepsilon_{n,i}^2 \mathbb{E}_{P_{n,i}} \left(\left(f_{n,i}\mathbf{1}_{A_{n,i}} - 1 \right)^2 \right) \right] \\ &= \sum_{i=1}^n \left[\frac{\varepsilon_{n,i}}{2} \mu_{n,i}\left(A_{n,i}^c \right) + \varepsilon_{n,i}^2 \mathbb{E}_{P_{n,i}} \left(f_{n,i}^2 \mathbf{1}_{A_{n,i}} - 1 - 2 f_{n,i}\mathbf{1}_{A_{n,i}} + 2 \right) \right]. \end{aligned}$$

■

Proof of Corollary 4.35. By Lemma 4.3 and Corollary 4.25 it remains to prove that (iii) is sufficient for (ii). Assume (iii). By Remark 4.24 the conditions (**B2a normal**)-(**B2c normal**) are equivalent. Thus, (**B2c normal**) holds for $y = y_0$. Finally, (ii) follows immediately from Lemma 4.37. ■

If for some $y > 0$ one of the sums in (**B1b**), (**B2b**), (**B2c**) converges to ∞ or one of the probabilities in (**B1a**), (**B2a**) converges to 1 then we obtain $\eta_j(y, \infty) = \infty$ for $j = 1$ or $j = 2$ heuristically. In this case $\{P_{(k_n)}, Q_{(k_n)}\}$ converges weakly to $\{\epsilon_{-\infty}, \epsilon_\infty\}$ by the following corollary.

Corollary 4.38. *Let $y_0 > 0$. Under Assumption 2.1*(i) *the following conditions* (i)-(iii) *are equivalent for a subsequence $\{k_n : n \in \mathbb{N}\} \subseteq \mathbb{N}$.*

(i) The binary experiment $\{P_{(k_n)}, Q_{(k_n)}\}$ converges weakly to the full informative experiment $\{\epsilon_{-\infty}, \epsilon_\infty\}$ as $n \to \infty$.

(ii) The sum D_{k_n} given by (4.2) *converges to ∞ as $n \to \infty$.*

(iii) One of the following conditions (**D1**)-(**D4**) *holds:*

$$\lim_{n\to\infty} \sum_{i=1}^{k_n} \varepsilon_{k_n,i}^2 \, \mathbb{E}_{P_{k_n,i}} \left(f_{k_n,i}^2 \mathbf{1}_{\{\varepsilon_{k_n,i} f_{k_n,i} \le y_0\}} - 1 \right) = \infty, \tag{\textbf{D1}}$$

$$\lim_{n\to\infty} \sum_{i=1}^{k_n} \varepsilon_{k_n,i} \, \mu_{k_n,i} \left(\varepsilon_{k_n,i} f_{k_n,i} > y_0 \right) = \infty, \tag{\textbf{D2}}$$

$$\lim_{n\to\infty} \sum_{i=1}^{k_n} Q_{k_n,i} \left(\varepsilon_{k_n,i} f_{k_n,i} > y_0 \right) = \infty, \tag{\textbf{D3}}$$

$$\lim_{n\to\infty} Q_{(k_n)} \left(\max_{1\le i\le k_n} \{ \varepsilon_{k_n,i} f_{k_n,i}\left(X_{k_n,i}\right)\} \le y_0 \right) = 0. \tag{\textbf{D4}}$$

Remark 4.39. (i) Because $y_0 > 0$ can be chosen arbitrarily, condition (iii) is fulfilled for all $y_0 > 0$ if it is for one.

(ii) Suppose that (**D3**) or (**D4**) is fulfilled, where $Q_{k_n,i}$ is replaced by $P_{k_n,i}$. Lemma 4.20 yields that (iii) holds and so does (i). □

Proof. By Lemma 4.3 it is sufficient to show that (ii) and (iii) are equivalent.
"(ii) ⇒ (iii)": Conversely, suppose that none of the conditions (**D1**)-(**D4**) holds. Combining Lemma 4.37 and the fact, that (**D1**) and (**D2**) do not hold, yields

$$\liminf_{n\to\infty} D_{k_n} \ < \ \infty.$$

"(iii)⇒ (ii)": First, suppose one of the conditions (**D2**)-(**D4**). By Lemma 4.16 and (4.27) of Lemma 4.20 we obtain (**D2**) in all three cases. Define $A_{n,i}$ as in (4.59). Note that for all sufficiently large $n \in \mathbb{N}$, every $i = 1, \ldots, k_n$ and all $\omega \in A_{k_n,i}^c$

$$g_{k_n,i}(\omega) \ = \ \varepsilon_{k_n,i} f_{k_n,i}(\omega) - \varepsilon_{k_n,i} + 1 \ > \ y_0 - \varepsilon_{k_n:k_n} + 1 \ > \ \frac{y_0}{2} + 1 \ > \ 1. \tag{4.61}$$

Set

$$C := \left(\sqrt{\frac{y_0}{2}+1}+1\right)^{-1}.$$

By (4.61) and the third binomial formula

$$\begin{aligned} W_{k_n,i}\mathbf{1}_{A^c_{k_n,i}} &= \left(\sqrt{g_{k_n,i}}-1\right)\mathbf{1}_{A^c_{k_n,i}} \\ &= \frac{g_{k_n,i}-1}{\sqrt{g_{k_n,i}}+1}\mathbf{1}_{A^c_{k_n,i}} \quad \leq \quad C\left(g_{k_n,i}-1\right)\mathbf{1}_{A^c_{k_n,i}} \end{aligned} \tag{4.62}$$

for all sufficiently large $n \in \mathbb{N}$ and every $1 \leq i \leq k_n$. Note that $C < \frac{1}{2}$. From (**D2**), (4.26) of Lemma 4.20, (4.37), (4.44) of Lemma 4.27 and (4.62) we obtain

$$\begin{aligned} 2\,D_{k_n} &\geq \sum_{i=1}^{k_n} \mathbb{E}_{P_{k_n,i}}\left(W^2_{k_n,i}\mathbf{1}_{A^c_{k_n,i}}\right) \\ &= \sum_{i=1}^{k_n} \mathbb{E}_{P_{k_n,i}}\left(\left(g_{k_n,i}-1-2W_{k_n,i}\right)\mathbf{1}_{A^c_{k_n,i}}\right) \\ &\geq (1-2C)\sum_{i=1}^{k_n} \mathbb{E}_{P_{k_n,i}}\left(\left(g_{k_n,i}-1\right)\mathbf{1}_{A^c_{k_n,i}}\right) \\ &= (1-2C)\sum_{i=1}^{k_n} \mathbb{E}_{P_{k_n,i}}\left(\varepsilon_{k_n,i}\left(f_{k_n,i}-1\right)\mathbf{1}_{A^c_{k_n,i}}\right) \\ &= (1-2C)\left[\sum_{i=1}^{k_n} \varepsilon_{k_n,i}\mu_{k_n,i}\left(A^c_{k_n,i}\right) - \varepsilon_{k_n,i}P_0\left(A^c_{k_n,i}\right)\right] \\ &\geq (1-2C)\left(1-\frac{\varepsilon_{k_n:k_n}}{y_0}\right)\sum_{i=1}^{k_n} \varepsilon_{k_n,i}\mu_{k_n,i}\left(A^c_{k_n,i}\right) \to \infty \quad \text{as } n \to \infty. \end{aligned}$$

Now assume (**D1**). Again, we want to verify (ii). For this purpose it is sufficient to show that for every subsequence $\{k_{n,1} : n \in \mathbb{N}\}$ of $\{k_n : n \in \mathbb{N}\}$ there is a further subsequence $\{k_{n,2} : n \in \mathbb{N}\} \subseteq \{k_{n,1} : n \in \mathbb{N}\}$ such that $D_{k_{n,2}}$ tends to ∞ as $n \to \infty$. Let $\{k_{n,1} : n \in \mathbb{N}\}$ be an arbitrary subsequence of $\{k_n : n \in \mathbb{N}\}$. If (**D3**) holds for $\{k_{n,1} : n \in \mathbb{N}\}$ then we set $k_{n,2} := k_{n,1}$ for all $n \in \mathbb{N}$ and due to the previous case (ii) holds for $\{k_{n,2} : n \in \mathbb{N}\}$. If (**D3**) does not hold for $\{k_{n,1} : n \in \mathbb{N}\}$ then there exists a subsequence $\{k_{n,2} : n \in \mathbb{N}\}$ of $\{k_{n,1} : n \in \mathbb{N}\}$ such that

$$\lim_{n\to\infty} \sum_{i=1}^{k_{n,2}} Q_{k_{n,2},i}\left(\varepsilon_{k_{n,2},i}f_{k_{n,2},i} > y_0\right) \in [0,\infty).$$

By this and Lemma 4.20 the inequality (4.24) holds for $\{k_{n,2} : n \in \mathbb{N}\}$. Hence, combining (**D1**), Remark 4.17, (4.37), (4.41) and Lemma 4.27 there exists some $C_1 > 0$ such that

$$\begin{aligned} 2\, D_{k_{n,2}} &\geq \sum_{i=1}^{k_{n,2}} \varepsilon_{k_{n,2},i} \left(W^2_{k_{n,2},i} \mathbf{1}_{\{\varepsilon_{k_{n,2},i} f_{k_{n,2},i} \leq y_0\}} \right) \\ &\geq C_1 \sum_{i=1}^{k_{n,2}} \varepsilon^2_{k_{n,2},i} \left(\left(f_{k_{n,2},i} - 1 \right)^2 \mathbf{1}_{\{\varepsilon_{k_{n,2},i} f_{k_{n,2},i} \leq y_0\}} \right) \\ &= C_1 \sum_{i=1}^{k_{n,2}} \varepsilon^2_{k_{n,2},i} \left(f^2_{k_{n,2},i} \mathbf{1}_{\{\varepsilon_{k_{n,2},i} f_{k_{n,2},i} \leq y_0\}} - 1 \right) + o(1) \quad \to \quad \infty \quad \text{as } n \to \infty. \quad \blacksquare \end{aligned}$$

The following corollary is an immediate consequence of Corollary 4.25, Corollary 4.35, Corollary 4.38 and $\mu_{k_n,i} \ll P_{k_n,i}$, see (2.2).

Corollary 4.40. *Suppose Assumption 2.1*(i)*. Let $\{k_n : n \in \mathbb{N}\}$ be a subsequence of $\mathbb{N}$. Assume that for every fixed $y > 0$*

$$\max_{1 \leq i \leq k_n} \left\{ P_{k_n,i} \left(\varepsilon_{k_n,i} f_{k_n,i} > y \right) \right\} = 0 \tag{4.63}$$

for all $n \geq n_0(y)$ for some $n_0(y) \in \mathbb{N}$. Moreover, suppose that

$$\lim_{n \to \infty} \sum_{i=1}^{k_n} \varepsilon^2_{k_n,i} \, \mathbb{E}_{P_{k_n,i}} \left(f^2_{k_n,i} - 1 \right) =: \sigma^2 \in [0, \infty]. \tag{4.64}$$

Then we have:

*(i) (**A normal**) holds for σ^2 if and only if $\sigma^2 \in [0, \infty)$.*

(ii) The binary experiment $\{P_{(k_n)}, Q_{(k_n)}\}$ converges weakly to the full informative experiment $\{\epsilon_{-\infty}, \epsilon_\infty\}$ if and only if $\sigma^2 = \infty$.

(iii) The binary experiment $\{P_{(k_n)}, Q_{(k_n)}\}$ converges weakly to the uninformative experiment $\{\epsilon_0, \epsilon_0\}$ if and only if $\sigma^2 = 0$.

Remark 4.41. (i) If

$$\max_{1 \leq i \leq k_n} \left\{ \varepsilon_{k_n,i} \sup_{x \in \Omega} f_{k_n,i}(x) \right\} \to 0 \quad \text{as } n \to \infty \tag{4.65}$$

then (4.63) is fulfilled.

(ii) Due to Corollary 4.38 the equivalence in (ii) is still valid if (4.63) holds only for some $y_0 > 0$. Furthermore, the equivalence in (iii) even holds if (4.63) does not hold for any $y > 0$, see the following Lemma 4.42. □

Calculations with different distributions, e.g., the Gumbel distribution, showed us that, in particular for the dense case,

$$\max_{1\leq i\leq k_n} \{f_{k_n,i}(x)\} \leq C < \infty \text{ for every } x \in \Omega \text{ and all sufficiently large } n \in \mathbb{N}.$$

Under these circumstances, (4.63) and (4.65) hold obviously. Therefore, by Corollary 4.40 every accumulation point of $\{P_{(n)}, Q_{(n)}\}$ is either a Gaussian experiment or a trivial experiment. If (4.63) does not hold for all $y > 0$ then (4.64) can still be a useful condition:

Lemma 4.42. *Suppose that Assumption 2.1*(i) *and* (4.64) *hold for some* $\sigma^2 \in [0, \infty)$ *and a subsequence* $\{k_n : n \in \mathbb{N}\}$ *of* $\mathbb{N}$. *Then no accumulation point of* $\{P_{(k_n)}, Q_{(k_n)}\}$ *is equivalent to the full informative experiment* $\{\epsilon_{-\infty}, \epsilon_\infty\}$.
Additionally, if $\sigma^2 = 0$ *then* $\{P_{(k_n)}, Q_{(k_n)}\}$ *converges weakly to* $\{\epsilon_0, \epsilon_0\}$.

Proof. Let D_n be defined as in (4.2). By the third binomial formula, the definition of $g_{n,i}$, see (2.4), and (4.64) we have

$$\begin{aligned} 2D_{k_n} &= \sum_{i=1}^{k_n} \int \left(\sqrt{g_{k_n,i}} - 1\right)^2 \,\mathrm{d}P_{k_n,i} &&= \sum_{i=1}^{k_n} \int \frac{(g_{k_n,i} - 1)^2}{(g_{k_n,i} + 1)^2} \,\mathrm{d}P_{k_n,i} \\ &\leq \sum_{i=1}^{k_n} \int \frac{\varepsilon_{k_n,i}^2 \,(f_{k_n,i} - 1)^2}{(0+1)^2} \,\mathrm{d}P_{k_n,i} &&= \sum_{i=1}^{k_n} \varepsilon_{k_n,i}^2 \left[\int f_{k_n,i}^2 \,\mathrm{d}P_{k_n,i} - 1\right] \\ &\to \ \sigma^2 \quad \text{as } n \to \infty. \end{aligned} \tag{4.66}$$

By Corollary 4.38 there exists no subsequence $\{k_{n,1} : n \in \mathbb{N}\}$ of k_n such that $\{P_{(k_{n,1})}, Q_{(k_{n,1})}\}$ converges weakly to $\{\epsilon_{-\infty}, \epsilon_\infty\}$. If $\sigma^2 = 0$ then combining (4.66) and Corollary 4.35 yields that $\{P_{(k_n)}, Q_{(k_n)}\}$ converges weakly to $\{\epsilon_0, \epsilon_0\}$. ■

Remark 4.43. There are cases in which (4.64) holds for some $\sigma^2 \in (0, \infty)$ and at the same time $\{P_{(k_n)}, Q_{(k_n)}\}$ converges weakly to $\{\epsilon_0, \epsilon_0\}$. Moreover, there are cases in which (4.64) holds for $\sigma^2 = \infty$ and at the same time $\{P_{(k_n)}, Q_{(k_n)}\}$ converges weakly to a non-trivial binary experiment. □

Suppose that Corollary 4.38(i) holds. Heuristically, we would expect that Corollary 4.38(i) still holds if we only increase the signal probabilities $\varepsilon_{k_n,1}, \ldots, \varepsilon_{k_n,k_n}$ and keep the measures

$P_{n,i}$ and $\mu_{n,i}$ unchanged for all $1 \leq i \leq n \in \mathbb{N}$. We would also expect an analogous result to Corollary 4.35(i) if we decrease the signal probabilities. Both implications are already mentioned and verified in [12], see Lemma 8 and Remark 4 in their appendix.

Theorem 4.44. *Suppose Assumption 2.1*(i). *Let* $(\widetilde{\varepsilon}_{n,i})_{1\leq i\leq n\in\mathbb{N}}$ *be a further triangular array of real numbers in* $[0,1]$ *such that*

$$\max_{1\leq i\leq n} \widetilde{\varepsilon}_{n,i} = o(1) \text{ as } n \to \infty.$$

Define for all $1 \leq i \leq n \in \mathbb{N}$

$$\widetilde{Q}_{n,i} := (1-\widetilde{\varepsilon}_{n,i})P_{n,i} + \widetilde{\varepsilon}_{n,i}\mu_{n,i} \text{ and } \widetilde{Q}_{(n)} := \bigotimes_{i=1}^{n} \widetilde{Q}_{n,i}.$$

Let $\{k_n : n \in \mathbb{N}\}$ *be a subsequence of* $\mathbb{N}$.

(i) If $\varepsilon_{k_n,i} \leq \widetilde{\varepsilon}_{k_n,i}$ *for all* $1 \leq i \leq k_n$, $n \in \mathbb{N}$, *and if* $\{P_{(k_n)}, Q_{(k_n)}\}$ *converges weakly to the full informative experiment* $\{\epsilon_{-\infty}, \epsilon_{\infty}\}$ *then* $\left\{P_{(k_n)}, \widetilde{Q}_{(k_n)}\right\}$ *does so as well.*

(ii) If $\varepsilon_{k_n,i} \geq \widetilde{\varepsilon}_{k_n,i}$ *for all* $1 \leq i \leq k_n$, $n \in \mathbb{N}$, *and if* $\{P_{(k_n)}, Q_{(k_n)}\}$ *converges weakly to the uninformative experiment* $\{\epsilon_0, \epsilon_0\}$ *then* $\left\{P_{(k_n)}, \widetilde{Q}_{(k_n)}\right\}$ *does so as well.*

Proof. It is easy to show that

$$\varepsilon \mapsto d^2\Big(P\ ,\ (1-\varepsilon)P + \varepsilon Q\Big)$$

is non-decreasing in $[0,1]$ for every pair (P,Q) of probability measures on some measurable space $(\Omega, \mathcal{A})$, see Lemma 8 in [12]. By carefully reading the proof one notes that the word *decreasing* should be replaced by *non-decreasing* in their Lemma 8. Combining this monotonicity, Corollary 4.35 and Corollary 4.38 completes the proof. ■

Simplifications for exponential families

For one-parametric exponential families introduced in Section 2.3.1 we can simplify (4.64) as follows:

Lemma 4.45. *Under Assumption 2.11 the condition* (4.64) *for* $\sigma^2 \in [0,\infty]$ *holds if and only if*

$$\lim_{n\to\infty} \sum_{i=1}^{k_n} \varepsilon_{k_n,i}^2 \left[\frac{C(\vartheta_{k_n})^2}{C(2\vartheta_{k_n})} - 1\right] = \sigma^2. \tag{4.67}$$

Proof. The statement follows from (2.8) and basic calculations:

$$\int f_{k_n,i}^2 - 1 \,\mathrm{d}P_{k_n,i} = \int C\left(\vartheta_{k_n}\right)^2 \exp\left(-2\vartheta_{k_n} h\right) \,\mathrm{d}\mathcal{Q}_0 - 1$$
$$= \frac{C\left(\vartheta_{k_n}\right)^2}{C\left(2\vartheta_{k_n}\right)} \int \frac{\mathrm{d}\mathcal{Q}_{2\vartheta_{k_n}}}{\mathrm{d}\mathcal{Q}_0} \,\mathrm{d}\mathcal{Q}_0 - 1 \;=\; \frac{C\left(\vartheta_{k_n}\right)^2}{C\left(2\vartheta_{k_n}\right)} - 1. \qquad \blacksquare$$

Due to the previous lemma we are interested in the asymptotic behaviour of $\vartheta \to C(\vartheta)$. If $C\left(\vartheta_{k_n}\right)$ is known for sufficiently large n then by the following corollary it is possible to determine ϑ_{k_n} and $\varepsilon_{k_n,i}$ such that the limit of $\{P_{(k_n)}, Q_{(k_n)}\}$ is Gaussian or equal to the uninformative experiment $\{\epsilon_0, \epsilon_0\}$. If the asymptotic behaviour of $C\left(\vartheta_{k_n}\right)$ is unknown then the results of Section 2.3.2 may be used to determine it.

Corollary 4.46. *Let Assumption 2.11 and* (4.67) *be fulfilled for some* $\sigma^2 \in [0,\infty)$. *Moreover, assume that*

$$C\left(\vartheta_{k_n}\right) \varepsilon_{k_n:k_n} \to 0 \quad \textit{as } n \to \infty \tag{4.68}$$

$$\textit{and} \quad \exp\left[-\vartheta_{k_n} h(x)\right] \leq M \in (0,\infty) \textit{ for all } n \in \mathbb{N},\, x \in \Omega. \tag{4.69}$$

*Then (**A normal**) holds for* σ^2.

Proof. It is easy to see that (4.65) holds if (4.68) and (4.69) are fulfilled. Hence, applying Corollary 4.40, Remark 4.41(i) and Lemma 4.45 completes the proof. ■

4.4. Connection to extreme value theory

In the literature it is known that there is a strong connection between extreme value theory and convergence to infinitely divisible distributions, see, e.g., Janssen [36] or the Extrema Criterion 23.4C of Lóeve [49]. In this section we explain the mentioned connection in our setting and how to use it for verifying (**B1a**). Consequently, we can use this connection to find parameters and measures such that (**A**) is possibly fulfilled. To prove that (**A**) actually holds, Theorem 4.19 can be applied.

Let $Y_1, \ldots, Y_n$ be i.i.d. real-valued random variables on some probability space $(\Omega, \mathcal{A}, \mathcal{P})$ with non-degenerate cumulative distribution function F. A distribution function H is said to be non-degenerate if $H(s) \in (0,1)$ for some $s \in \mathbb{R}$. The extreme value theory deals with the asymptotic behaviour of the maximum statistic $Y_{n:n}$. It is easy to see that this statistic converges in probability to the endpoint $x^* = \inf\{x \in \mathbb{R} : F(x) = 1\}$ of the distribution of Y_1, where the convention $\inf \emptyset = \infty$ is used. Hence, a normalization is needed to get a non-degenerate limit distribution. In this context the linear normalisation $a_n^{-1}(Y_{n:n} - b_n)$ with $a_n > 0$ and $b_n \in \mathbb{R}$ for every $n \in \mathbb{N}$ is well understood and leads to a rich theory. For a thorough treatment of this theory we refer the reader to Resnick [53], [54] and Haan and Ferreira [28]. Note that $a_n^{-1}(Y_{n:n} - b_n)$ converges in distribution to a non-degenerate random variable Z if and only if we have for all

$$\lim_{n\to\infty} \mathcal{P}\left(\frac{Y_{n:n} - b_n}{a_n} \leq t\right) = \lim_{n\to\infty} (F(a_n t + b_n))^n = G(t) \text{ for all } t \in C(G), \tag{4.70}$$

where G is the non-degenerate cumulative distribution function of Z.

We have to admit that the connection between our problem and extreme value theory can only be used for a restricted version of our model, see the following remark for details.

Remark 4.47. Suppose that Assumption 2.1(iii) holds, where $P_n = P_1 =: P_0$ and f_n is non-decreasing for all $n \in \mathbb{N}$. Furthermore, let $Y_1, \ldots, Y_n$ be i.i.d. real-valued random variables on $(\Omega, \mathcal{A}, \mathcal{P})$ with $Y_1 \sim P_0$. Then for all $n \in \mathbb{N}$

$$\max_{1\leq i\leq k_n} \{\varepsilon_{k_n,i} f_{k_n,i}(Y_i)\} = \varepsilon_{k_n} f_{k_n}(Y_{k_n:k_n}) =: R_{k_n}. \tag{4.71}$$

(i) By (4.71) R_{k_n} converges in distribution to some real-valued random variable R on $(\Omega, \mathcal{A}, \mathcal{P})$ with

$$\mathcal{P}(R \leq t) > 0 \text{ for every } t > 0$$

if and only if (**B1a**) holds for all $x \in C(\eta_1) \cap (0, \infty)$, where the measure η_1 is given by

$$\eta_1(y, \infty) = -\log\Big(\mathcal{P}(R \leq e^y - 1)\Big) \text{ for all } y \in C(\eta_1) \cap (0, \infty).$$

The case, that R_{k_n} converges in probability to 0, is of specific interest because it leads to $\eta_1 \equiv 0$ and so to Gaussian limit experiments of $\{P_{(k_n)}, Q_{(k_n)}\}$.

(ii) Assume that

$$\lim_{n\to\infty} \mathcal{P}(R_{k_n} \leq e^x - 1) = 0 \tag{4.72}$$

for some $x > 0$. Then the probability in (**B1a**) converges to 0 and so by Lemma 4.16 the sum in (**B1b**) converges to ∞ for x. Consequently, by Lemma 4.20 the condition

(**D2**) of Corollary 4.38 is fulfilled for x. Finally, we conclude from Corollary 4.38 that $\{P_{(k_n)}, Q_{(k_n)}\}$ converges weakly to $\{\epsilon_{-\infty}, \epsilon_{\infty}\}$.
(4.72) is fulfilled, e.g., if R_{k_n} converges in distribution to a random variable, which is bounded away from zero ($\mathcal{P}$-a.s.).

(iii) Let (4.70) be fulfilled for some non-degenerate distribution function G, some sequence $(a_n)_{n\in\mathbb{N}}$ in $(0,\infty)$ and some sequence $(b_n)_{n\in\mathbb{N}}$ in $\mathbb{R}$. Let $Z \sim G$. We can rewrite (4.71) into

$$R_{k_n} = \psi_{k_n}\left(\frac{Y_{k_n:k_n} - b_{k_n}}{a_{k_n}}\right) \tag{4.73}$$
$$\text{with } \ \psi_{k_n}(x) := \varepsilon_{k_n} f_{k_n}\left(a_{k_n}x + b_{k_n}\right), \ x \in \mathbb{R}.$$

Suppose that $\psi_{k_n} : \mathbb{R} \to \mathbb{R}$ converges pointwisely to a continuous function $\psi : \mathbb{R} \to \mathbb{R}$. Since f_n is non-decreasing and $a_n > 0$, the function ψ_{k_n} is non-decreasing for all $n \in \mathbb{N}$ and so is ψ. Thus, ψ_{k_n} even converges to ψ uniformly on every compact interval. By Theorem 5.5 of Billingsley [7], an extension of the continuous mapping theorem, R_{k_n} converges in distribution to $R := \psi(Z)$.

(iv) We get similar results to (i) and (iii) if f_n is non-increasing for all $n \in \mathbb{N}$. In this case $Y_{k_n:k_n}$ in (4.71) and (4.73) is replaced by the minimum statistic $Y_{1:k_n}$. Because

$$Y_{1:k_n} = -\max_{1\le i\le k_n} \{-Y_i\}$$

the results about the asymptotic behaviour of maximum statistics can be transferred to the minimum statistic. □

At the end of this section we give an example how the above-mentioned remark can be applied.

Example 4.48. Consider the Gumbel model introduced in Example 2.8, i.e.,

$$f_n(x) = e^{\vartheta_n} \exp\left(-e^{-x}\left[e^{\vartheta_n} - 1\right]\right), \ x, \vartheta_n \in \mathbb{R}, \ n \in \mathbb{N}.$$

Let $\{k_n : n \in \mathbb{N}\}$ be a subsequence of $\mathbb{N}$ and $Y_1, \ldots, Y_{k_n}$ be i.i.d. P_0-distributed for all $n \in \mathbb{N}$. It is known that (4.70) holds for

$$a_n := 1, \ b_n := \log(n) \text{ and } G = \Lambda \text{ for all } n \in \mathbb{N}.$$

Clearly, f_n is non-decreasing for all $n \in \mathbb{N}$ and (4.73) holds for $\psi_n : \mathbb{R} \to \mathbb{R}$ given by

$$\psi_n(x) = \varepsilon_n e^{\vartheta_n} \exp\left(e^{-x} \frac{1}{n} \left(1 - e^{\vartheta_n}\right)\right), \ x \in \mathbb{R}, \ n \in \mathbb{N}.$$

By using Remark 4.47(i) and (iii) we want to determine sequences of parameters $(\vartheta_{k_n})_{n\in\mathbb{N}}$ and $(\varepsilon_{k_n})_{n\in\mathbb{N}}$ in $[0,\infty)$ and $[0,1]$, respectively, such that (**B1**) is fulfilled.

(i) Suppose that

$$\lim_{n\to\infty} \varepsilon_{k_n} e^{\vartheta_{k_n}} \ = \ 0. \tag{4.74}$$

Because $\vartheta_{k_n} \geq 0$ we have

$$\psi_{k_n}(x) \ \leq \ \varepsilon_{k_n} e^{\vartheta_{k_n}} \text{ for all } x \in \mathbb{R}.$$

Combining this, Remark 4.47(i) and (iii) yields (**B1**) for $\eta_1 \equiv 0$. Note that (4.74) is always fulfilled in the dense case, i.e., $\vartheta_{k_n} = o(1)$ as $n \to \infty$. In Section 5.2.1 and Section 5.2.2 we show that under some additional assumptions $\{P_{(k_n)}, Q_{(k_n)}\}$ converges weakly to some Gaussian experiment, see Corollaries 5.10(ii) and 5.13(ii) for details.

(ii) Suppose that

$$\lim_{n\to\infty} \varepsilon_{k_n} e^{\vartheta_{k_n}} = 1 \ \text{ and } \ \lim_{n\to\infty} k_n \, \mathrm{e}^{-\vartheta_{k_n}} = 1.$$

Clearly, it follows that $\lim_{n\to\infty} k_n \varepsilon_{k_n} = 1$ and that ψ_{k_n} converges pointwisely to function $\psi : \mathbb{R} \to \mathbb{R}$ given by

$$\psi(x) \ = \ \exp\left(-e^{-x}\right) \text{ for every } x \in \mathbb{R}.$$

Let Z be standard Gumbel distributed on $(\Omega, \mathcal{A}, \mathcal{P})$, i.e., the distribution function Λ of Z is equal to ψ. Thus, by Remark 4.47(iii)

$$R_{k_n} \xrightarrow{\mathrm{d}} R := \psi(Z) \sim \mathcal{U}(0,1) \text{ as } n \to \infty.$$

From this and Remark 4.47(i) we deduce (**B1**) for η_1 given by

$$\eta_1(x,\infty) = -\log\left(e^x - 1\right) \mathbf{1}_{(0,\log(2))}(x), \ x > 0.$$

Hence,

$$\frac{\mathrm{d}\eta_1}{\mathrm{d}\lambda\!\!\lambda}(x) = \frac{e^x}{e^x - 1}\mathbf{1}_{(0,\log(2))}(x),\ x \in \mathbb{R} \setminus \{0\}.$$

In Section 5.2.1 we show that under the above-mentioned assumptions $\{P_{(k_n)}, Q_{(k_n)}\}$ converges weakly to some non-trivial and non-Gaussian experiment, see Corollary 5.10(iii) for details. □

4.5. Outlook: further non-trivial accumulation points

During the study we recognised that there are accumulation points of $\{P, Q\}$ for certain models which are neither trivial nor fulfilling (**A**). While proof-reading this thesis we finally found the answer how to determine these accumulation points. In this section we present and discuss briefly some first results for this issue which complete the theory of Sections 4.2 and 4.3. Moreover, we give an example in the following chapter, see Remark 5.8(i). We refer the reader to an upcoming paper, which we plan to write after finishing this thesis, for a fuller treatment of this issue and for the proofs which are omitted here.

Using the ideas of the proof of Theorem 4.15 and modifying them slightly we can show the following lemma.

Lemma 4.49. *Suppose that Assumption 2.1*(i) *holds.* (**A1**) *holds for some real-valued random variable ξ_1 and some subsequence $\{k_n : n \in \mathbb{N}\}$ of $\mathbb{N}$ if and only if there exist some measure η on $(\mathbb{R} \setminus \{0\}, \mathcal{B}(\mathbb{R} \setminus \{0\}))$, some function $\psi : (0, \infty) \to [0, \infty)$ and some constant $\sigma^2 \in [0, \infty)$ such that* (i)-(iii) *hold.*

(i) We have

$$\eta(-\infty, 0) \ = \ 0 \ \text{ and } \ \eta(x, \infty) < \infty \text{ for all } x > 0.$$

*(ii) (**B1**) and (**B3**) are fulfilled for $\eta_1 := \eta$ and σ^2.*

(iii) For all sufficiently small $y \in C(\eta) \cap (0, \infty)$ we have

$$\lim_{n\to\infty} \sum_{i=1}^{k_n} \varepsilon_{k_n,i}\mu_{k_n,i}\Big(\varepsilon_{k_n,i} f_{k_n,i} > e^y - 1\Big) \ = \ \psi(y).$$

Moreover, if both above-mentioned, equivalent conditions hold then we obtain (a)-(c).

(a) ξ_1 *is infinitely divisible with Lévy characteristic* (γ, σ^2, η)*, where*

$$\gamma \;=\; -\frac{\sigma^2}{2} + \lim_{C(\eta)\ni\tau\searrow 0}\left(-\psi(\tau) + \int_{(\tau,\infty)} \frac{x}{1+x^2}\; \mathrm{d}\eta(x)\right).$$

(b) $(\mathbf{A2})$ *holds for some random variable* $\xi_2 : (\Omega, \mathcal{A}, \mathcal{P}) \to (\bar{\mathbb{R}}, \bar{\mathcal{B}})$*. Denote by* ν_j *the distribution of* ξ_j *for* $j = 1, 2$*. Then* $\nu_1 \ll \nu_2$*,*

$$\frac{\mathrm{d}\nu_1}{\mathrm{d}\nu_2}(x) \;=\; \exp(-x) \textit{ for all } x \in \mathbb{R} \textit{ and } \nu_2\left(\{-\infty\}\right) = 1 - \int \exp\; \mathrm{d}\nu_1.$$

If $\nu_2\left(\{-\infty\}\right) = 0$ *then* ξ_2 *is also real-valued (*$\mathcal{P}$*-a.s.).*

(c) *We have* $P_{(k_n)} \lhd Q_{(k_n)}$*.*

Combining Lemma 4.20, Corollary 4.38, Lemma 4.49 and subsequence arguments we can verify the following theorem.

Theorem 4.50. *Let Assumption 2.1 be fulfilled and* $\{k_n : n \in \mathbb{N}\}$ *be a subsequence of* $\mathbb{N}$*. Moreover, assume that* $\{P_{(k_n)}, Q_{(k_n)}\}$ *converges weakly to* $\{P, Q\}$ *as* $n \to \infty$*. Then either* $\{P, Q\}$ *is a full informative experiment, i.e., its standard form is* $\{\epsilon_{-\infty}, \epsilon_\infty\}$*, or* $(\mathbf{A1})$ *holds for some real-valued random variable* ξ_1*.*

The theory developed and presented in Sections 4.2 and 4.3 is not rich enough to explain all cases or, in other words, to determine all possible accumulation points of $\{P_{(n)}, Q_{(n)}\}$, see Section 5.2.1 for a counter example. By Theorem 4.50 every non-trivial accumulation point of $\{P_{(k_n)}, Q_{(k_n)}\}$ fulfils $(\mathbf{A1})$ for some real-valued ξ_1. For these accumulation points Lemma 4.49 explains how to determine the distributions of ξ_1 and ξ_2, where the latter one is in general a distribution on $(\bar{\mathbb{R}}, \bar{\mathcal{B}})$. The results presented in this section complete the theory of Sections 4.2 and 4.3 in the sense that we are now able to determine every accumulation point of $\{P_{(n)}, Q_{(n)}\}$.

5. Application to practical detection models

In this chapter we apply the theoretical results proved in Chapter 4 to some detection models, among others the ones introduced in Chapter 2. Donoho and Jin [20] suggested to use the so called *phase diagram* and *detection boundary* to illustrate their results concerning the heterogeneous normal mixture model. In the phase diagram a certain parametrisation plane (the signal probability $\varepsilon_n = \varepsilon_{n,i}$ and the signal strength ϑ_n are parametrised in some way, see, e.g., (1.2)) is split into two areas by the detection boundary. One of the areas is the detectable area, in which $\{P_{(n)}, Q_{(n)}\}$ converges weakly to the full informative experiment or, in other words, in which LLRT can completely separate the null and the alternative. The other one is the undetectable area, in which $\{P_{(n)}, Q_{(n)}\}$ converges weakly to the uninformative experiment or, in other words, LLRT cannot successfully separate the null and the alternative. An example of a phase diagram and a detection boundary can be found in Chapter 1, see Figure 1.1 and Equation (1.3). Beside determining the detection boundary we are interested in the question, what happens on the detection boundary. The answer for this question is already known in the literature for the heterogeneous and heteroscedastic normal mixtures, see [10, 32, 41]. In Section 5.1 we present the known results concerning the detection boundary and the asymptotic behaviour of LLR_n on it for these normal mixtures by using our notation. Moreover, we give the proof for the dense case, which was omitted in [10]. Note that we introduced the terms *sparse case* and *dense case* in Notation 2.5. In Section 5.2 and Section 5.3 we present results about the exponential families and the h-model introduced in Sections 2.3 and 2.4, respectively. For both we first prove a general result about trivial and non-trivial accumulation points of $\{P_{(n)}, Q_{(n)}\}$ for the case, that $\varepsilon_{n,i}$ may depend on i. Using this result we can determine the detection boundary for the case $\varepsilon_{n,i} = \varepsilon_n$ and, moreover, we discuss the asymptotic behaviour of $\{P_{(n)}, Q_{(n)}\}$ on it. One interesting observation is that the limits on the boundary are non-trivial but not always Gaussian. Note that in the literature it was already discovered that there are non-trivial and non-Gaussian limit experiments of $\{P_{(n)}, Q_{(n)}\}$, see [10, 32, 41] for the normal mixtures and see also [9] for a closely related detection problem. The results of Sections 5.1 to 5.3 are illustrated by phase diagrams, see Figures 5.1 to 5.3. At the end of this chapter we apply our results to the general class of exponential families suggested by Cai and Wu [12]. By doing this we slightly extend their results, see Section 5.4.

5.1. Heterogeneous and heteroscedastic normal mixtures

Ingster [32], and Donoho and Jin [20] and Jin [41], determined the detection boundary and discussed the asymptotic behaviour of LLR_n on it for the (sparse) heterogeneous normal mixture model, i.e., the model discussed in Example 2.6 with $\tau = 1$. Cai et al. [10] extended the results to the (sparse) heteroscedastic normal mixture model, i.e., to the case of general $\tau > 0$. We want to emphasise that they presented results not only for the sparse case but also for the dense case. However, they omitted the proofs for the results concerning the dense case. Moreover, they only mentioned that there are non-trivial limits on the boundary in the dense case without giving more details about them. In this section we present the above-mentioned results and fill the gap for the dense case. We reformulate the results of Cai et al. [10] in terms of binary experiments. We want to mention that these results can also be shown by using our techniques, see for example Appendix A.5. The results are illustrated by the phase diagrams for the dense and the sparse case, see Figure 5.1.

Theorem 5.1 (Detection boundary for the sparse case, see [10]). *Suppose that the heteroscedastic normal mixture model, see Example 2.6, is given, where*

$$\varepsilon_n := n^{-\beta} \text{ and } \vartheta_n := \sqrt{2r\log n} \text{ for all } n \in \mathbb{N} \text{ and some } \beta \in \left(\frac{1}{2}, 1\right),\ r \in (0,1).$$

Moreover, let the detection boundary be defined by

$$\rho^*(\beta,\tau) := \begin{cases} (2-\tau^2)\left(\beta - \frac{1}{2}\right) & \text{if } \frac{1}{2} < \beta \leq 1 - \frac{\tau^2}{4},\ \tau \in (0,\sqrt{2}). \\ (1-\tau\sqrt{1-\beta})^2 & \text{if } 1 - \frac{\tau^2}{4} < \beta < 1,\ \tau \in (0,\sqrt{2}). \\ 0 & \text{if } \frac{1}{2} < \beta \leq 1 - \frac{1}{\tau^2},\ \tau \geq \sqrt{2}. \\ (1-\tau\sqrt{1-\beta})^2 & \text{if } 1 - \frac{1}{\tau^2} < \beta < 1,\ \tau \geq \sqrt{2}. \end{cases} \tag{5.1}$$

(i) If $r < \rho^*(\beta,\tau)$ then $\{P_{(n)}, Q_{(n)}\}$ converges weakly to $\{\epsilon_0, \epsilon_0\}$.

(ii) If $r > \rho^*(\beta,\tau)$ then $\{P_{(n)}, Q_{(n)}\}$ converges weakly to $\{\epsilon_{-\infty}, \epsilon_\infty\}$.

(iii) Assume that $\beta \in \left(\frac{1}{2}, 1 - \frac{\tau^2}{4}\right]$, $\tau^2 < 2$ *and* $r = \rho^*(\beta, \tau)$. *Then (**A normal**) holds for*

$$\sigma^2 := \begin{cases} \left(\tau\sqrt{2-\tau^2}\right)^{-1} & \text{if } \beta < 1 - \frac{\tau^2}{4}. \\ \frac{1}{2}\left(\tau\sqrt{2-\tau^2}\right)^{-1} & \text{if } \beta = 1 - \frac{\tau^2}{4}. \end{cases}$$

(iv) Suppose

$$(\beta, \tau) \in \left(1 - \frac{\tau^2}{4}, 1\right) \times \left(0, \sqrt{2}\right) \cup \left(1 - \frac{1}{\tau^2}, 1\right) \times \left[\sqrt{2}, \infty\right) \tag{5.2}$$

and $r = \rho^*(\beta, \tau)$. *Moreover, replace* $\varepsilon_n = n^{-\beta}$ *by*

$$\varepsilon_n := n^{-\beta} \left(\log(n)\right)^{\frac{1}{2}\left(1 - \frac{1}{\tau}\sqrt{1-\beta}\right)} \quad \text{for all } n \in \mathbb{N}. \tag{5.3}$$

*Then (**A**) holds for real-valued, infinitely divisible random variables* ξ_1 *and* ξ_2. *Furthermore, the Lévy characteristic of* ξ_j *is given by* $(\gamma_j, 0, \eta_j)$, $j \in \{1, 2\}$, *where the Lévy measures are given by their* $\lambda\!\!\lambda$*-densities*

$$\frac{\mathrm{d}\eta_1}{\mathrm{d}\lambda\!\!\lambda}(x) = \frac{1}{c_1}\left(e^x - 1\right)^{c_2 - 3} e^x \text{ and } \frac{\mathrm{d}\eta_2}{\mathrm{d}\lambda\!\!\lambda}(x) = e^x \frac{\mathrm{d}\eta_1}{\mathrm{d}\lambda\!\!\lambda}(x),\ x > 0,$$
$$\text{with } c_1 := 2\sqrt{\pi}\tau^{c_3}\left(\tau - \sqrt{1-\beta}\right),\ c_2 := \frac{\tau - 2\sqrt{1-\beta}}{\tau - \sqrt{1-\beta}},\ c_3 := \frac{\sqrt{1-\beta}}{\tau - \sqrt{1-\beta}},$$

and where the constants γ_1, γ_2 *are given by* (4.16) *with* $\sigma^2 = 0$.

Remark 5.2. (i) If we do not add the logarithmic factor in the definition of ε_n, see (5.3), then $\{P_{(n)}, Q_{(n)}\}$ converges weakly to $\{\epsilon_0, \epsilon_0\}$ under the assumptions of (iv).

(ii) By carefully reading the proof in [10], see in particular the top of page 658, there must be an additional factor $\frac{1}{2}$ in the exponent of the logarithmic term in their definition of ε_n. The definition of ε_n in (5.3) is the corrected version. □

Proof. The statements of (i)-(iii) were proved by Cai et al. [10], see their Theorems 3 to 5. It remains to prove the statement of (iv). Let φ_{ξ_j}, $j \in \{1, 2\}$, be the characteristic function of ξ_j and ψ_{ξ_j} be the function in the exponent, i.e.,

$$\varphi_{\xi_j} = \exp\left(\psi_{\xi_j}\right).$$

By carefully completing the omitted parts of the proofs of Theorems 5 and 6 in [10] it is sufficient to show that for all $t \in \mathbb{R}$

$$\begin{aligned}
&\psi_{\xi_1}(t) = \psi^0_{\beta,\tau}(t) \text{ and } \psi_{\xi_2}(t) = \psi^0_{\beta,\tau}(t) + \psi^1_{\beta,\tau}(t), \\
&\text{where } \psi^0_{\beta,\tau}(t) := \frac{1}{c_1}\int_{-\infty}^{\infty}\Big[\exp\left(\mathbf{i}t\log\left[1+e^y\right]\right) - 1 - \mathbf{i}te^y\Big]e^{(c_2-2)y}\,\mathrm{d}y \\
&\text{and} \quad \psi^1_{\beta,\tau}(t) := \frac{1}{c_1}\int_{-\infty}^{\infty}\Big[\exp\left(\mathbf{i}t\log\left[1+e^y\right]\right) - 1\Big]e^{(c_2-1)y}\,\mathrm{d}y.
\end{aligned} \tag{5.4}$$

By using the substitution $x = \log\left[1+e^y\right]$ and considering (4.16) it follows that for all $t \in \mathbb{R}$

$$\begin{aligned}
\psi^0_{\beta,\tau}(t) &= \frac{1}{c_1}\int_{(0,\infty)}\Big(e^{\mathbf{i}xt} - 1 - \mathbf{i}t\left(e^x - 1\right)\Big)\left(e^x - 1\right)^{(c_2-3)}e^x\,\mathrm{d}x \\
&= \int_{(0,\infty)}\left(e^{\mathbf{i}xt} - 1 - \frac{\mathbf{i}xt}{1+x^2} + \frac{\mathbf{i}xt}{1+x^2} - \mathbf{i}t\left(e^x-1\right)\right)\frac{\mathrm{d}\eta_1}{\mathrm{d}\lambda}(x)\,\mathrm{d}\lambda(x) \\
&= \mathbf{i}\gamma_1 t + \int_{(0,\infty)}\left(e^{\mathbf{i}xt} - 1 - \frac{\mathbf{i}xt}{1+x^2}\right)\frac{\mathrm{d}\eta_1}{\mathrm{d}\lambda}(x)\,\mathrm{d}\lambda(x) \\
&= \psi_{\xi_1}(t).
\end{aligned}$$

Analogously, the second equation in (5.4) can be proved. ■

Theorem 5.3 (Detection boundary for the dense case, see [10]). *Suppose that the heteroscedastic normal mixture model, see Example 2.6, is given, where*

$$\varepsilon_n = n^{-\beta} \text{ and } \vartheta_n := n^{-r} \text{ for all } n \in \mathbb{N} \text{ and some } \beta \in \left(0, \frac{1}{2}\right),\ r \in \left(0, \frac{1}{2}\right).$$

Moreover, let the detection boundary be defined by

$$\rho^*_{dense}(\beta,\tau) : = \begin{cases} \infty & \text{if } \tau \neq 1. \\ \frac{1}{2} - \beta & \text{if } \tau = 1. \end{cases} \tag{5.5}$$

(i) If $r > \rho^*_{dense}(\beta,\tau)$ then $\{P_{(n)}, Q_{(n)}\}$ converges weakly to $\{\epsilon_0, \epsilon_0\}$.

(ii) If $r < \rho^*_{dense}(\beta,\tau)$ then $\{P_{(n)}, Q_{(n)}\}$ converges weakly to $\{\epsilon_{-\infty}, \epsilon_\infty\}$.

*(iii) If $r = \rho^*_{dense}(\beta,\tau)$ then (**A normal**) holds for $\{k_n : n \in \mathbb{N}\} = \mathbb{N}$ and $\sigma^2 = 1$.*

Remark 5.4. The statements of (i) and (ii) are equivalent to the ones in Theorems 2.4 and 2.5 in [10]. Furthermore, Cai et al. [10] mentioned that there are non-trivial accumulation points of $\{P_{(n)}, Q_{(n)}\}$ in the dense case but they did not present any details about these points. The statement of (iii) fills this gap. □

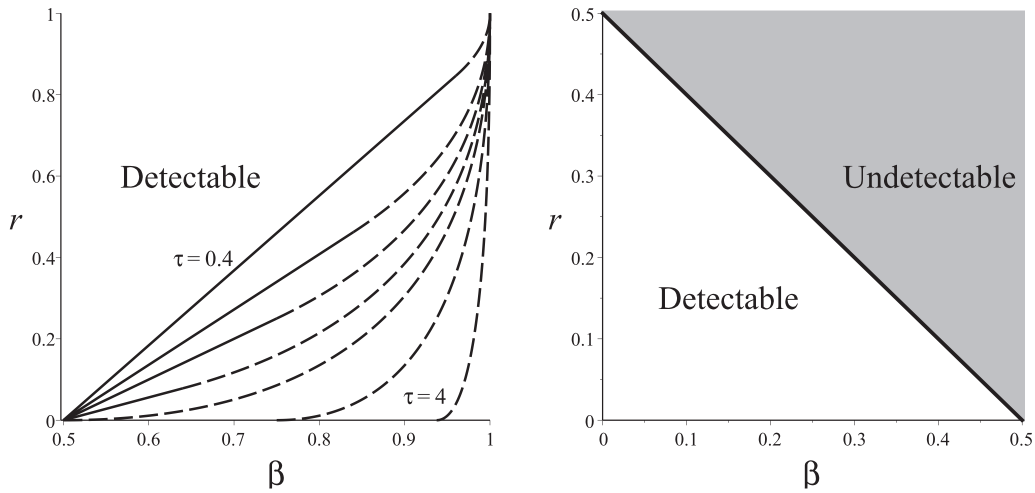

Figure 5.1.: Detection boundaries for the sparse (left) and the dense (right) heteroscedastic normal mixture model. Left: The function $\beta \mapsto \rho^*(\beta, \tau)$, see (5.1), is plotted for $\tau \in \{0.4, 0.8, 1, 1.2, \sqrt{2}, 2, 4\}$. If $r > \rho^*(\beta, \tau)$ the limit experiment of $\{P_{(n)}, Q_{(n)}\}$ is $\{\epsilon_{-\infty}, \epsilon_\infty\}$ (Detectable). If $r < \rho^*(\beta, \tau)$ it is $\{\epsilon_0, \epsilon_0\}$. If $r = \rho^*(\beta, \tau)$ it is Gaussian on the linear part (solid) and non-Gaussian on the quadratic part (dashed). Right: $\beta \mapsto \rho^*_{\text{dense}}(\beta, \tau)$, see (5.5), is plotted for $\tau = 1$. If $r > \rho^*_{\text{dense}}(\beta, 1)$ the limit experiment is $\{\epsilon_0, \epsilon_0\}$ (Undetectable). If $r < \rho^*_{\text{dense}}(\beta, 1)$ it is $\{\epsilon_{-\infty}, \epsilon_\infty\}$ (Detectable). If $r = \rho^*_{\text{dense}}(\beta, 1)$ it is Gaussian.

Proof. (i) follows immediately from Remark 4.1(i) by Theorem 8.8, which we will show in Part II, because if any test can completely separate the null and the alternative then, clearly, LLRT can do so. Hence, we only need to discuss the case $\tau = 1$. In Section 5.2.1 we discuss the dense case for exponential family models including the heterogeneous normal mixture model, i.e., the case $\tau = 1$. We can deduce (i)-(iii) for $\tau = 1$ immediately from Corollary 5.13, which we prove there. ∎

5.2. Exponential families

5.2.1. Sparse case

In this section we discuss the behaviour of LLR_n for the model introduced in Assumption 2.16. First, we present and prove a result for the general case that the signal probability may depend on i. After that we discuss the case $\varepsilon_{n,i} = \varepsilon_n$ in the context of the detection boundary. As in the previous section we show that the limits on the detection boundary are non-trivial. In particular, we obtain Gaussian but also non-Gaussion limits on the boundary. The results are illustrated by a phase diagram, which can be found in the next section, see Figure 5.2 on p. 79. At the end of this section we discuss the specific models introduced in Examples 2.8 to 2.10.

Theorem 5.5. *Let $p \geq 0$ and $L : (0, \infty) \to (0, \infty)$ be a slowly varying function at infinity. If $p = 0$ we assume additionally that $L(\vartheta)$ converges to 0 as $\vartheta \to \infty$. Suppose that Assumption 2.16 holds and that for some subsequence $\{k_n : n \in \mathbb{N}\}$ of $\mathbb{N}$*

$$C(\vartheta)^{-1} = \omega(\vartheta) \sim_{asy} \vartheta^{-p} L(\vartheta) \quad \text{as } \vartheta \to \infty \tag{5.6}$$

$$\text{and} \quad \sum_{i=1}^{k_n} \varepsilon_{k_n,i}^2 \frac{\vartheta_{k_n}^p}{L(\vartheta_{k_n})} \to M \in [0, \infty] \quad \text{as } n \to \infty. \tag{5.7}$$

(i) If $M = 0$ then $\{P_{(k_n)}, Q_{(k_n)}\}$ converges weakly to $\{\epsilon_0, \epsilon_0\}$.

(ii) If $M \in (0, \infty)$ and

$$\varepsilon_{k_n:k_n} \frac{\vartheta_{k_n}^p}{L(\vartheta_{k_n})} \to 0 \quad \text{as } n \to \infty \tag{5.8}$$

*then (**A normal**) is fulfilled for $\sigma^2 := M 2^{-p}$.*

(iii) If $M = \infty$ and (5.8) hold then $\{P_{(k_n)}, Q_{(k_n)}\}$ converges weakly to $\{\epsilon_{-\infty}, \epsilon_{\infty}\}$.

Remark 5.6. (i) By Lemma 2.15 and the fact, that $L(\vartheta)$ converges to 0 as $\vartheta \to \infty$ if $p = 0$, we have

$$\frac{\vartheta_{k_n}^p}{L(\vartheta_{k_n})} \to \infty \text{ as } n \to \infty.$$

Hence, if (5.7) holds for some $M \in [0, \infty)$ then

$$\sum_{i=1}^{k_n} \varepsilon_{k_n,i}^2 \to 0 \quad \text{as } n \to \infty. \tag{5.9}$$

(ii) If

$$\lim_{n\to\infty} \sum_{i=1}^{k_{n,1}} \varepsilon_{k_{n,1},i}^2 \;>\; 0$$

for some subsequence $\{k_{n,1} : n \in \mathbb{N}\}$ of $\{k_n : n \in \mathbb{N}\}$ then by Remark 4.5(ii) $\{P_{(k_{n,1})}, Q_{(k_{n,1})}\}$ converges weakly to $\{\epsilon_{-\infty}, \epsilon_\infty\}$.

(iii) If $\omega(\vartheta)$ is unknown for large ϑ then Theorem 2.19 can be applied to determine the asymptotic behaviour of it. If, e.g., (2.14) and (2.15) hold for some slowly varying function $\widetilde{L}$, $p \geq 0$ and $c, \nu > 0$ then (5.6) holds for the same p and L given by

$$L(\vartheta) := c^{-p}\Gamma(p+1)\widetilde{L}(\vartheta^{\frac{1}{\nu}}) \quad \text{for all } \vartheta > 0.$$

(iv) If $M \in (0,\infty)$ then (5.8) is equivalent to

$$\frac{\varepsilon_{k_n:k_n}}{\sum_{i=1}^{k_n} \varepsilon_{k_n,i}^2} \;\to\; 0 \quad \text{as } n \to \infty.$$

(v) Let $\varepsilon_{n,i} := \varepsilon_n := n^{-\beta+o(1)}$ for some $\beta \in (\frac{1}{2}, 1)$ and all $1 \leq i \leq n \in \mathbb{N}$. Then

$$\frac{\varepsilon_{k_n:k_n}}{\sum_{i=1}^{k_n} \varepsilon_{k_n,i}^2} = k_n^{\beta-1+o(1)} = o(1) \quad \text{as } n \to \infty.$$

Hence, we obtain (5.8) from (iv). □

Proof of Theorem 5.5. First, observe that due to Remark 5.6(i) and (ii) we can assume (5.9) without loss of generality. We deduce from (2.8) of Assumption 2.11, Assumption 2.16(iii), (5.6) and (5.8) that under (ii) and under (iii), respectively,

$$\max_{1\leq i\leq k_n} \left\{ \varepsilon_{k_n,i} \sup_{x\geq a} f_{k_n,i}(x) \right\} \;\leq\; \varepsilon_{k_n:k_n} C(\vartheta_{k_n}) \;\sim_{\text{asy}}\; \varepsilon_{k_n:k_n} \frac{\vartheta_{k_n}^p}{L(\vartheta_{k_n})} \;=\; o(1)$$

as $n \to \infty$. Consequently, by Corollary 4.40, Remark 4.41(i), Lemma 4.42 and Lemma 4.45 it is sufficient for the whole proof of Theorem 5.5 to show that

$$\sum_{i=1}^{k_n} \varepsilon_{k_n,i}^2 \frac{C(\vartheta_{k_n})^2}{C(2\vartheta_{k_n})} \to \frac{M}{2^p} \quad \text{as } n \to \infty.$$

Finally, observe that by (5.6)

$$\frac{C(\vartheta_{k_n})^2}{C(2\vartheta_{k_n})} \sim_{\text{asy}} \frac{\vartheta_{k_n}^{2p}}{2^p \vartheta_{k_n}^p} \frac{L(2\vartheta_{k_n})}{L(\vartheta_{k_n})^2} \sim_{\text{asy}} \frac{\vartheta_{k_n}^p}{2^p L(\vartheta_{k_n})} \quad \text{as } n \to \infty.$$ ■

Now we consider the more specific case that $\varepsilon_{n,i}$ does not depend on i.

Corollary 5.7. *Let $p \geq 0$ and $L : (0, \infty) \to (0, \infty)$ be a slowly varying function at infinity. If $p = 0$ then we assume additionally that $L(\vartheta)$ converges to 0 as $\vartheta \to \infty$. Suppose that Assumption 2.16 and (5.6) hold. For some $r > 0$ and $\beta > \frac{1}{2}$ let*

$$\vartheta_n \sim_{asy} n^r \quad \text{as } n \to \infty \quad \text{and} \quad \varepsilon_{n,i} := \varepsilon_n := n^{-\beta} \text{ for all } 1 \leq i \leq n \in \mathbb{N}.$$

Define the detection boundary for β by

$$\beta^{\#}_{Exp}(r,p) := \frac{(rp \wedge 1) + 1}{2}. \tag{5.10}$$

(i) If $\beta > \beta^{\#}_{Exp}(r,p)$ then $\{P_{(n)}, Q_{(n)}\}$ converges weakly to $\{\epsilon_0, \epsilon_0\}$.

(ii) If $\beta > 1$, $\beta = \beta^{\#}_{Exp}(r,p)$ and we replace $\varepsilon_n = n^{-\beta}$ by

$$\varepsilon_n = n^{-\beta}\sqrt{L(n^r)} \text{ for all } n \in \mathbb{N} \tag{5.11}$$

*then **(A normal)** is fulfilled for $\sigma^2 := 2^{-p}$.*

*(iii) Suppose that L is a constant function equal to some constant $K > 0$. If $p > 0$, $r = \frac{1}{p}$ and $\beta = \beta^{\#}_{Exp}(\frac{1}{p}, p)$ then **(A)** holds for some ξ_1 and ξ_2. Moreover, the Lévy characteristic of ξ_j, $j \in \{1,2\}$, equals $(\gamma_j, 0, \eta_j)$, where the Lévy measure η_j is uniquely determined by its $\lambda\!\!\!\lambda$-density*

$$\frac{\mathrm{d}\eta_j}{\mathrm{d}\lambda\!\!\!\lambda}(x) \;=\; \frac{K}{\Gamma(p)}\frac{e^{jx}}{e^x - 1}\Big(-\log\left[(e^x - 1)K\right]\Big)^{p-1}\mathbf{1}_{(0,\log(K^{-1}+1))}(x),\ x \in \mathbb{R}\setminus\{0\},$$

and γ_j is given by (4.16) with $\sigma^2 = 0$.

*(iv) Suppose $p > 0$, $r > \frac{1}{p}$ and $\beta = \beta^{\#}_{Exp}(\frac{1}{p}, r)$. Then **(B1)** holds for $\eta_1 \equiv 0$, **(B3)** is fulfilled for $\sigma^2 = 0$ and*

$$\lim_{n\to\infty} n\varepsilon_n\mu_n\Big(\varepsilon_n f_n > y\Big) \;=\; 1 \text{ for all } y > 0.$$

(v) If $\beta < \beta^{\#}_{Exp}(r,p)$ then $\{P_{(n)}, Q_{(n)}\}$ converges weakly to $\{\epsilon_{-\infty}, \epsilon_{\infty}\}$.

Remark 5.8. (i) Suppose that the assumptions of (iv) hold. By Lemma A.23 there exist a subsequence $\{k_n : n \in \mathbb{N}\} \subseteq \mathbb{N}$ and a binary experiment $\{P, Q\}$ such that $\{P_{(k_n)}, Q_{(k_n)}\}$ converges weakly to $\{P, Q\}$ as $n \to \infty$. But we can conclude from

Theorem 4.19 and Corollary 4.38 that $\{P, Q\}$ is neither full informative nor uninformative nor fulfilling Condition (**A**). In Section 4.5 we briefly discussed this issue in general. Applying Lemma 4.49 yields (**A1**) for $\xi_1 : (\Omega, \mathcal{A}, \mathcal{P}) \to (\mathbb{R}, \mathcal{B})$ with Lévy characteristic $(-1, 0, 0)$, i.e., $\xi_1 = -1$ ($\mathcal{P}$-a.s.). Finally, by Lemma 4.49

$$\left\{P_{(n)}, Q_{(n)}\right\} \text{ converges weakly to } \left\{\epsilon_{-1}\,,\, e^{-1}\epsilon_{-1} + (1 - e^{-1})\epsilon_\infty\right\}$$

as $n \to \infty$. A similar result can be proven, e.g., for the sparse heteroscedastic normal mixture model discussed in Theorem 5.1 if $r > 1$ and $\beta = 1$.

(ii) The above-mentioned result can also be formulated by using the detection boundary $\rho^*_{Exp,s}$ for the parameter r given by

$$\rho^*_{Exp,s}(\beta, p) := \begin{cases} \frac{1}{p}(2\beta - 1) & \text{if } p > 0 \text{ and } \beta \in \left(\frac{1}{2}, 1\right]. \\ \infty & \text{if } p = 0 \text{ or } \beta > 1. \end{cases} \tag{5.12}$$

□

Proof. First, we prove (i), (ii) and (v) by applying Theorem 5.5. Second, we give the proof of (iii). Finally, we verify (iv).

Observe that by Remark 4.5(i) $\{P_{(n)}, Q_{(n)}\}$ converges weakly to $\{\epsilon_0, \epsilon_0\}$ for all $\beta > 1$. Moreover, if $\beta \leq 1$ and $\beta^{\#}_{Exp}(r, p) < \beta$ then by Lemma 2.15

$$n\varepsilon_n^2 \frac{\vartheta_n^p}{L(\vartheta_n)} \sim_{\text{asy}} n^{1-2\beta+pr+o(1)} = n^{2\beta^{\#}_{Exp}(r,p)-2\beta+o(1)} \to 0 \text{ as } n \to \infty.$$

Consequently, applying Theorem 5.5(i) proves (i). Under (ii) we have

$$n\varepsilon_n^2 \frac{\vartheta_n^p}{L(\vartheta_n)} \to 1 \quad \text{as } n \to \infty. \tag{5.13}$$

From Lemma 2.15 and Remark 5.6(v) we obtain (5.8). Combining this, (5.13), Theorem 5.5(ii) and Remark 5.6(v) yields (ii).

Now suppose $\beta < \beta^{\#}_{Exp}(r, p)$. Hence, $\beta < 1$. By Lemma 2.15

$$n\varepsilon_n^2 \frac{\vartheta_n^p}{L(\vartheta_n)} \sim_{\text{asy}} n^{1-2\beta+pr+o(1)} \geq n^{2\beta^{\#}_{Exp}(r,p)-2\beta+o(1)} \to \infty \text{ as } n \to \infty.$$

Thus, we can conclude (v) from Theorem 5.5(iii) and Remark 5.6(v).

Now, consider $p > 0$, $r = \frac{1}{p}$, $\beta = \beta^{\#}_{Exp}(r, p)$ and $L \equiv K$. Note that $\beta = 1$. By Theorem 4.19

it is sufficient for (iii) to show (**B3**) for $\sigma^2 := 0$ and

$$\lim_{n\to\infty} n\varepsilon_n\mu_n\Big(x \geq a : \varepsilon_n f_n(x) > y\Big) \;=\; (\eta_2 - \eta_1)\Big(\log(1+y), \infty\Big) \tag{5.14}$$

for all $y \in (0,\infty) \setminus \{K^{-1}\}$, where by substituting $e^{-z} = [e^x - 1]K$

$$\begin{aligned}
&(\eta_2 - \eta_1)\Big(\log(1+y), \infty\Big)\\
&= \frac{1}{\Gamma(p)} \int_{(\log(1+y),\infty)} K\frac{e^{2x} - e^x}{(e^x - 1)} \Big(-\log\left[(e^x - 1)K\right]\Big)^{p-1} \mathbf{1}_{(0,\log(\frac{1}{K}+1))}(x)\,\mathrm{d}x\\
&= \frac{1}{\Gamma(p)} \int_{(\log(1+y),\log(K^{-1}+1))} \Big(-\log\left[(e^x - 1)K\right]\Big)^{p-1} K e^x\,\mathrm{d}x\,\mathbf{1}_{(0,K^{-1})}(y)\\
&= \frac{1}{\Gamma(p)} \int_0^{-\log(yK)} z^{p-1}e^{-z}\,\mathrm{d}z\,\mathbf{1}_{(0,K^{-1})}(y).
\end{aligned}$$

Define for all $y \in (0,\infty) \setminus \{K^{-1}\}$ and every $n \in \mathbb{N}$

$$A_{n,y} := \{x \geq a : \varepsilon_n f_n(x) \leq y\} = \Big\{x \geq a : n^{-1}C(\vartheta_n)\exp\left[-\vartheta_n h(x)\right] \leq y\Big\}.$$

By (5.6)

$$n^{-1}C(\vartheta_n) \;\sim_{\text{asy}}\; n^{-1+rp}K^{-1} \;=\; K^{-1} \text{ as } n \to \infty. \tag{5.15}$$

Thus, there exists a sequence $(\alpha_n)_{n\in\mathbb{N}}$ with $\lim_{n\to\infty}\alpha_n = 0$ such that

$$A_{n,y} \;=\; \Big\{x \geq a : \vartheta_n h(x) \geq -\left[1 + \alpha_n\right]\log\left(yK\right)\Big\}$$

for all $y \in (0,\infty) \setminus \{K^{-1}\}$ and every $n \in \mathbb{N}$. Clearly, for all fixed $y \in (0, K^{-1})$

$$G_{n,y} := -\left[1 + \alpha_n\right]\log\left(yK\right) > 0$$

if $n \in \mathbb{N}$ is sufficiently large. Moreover, for all fixed $y \in (K^{-1},\infty)$

$$G_{n,y} := -\left[1 + \alpha_n\right]\log\left(yK\right) < 0 \quad\text{and}\quad A_{n,y} = [a,\infty) \tag{5.16}$$

if $n \in \mathbb{N}$ is sufficiently large. Note that from Theorem 2.19(i) we obtain

$$\mathcal{Q}_0^h[0,t] \;\sim_{\text{asy}}\; t^p\,\frac{K}{p\Gamma(p)} \qquad \text{as } t \searrow 0,$$

where Γ is defined in (2.11). In other words, (2.13) holds for p and the slowly varying function $\widetilde{L} \equiv K$. In the following we apply Lemma 2.21 with $M = -2\log(yK)$,

$$\psi_n(x) := e^{-2x}\mathbf{1}_{[0,G_{n,y})}(x),\ x \geq 0,\ n \in \mathbb{N}, \text{ and } \psi_0(x) := e^{-2x}\mathbf{1}_{[0,-\log(yK))}(x),\ x \geq 0,$$

where $y \in (0, K^{-1})$. Note that in this case the set E from Lemma 2.21 equals $\{-\log(yK)\}$. Hence, by Lemma 2.21, (5.6) and (5.15) we have for all $y \in (0, K^{-1})$

$$\begin{aligned}
&n\varepsilon_n^2 \mathbb{E}_{\mathcal{Q}_0}\left(f_n^2 \mathbf{1}_{A_{n,y}}\right) \\
=\ & n\varepsilon_n^2 C(\vartheta_n)^2 \int \exp\Big(-2\vartheta_n h(x)\Big)\, \mathbf{1}_{[G_{n,y},\infty)}\left(\vartheta_n h(x)\right)\, \mathrm{d}\mathcal{Q}_0(x) \\
\sim_{\text{asy}}\ & nK^{-2}\left[\omega(2\vartheta_n) - \int \psi_n\left(\vartheta_n h(x)\right)\, \mathrm{d}\mathcal{Q}_0(x)\right] \\
=\ & nK^{-2}\left[2^{-p}n^{-rp}K(1+o(1)) - n^{-rp}K\Gamma(p)^{-1}\int_0^{-\log(yK)} e^{-2x}x^{p-1}\, \mathrm{d}x\, (1+o(1))\right] \\
\sim_{\text{asy}}\ & K^{-1}\Gamma(p)^{-1}2^{-p}\left[\Gamma(p) - \int_0^{-\frac{1}{2}\log(yK)} \exp\left(-z\right)\, z^{p-1}\, \mathrm{d}z\right] \\
=\ & K^{-1}\Gamma(p)^{-1}2^{-p}\int_{-\frac{1}{2}\log(yK)}^{\infty} e^{-z}\, z^{p-1}\, \mathrm{d}z \ \text{ as } n \to \infty.
\end{aligned} \tag{5.17}$$

Letting $y \searrow 0$ and considering $n\varepsilon_n^2 = o(1)$ yield (**B3**) for $\sigma^2 = 0$. From (5.16) we obtain (5.14) for all $y > K^{-1}$. Analogously to (5.17) we have for every $y \in (0, K^{-1})$

$$\begin{aligned}
n\varepsilon_n\mu_n\left(A_{n,y}^c\right) &\sim_{\text{asy}} nK^{-1}\int \exp\left(-\vartheta_n h(x)\right)\mathbf{1}_{[0,G_{n,y})}\left(\vartheta_n h(x)\right)\, \mathrm{d}\mathcal{Q}_0(x) \\
&\sim_{\text{asy}} nK^{-1}\, n^{-rp}K\Gamma(p)^{-1}\int \exp\left(-x\right)\, \mathbf{1}_{[0,-\log(yK))}(x)\, x^{p-1}\, \mathrm{d}x \\
&\sim_{\text{asy}} \Gamma(p)^{-1}\int_0^{-\log(yK)} \exp\left(-x\right)\, x^{p-1}\, \mathrm{d}x \text{ as } n \to \infty.
\end{aligned}$$

Now we verify (iv). Note that $\beta = 1$. By Lemma 2.15 we have for all $y > 0$

$$\begin{aligned}
\mu_n\Big(\varepsilon_n f_n \leq y\Big) &= \int f_n \mathbf{1}_{\{\varepsilon_n f_n \leq y\}}\, \mathrm{d}\mathcal{Q}_0 \\
&\leq \varepsilon_n^{-\frac{1}{2}} y^{\frac{1}{2}} \int \sqrt{C(\vartheta_n)}\, \exp\left(-\frac{1}{2}\vartheta_n h\right)\, \mathrm{d}\mathcal{Q}_0 \\
&= \varepsilon_n^{-\frac{1}{2}} y^{\frac{1}{2}} \sqrt{C(\vartheta_n)}\, \omega\left(\frac{1}{2}\vartheta_n\right) \\
&\sim_{\text{asy}} y^{\frac{1}{2}}\, n^{\frac{1}{2}+\frac{1}{2}rp-rp+o(1)} \ = \ o(1) \text{ as } n \to \infty.
\end{aligned}$$

Thus, for every $y > 0$

$$n\varepsilon_n\mu_n\Big(\varepsilon_n f_n > y\Big) \sim_{\text{asy}} n\varepsilon_n = 1 \text{ as } n \to \infty.$$

Moreover, for every $y > 0$

$$n\varepsilon_n^2 \int f_n^2 \mathbf{1}_{\{\varepsilon_n f_n \leq y\}} \,\mathrm{d}\mathcal{Q}_0 \;\leq\; y \int f_n \mathbf{1}_{\{\varepsilon_n f_n \leq y\}} \,\mathrm{d}\mathcal{Q}_0 \;=\; o(1) \text{ as } n \to \infty.$$

Note that by Lemma 2.15, Theorem 2.19(i) and (5.6)

$$\begin{aligned} &\mathcal{Q}_0[0,t] \;=\; t^{p+o(1)} \text{ as } t \searrow 0 \\ \text{and } &\frac{1}{\vartheta_n} \log\left(\frac{C(\vartheta_n)}{n}\right) \;=\; n^{-r+o(1)} \log\left(n^{rp-1+o(1)}\right) \;=\; n^{-r+o(1)} \text{ as } n \to \infty. \end{aligned}$$

Finally, we have for all $y > 0$

$$\begin{aligned} nP_n\Big(\varepsilon_n f_{n,i} \geq y\Big) \;&=\; n\mathcal{Q}_0\left(\frac{1}{n}C\left(\vartheta_n\right)\exp\left[-\vartheta_n h\right] \;\geq\; y\right) \\ &=\; n\mathcal{Q}_0^h\left[0\,,\, -\frac{1}{\vartheta_n}\log\left(\frac{n}{C(\vartheta_n)}\right)\right] \\ &=\; n^{1-rp+o(1)} \;\to\; 0 \text{ as } n \to \infty. \end{aligned}$$

■

In the following we have a look at the specific models introduced in Examples 2.8 to 2.10, which fulfil Assumption 2.16 obviously. Before we present the result, we want to make the following remark to show that there is a strong connection between these examples.

Remark 5.9. Let $\{P_{n,0,G}, \mu_{n,\vartheta_n,G}\}$, $\{P_{n,0,F}, \mu_{n,\vartheta_n,F}\}$ and $\{P_{n,0,E}, \mu_{n,\vartheta_n,E}\}$ be the binary experiments belonging to Examples 2.8 to 2.10. Let $\alpha > 0$ be the parameter from Example 2.9. Moreover, define $T_1 : \mathbb{R} \to (0,\infty)$ and $T_2 : (0,\infty) \to (0,\infty)$ by

$$T_1(x) := \exp(-x) \text{ for all } x \in \mathbb{R} \quad \text{ and } \quad T_2(y) := y^{-\alpha} \text{ for all } y \in (0,\infty).$$

Using the transformation formula for densities with the transformations T_1 and T_2 yields

$$\{P_{n,0,G}^{T_1}\,,\, \mu_{n,\log(\vartheta_n+1),G}^{T_1}\} \;=\; \{P_{n,0,E}\,,\, \mu_{n,\vartheta_n,E}\} \text{ and } \{P_{n,0,F}^{T_2}\,,\, \mu_{n,\vartheta_n,F}^{T_2}\} = \{P_{n,0,E}\,,\, \mu_{n,\vartheta_n^{\alpha}-1,E}\},$$

By Remark A.17(ii) and (iii) it is sufficient for later discussion to have a look at one of these models in detail, because the results can be transferred to the other ones. □

Corollary 5.10. *Let $r > 0$ and one of the following models* (a)-(c) *be given:*

(a) (Gumbel) the model introduced in Example 2.8 with $\vartheta_n \sim_{asy} r\log(n)$ as $n \to \infty$.

(b) (Fréchet) the model introduced in Example 2.9 with $\vartheta_n \sim_{asy} n^{\frac{r}{\alpha}}$ as $n \to \infty$.

(c) (Exponential) the model introduced in Example 2.10 with $\vartheta_n \sim_{asy} n^r$ as $n \to \infty$.

Let $\varepsilon_n := n^{-\beta}$ *for all* $n \in \mathbb{N}$ *and some* $\beta > \frac{1}{2}$*. Define*

$$\beta^{\#}_{GFE}(r) := \frac{(r \wedge 1) + 1}{2}.$$

(i) If $\beta > \beta^{\#}_{GFE}(r)$ *then* $\{P_{(n)}, Q_{(n)}\}$ *converges weakly to* $\{\epsilon_0, \epsilon_0\}$*.*

(ii) If $\beta < 1$ *and* $\beta = \beta^{\#}_{GFE}(r)$ *then* ***(A normal)*** *is fulfilled for* $\sigma^2 := \frac{1}{2}$*.*

(iii) If $\beta = 1$ *and* $r = 1$ *then* ***(A)*** *holds for real-valued random variables* ξ_1 *and* ξ_2*. Moreover, the Lévy characteristic of* ξ_j*,* $j \in \{1, 2\}$*, is equal to* $(\gamma_j, 0, \eta_j)$*, where the Lévy measure* η_j *is uniquely determined by its* $\lambda\!\!\lambda$*-density*

$$\frac{\mathrm{d}\eta_j}{\mathrm{d}\lambda\!\!\lambda}(x) \;=\; \frac{e^{jx}}{e^x - 1}\mathbf{1}_{(0,\log(2))}(x),\; x \in \mathbb{R} \setminus \{0\},$$

and γ_j *is given by* (4.16) *with* $\sigma^2 = 0$*.*

(iv) If $\beta < \beta^{\#}_{GFE}(r)$ *then* $\{P_{(n)}, Q_{(n)}\}$ *converges weakly to* $\{\epsilon_{-\infty}, \epsilon_\infty\}$*.*

Proof. By Remark 5.9 we can assume without loss of generality that the Exponential model is given. Note that in this case the assumptions of Corollary 5.7 are fulfilled with $L \equiv 1$ and $p = 1$. ■

5.2.2. Dense case

Here, we focus on the dense case, i.e., $\vartheta_n \to 0$ as $n \to \infty$. To be more specific, we suppose Assumption 2.23. The main result of this section is that the asymptotic behaviour of

$$B_n := \sum_{i=1}^{n} \varepsilon_{n,i}^2 \vartheta_n^2\,,\;\; n \in \mathbb{N}, \tag{5.18}$$

characterises the asymptotic behaviour of $\{P_{(n)}, Q_{(n)}\}$ uniquely and independently of the special shape of h. We obtain a dichotomy: every accumulation point of $\{P_{(n)}, Q_{(n)}\}$ is either equivalent to the full informative experiment or a Gaussian experiment. Note that the uninformative experiment is also a Gaussian experiment, see Remark 4.26(ii).

The section is structured as follows. At the beginning we present the main results for the general model, where the signal probability $\varepsilon_{n,i}$ can depend on i. After that we determine the detection boundary for the case that $\varepsilon_{n,i} = \varepsilon_n$ for all $1 \le i \le n$. Note that the corresponding phase diagram is visualised in Figure 5.2. The remaining part of the section consists of the proof of the main result including two technical lemmas.

Theorem 5.11. *Suppose Assumption 2.23 and* $Var_{\mathcal{Q}_0}(h) \in (0, \infty)$.

(i) We have

$$B_{k_n} := \sum_{i=1}^{k_n} \varepsilon_{k_n,i}^2 \vartheta_{k_n}^2 \ \to \ K \in [0, \infty) \text{ as } n \to \infty \tag{5.19}$$

for some subsequence $\{k_n : n \in \mathbb{N}\}$ *of* $\mathbb{N}$ *if and only if (**A normal**) holds for*

$$\sigma^2 := K \, Var_{\mathcal{Q}_0}(h) \,. \tag{5.20}$$

Furthermore, (5.19) *holds for* $K = 0$ *and a subsequence* $\{k_n : n \in \mathbb{N}\}$ *of* $\mathbb{N}$ *if and only if* $\{P_{(k_n)}, Q_{(k_n)}\}$ *converges weakly to* $\{\epsilon_0, \epsilon_0\}$.

(ii) For some subsequence $\{k_n : n \in \mathbb{N}\}$ *of* $\mathbb{N}$ *we have*

$$B_{k_n} := \sum_{i=1}^{k_n} \varepsilon_{k_n,i}^2 \vartheta_{k_n}^2 \ \to \ \infty \text{ as } n \to \infty \tag{5.21}$$

if and only if $\{P_{(k_n)}, Q_{(k_n)}\}$ *converges weakly to* $\{\epsilon_{-\infty}, \epsilon_\infty\}$.

Remark 5.12. (i) The assumption $\mathrm{Var}_{\mathcal{Q}_0}(h) > 0$ is not an actual restriction because otherwise $h \equiv c \in \mathbb{R}$ ($\mathcal{Q}_0$-a.s.) and thus $\mathcal{Q}_0 = \mathcal{Q}_\vartheta$ for all $\vartheta \in \Theta$.

(ii) We can conclude as an immediate consequence of Theorem 5.11 that every non-trivial accumulation point of $\{P_{(n)}, Q_{(n)}\}$ is already Gaussian. This can easily be shown, e.g., by a proof by contradiction. □

Corollary 5.13. *Suppose Assumption 2.23 and* $Var_{\mathcal{Q}_0}(h) > 0$. *Moreover, let*

$$\varepsilon_{n,i} := \varepsilon_n \sim_{asy} n^{-\beta} \quad \text{and} \quad \vartheta_n \sim_{asy} K n^{-r} \quad \text{as } n \to \infty$$

for some $K, r > 0$ *and some* $\beta \in (0, \frac{1}{2})$. *Define the detection boundary for* r *by*

$$\rho^*_{Exp,d}(\beta) = \frac{1}{2} - \beta. \tag{5.22}$$

(i) If $r < \rho^*_{Exp,d}(\beta)$ *then* $\{P_{(n)}, Q_{(n)}\}$ *converges weakly to* $\{\epsilon_0, \epsilon_0\}$.

(ii) If $r = \rho^*_{Exp,d}(\beta)$ *then (**A normal**) is fulfilled for* σ^2 *given by* (5.20).

(iii) If $r > \rho^*_{Exp,d}(\beta)$ *then* $\{P_{(n)}, Q_{(n)}\}$ *converges weakly to* $\{\epsilon_{-\infty}, \epsilon_\infty\}$.

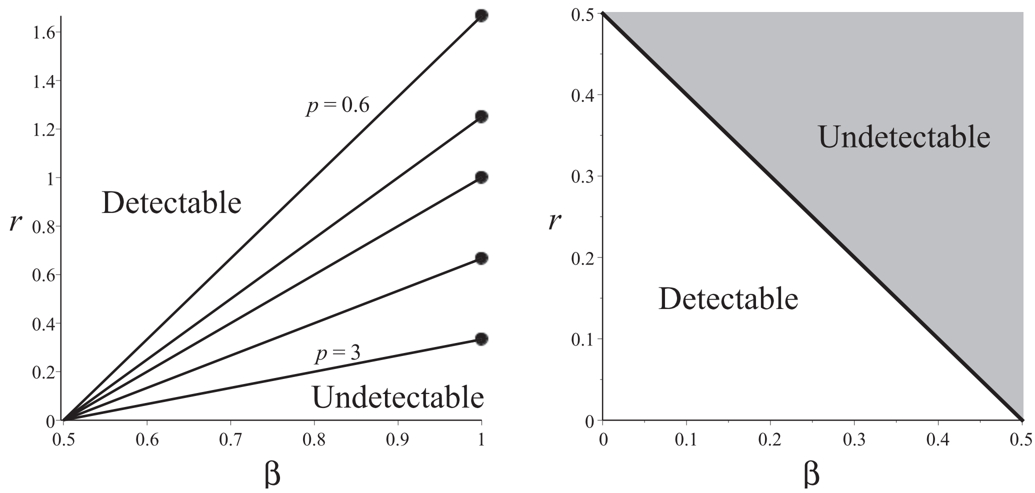

Figure 5.2.: Detection boundaries for the sparse (left) and the dense (right) exponential family mixture model. Left: $\beta \mapsto \rho^*_{Exp,s}(\beta, p)$, see (5.12), is plotted for $p \in \{0.6, 0.8, 1, 1.5, 3\}$. If $r > \rho^*_{Exp,s}(\beta, p)$ and $\beta < 1$ the limit experiment of $\{P_{(n)}, Q_{(n)}\}$ is $\{\epsilon_{-\infty}, \epsilon_\infty\}$ (Detectable). If $r < \rho^*_{Exp,s}(\beta, p)$ it is $\{\epsilon_0, \epsilon_0\}$ (Undetectable). If $\beta < 1$ and $r = \rho^*_{Exp,s}(\beta, p)$ it is Gaussian (solid line). If $\beta = 1$, $r = \rho^*_{Exp,s}(\beta, p)$ and $L \equiv K \in \mathbb{R}$ it is non-Gaussian (solid circle). Right: $\beta \mapsto \rho^*_{Exp,d}(\beta)$, see (5.22), is plotted. If $r > \rho^*_{Exp,d}(\beta)$ the limit experiment is $\{\epsilon_0, \epsilon_0\}$ (Undetectable). If $r < \rho^*_{Exp,d}(\beta)$ it is $\{\epsilon_{-\infty}, \epsilon_\infty\}$ (Detectable). If $r = \rho^*_{Exp,d}(\beta)$ it is Gaussian.

In regard to Corollary 4.40, Lemmas 4.42 and 4.45 the asymptotic behaviour of

$$\frac{C\left(\vartheta_{k_n}\right)^2}{C\left(2\vartheta_{k_n}\right)} - 1 \tag{5.23}$$

is of great interest to prove the previous results. We want to mention that in general (4.63) does not hold under Assumption 2.23 and, thus, Corollary 4.40 is not applicable. But we can fix this issue. In the following lemma we determine the convergence rate of (5.23).

Lemma 5.14. *Under the assumptions of Theorem 5.11*

$$\frac{C\left(\vartheta_{k_n}\right)^2}{C\left(2\vartheta_{k_n}\right)} - 1 \sim_{asy} \vartheta_{k_n}^2 \, Var_{\mathcal{Q}_0}(h) \quad \text{as } n \to \infty.$$

Proof. By Taylor's formula and Remark 2.25

$$\begin{aligned} \omega(2t) = \;& \omega(0) + (2t - 0)\,\omega^{(1)}(0) + \frac{(2t-0)^2}{2}\,\omega^{(2)}(0) + o\left(t^2\right) \\ = \;& 1 - 2t\,\mathbb{E}_{\mathcal{Q}_0}(h) + 2t^2\,\mathbb{E}_{\mathcal{Q}_0}\left(h^2\right) + o\left(t^2\right) \quad \text{as } t \to 0. \end{aligned}$$

Moreover,

$$\begin{aligned}
\omega(t)^2 = &\ \omega(0)^2 + t\,2\,\omega(0)\,\omega^{(1)}(0) + \frac{t^2}{2}\,2\left[\omega^{(1)}(0)^2 + \omega(0)\,\omega^{(2)}(0)\right] + o\left(t^2\right)\\
= &\ 1 - 2t\,\mathbb{E}_{\mathcal{Q}_0}(h) + t^2\left[\mathbb{E}_{\mathcal{Q}_0}(h)^2 + \mathbb{E}_{\mathcal{Q}_0}\left(h^2\right)\right] + o\left(t^2\right)\\
= &\ \omega(2t) - t^2\,\mathrm{Var}_{\mathcal{Q}_0}(h) + o\left(t^2\right) \quad \text{as } t \to 0. \qquad (5.24)
\end{aligned}$$

Note that ω is continuous at 0, see Lemma 2.24. Combining this with (5.24) shows

$$\begin{aligned}
\frac{C\left(\vartheta_{k_n}\right)^2}{C\left(2\vartheta_{k_n}\right)} - 1 = &\ \omega\left(\vartheta_{k_n}\right)^{-2}\left[\omega\left(2\vartheta_{k_n}\right) - \omega\left(\vartheta_{k_n}\right)^2\right]\\
= &\ \omega\left(\vartheta_{k_n}\right)^{-2}\left[\vartheta_{k_n}^2\,\mathrm{Var}_{\mathcal{Q}_0}(h) + o\left(\vartheta_{k_n}^2\right)\right]\\
\sim_{\text{asy}} &\ \vartheta_{k_n}^2\,\mathrm{Var}_{\mathcal{Q}_0}(h) \quad \text{as } n \to \infty. \qquad \blacksquare
\end{aligned}$$

We also need the following technical lemma in order to prove Theorem 5.11.

Lemma 5.15. *Under the assumptions of Theorem 5.11 we have for every $\lambda, y > 0$*

$$\max_{1\le i\le k_n} \mathcal{Q}_{2\vartheta_{k_n}}\left(\varepsilon_{k_n,i}\,C\left(\vartheta_{k_n}\right)\exp\left(-\vartheta_{k_n}h\right) > y\right) = o\left(\vartheta_{k_n}^{\lambda}\right) \quad \textit{as } n \to \infty. \qquad (5.25)$$

Proof. Obviously, it is sufficient to show (5.25) for all $\lambda > 0$ of the shape $\lambda = 2m - 1$, $m \in \mathbb{N}$. Let $y > 0$ and $m \in \mathbb{N}$ be fixed. We deduce from the continuity of the Laplace transform ω at 0, see Lemma 2.24, and $\varepsilon_{k_n:k_n} = o(1)$, see (2.1), that

$$\begin{aligned}
\varepsilon_{k_n:k_n}\,C\left(\vartheta_{k_n}\right) = &\ \frac{\varepsilon_{k_n:k_n}}{\omega\left(\vartheta_{k_n}\right)} = o(1) \quad \text{as } n \to \infty\\
\text{and } \varepsilon_{k_n:k_n}\,C\left(\vartheta_{k_n}\right) \le &\ e^{-2}\,y \text{ for all sufficiently large } n \in \mathbb{N}. \qquad (5.26)
\end{aligned}$$

Clearly, $\vartheta_{k_n} \in (-\varepsilon, \varepsilon)$ for all sufficiently large $n \in \mathbb{N}$. By Lemma 2.24 and (5.26)

$$\begin{aligned}
&\max_{1\le i\le k_n} \mathcal{Q}_{2\vartheta_{k_n}}\left(\varepsilon_{k_n,i}\,C\left(\vartheta_{k_n}\right)\exp\left(-\vartheta_{k_n}h\right) > y\right)\\
\le &\ C\left(2\vartheta_{k_n}\right)\int_{\{\exp(-\vartheta_{k_n}h-2)>1\}} \exp\left(-2\vartheta_{k_n}h\right)\,\mathrm{d}\mathcal{Q}_0\\
\le &\ \omega\left(2\vartheta_{k_n}\right)^{-1}\int_{\{-\vartheta_{k_n}h>2\}} \left(-\frac{\vartheta_{k_n}h}{2}\right)^{\lambda+1} \exp\left(-2\vartheta_{k_n}h\right)\,\mathrm{d}\mathcal{Q}_0\\
= &\ \frac{\vartheta_{k_n}^{2m}}{\omega\left(2\vartheta_{k_n}\right)2^{2m}}\int_{\{-\vartheta_{k_n}h>2\}} h^{2m}\exp\left(-2\vartheta_{k_n}h\right)\,\mathrm{d}\mathcal{Q}_0\\
\le &\ \frac{\vartheta_{k_n}^{2m}\,\omega^{(2m)}\left(2\vartheta_{k_n}\right)}{\omega\left(2\vartheta_{k_n}\right)2^{2m}}\\
\sim_{\text{asy}} &\ \vartheta_{k_n}^{2m}\,2^{-2m}\,\omega^{(2m)}(0) = o\left(\vartheta_{k_n}^{\lambda}\right) \quad \text{as } n \to \infty. \qquad \blacksquare
\end{aligned}$$

Proof of Theorem 5.11. Regarding Corollaries 4.25 and 4.38 we first determine the asymptotic behaviour of the quantities from (**B2c normal**) and (**B3 normal**) (from (**D1**) and (**D3**), respectively). Let $\{k_n : n \in \mathbb{N}\}$ be some subsequence of $\mathbb{N}$. From the continuity of ω at 0, see Lemma 2.24, Lemmas 5.14 and 5.15 we obtain for all $y > 0$

$$\begin{aligned}
&\sum_{i=1}^{k_n} \varepsilon_{k_n,i}^2 \left[\int_{\{\varepsilon_{k_n,i} f_{k_n,i} \leq y\}} f_{k_n,i}^2 \, \mathrm{d}\mathcal{Q}_0 - 1 \right] \\
= &\quad \sum_{i=1}^{k_n} \varepsilon_{k_n,i}^2 \left[\int_{\{\varepsilon_{k_n,i} f_{k_n,i} \leq y\}} C\left(\vartheta_{k_n}\right)^2 \exp\left(-2\vartheta_{k_n} h\right) \, \mathrm{d}\mathcal{Q}_0 - 1 \right] \\
= &\quad \sum_{i=1}^{k_n} \varepsilon_{k_n,i}^2 \left[\frac{C\left(\vartheta_{k_n}\right)^2}{C\left(2\vartheta_{k_n}\right)} \mathcal{Q}_{2\vartheta_{k_n}} \Big(\varepsilon_{k_n,i} \, C\left(\vartheta_{k_n}\right) \exp\left(-\vartheta_{k_n} h\right) \leq y \Big) - 1 \right] \\
= &\quad \sum_{i=1}^{k_n} \varepsilon_{k_n,i}^2 \left[\frac{C\left(\vartheta_{k_n}\right)^2}{C\left(2\vartheta_{k_n}\right)} - 1 - \frac{\omega\left(2\vartheta_{k_n}\right)}{\omega\left(\vartheta_{k_n}\right)^2} \mathcal{Q}_{2\vartheta_{k_n}} \Big(\varepsilon_{k_n,i} \, C\left(\vartheta_{k_n}\right) \exp\left(-\vartheta_{k_n} h\right) > y \Big) \right] \\
= &\quad \left[\frac{C\left(\vartheta_{k_n}\right)^2}{C\left(2\vartheta_{k_n}\right)} - 1 + o(\vartheta_{k_n}^2) \right] \sum_{i=1}^{k_n} \varepsilon_{k_n,i}^2 \quad \sim_{\text{asy}} \quad \mathrm{Var}_{\mathcal{Q}_0}\left(h\right) B_{k_n} \quad \text{as } n \to \infty. \qquad (5.27)
\end{aligned}$$

From Lemmas 2.24 and 5.15 we can conclude that for all $y > 0$

$$\begin{aligned}
&\sum_{i=1}^{k_n} \int_{\{\varepsilon_{k_n,i} f_{k_n,i} > y\}} \varepsilon_{k_n,i} f_{k_n,i} \, \mathrm{d}\mathcal{Q}_0 \;\leq\; \sum_{i=1}^{k_n} \int_{\{\varepsilon_{k_n,i} f_{k_n,i} > y\}} \frac{\varepsilon_{k_n,i}^2 f_{k_n,i}^2}{y} \, \mathrm{d}\mathcal{Q}_0 \\
= &\quad y^{-1} \sum_{i=1}^{k_n} \varepsilon_{k_n,i}^2 \frac{C\left(\vartheta_{k_n}\right)^2}{C\left(2\vartheta_{k_n}\right)} \mathcal{Q}_{2\vartheta_{k_n}} \Big(\varepsilon_{k_n,i} \, C\left(\vartheta_{k_n}\right) \exp\left(-\vartheta_{k_n} h\right) > y \Big) \\
\leq &\quad y^{-1} \left[\sum_{i=1}^{k_n} \varepsilon_{k_n,i}^2 \right] \frac{\omega\left(2\vartheta_{k_n}\right)}{\omega\left(\vartheta_{k_n}\right)^2} \max_{1 \leq i \leq k_n} \Big\{ \mathcal{Q}_{2\vartheta_{k_n}} \Big(\varepsilon_{k_n,i} \, C\left(\vartheta_{k_n}\right) \exp\left(-\vartheta_{k_n} h\right) > y \Big) \Big\} \\
= &\quad y^{-1} \left[\sum_{i=1}^{k_n} \varepsilon_{k_n,i}^2 \right] o\left(\vartheta_{k_n}^2\right) \;=\; o\left(B_{k_n}\right) \quad \text{as } n \to \infty. \qquad (5.28)
\end{aligned}$$

The equivalence in (i) follows immediately from Corollary 4.25, (5.27) and (5.28).
Suppose that (5.21) holds for some subsequence $\{k_n : n \in \mathbb{N}\}$ of $\mathbb{N}$. Then from (5.27) we obtain (**D1**) of Corollary 4.38(iii). Hence, by Corollary 4.38 the binary experiment $\{P_{(k_n)}, Q_{(k_n)}\}$ converges weakly to $\{\epsilon_{-\infty}, \epsilon_{\infty}\}$.
Now, suppose that $\{P_{(k_n)}, Q_{(k_n)}\}$ converges weakly to $\{\epsilon_{-\infty}, \epsilon_{\infty}\}$ for some subsequence $\{k_n : n \in \mathbb{N}\}$ of $\mathbb{N}$. Contrary to (ii) assume (5.19) for a subsequence $\{k_{n,1} : n \in \mathbb{N}\}$ of $\{k_n : n \in \mathbb{N}\}$. By the equivalence in (i) $\{P_{(k_{n,1})}, Q_{(k_{n,1})}\}$ converges weakly to $\{\epsilon_0, \epsilon_0\}$ or a non-trivial Gaussian experiment. This contradicts our assumption that $\{P_{(k_n)}, Q_{(k_n)}\}$ converges weakly to $\{\epsilon_{-\infty}, \epsilon_{\infty}\}$ and so does $\{P_{(k_{n,1})}, Q_{(k_{n,1})}\}$. ∎

5.3. The h-Model

In this section we focus on the semi-parametric structure model introduced in Section 2.4, namely the h-model. We determine the detectable and the undetectable areas, which do not depend on the special shape of the function h. The behaviour of $\{P_{(n)}, Q_{(n)}\}$ on the detection boundary is comparable with the one in Section 5.2.1 for the sparse exponential family model. The limit experiment $\{P, Q\}$ of $\{P_{(n)}, Q_{(n)}\}$ is non-Gaussian at the endpoint of the detection boundary and Gaussian on the remaining part of the detection boundary. In the Gaussian case $\{P, Q\}$ only depends on c_2 and not on the special shape of h. But in the non-Gaussian case $\{P, Q\}$ does depend on the special shape of h.
We start by presenting the main results for the general case. After that we determine the detection boundary using this result for the more specific model that neither $\varepsilon_{n,i}$ nor $\tau_{n,i}$ depends on i. The results are illustrated by a phase diagram, see Figure 5.3. The remaining part of this section consists of a technical lemma and the proofs.

Theorem 5.16. *Suppose Assumption 2.26. Let $\{k_n : n \in \mathbb{N}\}$ be a subsequence of $\mathbb{N}$,*

$$\sum_{i=1}^{k_n} \frac{\varepsilon_{k_n,i}^2}{\tau_{k_n,i}} \;\to\; K \in [0,\infty] \quad \text{and} \quad \sum_{i=1}^{k_n} \varepsilon_{k_n,i}^2 \;\to\; 0 \quad \text{as } n \to \infty. \tag{5.29}$$

(i) If $K = 0$ then $\{P_{(k_n)}, Q_{(k_n)}\}$ converges weakly to $\{\epsilon_0, \epsilon_0\}$ as $n \to \infty$.

(ii) If $K \in (0,\infty)$ and

$$\lim_{n\to\infty} \max_{1\le i\le k_n} \frac{\varepsilon_{k_n,i}}{\tau_{k_n,i}} \;=\; 0 \tag{5.30}$$

*then (**A normal**) is fulfilled for $\sigma^2 := Kc_2$.*

(iii) If $K = \infty$ and

$$\limsup_{n\to\infty} \max_{1\le i\le k_n} \frac{\varepsilon_{k_n,i}}{\tau_{k_n,i}} \;<\; \infty$$

then $\{P_{(k_n)}, Q_{(k_n)}\}$ converges weakly to $\{\epsilon_{-\infty}, \epsilon_\infty\}$ as $n \to \infty$.

(iv) Assume for all $n \in \mathbb{N}$ that

$$\frac{\varepsilon_{k_n,1}}{\tau_{k_n,1}} \;\le\; \frac{\varepsilon_{k_n,2}}{\tau_{k_n,2}} \;\le\; \ldots \;\le\; \frac{\varepsilon_{k_n,k_n}}{\tau_{k_n,k_n}}. \tag{5.31}$$

Let $r_n \in \{1,\ldots,k_n\}$ for all $n \in \mathbb{N}$ such that

$$\sum_{i=r_n}^{k_n} \varepsilon_{k_n,i} \;\to \infty \quad \text{and} \quad \sum_{i=1}^{r_n} \frac{\varepsilon_{k_n,i}^2}{\tau_{k_n,i}} \;\to\; \infty \quad \text{as } n \to \infty. \tag{5.32}$$

Then $\{P_{(k_n)}, Q_{(k_n)}\}$ converges weakly to $\{\epsilon_{-\infty}, \epsilon_\infty\}$ as $n \to \infty$.

Remark 5.17. (i) There is no restriction of generality in assuming that (5.31) holds since the standard form of $\{P_{(k_n)}, Q_{(k_n)}\}$ and its asymptotic behaviour are not affected by reordering, see Remark A.17(iii).

(ii) If $\varepsilon_{k_n,i} = \varepsilon_{k_n}$ and $\tau_{k_n,i} = \tau_{k_n}$ do not depend on i then choosing $r_n = \frac{1}{2}k_n$ yields that (5.32) is equivalent to

$$k_n\varepsilon_{k_n} = \sum_{i=1}^{k_n} \varepsilon_{k_n,i} \to \infty \text{ and } k_n\frac{\varepsilon_{k_n}^2}{\tau_{k_n}} = \sum_{i=1}^{k_n} \frac{\varepsilon_{k_n,i}^2}{\tau_{k_n,i}} \to \infty \quad \text{as } n \to \infty.$$

Note that by Remark 4.5(i) "$\sum_{i=1}^{k_n} \varepsilon_{k_n,i} \to \infty$" is a necessary condition for weak convergence to $\{\epsilon_{-\infty}, \epsilon_\infty\}$. □

Corollary 5.18. *Suppose that Assumption 2.26 holds, where we assume additionally that*

$$\varepsilon_{n,i} = \varepsilon_n := n^{-\beta} \text{ and } \tau_{n,i} = \tau_n := n^{-r}$$

for some $\beta \in (\frac{1}{2}, 1]$, $r \in (0,1]$ and all $1 \le i \le n \in \mathbb{N}$. Define the detection boundary by

$$\beta_h^{\#}(r) := \frac{r+1}{2}.$$

(i) If $\beta > \beta_h^{\#}(r)$ then $\{P_{(n)}, Q_{(n)}\}$ converges weakly to $\{\epsilon_0, \epsilon_0\}$.

(ii) If $\beta < \beta_h^{\#}(r)$ then $\{P_{(n)}, Q_{(n)}\}$ converges weakly to $\{\epsilon_{-\infty}, \epsilon_\infty\}$.

*(iii) If $\beta = \beta_h^{\#}(r)$ and $r < 1$ then (**A normal**) is fulfilled for $\sigma^2 := c_2$.*

(iv) Define $T : (0,\infty) \to (0,\infty)$ by

$$T(x) := \exp(x) - 1 \text{ for all } x > 0.$$

*If $\beta = \beta_h^{\#}(1)$ and $r = 1$ then (**A**) holds for some random variables ξ_1 and ξ_2. Moreover, the Lévy characteristic of ξ_j, $j \in \{1,2\}$, is equal to $(\gamma_j, 0, \eta_j)$, where the Lévy measure η_1 is uniquely determined by $\eta_1(-\infty, 0) = 0$ and*

$$\int_{(x,\infty)} T \,\mathrm{d}\eta_1 = \int_{(x,\infty)} T \,\mathrm{d}P_0^{T^{-1}\circ h} + \mathbf{1}_{(0,\log 2)}(x) \int_{\{x\}} T \,\mathrm{d}P_0^{T^{-1}\circ h} \tag{5.33}$$

for all $x > 0$, η_2 is uniquely determined by its η_1-density $\frac{\mathrm{d}\eta_2}{\mathrm{d}\eta_1} = \exp$ and γ_j is given by (4.16) with $\sigma^2 = 0$.

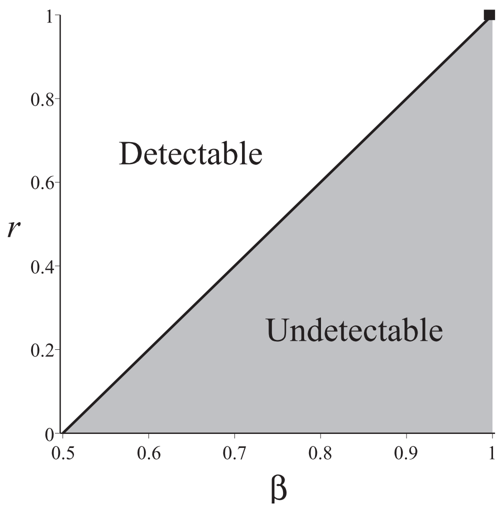

Figure 5.3.: The detection boundary $\beta \mapsto \rho_h^*(\beta)$, see Remark 5.19(ii), is plotted. If r exceeds $\rho_h^*(\beta)$ the limit experiment of $\{P_{(n)}, Q_{(n)}\}$ is $\{\epsilon_{-\infty}, \epsilon_\infty\}$ (Detectable). If $r < \rho_h^*(\beta)$ it is $\{\epsilon_0, \epsilon_0\}$ (Undetectable). If $r = \rho^*(\beta)$ and $\beta < 1$ it is Gaussian (solid line). If $r = \rho_h^*(\beta)$ and $\beta = 1$ it is non-Gaussian (solid box).

Remark 5.19. (i) Suppose that the assumptions of Corollary 5.18(iv) hold. Since $T > 0$ we deduce from (5.33) that

$$\eta_{1|(\log(2),\infty)} = \left(P_0^{T^{-1}\circ h}\right)_{|(\log(2),\infty)} .$$

Moreover, if

$$P_0^{T^{-1}\circ h}(\{x\}) = \lambda\!\!\lambda\Big(u \in (0,1) : h(u) = e^x - 1\Big) = 0$$

for all $x \in (0, \log 2)$ then

$$\eta_1 = P_0^{T^{-1}\circ h} .$$

(ii) Clearly, Corollary 5.18 can also be formulated by using the detection boundary ρ_h^*, which is plotted in Figure 5.3, for the parameter r given by

$$\rho_h^*(\beta) := 2\beta - 1 \text{ for all } \beta \in \left(\frac{1}{2}, 1\right] . \qquad \square$$

We prove Theorem 5.16 and Corollary 5.18 by applying Theorem 4.19, Corollary 4.25 and Corollary 4.38. For this purpose we need to determine the sums from (**B2c**) and (**B3**). In the following lemma we simplify these sums for the h-model.

Lemma 5.20. *Suppose Assumption 2.26. Define for all $\varepsilon \in (0,\infty)$ and $1 \le i \le n \in \mathbb{N}$*

$$\begin{aligned} A_{n,i,\varepsilon} &:= \{u \in (0,1) : \varepsilon_{n,i} f_{n,i}(u) \le \varepsilon\} \\ B_{n,i,\varepsilon} &:= \left\{ y \in (0,1) : \frac{\varepsilon_{n,i}(1-\tau_{n,i})}{\tau_{n,i}} h(y) \le \varepsilon - \varepsilon_{n,i} \right\}, \\ A^c_{n,i,\varepsilon} &:= (0,1) \setminus A_{n,i,\varepsilon} \quad \text{and} \quad B^c_{n,i,\varepsilon} := (0,1) \setminus B_{n,i,\varepsilon}. \end{aligned}$$

If $n \in \mathbb{N}$ is sufficiently large we have for all $\varepsilon > 0$ and every $1 \le i \le n$

$$\mu_{n,i}\left(A^c_{n,i,\varepsilon}\right) = \tau_{n,i} \lambda\!\!\lambda \left(B^c_{n,i,\varepsilon}\right) + (1-\tau_{n,i}) \int_{B^c_{n,i,\varepsilon}} h(y)\,\mathrm{d}y. \tag{5.34}$$

If

$$\lim_{n\to\infty} \sum_{i=1}^{k_n} \varepsilon^2_{k_n,i} = 0 \tag{5.35}$$

for some subsequence $\{k_n : n \in \mathbb{N}\}$ of $\mathbb{N}$ then for all $\varepsilon > 0$

$$\sum_{i=1}^{k_n} \varepsilon^2_{k_n,i} \int_0^1 \left(f^2_{k_n,i} \mathbf{1}_{A_{k_n,i,\varepsilon}} - 1 \right) \mathrm{d}\lambda\!\!\lambda = \left[\sum_{i=1}^{k_n} \frac{\varepsilon^2_{k_n,i}}{\tau_{k_n,i}} \int_{B_{k_n,i,\varepsilon}} h^2 \,\mathrm{d}\lambda\!\!\lambda \right] + o(1) \text{ as } n \to \infty. \tag{5.36}$$

Proof. Fix $\varepsilon > 0$. By (2.17)-(2.19) of Assumption 2.26

$$\begin{aligned} A^c_{n,i,\varepsilon} &= \left\{ u \in (0,\tau_{n,i}) : \varepsilon_{n,i} + \varepsilon_{n,i} \frac{1-\tau_{n,i}}{\tau_{n,i}} h\left(\frac{u}{\tau_{n,i}}\right) > \varepsilon \right\} \\ &\quad \cup \{ u \in [\tau_{n,i}, 1) : \varepsilon_{n,i}(1-c_1) > \varepsilon \} \\ &= \left\{ u \in (0,\tau_{n,i}) : \varepsilon_{n,i} \frac{1-\tau_{n,i}}{\tau_{n,i}} h\left(\frac{u}{\tau_{n,i}}\right) > \varepsilon - \varepsilon_{n,i} \right\}, \\ A_{n,i,\varepsilon} &= \left\{ u \in (0,\tau_{n,i}) : \varepsilon_{n,i} \frac{1-\tau_{n,i}}{\tau_{n,i}} h\left(\frac{u}{\tau_{n,i}}\right) \le \varepsilon - \varepsilon_{n,i} \right\} \cup [\tau_{n,i}, 1) \end{aligned} \tag{5.37}$$

for all $1 \le i \le n$ if n is sufficiently large. By this, (2.18), (2.19) and substituting $y = \tau^{-1}_{n,i} u$

$$\begin{aligned} \mu_{n,i}\left(A^c_{n,i,\varepsilon}\right) &= \tau_{n,i} \int_{B^c_{n,i,\varepsilon}} f_{n,i}(\tau_{n,i} y)\,\mathrm{d}y \\ &= \tau_{n,i} \lambda\!\!\lambda \left(B^c_{n,i,\varepsilon}\right) + (1-\tau_{n,i}) \int_{B^c_{n,i,\varepsilon}} h(y)\,\mathrm{d}y \end{aligned} \tag{5.38}$$

for all $1 \le i \le n$ if $n \in \mathbb{N}$ is sufficiently large.

By (5.37) and substituting $y = \tau_{n,i}^{-1} u$

$$
\begin{aligned}
&\int_0^1 f_{n,i}^2 \mathbf{1}_{A_{n,i,\varepsilon}} \,\mathrm{d}\lambda\!\!\lambda \\
=\;& \int \left((1-c_1)^2 \mathbf{1}_{A_{n,i,\varepsilon}\cap[\tau_{n,i},1)}(u) + \left[1 + \frac{1-\tau_{n,i}}{\tau_{n,i}} h\left(\frac{u}{\tau_{n,i}}\right)\right]^2 \mathbf{1}_{(0,\tau_{n,i})\cap A_{n,i,\varepsilon}}(u) \right) \mathrm{d}u \\
=\;& (1-c_1)^2(1-\tau_{n,i}) + \tau_{n,i}\lambda\!\!\lambda\,(B_{n,i,\varepsilon}) + 2(1-\tau_{n,i}) \int_{B_{n,i,\varepsilon}} h(y)\,\mathrm{d}y + \frac{(1-\tau_{n,i})^2}{\tau_{n,i}} \int_{B_{n,i,\varepsilon}} h(y)^2\,\mathrm{d}y \\
=\;& c_1^2 - c_1^2\tau_{n,i} + 1 - \tau_{n,i}\lambda\!\!\lambda \left(B_{n,i,\varepsilon}^c\right) - 2(1-\tau_{n,i}) \int_{B_{n,i,\varepsilon}^c} h\,\mathrm{d}\lambda\!\!\lambda + \frac{(1-\tau_{n,i})^2}{\tau_{n,i}} \int_{B_{n,i,\varepsilon}} h^2\,\mathrm{d}\lambda\!\!\lambda
\end{aligned}
$$

for all $1 \le i \le n$ if $n \in \mathbb{N}$ is sufficiently large. Combining this and (5.35) yields (5.36). ■

Proof of Theorem 5.16. First, we prove (i) and (ii). By Corollary 4.25 and Remark 4.26(iii) it is sufficient to show that for every $\varepsilon > 0$ (**B2c normal**) and (**B3 normal**) are fulfilled for $y = y_0 = \varepsilon$ and $\sigma^2 = 0$ or $\sigma^2 = c_2 K$, respectively. For this purpose we apply Lemma 5.20. Let $\varepsilon > 0$ be fixed. By (2.16) and (2.17) of Assumption 2.26

$$
\begin{aligned}
\sum_{i=1}^{k_n} \varepsilon_{k_n,i}\tau_{k_n,i}\lambda\!\!\lambda\left(B_{k_n,i,\varepsilon}^c\right) &\le \sum_{i=1}^{k_n} \varepsilon_{k_n}\tau_{k_n,i}\lambda\!\!\lambda\left(x \in (0,1) : \frac{2\varepsilon_{k_n,i}}{\varepsilon\,\tau_{k_n,i}} h(x) > 1\right) \\
&\le \frac{4}{\varepsilon^2} \sum_{i=1}^{k_n} \frac{\varepsilon_{k_n,i}^3}{\tau_{k_n,i}} \int_0^1 h^2\,\mathrm{d}\lambda\!\!\lambda \\
&\le \frac{4\,c_2\,\varepsilon_{k_n:k_n}}{\varepsilon^2} \sum_{i=1}^{k_n} \frac{\varepsilon_{k_n,i}^2}{\tau_{k_n,i}} \qquad (5.39)
\end{aligned}
$$

for all sufficiently large $n \in \mathbb{N}$. Similarly, for all sufficiently large $n \in \mathbb{N}$

$$
\begin{aligned}
&\sum_{i=1}^{k_n} \varepsilon_{k_n,i}(1-\tau_{k_n,i}) \int h\,\mathbf{1}_{B_{k_n,i,\varepsilon}^c}\,\mathrm{d}\lambda\!\!\lambda \\
\le\;& \sum_{i=1}^{k_n} \varepsilon_{k_n,i} \int h\,\mathbf{1}_{\left\{x\in(0,1):\frac{\varepsilon_{k_n,i}}{\tau_{k_n,i}}h(x)>\frac{\varepsilon}{2}\right\}}\,\mathrm{d}\lambda\!\!\lambda \\
\le\;& \frac{2}{\varepsilon}\left[\sum_{i=1}^{k_n} \frac{\varepsilon_{k_n,i}^2}{\tau_{k_n,i}}\right] \int h^2\,\mathbf{1}_{\left\{x\in(0,1):\max\limits_{1\le i\le k_n}\left\{\frac{\varepsilon_{k_n,i}}{\tau_{k_n,i}}\right\}h(x)>\frac{\varepsilon}{2}\right\}}\,\mathrm{d}\lambda\!\!\lambda. \qquad (5.40)
\end{aligned}
$$

Obviously, from (5.34) of Lemma 5.20 we obtain (**B2c normal**) under (i). Moreover, applying additionally Lebesgue's dominated convergence theorem to (5.40) shows that

(**B2c normal**) holds under (ii). By Lebesgue's Theorem we have under (5.30)

$$\sum_{i=1}^{k_n} \frac{\varepsilon_{k_n,i}^2}{\tau_{k_n,i}} \int_{B_{k_n,i,\varepsilon}} h^2 \, \mathrm{d}\lambda\!\!\lambda \quad \geq \quad \left[\sum_{i=1}^{k_n} \frac{\varepsilon_{k_n,i}^2}{\tau_{k_n,i}}\right] \int h^2 \, \mathbf{1}_{\left\{x\in(0,1): \max\limits_{1\le i\le k_n} \left\{\frac{\varepsilon_{k_n,i}}{\tau_{k_n,i}}\right\} h(x) \le \frac{\varepsilon}{2}\right\}} \, \mathrm{d}\lambda\!\!\lambda$$
$$\sim_{\text{asy}} \sum_{i=1}^{k_n} \frac{\varepsilon_{k_n,i}^2}{\tau_{k_n,i}} c_2 \tag{5.41}$$

as $n \to \infty$. Moreover, note that for all $n \in \mathbb{N}$

$$\sum_{i=1}^{k_n} \frac{\varepsilon_{k_n,i}^2}{\tau_{k_n,i}} \int_{B_{k_n,i,\varepsilon}} h^2 \, \mathrm{d}\lambda\!\!\lambda \quad \leq \quad \sum_{i=1}^{k_n} \frac{\varepsilon_{k_n,i}^2}{\tau_{k_n,i}} c_2. \tag{5.42}$$

Combining (5.41), (5.42) and (5.36) of Lemma 5.20 completes the proof of (i) and (ii). Now we verify (iv). Suppose that the assumptions of (iv) hold. We split the proof into three cases.

First case: Suppose that

$$\frac{\varepsilon_{k_n,r_n}}{\tau_{k_n,r_n}} \to C \in [0,\infty) \quad \text{as } n \to \infty. \tag{5.43}$$

By (5.31), (5.36) of Lemma 5.20 and (5.43)

$$\begin{aligned}
&\sum_{i=1}^{k_n} \varepsilon_{k_n,i}^2 \int_0^1 \left(f_{k_n,i}^2 \mathbf{1}_{A_{k_n,i,\varepsilon}} - 1 \right) \mathrm{d}\lambda\!\!\lambda \\
\geq\; & \sum_{i=1}^{r_n} \frac{\varepsilon_{k_n,i}^2}{\tau_{k_n,i}} \int h^2 \mathbf{1}_{\left\{y\in(0,1): \frac{\varepsilon_{k_n,i}(1-\tau_{k_n,i})}{\tau_{k_n,i}} h(y) \le \varepsilon - \varepsilon_{k_n,i}\right\}} \, \mathrm{d}\lambda\!\!\lambda + o(1) \\
\geq\; & \left[\sum_{i=1}^{r_n} \frac{\varepsilon_{k_n,i}^2}{\tau_{k_n,i}}\right] \int h^2 \mathbf{1}_{\left\{y\in(0,1): (C+1)h(y) \le \frac{1}{2}\varepsilon\right\}} \, \mathrm{d}\lambda\!\!\lambda + o(1) \text{ as } n \to \infty
\end{aligned} \tag{5.44}$$

for all $\varepsilon > 0$. Note that by Lebesgue's theorem

$$\lim_{n\to\infty} \int h^2 \, \mathbf{1}_{\{y\in(0,1): h(y)\le n\}} \, \mathrm{d}\lambda\!\!\lambda = \int h^2 \, \mathrm{d}\lambda\!\!\lambda = c_2 > 0.$$

Hence, for sufficiently large $\varepsilon > 0$

$$\int h^2 \, \mathbf{1}_{\left\{y\in(0,1): (C+1)h(y) \le \frac{1}{2}\varepsilon\right\}} \, \mathrm{d}\lambda\!\!\lambda \;\geq\; \frac{1}{2} c_2 \;>\; 0.$$

We deduce (**D1**) from (5.32) and (5.44). Applying Corollary 4.38 yields (iv).

<u>*Second case:*</u> Suppose that

$$\frac{\varepsilon_{k_n,r_n}}{\tau_{k_n,r_n}} \to \infty \quad \text{as } n \to \infty. \tag{5.45}$$

Set $\varepsilon := \frac{1}{2}$. By Lemma 5.20, (5.31), (5.32) and Lebesgue's theorem

$$\begin{aligned}\sum_{i=1}^{k_n} \varepsilon_{k_n,i}\mu_{k_n,i}\left(A^c_{k_n,i,\varepsilon}\right) &\geq \left[\sum_{i=r_n}^{k_n} \varepsilon_{k_n,i}\right] \frac{1}{2} \int_{\left\{y\in(0,1):\frac{\varepsilon_{k_n,r_n}}{\tau_{k_n,r_n}}h(y)>1\right\}} h(x)\,\mathrm{d}x \\ &\sim_{\text{asy}} \sum_{i=r_n}^{k_n} \varepsilon_{k_n,i}\frac{c_1}{2} \longrightarrow \infty \quad \text{as } n \to \infty.\end{aligned}$$

Hence, (**D2**) holds. Applying Corollary 4.38 yields the statement of (iv).

<u>*Third case:*</u> If neither (5.43) nor (5.45) holds then we use a typical subsequence argument. Let $\{k_{n,1} : n \in \mathbb{N}\}$ be a subsequence of $\{k_n : n \in \mathbb{N}\}$. Then there exists a further subsequence $\{k_{n,2} : n \in \mathbb{N}\}$ of $\{k_{n,1} : n \in \mathbb{N}\}$ such that either (5.43) or (5.45) holds for it. Thus, by the first and second case $\{P_{(k_{n,2})}, Q_{(k_{n,2})}\}$ converges weakly to the full informative experiment $\{\epsilon_{-\infty}, \epsilon_\infty\}$. Because the subsequence $\{k_{n,1} : n \in \mathbb{N}\}$ was chosen arbitrarily we can conclude that all accumulation points of $\{P_{(k_n)}, Q_{(k_n)}\}$ are equal to $\{\epsilon_{-\infty}, \epsilon_\infty\}$.

By following the argumentation of the first case discussed above (iii) can be proven, where r_n is replaced by k_n for all $n \in \mathbb{N}$. The details are left to the reader. ■

Remark 5.21. Suppose (5.29) for some $K \in (0, \infty)$. It can be concluded from (5.34), (5.36), (5.39)-(5.42) and subsequence arguments that for every subsequence $\{k_{n,1} : n \in \mathbb{N}\}$ of $\{k_n : n \in \mathbb{N}\}$ there exists a further subsequence $\{k_{n,2} : n \in \mathbb{N}\}$ of $\{k_{n,1} : n \in \mathbb{N}\}$ such that (**B'**) holds for it. From Theorem 4.19 and Corollary 4.12(i) we deduce that (**A**) holds for $\{k_{n,2} : n \in \mathbb{N}\}$ and $Q_{(k_{n,2})} \lhd\rhd P_{(k_{n,2})}$. Hence, it is easy to see that (**A**) holds for all accumulation points of $\{P_{(k_n)}, Q_{(k_n)}\}$, and that $Q_{(k_n)} \lhd\rhd P_{(k_n)}$. The detailed verification of the above-mentioned argumentation and statements are left to the reader. □

Proof of Corollary 5.18. Note that

$$\sum_{i=1}^{n} \frac{\varepsilon^2_{n,i}}{\tau_{n,i}} = n^{1-2\beta+r} \text{ and } \sum_{i=1}^{n} \varepsilon_{n,i} = n^{1-\beta} \text{ for all } n \in \mathbb{N}.$$

Regarding Remark 5.17(ii) we deduce (i)-(iii) from Theorem 5.16.

Now suppose that the assumptions of (iv) are fulfilled. By Theorem 4.19 it is sufficient to verify (**B3**) and (**B2c**) for $\sigma^2 = 0$, η_1, η_2 and all $x > 0$.

First, we show that (**B2c**) is fulfilled. Fix $x > 0$. By Lemma 5.20 it is sufficient to determine the two summands from the right side of (5.34) for $\varepsilon := T(x) = e^x - 1$ multiplied by $n\varepsilon_n$. Because $B_{n,1,\varepsilon} \subseteq (0,1)$ we have

$$n\varepsilon_n \tau_n \lambda\!\!\lambda \left(B^c_{n,1,\varepsilon}\right) \;\le\; n\frac{1}{n^2}\cdot 1 \;=\; o(1) \quad \text{as } n \to \infty.$$

By (5.33), Lebesgue's theorem and the transformation formula for image measures

$$\begin{aligned}
& n\varepsilon_n(1-\tau_n)\int_{B^c_{n,1,\varepsilon}} h(y)\,\mathrm{d}y \\
= \;& \frac{n-1}{n}\int h\,\mathbf{1}_{\left\{y\in(0,1):\left[1-\frac{1}{n}\right]h(y)>T(x)-\frac{1}{n}\right\}}\,\mathrm{d}P_0 \\
= \;& \frac{n-1}{n}\int h\,\mathbf{1}_{\left\{y\in(0,1):h(y)>T(x)+\frac{e^x-2}{n-1}\right\}}\,\mathrm{d}P_0 \\
\sim_{\text{asy}} & \int h\,\left[\mathbf{1}_{\{y\in(0,1):h(y)>T(x)\}} + \mathbf{1}_{(-1,0)}\,(e^x-2)\,\mathbf{1}_{\{y\in(0,1):h(y)=T(x)\}}\right]\,\mathrm{d}P_0 \\
= \;& \int T\circ T^{-1}\circ h\,\left[\mathbf{1}_{(x,\infty)}\left(T^{-1}\circ h\right) + \mathbf{1}_{(0,\log 2)}(x)\,\mathbf{1}_{\{x\}}\left(T^{-1}\circ h\right)\right]\,\mathrm{d}P_0 \\
= \;& \int_{(x,\infty)} T\,\mathrm{d}\eta_1 \;=\; \int_{(x,\infty)}\left(\frac{\mathrm{d}\eta_2}{\mathrm{d}\eta_1}-1\right)\mathrm{d}\eta_1 \;=(\eta_2-\eta_1)(x,\infty) \quad \text{as } n\to\infty.
\end{aligned}$$

Consequently, (**B2c**) follows. Moreover, for every $\varepsilon > 0$ and all $n \in \mathbb{N}$

$$\begin{aligned}
0 \;\le\; n\frac{\varepsilon_n^2}{\tau_n}\int_{B_{n,1,\varepsilon}} h^2\,\mathrm{d}\lambda\!\!\lambda \;&=\; \int_{\{y\in(0,1):(1-n^{-1})h(y)\le\varepsilon-n^{-1}\}} h^2\,\mathrm{d}\lambda\!\!\lambda \\
&\le \left(\varepsilon-\frac{1}{n}\right)^2\frac{n^2}{(n-1)^2}\,.
\end{aligned}$$

Combining this and (5.36) of Lemma 5.20 yields (**B3**) with $\sigma^2 = 0$. ■

5.4. Extensions of the results of Cai and Wu

Cai and Wu [12] suggested how to determine the detection boundary for each member of a general class of distributions including the heterogeneous and heteroscedastic normal mixtures. For their proofs they used the Hellinger distance, see Definition and Lemma A.12(iii), and the simplification of it for product measures, see Lemma A.15. In contrast to [10, 20], Cai and Wu [12] determined thresholds for the parameter β belonging to ε_n and not for the parameter r of the signal strength μ_n, see (5.3) for a possible parametrisation. Obviously, both approaches are equivalent and can be transferred to each other, see Remark 5.8(ii) and Remark 5.19(ii). Note that Cai and Wu [12] only

presented results regarding the sparse case. They explained shortly why the dense case is more difficult in some sense. It is still an open question how or whether the results and ideas in [12] can be extended to the dense case.
We noticed that the model given in Corollary 5.7 does not fulfil the assumptions in [12]. But the results of Corollary 5.7 are exactly what we could deduce from Theorem 3 in [12] if the assumptions of this theorem would be fulfilled for the model. In this section we present a slight extension of Theorem 3 in [12] that can be applied to the model given in Corollary 5.7. After this general result we present a simplification of it in the context of detection boundaries. Afterwards we introduce Theorem 3 of [12] to explain the benefits of our extension. At the end of this section we present an extension of Theorem 1 in [12] for the case that the null distribution $P_{n,i}$ is a standard normal distribution for all $1 \leq i \leq n$. Theorem 5.1(i) and (ii), the results concerning heteroscedastic normal mixtures, could be concluded from their Theorem 1, see [12] for details, and so from our extension of it.
The proof of our extension is presented in the following subsection. It is inspired by the ones in [12]. But instead of the Hellinger distance as criterion for the trivial limit experiments, we use our results from Section 4.3. Doing this we want to show that our criterion can compete.

First, we present the assumptions for the model.

Assumption 5.22. *(i) Suppose that Assumption 2.1*(i) *hold with continuous measures* $P_{n,i}$ *and* $\mu_{n,i}$ *for all* $1 \leq i \leq n \in \mathbb{N}$. *Moreover, let* $T_{n,i} : (\Omega, \mathcal{A}) \to ([0,1], \mathcal{B}[0,1])$ *and* $T_{n,i}^{-1} : ((0,1), \mathcal{B}(0,1)) \to (\Omega, \mathcal{A})$ *be measurable mappings such that*

$$P_{n,i}^{T_{n,i}} = \mathcal{U}(0,1) \tag{5.46}$$

$$\text{and} \quad P_{n,i}\Big(\omega \in \Omega : \ T_{n,i}(\omega) \in (0,1),\, T_{n,i}^{-1}\left(T_{n,i}(\omega)\right) = \omega\Big) = 1 \tag{5.47}$$

for all $1 \leq i \leq n \in \mathbb{N}$. *Moreover, define for all* $1 \leq i \leq n \in \mathbb{N}$

$$l_{n,i} := \log\left(f_{n,i}\right).$$

(ii) Suppose that (i) *and Assumption 2.1*(ii) *hold simultaneously. Moreover, suppose that* $T_{n,i} = T_{n,1}$ *and* $T_{n,i}^{-1} = T_{n,1}^{-1}$ *for all* $1 \leq i \leq n$. *Set* $T_n := T_{n,1}$, $T_n^{-1} := T_{n,1}^{-1}$ *and* $l_n := l_{n,1}$ *for all* $n \in \mathbb{N}$.

(iii) Suppose that (ii) *and Assumption 2.1*(iii) *hold simultaneously.*

Now we present our extension of Theorem 3 in [12].

Theorem 5.23. *Suppose Assumption 5.22*(i). *Define*

$$h_{n,1,i}(s) := l_{n,i}\left(T_{n,i}^{-1}\left(n^{-s}\right)\right), \; h_{n,2,i}(s) := l_{n,i}\left(T_{n,i}^{-1}\left(1-n^{-s}\right)\right) \tag{5.48}$$

$$\text{and } \; h_{n,i}(s) := \max\left\{h_{n,1,i}(s), h_{n,2,i}(s)\right\} \tag{5.49}$$

for all $1 \le i \le n \in \mathbb{N}$ *and* $s > 0$. *Let* $\{k_n : n \in \mathbb{N}\}$ *be a subsequence of* $\mathbb{N}$ *and* $\beta^\# \in \mathbb{R}$.

(i) Suppose that for all sufficiently large $n \in \mathbb{N}$ *and some* $\delta > 0$

$$\sum_{i=1}^{k_n} \varepsilon_{k_n,i}^2 \le k_n^{1-2\beta^\#-\delta} \tag{5.50}$$

$$\text{and } \; \max_{1\le i\le k_n}\left\{ \lambda\!\!\lambda\left(s \ge \frac{\log(2)}{\log(k_n)} : \beta^\# - \frac{1}{2} \le \frac{h_{k_n,i}(s)}{\log(k_n)} - s + \frac{s\wedge 1}{2}\right)\right\} = 0. \tag{5.51}$$

Let $M \in (1,\infty)$. *If for all sufficiently large* $n \in \mathbb{N}$

$$\max_{1\le i\le k_n}\left\{ \lambda\!\!\lambda\left(s \ge M : \beta^\# - 1 \le \frac{h_{k_n,i}(s)}{\log(k_n)} - \left[1 - \frac{\log\log(k_n)}{\log(k_n)}\right] s\right)\right\} = 0 \tag{5.52}$$

$$\text{or if } \; \lim_{n\to\infty} \max_{1\le i\le n}\left\{ \sup_{s\ge M}\left| \frac{h_{n,i}(s)}{\log(n)} - \gamma(s)\right|\right\} = 0 \tag{5.53}$$

for some measurable $\gamma : [M,\infty) \to \mathbb{R}$ *then* $\{P_{(k_n)}, Q_{(k_n)}\}$ *converges weakly to* $\{\epsilon_0, \epsilon_0\}$.

(ii) If for some $\delta > 0$

$$\lim_{n\to\infty} \sum_{i=1}^{k_n} \varepsilon_{k_n,i}^2 = 0, \quad \sum_{i=1}^{k_n} \varepsilon_{k_n,i} \ge k_n^{1-\beta^\#+\delta} \quad \text{for all } n \in \mathbb{N} \tag{5.54}$$

$$\text{and } \; \liminf_{n\to\infty} \min_{1\le i\le k_n} \lambda\!\!\lambda\left(s \ge \frac{\log(2)}{\log(k_n)} : \beta^\# - \frac{1}{2} \le \frac{h_{k_n,i}(s)}{\log(k_n)} - s + \frac{s\wedge 1}{2}\right) > 0 \tag{5.55}$$

then $\{P_{(k_n)}, Q_{(k_n)}\}$ *converges weakly to* $\{\epsilon_{-\infty}, \epsilon_\infty\}$.

Remark 5.24. (i) Since $P_{n,i}$ is a continuous measure (5.46) and (5.47) are always fulfilled for the distribution function $T_{n,i} = F_{n,i}$ and the left-continuous quantile function $T_{n,i}^{-1} = F_{n,i}^{-1}$, see (2.5). Another possible choice is $T_{n,i} = 1 - F_{n,i}$ and $T_{n,i}^{-1}(u) = F_{n,i}^{-1}(1-u)$ for all $u \in (0,1)$. In this section the choices $T_{n,i} = F_{n,i}$ and $T_{n,i} = 1 - F_{n,i}$ lead to the same results because the conditions of Theorem 5.23 are symmetric. But in Part II, to be more specific in Section 8.3, we get different results for these two cases.

(ii) Suppose (5.50) for some $\beta^\# \ge 1$. By Lemma A.30 we have $\sum_{i=1}^{k_n} \varepsilon_{k_n} = o(1)$ as $n \to \infty$. Hence, by Remark 4.5 $\{P_{(k_n)}, Q_{(k_n)}\}$ converges weakly to $\{\epsilon_0, \epsilon_0\}$. □

An immediate consequence of Theorem 5.23 is given in the following for the case that $\varepsilon_{n,i}$ does not depend on i. This result should give the reader an idea how our extension can be applied to determine, e.g., the detection boundary for the parameter of ε_n.

Corollary 5.25. *Suppose that the assumptions of Theorem 5.23 are fulfilled with*

$$\varepsilon_{n,i} = \varepsilon_{n,1} =: \varepsilon_n := n^{-\beta} \text{ for all } n \in \mathbb{N} \text{ and for some } \beta \in \left(\frac{1}{2}, \infty\right). \tag{5.56}$$

Moreover, assume that there exists some $\beta^{\#} \in \mathbb{R}$ such that for every $\delta > 0$

$$\max_{1\leq i\leq n} \left\{ \lambda\!\!\lambda \left(s \geq \frac{\log(2)}{\log(n)} : \beta^{\#} + \delta - \frac{1}{2} \leq \frac{h_{n,i}(s)}{\log(n)} - s + \frac{s \wedge 1}{2} \right) \right\} = 0 \tag{5.57}$$

if $n \in \mathbb{N}$ is sufficiently large, and that for every $\delta > 0$

$$\liminf_{n\to\infty} \min_{1\leq i\leq n} \left\{ \lambda\!\!\lambda \left(s \geq \frac{\log(2)}{\log(n)} : \beta^{\#} - \delta - \frac{1}{2} \leq \frac{h_{n,i}(s)}{\log(n)} - s + \frac{s \wedge 1}{2} \right) \right\} > 0.$$

Let $M > 1$. Additionally, suppose that for every $\delta > 0$

$$\max_{1\leq i\leq n} \left\{ \lambda\!\!\lambda \left(s \geq M : \beta^{\#} + \delta - 1 \leq \frac{h_{n,i}(s)}{\log(n)} - \left[1 - \frac{\log\log(n)}{\log(n)}\right] s \right) \right\} = 0 \tag{5.58}$$

if $n \in \mathbb{N}$ is sufficiently large, or that for some measurable function $\gamma : [M, \infty) \to \mathbb{R}$

$$\lim_{n\to\infty} \max_{1\leq i\leq n} \left\{ \sup_{s\geq M} \left| \frac{h_{n,i}(s)}{\log(n)} - \gamma(s) \right| \right\} = 0.$$

(i) If β exceeds $\beta^{\#}$ then $\{P_{(n)}, Q_{(n)}\}$ converges weakly to $\{\epsilon_0, \epsilon_0\}$.

(ii) If β is smaller than $\beta^{\#}$ then $\{P_{(n)}, Q_{(n)}\}$ converges weakly to $\{\epsilon_{-\infty}, \epsilon_\infty\}$.

In the following we present Theorem 3 of Cai and Wu [12], which follows immediately from Corollary 5.25. Note that there is a typographically error in their theorem: F_n and z_n must be the distribution function and the quantile function of Q_n and not of G_n.

Corollary 5.26. *Suppose Assumption 5.22*(iii)*. Let $(\varepsilon_n)_{n\in\mathbb{N}}$ be given by* (5.56)*. Define*

$$h_{n,1}(s) := l_n\left(T_n^{-1}\left(n^{-s}\right)\right), \qquad h_{n,2}(s) := l_n\left(T_n^{-1}\left(1 - n^{-s}\right)\right),$$
$$h_n(s) := \max\left\{h_{n,1}(s), h_{n,2}(s)\right\}$$

for all $s > 0$ and for all $n \in \mathbb{N}$. Assume that

$$\lim_{n\to\infty} \sup_{s\in\left[\frac{\log(2)}{\log(n)},\infty\right)} \left|\frac{h_n(s)}{\log(n)} - \gamma(s)\right| = 0$$

for some measurable $\gamma : (0,\infty) \to \mathbb{R}$ and define

$$\beta^{\#} := \frac{1}{2} + \operatorname*{ess\,sup}_{s>0} \left\{\gamma(s) - s + \frac{s \wedge 1}{2}\right\}.$$

Suppose $\beta^{\#} \in \mathbb{R}$.

(i) If β exceeds $\beta^{\#}$ then $\{P_{(n)}, Q_{(n)}\}$ converges weakly to $\{\epsilon_0, \epsilon_0\}$.

(ii) If β is smaller than $\beta^{\#}$ then $\{P_{(n)}, Q_{(n)}\}$ converges weakly to $\{\epsilon_{-\infty}, \epsilon_\infty\}$.

As mentioned in the introduction Theorem 3 of [12] can not be applied to the model discussed in Corollary 5.7. The main reason is that $\frac{h_n}{\log(n)}$ does not converge uniformly for this model. Note that $\frac{h_n(s)}{\log(n)}$ even converges to $-\infty$ for some $s > 0$. But the assumptions of our extension are fulfilled. In the following we give an alternative proof of Corollary 5.7(i) and (v). By doing this we want to display the benefits of our extension and its opportunities to be applied. We refer the reader to [12] for more examples.

Alternative proof of Corollary 5.7 (i) *and* (v). Suppose that the assumptions of Corollary 5.7 are fulfilled. It remains to show that the assumptions of Theorem 5.23 and the one of Corollary 5.25 hold for

$$T_n = F_0 =: T \text{ and } T_n^{-1} = F_0^{-1} =: T^{-1} \text{ for all } n \in \mathbb{N},$$

where F_0 and F_0^{-1} are the distribution function and the left continuous quantile function of $\mathcal{Q}_0$, see (2.5). By Lemma 2.15, Assumption 2.16(iii) and the monotonicity of F_0^{-1} we can deduce that

$$\begin{aligned}\frac{h_n(s)}{\log(n)} &= \frac{\log(C(\vartheta_n))}{\log(n)} + \frac{\vartheta_n}{\log(n)} \max\left\{-h\left(F_0^{-1}\left(n^{-s}\right)\right), -h\left(F_0^{-1}\left(1-n^{-s}\right)\right)\right\} \\ &= \frac{\log\left(n^{rp+o(1)}\right)}{\log(n)} - \frac{n^r}{\log(n)}(1+o(1))\, h\left(F_0^{-1}\left(n^{-s}\right)\right) \\ &= rp + o(1) - n^{r+o(1)} h\left(F_0^{-1}\left(n^{-s}\right)\right) \text{ as } n \to \infty \end{aligned} \tag{5.59}$$

for all $s > 0$, where the $o(1)$-terms are independent of s.

Consider $p > 0$. Set $\beta^{\#} := \beta^{\#}_{Exp}(r,p)$, see (5.10). By Corollary 5.25 it remains to show that for all $\kappa \in (0, rp)$

$$\lambda\!\!\!\lambda\left(s > 0 : \beta^{\#} + 2\kappa - \frac{1}{2} \le \frac{h_n(s)}{\log(n)} - \left(1 - \frac{\log\log(n)}{\log(n)}\right) s + \frac{s \wedge 1}{2}\right) = 0 \tag{5.60}$$

$$\text{and} \min_{1 \le i \le k_n} \lambda\!\!\!\lambda\left(s > 0 : \beta^{\#} - 4\kappa - \frac{1}{2} \le \frac{h_n(s)}{\log(n)} - s + \frac{s \wedge 1}{2}\right) \ge \kappa \tag{5.61}$$

if $n \in \mathbb{N}$ is sufficiently large. Note that from (5.60) we obtain (5.57) and (5.58). Let $\kappa \in (0, rp)$ be fixed. By Assumption 2.16(iii) $h_{|[a,a+\delta]}$ is strictly increasing and continuous. Hence, it is invertible and its inverse $h^{-1} : [0, h(a+\delta)] \to [a, a+\delta]$ is also strictly increasing and continuous. By this, Theorem 2.19(i) and (5.6)

$$(F_0 \circ h^{-1})(t) \;=\; \mathcal{Q}^h_0[0,t] \;\sim_{\text{asy}}\; t^p L_1\left(\frac{1}{t}\right) \qquad \text{as } t \searrow 0 \tag{5.62}$$

for some slowly varying function L_1. Let $F : \mathbb{R} \to \mathbb{R}$ be a distribution function such that

$$F_{|(0,\varepsilon)} = (F_0 \circ h^{-1})_{|(0,\varepsilon)} \text{ for some } \varepsilon > 0.$$

Denote by F^{-1} its left continuous quantile function. It is easy to see that

$$F^{-1}(u) \;=\; \left(h \circ F_0^{-1}\right)(u)$$

for every sufficiently small $u \in (0,1)$. Combining Lemma 2.15, (5.62) and Theorem 1.5.12 of Bingham et al. [8], see Lemma A.32 for a more detailed verification, yields

$$\left(h \circ F_0^{-1}\right)(u) \;=\; F^{-1}(u) \;\sim_{\text{asy}}\; u^{\frac{1}{p}+o(1)} \quad \text{as } u \searrow 0. \tag{5.63}$$

Hence, by (5.59)

$$\begin{aligned}\sup_{s \le rp-\kappa}\left\{\frac{h_n(s)}{\log(n)} - \left(1 - \frac{\log\log(n)}{\log(n)}\right)s + \frac{s \wedge 1}{2}\right\} &\le 2rp - n^{r+o(1)} h\left(F_0^{-1}\left(n^{-rp+\kappa}\right)\right)(1 + o(1))\\ &\le 2rp - n^{r+o(1)-\frac{rp-\kappa}{p}} \;\;\to\;\; -\infty\end{aligned}$$

as $n \to \infty$. Since h is non-negative, see 2.16(iii), we deduce from (5.59) that

$$\begin{aligned}&\sup_{s \ge rp-\kappa}\left\{\frac{h_n(s)}{\log(n)} - \left(1 - \frac{\log\log(n)}{\log(n)}\right)s + \frac{s \wedge 1}{2}\right\}\\ &\le rp(1 + o(1)) - \left(1 - \frac{\log\log(n)}{\log(n)}\right)(rp - \kappa) + \frac{(rp - \kappa) \wedge 1}{2}\\ &\le \beta^{\#} - \frac{1}{2} + \frac{3}{2}\kappa \;\text{ as } n \to \infty.\end{aligned} \tag{5.64}$$

Consequently, (5.60) is verified. From (5.59) and (5.63) we see that

$$\begin{aligned}
&\sup_{s\in(rp+\kappa,rp+2\kappa)}\left\{\frac{h_n(s)}{\log(n)}-s+\frac{s\wedge 1}{2}\right\}\\
&\geq rp(1+o(1))-n^{r+o(1)}h\left(F_0^{-1}\left(n^{-rp-\kappa}\right)\right)-rp-2\kappa+\frac{(rp+2\kappa)\wedge 1}{2}\\
&\geq o(1)-n^{r+o(1)-\frac{rp+\kappa}{p}}+\beta^{\#}_{Exp}(r,p)-\frac{1}{2}-2\kappa\\
&\geq \beta^{\#}-\frac{1}{2}-3\kappa \quad \text{as } n\to\infty.
\end{aligned}$$

Hence, (5.61) holds.

Now consider $p=0$ and $\beta>\beta^{\#}_{Exp}(r,0)=\frac{1}{2}$. Since h is non-negative we deduce from (5.59) analogously to (5.64) that

$$\sup_{s>0}\left\{\frac{h_n(s)}{\log(n)}-\left(1-\frac{\log\log(n)}{\log(n)}\right)s+\frac{s\wedge 1}{2}\right\}\leq o(1)<\beta-\frac{1}{2}-\kappa \text{ as } n\to\infty$$

for some $\kappa\in(0,\beta-\frac{1}{2})$. Consequently, applying Theorem 5.23(i) with $\beta^{\#}:=\beta-\kappa$ shows that $\{P_{(n)},Q_{(n)}\}$ converges weakly to $\{\epsilon_0,\epsilon_0\}$. ■

Cai and Wu [12] discussed the case $P_{n,i}=N(0,1)$ separately. In the following we present our extension of it.

Theorem 5.27 (Extension of Theorem 1 in [12]). *Suppose Assumption 5.22*(i). *Let*

$$\widetilde{h}_{n,i}(x):=l_{n,i}\left(x\sqrt{2\log(n)}\right)$$

for all $x\in\mathbb{R}$ and $i\in\{1,\dots,n\}$. Let $\{k_n:n\in\mathbb{N}\}$ be a subsequence of $\mathbb{N}$ and $\beta^{\#}\in(\frac{1}{2},\infty)$.

(i) Suppose that for all sufficiently large $n\in\mathbb{N}$ and some $\delta>0$

$$\sum_{i=1}^{k_n}\varepsilon^2_{k_n,i}\leq k_n^{1-2\beta^{\#}-\delta} \quad \textit{and} \quad \max_{1\leq i\leq k_n}\lambda\!\!\lambda\left(x\in\mathbb{R}:\beta^{\#}-\frac{1}{2}\leq\frac{\widetilde{h}_{k_n,i}(x)}{\log(k_n)}-x^2+\frac{x^2\wedge 1}{2}\right)=0.$$

Let $M\in(1,\infty)$. If for all sufficiently large $n\in\mathbb{N}$

$$\begin{aligned}
&\max_{1\leq i\leq k_n}\lambda\!\!\lambda\left(x\in\mathbb{R}:|x|\geq M,\ \beta^{\#}-1\leq\frac{\widetilde{h}_{k_n,i}(x)}{\log(k_n)}-\left(1-\frac{\log\log(k_n)}{\log(k_n)}\right)x^2\right)=0\\
&\textit{or}\ \lim_{n\to\infty}\max_{1\leq i\leq n}\left\{\sup_{x\in\mathbb{R}:|x|\geq M}\left|\frac{\widetilde{h}_{n,i}(x)}{\log(n)}-\alpha(x)\right|\right\}=0
\end{aligned}$$

for some measurable $\alpha:\mathbb{R}\to\mathbb{R}$ then $\{P_{(k_n)},Q_{(k_n)}\}$ converges weakly to $\{\epsilon_0,\epsilon_0\}$.

(ii) If for some $\delta > 0$

$$\lim_{n\to\infty} \sum_{i=1}^{k_n} \varepsilon_{k_n,i}^2 = 0, \quad \sum_{i=1}^{k_n} \varepsilon_{k_n,i} \geq k_n^{1-\beta^\# + \delta} \quad \textit{for all } n \in \mathbb{N} \tag{5.65}$$

$$\textit{and} \quad \liminf_{n\to\infty} \min_{1\leq i\leq k_n} \lambda\!\!\lambda \left(x \in \mathbb{R} : \beta^\# - \frac{1}{2} \leq \frac{h_{k_n,i}(x)}{\log(k_n)} - x^2 + \frac{x^2 \wedge 1}{2} \right) > 0$$

then $\{P_{(k_n)}, Q_{(k_n)}\}$ converges weakly to $\{\epsilon_{-\infty}, \epsilon_\infty\}$.

Remark 5.28. Note that by Lemma A.29 we have for all $s > 0$ and $n \in \mathbb{N}$

$$-\Phi^{-1}(n^{-s}) = \Phi^{-1}(1 - n^{-s}) = \sqrt{2s\log(n)} - \widetilde{\Psi}_n(s), \tag{5.66}$$

where $\widetilde{\Psi}_n(s)$ is some remainder. If we would ignore $\widetilde{\Psi}$ then we could deduce Theorem 5.27 from Theorem 5.23 with $F_{n,i} = \Phi$ and

$$\begin{aligned} h_{n,i}(s^2) &= \max\left\{ l_{n,i}\left(F_{n,i}^{-1}\left(1 - n^{-s^2}\right)\right), \; l_{n,i}\left(F_{n,i}^{-1}\left(n^{-s^2}\right)\right)\right\} \\ &\approx \left\{ l_{n,i}\left(s\sqrt{2\log(n)}\right), l_{n,i}\left(-s\sqrt{2\log(n)}\right)\right\} = \max\left\{\widetilde{h}_{n,i}(s), \widetilde{h}_{n,i}(-s)\right\}. \end{aligned}$$

But due to the remainder $\widetilde{\Psi}$ we actually cannot conclude Theorem 5.27 from Theorem 5.23. □

We paid more attention to the extension of Theorem 3 in [12] since in Part II we show a new result concerning it. Due to this and the similarity of the corresponding proof and the one of Theorem 5.23 we leave the proof of Theorem 5.27 to the reader and do not present the simplification of it for the case $\varepsilon_{n,i} = \varepsilon_n$.

5.4.1. Proof of Theorem 5.23

As Cai and Wu [12] did we use the random variables $Y_{n,i}$, U, S and S_n defined in the following lemma. The statements about the distributions of these random variables, see Lemma 5.29, are a consequence of (5.46) and basic calculations.

Lemma 5.29. *Suppose that the assumptions of Theorem 5.23 hold. Let $\mathcal{P}$ be a probability measure on $(\Omega, \mathcal{A})$ and $Y_{n,i}$, U, S, S_n be real-valued random variables on $(\Omega, \mathcal{A}, \mathcal{P})$ for all $1 \leq i \leq n \in \mathbb{N}$ such that*

$$Y_{n,i} \sim P_{n,i}, \; U \sim \mathcal{U}(0,1), \; S = -\log(U) \textit{ and } S_n = \log(n)^{-1} S.$$

Then for all $1 \leq i \leq n \in \mathbb{N}$

$$\begin{aligned} S &\sim Exp(1), & S_n &\sim Exp(\log(n)), \\ Y_{n,i} &\overset{D}{=} T_{n,i}^{-1}(U) & &= T_{n,i}^{-1}(n^{-S_n}), \\ \textit{and} \quad Y_{n,i} &\overset{D}{=} T_{n,i}^{-1}(1-U) &&= T_{n,i}^{-1}(1-n^{-S_n}). \end{aligned}$$

The proof of the following lemma is almost identical with the one of Lemma 2 in [12].

Lemma 5.30. *Suppose the assumptions of Theorem 5.23. Let* (5.53) *be fulfilled for some* $M > 1$ *and some measurable* γ*. Then for all* $\kappa > 0$

$$\int_M^\infty n^{\gamma(u)-u} \, \mathrm{d}u \leq 2n^\kappa \tag{5.67}$$

if $n \in \mathbb{N}$ *is sufficiently large. Moreover,*

$$\lambda\!\!\lambda \left(u \geq M : \gamma(u) - u > 0\right) = 0. \tag{5.68}$$

Proof. Let $\kappa > 0$ and $Y_{n,i}$, U, S, S_n be defined as in Lemma 5.29. We deduce (5.67) from Lemma 5.29, (5.48), (5.49), (5.53) since for all sufficiently large $n \in \mathbb{N}$

$$\begin{aligned} \int_M^\infty n^{\gamma(u)-u} \, \mathrm{d}u = &\ \log(n)^{-1} \int_M^\infty n^{\gamma(s)} \, \mathrm{d}\mathcal{P}^{S_n}(s) \\ &\leq \int_M^\infty \exp(h_{n,n}(s) + \log(n)\kappa) \, \mathrm{d}\mathcal{P}^{S_n}(s) \\ &\leq n^\kappa \int f_{n,n}\left(T_{n,n}^{-1}\left[n^{-S_n}\right]\right) \mathrm{d}\mathcal{P} \\ &\quad + n^\kappa \int f_{n,n}\left(T_{n,n}^{-1}\left[1 - n^{-S_n}\right]\right) \mathrm{d}\mathcal{P} \\ &= 2\, n^\kappa \int f_{n,n}(Y_{n,n}) \, \mathrm{d}\mathcal{P} \ = \ 2n^\kappa \int f_{n,n} \, \mathrm{d}P_{n,n}. \end{aligned}$$

Suppose, contrary to (5.68), that there exist $\kappa_1, \kappa_2 > 0$ such that

$$\lambda\!\!\lambda\Big(u \geq M : \gamma(u) - u > \kappa_1\Big) \geq \kappa_2.$$

Thus, for all sufficiently large $n \in \mathbb{N}$

$$\int_M^\infty n^{\gamma(u)-u} \, \mathrm{d}u \geq \kappa_2 n^{\kappa_1}.$$

This contradicts (5.67). Consequently, (5.68) holds. ∎

Proof of Theorem 5.23. The statement of (ii) follows immediately from Theorem 8.10(i). It remains to verify (i). In order to do this we apply Corollary 4.35. Let the random variables $Y_{n,i}$, U, S and S_n on $(\Omega, \mathcal{A}, \mathcal{P})$, as introduced in Lemma 5.29, be given for all $1 \leq i \leq n \in \mathbb{N}$. Suppose for the first part of the proof that $x > 0$, $n \in \mathbb{N}$ and $i \in \{1, \ldots, k_n\}$ are fixed. Because $U \overset{D}{=} 1 - U$

$$\begin{aligned}
&\mu_{k_n,i}\Big(\varepsilon_{k_n,i} f_{k_n,i} > x\Big) \\
&= \int_{\{\varepsilon_{k_n,i} f_{k_n,i} > x\}} \exp\left(l_{k_n,i}\right) \, \mathrm{d}P_{k_n,i} \\
&= \int_{\{\varepsilon_{k_n,i} f_{k_n,i}(Y_{k_n,i}) > x\}} \exp\left(l_{k_n,i}(Y_{k_n,i})\right) \, \mathrm{d}\mathcal{P} \\
&= \int_{\left\{\varepsilon_{k_n,i} f_{k_n,i}\left(T^{-1}_{k_n,i}(U)\right) > x\right\}} \exp\left(l_{k_n,i}\left(T^{-1}_{k_n,i}(U)\right)\right) \, \mathrm{d}\mathcal{P} \\
&= \int_{\left\{U \leq \frac{1}{2},\, \varepsilon_{k_n,i} f_{k_n,i}\left(T^{-1}_{k_n,i}(U)\right) > x\right\}} \exp\left(l_{k_n,i}\left(T^{-1}_{k_n,i}(U)\right)\right) \, \mathrm{d}\mathcal{P} \\
&\quad + \int_{\left\{U \leq \frac{1}{2},\, \varepsilon_{k_n,i} f_{k_n,i}\left(T^{-1}_{k_n,i}(1-U)\right) > x\right\}} \exp\left(l_{k_n,i}\left(T^{-1}_{k_n,i}(1-U)\right)\right) \, \mathrm{d}\mathcal{P}.
\end{aligned} \tag{5.69}$$

Note that

$$U \leq \frac{1}{2} \quad \Leftrightarrow \quad S_{k_n} \geq \frac{\log(2)}{\log(k_n)}.$$

Due to this and (5.69)

$$\begin{aligned}
&\mu_{k_n,i}\Big(\varepsilon_{k_n,i} f_{k_n,i} > x\Big) \\
&= \sum_{j=1}^{2} \int_{\left\{s \geq \frac{\log(2)}{\log(k_n)} \,:\, \varepsilon_{k_n,i} \exp(h_{k_n,j,i}(s)) > x\right\}} \exp\left(h_{k_n,j,i}(s)\right) \, \mathrm{d}\mathcal{P}^{S_{k_n}}(s).
\end{aligned} \tag{5.70}$$

By using the same arguments

$$\begin{aligned}
&\mathbb{E}_{P_{k_n,i}}\left(f^2_{k_n,i} \mathbf{1}_{\{\varepsilon_{k_n,i} f_{k_n,i} \leq x\}}\right) \\
&= \sum_{j=1}^{2} \int_{\left\{s \geq \frac{\log(2)}{\log(k_n)} \,:\, \varepsilon_{k_n,i} \exp(h_{k_n,j,i}(s)) \leq x\right\}} \exp\left(2\, h_{k_n,j,i}(s)\right) \, \mathrm{d}\mathcal{P}^{S_{k_n}}(s).
\end{aligned} \tag{5.71}$$

Let $y_0 > 0$,

$$I_{n,1} := \log(k_n) \sum_{i=1}^{k_n} \varepsilon_{k_n,i}^2 \int_{\left\{\frac{\log(2)}{\log(k_n)} \le s \le 1\right\}} \exp\left(\log(k_n)\left(\frac{2h_{k_n,i}(s)}{\log(k_n)} - s\right)\right) \,\mathrm{d}s$$

$$\text{and } I_{n,2} := \log(k_n) \sum_{i=1}^{k_n} \varepsilon_{k_n,i} \int_{\{s \ge 1\}} \exp\left(\log(k_n)\left(\frac{h_{k_n,i}(s)}{\log(k_n)} - s\right)\right) \,\mathrm{d}s$$

for all $n \in \mathbb{N}$. We conclude from (5.70) and (5.71) that for all $n \in \mathbb{N}$

$$\begin{aligned}
&\sum_{i=1}^{k_n} \varepsilon_{k_n,i} \mu_{k_n,i}\Big(\varepsilon_{k_n,i} f_{k_n,i} > y_0\Big) \\
&\le 2 \sum_{i=1}^{k_n} \varepsilon_{k_n,i} \int_{\left\{s \ge \frac{\log(2)}{\log(k_n)} \,:\, \varepsilon_{k_n,i} \exp(h_{k_n,i}(s)) > y_0\right\}} \exp\left(h_{k_n,i}(s)\right) \log(k_n) \exp\left(-\log(k_n) s\right) \,\mathrm{d}s \\
&\le 2 \log(k_n) \sum_{i=1}^{k_n} \varepsilon_{k_n,i} \int_{\{s \ge 1\}} \exp\Big(h_{k_n,i}(s) - s \log(k_n)\Big) \,\mathrm{d}s \\
&\quad + \frac{2}{y_0} \log(k_n) \sum_{i=1}^{k_n} \varepsilon_{k_n,i}^2 \int_{\left\{\frac{\log(2)}{\log(k_n)} \le s \le 1\right\}} \exp\Big(2h_{k_n,i}(s) - s \log(k_n)\Big) \,\mathrm{d}s \\
&= 2I_{n,2} + \frac{2}{y_0} I_{n,1}
\end{aligned}$$

$$\begin{aligned}
\text{and } &\sum_{i=1}^{k_n} \varepsilon_{k_n,i}^2 \, \mathbb{E}_{P_{k_n,i}}\left(f_{k_n,i}^2 \mathbf{1}_{\{\varepsilon_{k_n,i} f_{k_n,i} \le y_0\}} - 1\right) \\
&\le \sum_{j=1}^{2} \sum_{i=1}^{k_n} \varepsilon_{k_n,i}^2 \int_{\left\{\frac{\log(2)}{\log(k_n)} \le s \le 1\right\}} \exp\Big(2\, h_{k_n,j,i}(s)\Big) \,\mathrm{d}\mathcal{P}^{S_{k_n}}(s) \\
&\quad + y_0 \sum_{j=1}^{2} \sum_{i=1}^{k_n} \varepsilon_{k_n,i} \int_{\{s \ge 1\}} \exp\left(\, h_{k_n,j,i}(s)\right) \,\mathrm{d}P^{S_{k_n}}(s) \\
&\le 2I_{n,1} + 2\, y_0 I_{n,2}.
\end{aligned}$$

By Remark 4.36(ii) it remains to show that $I_{n,1}$ and $I_{n,2}$ converge to 0 as $n \to \infty$. By (5.50) and (5.51)

$$I_{n,1} \le k_n^{1-2\beta^{\#}-\frac{\delta}{2}} \int_{\frac{\log(2)}{\log(k_n)}}^{1} \exp\Big(\log(k_n)\left(2\beta^{\#} - 1\right)\Big) \,\mathrm{d}s \le k_n^{-\frac{\delta}{2}} = o(1)$$

as $n \to \infty$. Define for all $n \in \mathbb{N}$

$$\lambda_n := \frac{\log\log(k_n)}{\log(k_n)}.$$

Note that by (5.50) and Lemma A.30

$$\sum_{i=1}^{k_n} \varepsilon_{k_n,i} \;\leq\; \sqrt{k_n \sum_{i=1}^{k_n} \varepsilon_{k_n,i}^2} \;\leq\; k_n^{1-\beta^\#-\frac{\delta}{2}} \text{ for all } n \in \mathbb{N}. \tag{5.72}$$

First, suppose (5.52). We can conclude from (5.51) and (5.72) that

$$\begin{aligned} I_{n,2} \;&\leq\; k_n^{1-\beta^\#-\frac{1}{4}\delta} \int_1^M k_n^{\beta^\#-1}\,\mathrm{d}s + k_n^{1-\beta^\#-\frac{1}{4}\delta} \int_M^\infty \exp\Big(\log(k_n)\,\Big(\beta^\# - 1 - \lambda_n s\Big)\Big)\,\mathrm{d}s \\ &\leq\; k_n^{-\frac{1}{4}\delta} M + k_n^{-\frac{1}{4}\delta} \int_M^\infty \exp(-s\log\log(k_n))\,\mathrm{d}s \;=\; o(1) \text{ as } n \to \infty. \end{aligned}$$

Second, suppose (5.53). Let $x \mapsto \lfloor x \rfloor$, $x \in \mathbb{R}$, be the floor function, i.e.,

$$\lfloor x \rfloor \;:=\; \max\{m \in \mathbb{Z} \,:\, m \leq x\} \text{ for all } x \in \mathbb{R}.$$

From Lemma 5.30, (5.51), (5.53) and (5.72) we obtain

$$\begin{aligned} I_{n,2} \;&\leq\; o(1) + k_n^{1-\beta^\#-\frac{1}{4}\delta} \int_M^\infty \exp\Big(\log(k_n)\,\Big[(1-\lambda_n)(\beta^\# - 1) + \lambda_n(\gamma(s) - s)\Big]\Big)\,\mathrm{d}s \\ &\leq\; o(1) + k_n^{1-\beta^\#-\frac{1}{4}\delta+(1-\lambda_n)(\beta^\#-1)} \int_M^\infty (\log(k_n))^{\gamma(s)-s}\,\mathrm{d}s \\ &\leq\; o(1) + k_n^{-\frac{1}{8}\delta} \int_M^\infty \lfloor \log(k_n) \rfloor^{\gamma(s)-s}\,\mathrm{d}s \\ &\leq\; o(1) + k_n^{-\frac{1}{8}\delta} \lfloor \log(k_n) \rfloor \;=\; o(1) \text{ as } n \to \infty. \qquad \blacksquare \end{aligned}$$

Part II.

Power of the higher criticism test

6. Introduction and motivation

In Chapter 1 we already gave an overview of the literature concerning the higher criticism test, in short HC. We do not repeat the discussion here but rather focus on introducing the HC statistic.

Definition 6.1 (Definition of HC). *Let* $(p_{n,i})_{1\le i\le n\in\mathbb{N}}$ *be a triangular array of p-values on some probability space* $(\Omega,\mathcal{A},\mathcal{P})$*, i.e.,* $p_{n,i}\in(0,1)$ *for all* $1\le i\le n\in\mathbb{N}$*. Denote by*

$$p_{(n)} \;:=\; (p_{n,1},\ldots,p_{n,n})$$

the corresponding vector of p-values for every $n\in\mathbb{N}$*. Let* $\widehat{F}_{n,p}:(0,1)\to[0,1]$ *be the empirical distribution function of the p-values, i.e.,*

$$\widehat{F}_{n,p}(t)=\frac{1}{n}\sum_{i=1}^{n}\mathbf{1}_{(0,t]}\left(p_{n,i}\right)\ \text{for all } t\in(0,1),\, n\in\mathbb{N}. \tag{6.1}$$

For every $1\le i\le n\in\mathbb{N}$ *and all* $t\in(0,1)$ *we define*

$$Z_{n,p}(t)\;:=\;\sqrt{n}\,\frac{\left(\widehat{F}_{n,p}(t)-t\right)}{\sqrt{t(1-t)}}. \tag{6.2}$$

Let I_n *be a subinterval of* $(0,1)$ *for all* $n\in\mathbb{N}$*. The one-sided HC statistic and the two-sided HC statistic are given by*

$$HC_{I_n}^{n,p,1}:=\sup_{t\in I_n}\{Z_{n,p}(t)\}\quad\text{and}\quad HC_{I_n}^{n,p,2}:=\sup_{t\in I_n}\{\,|Z_{n,p}(t)|\,\} \tag{6.3}$$

for all $n\in\mathbb{N}$*, respectively.*

Remark 6.2. In order to improve the readability we omit, here and subsequently, the index of $p_{(n)}$ within the definition of $\widehat{F}_{n,p}$, $Z_{n,p}$, $HC_{I_n}^{n,p,1}$ and $HC_{I_n}^{n,p,2}$. □

As already mentioned in Chapter 1 our testing problem is connected to the theory of multiple testing problems, where numerous testing problems are treated simultaneously.

Multiple tests are often based on p-values and so does HC. A typical assumption of multiple testing problems is that the p-values $p_{n,1}, \dots, p_{n,n}$ are i.i.d. uniformly distributed on $(0,1)$. In this case it is well known that $Z_{n,p}(t)$ converges in distribution to a standard normally distributed random variable for each fixed $t \in (0,1)$. That is why Tukey [59] suggested to reject the (global) null if

$$Z_{n,p}(\beta) > u_{1-\alpha} \quad \text{or in} \quad |Z_{n,p}(\beta)| > u_{1-\frac{\alpha}{2}}, \quad \text{respectively,}$$

for some pre-chosen levels $\alpha, \beta \in (0,1)$, where u_δ is the δ-quantile of the standard normal distribution for every $\delta \in (0,1)$, i.e.,

$$N(0,1)\,(-\infty, u_\delta] = \delta.$$

Now it can be seen why HC is sometimes called a *second level test*. The statistician counts the number of p-values which are smaller than or equal to some pre-chosen (first) level $\beta \in (0,1)$ or, in other words, the statistician counts the number of $i \in \{1,\dots,n\}$ for which

$$\mathcal{H}_{0,i,n} : p_{n,i} \sim \mathcal{U}(0,1) \quad \text{or in our context} \quad \mathcal{H}_{0,i,n} : P_{n,i}$$

is rejected at level β. Finally, the statistician normalises this number according to (6.2) and compares the resulting value to the critical value of a one- or two-sided Gauss test, in practice often denoted as z-test, at pre-specified (second) level $\alpha \in (0,1)$.

This idea of standardisation maybe reminds the reader of the Anderson-Darling test statistic [2], which is an integral-type test, where in the simplest case the integrand is equal to $Z_{n,p}^2$. Instead of this integral-type statistic, Donoho and Jin [20] suggested to use a supremum-type statistic, which reminds us of the Kolmogorov-Smirnov statistic. To be more specific, they modified Tukey's idea by taking the supremum of $Z_{n,p}(\beta)$ or the absolute value of it, respectively, for all β between 0 and some tuning parameter $\beta_0 \in (0,1)$, i.e., $\beta \in (0,\beta_0)$, where the choice $\beta_0 = \frac{1}{2}$ is recommended in [20]. As mentioned in Chapter 1 the detectable area of HC and the one of LLRT coincide for several distributions, see, e.g., [10, 12, 20]. Donoho and Jin [20] also suggested a modified version, where the interval $(0,\beta_0)$ is replaced by $(\frac{1}{n},\beta_0)$. By simulations they showed that the performance of the modified version is better for moderate n. That is why the interval I_n in Definition 6.1 may depend on n.

The next chapter consists of the theoretical results. Among others, we present sufficient conditions for the two trivial cases: HC cannot successfully separate $\mathcal{H}_{0,n}$ and $\mathcal{H}_{1,n}$ and HC can completely separate them (asymptotically). As far as we know, there are no re-

sults to be found in the literature concerning the first mentioned case. The proof of the result of the second case is a generalisation of the ones in [10, 12, 20].
In the final chapter of this part we apply the theoretical results to the models which we already discussed within the context of determining the accumulation points of LLR_n and in particular the detection boundary. We show that the detectable areas of HC and LLRT are the same for all our examples. By doing this we solve an unanswered problem concerning the model of Cai and Wu [12]. They showed that the statement of their Theorem 1 (our extension of it is Theorem 5.27), where only the standard normal distribution is allowed as the null distribution, can be transferred to HC. But it was not clear if their Theorem 3 (our extension of it is Theorem 5.23) can also be transferred to HC. In Section 8.3 we display that this is actually possible. The following question was also unanswered.

How does HC behave on the detection boundary asymptotically?

Even for the heterogeneous normal mixture model the answer was unknown. Remember that the power of LLRT is non-trivial on the detection boundary for the models discussed in Sections 5.1 to 5.3. In Sections 8.1, 8.2 and 8.4 we show that

HC has no power on the detection boundary asymptotically

for all these models.

7. Theoretical results

In this section we present the theoretical results concerning HC. Due to the definition of HC, see Definition 6.1, we need to work with p-values here. First, we explain briefly that the observations can be transferred to p-values without affecting the result of Part I. After that we introduce the general assumption for Part II, which is almost equal to Assumption 5.22, where we already dealt with transformations to the interval $(0,1)$. Moreover, we present the convergence of the HC statistic under the null, which was first proven by Jaeschke [34]. Finally, we present conditions for the two trivial cases that the sum of type I and type II error probabilities of HC tends to 0 or to 1, respectively.

Remark 7.1 (Transformation into p-values). Suppose Assumption 5.22(i).

(i) Instead of $\{P_{(n)}, Q_{(n)}\}$ we can also analyse the transformed version of it

$$\left\{\widetilde{P}_{(n)}, \widetilde{Q}_{(n)}\right\} := \left\{\bigotimes_{i=1}^{n} P_{n,i}^{T_{n,i}}, \bigotimes_{i=1}^{n} Q_{n,i}^{T_{n,i}}\right\} = \left\{\bigotimes_{i=1}^{n} \mathcal{U}(0,1)\,, \bigotimes_{i=1}^{n} Q_{n,i}^{T_{n,i}}\right\} \tag{7.1}$$

for all $n \in \mathbb{N}$. By Remark A.17(ii) and (iii) the results of Chapter 3 are not affected by this transformation. Moreover, note that by Lemma A.31

$$\widetilde{f}_{n,i} := \frac{\mathrm{d}\mu_{n,i}^{T_{n,i}}}{\mathrm{d}P_{n,i}^{T_{n,i}}} = \frac{\mathrm{d}\mu_{n,i}}{\mathrm{d}P_{n,i}} \circ T_{n,i}^{-1} = f_{n,i} \circ T_{n,i}^{-1} \quad \text{for all } 1 \le i \le n \in \mathbb{N}.$$

(ii) In Sections 2.3 and 5.2 we discussed exponential families $(\mathcal{Q}_\vartheta)_{\vartheta\in\Theta}$. Suppose that $T_{n,i} = T_{1,1} =: T$ for all $1 \le i \le n \in \mathbb{N}$. We can conclude from Lemma A.31 that $(\mathcal{Q}_\vartheta^T)_{\vartheta\in\Theta}$ is also an exponential family with

$$\frac{\mathrm{d}\mathcal{Q}_\vartheta^T}{\mathrm{d}\mathcal{Q}_0^T} = C(\vartheta)\, \exp\Big(-\vartheta\, h \circ T^{-1}\Big) =: C(\vartheta)\, \exp\Big(-\vartheta\, \widetilde{h}\Big),\ \vartheta \in \Theta.$$

To sum up, there is no loss of generality in assuming that $P_{n,i}$, $\mu_{n,i}$, $Q_{n,i}$ are measures on $((0,1), \mathcal{B}((0,1)))$ and that $P_{n,i} = \mathcal{U}(0,1)$ for all $1 \le i \le n \in \mathbb{N}$.

In Assumption 5.22(i) and so in the second part of this thesis we only consider continuous measures. But also discrete measures $P_{n,i}, \mu_{n,i}$ are of interest for signal detection, see, e.g., [5]. Note that the results of Chapter 4 can also be applied to discrete measures. In the second part of this thesis we need uniformly distributed p-values under the null. That is why we restrict our model to continuous measures $P_{n,i}$. But we want to emphasise that it is also possible to transform the data to uniformly distributed p-values under the null in case of discrete measures, see, e.g., Lemma 1.5.4 of Reiss [52]. It is a possible future project to have a more detailed look at discrete models. □

In the following we present the general assumption for Part II. It is almost equal to Assumption 5.22, whereas it contains several additional definitions and notations, which are needed afterwards.

Assumption 7.2. *(i) (Parameters) Let $\alpha_0 \in (0,1)$ and define for all $n \in \mathbb{N}$*

$$a_n := \sqrt{2\log\log(n)} \text{ and } b_n := 2\log\log(n) + \frac{1}{2}\log\log\log(n) - \frac{1}{2}\log(\pi).$$

(ii) (Shape of the intervals I_n) Let α_0 be given as in (i)*. Moreover, let $\rho, \lambda_n \in (0,1)$ and $\rho_{n,1}, \rho_{n,2} \in \mathbb{R}$ for every $n \in \mathbb{N}$ such that for all $n \in \mathbb{N}$*

$$\lambda_n \geq -\frac{\log(\alpha_0)}{\log(n)} \text{ and } \rho_{n,1}, \rho_{n,2}, \lambda_n = o(1) \text{ as } n \to \infty. \tag{7.2}$$

Define for all $n \in \mathbb{N}$

$$r_n := n^{-1+\lambda_n}, \; s_n := n^{-\rho+\rho_{n,2}}, \; t_n := n^{-\rho+\rho_{n,1}} \text{ and } u_n := n^{-\lambda_n}.$$

Let $(I_n)_{n\in\mathbb{N}}$ be a sequence of subsets of $(0,1)$ such that

$$\text{either } (r_n, u_n) \subseteq I_n \subseteq (0, \alpha_0) \text{ for all } n \in \mathbb{N} \tag{7.3}$$

$$\text{or } (r_n, u_n) \cup (1-u_n, 1-r_n) \subseteq I_n \text{ for all } n \in \mathbb{N}. \tag{7.4}$$

(iii) Let $j \in \{1,2\}$, Assumption 5.22(i) *be fulfilled and the parameters and intervals of Assumption 7.2*(i) *and* (ii) *be given. Let $\mathcal{P}$ be a probability measure on $(\Omega, \mathcal{A}, \mathcal{P})$ and $U_1, \ldots, U_n$ be i.i.d. random variables on $(\Omega, \mathcal{A}, \mathcal{P})$ for all $n \in \mathbb{N}$ such that*

$$P_{n,i}^{T_{n,i}} = P_{1,1}^{T_{1,1}} = \mathcal{U}(0,1) =: P_0 \text{ and } \mathcal{L}(U_1|\mathcal{P}) = P_0 \text{ for all } n \in \mathbb{N}.$$

For all $1 \le i \le n \in \mathbb{N}$ *define the probability measure*

$$\widetilde{Q}_{n,i} := Q_{n,i}^{T_{n,i}}$$

and denote by $G_{n,i}$ *its distribution function, i.e.,*

$$G_{n,i}(u) \;=\; Q_{n,i}^{T_{n,i}}(0,u] \;=\; u + \varepsilon_{k_{n},i}\left(\mu_{k_{n},i}^{T_{k_{n},i}}(0,u] - u\right) \text{ for all } u \in [0,1].$$

Moreover, denote by $G_{n,i}^{-1}$ *its left-continuous quantile function, compare* (2.5)*. Define*

$$Y_{n,i} := G_{n,i}^{-1}\left(U_i\right) \tag{7.5}$$

for all $1 \le i \le n \in \mathbb{N}$*. Furthermore, set*

$$U_{(n)} \;:=\; (U_1, \ldots, U_n) \;\text{ and }\; Y_{(n)} \;:=\; (Y_{n,1}, \ldots, Y_{n,n})$$

for all $n \in \mathbb{N}$*. In the following we use* $\widehat{F}_{n,U}$*,* $\widehat{F}_{n,Y}$*,* $HC_I^{n,U,j}$*,* $HC_I^{n,Y,j}$*,* $Z_{n,U}$*,* $Z_{n,Y}$ *for the quantities defined in* (6.1)-(6.3)*, where* $p_{(n)}$ *is replaced by* $U_{(n)}$ *and* $Y_{(n)}$*, respectively. Moreover, let* $\widetilde{P}_{(n)}$*,* $\widetilde{Q}_{(n)}$*,* $\widetilde{f}_{n,i}$ *be defined as in Remark 7.1*(i) *and set* $\widetilde{\mu}_{n,i} := \mu_{n,i}^{T_{n,i}}$.

(iv) Suppose that (iii) *and Assumption 2.1*(ii) *hold simultaneously. Set for all* $n \in \mathbb{N}$

$$G_n := G_{n,1},\; G_n^{-1} := G_{n,1}^{-1},\; \widetilde{\mu}_n := \widetilde{\mu}_{n,1},\; \widetilde{Q}_n := \widetilde{Q}_{n,1} \text{ and } \widetilde{f}_n := \widetilde{f}_{n,1}.$$

(v) Suppose that (iv) *and Assumption 2.1*(iii) *hold simultaneously.*

Remark 7.3. (i) Despite λ_n the constants $\rho, \rho_{n,1}, \rho_{n,2}$ do not have any effect on the statistic or the model. Hence, they can be freely chosen as long as $\rho \in (0,1)$ and $\rho_{n,1}, \rho_{n,2} = o(1)$ as $n \to \infty$. Moreover, note that (7.3) and (7.4), respectively, still holds when we replace λ_n by $\widetilde{\lambda}_n \ge \lambda_n$ for all $n \in \mathbb{N}$.

(ii) Obviously, $Y_{(n)}$ is $\widetilde{Q}_{(n)}$-distributed for all $n \in \mathbb{N}$. Hence, $HC_{I_n}^{n,U,j}$ is a version of the HC statistic under $\mathcal{H}_{0,n}$ and $HC_{I_n}^{n,Y,j}$ is one under $\mathcal{H}_{1,n}$, both, of course, for the transformed experiment (7.1).

(iii) (Higher criticism test) Suppose Assumption 7.2(iii). Let $(c_n)_{n\in\mathbb{N}}$ be a sequence of critical values in $\mathbb{R}$. For some fixed $n \in \mathbb{N}$ let $\widetilde{X}_{(n)} = \left(\widetilde{X}_{n,1}, \ldots, \widetilde{X}_{n,n}\right)$ be either $P_{(n)}$-distributed or $Q_{(n)}$- distributed. In practice our task is to decide if the distribution of the observation vector $\widetilde{X}_{(n)}$ corresponds to the null or to the alternative. Set

$$p_{n,i} \;:=\; T_{n,i}\left(\widetilde{X}_{n,i}\right) \text{ for all } 1 \le i \le n.$$

Clearly, $p_{(n)} := (p_{n,i})_{1\le i\le n}$ is either $\widetilde{P}_{(n)}$- or $\widetilde{Q}_{(n)}$-distributed. Then a HC test is given by

$$\varphi^{HC}_{n,I_n,c_n,j}\left(\widetilde{X}_{(n)}\right) := \begin{cases} 1 \\ 0 \end{cases} HC^{n,p,j}_{I_n} \begin{matrix} > \\ \le \end{matrix} \; c_n \;,\; n \in \mathbb{N}.$$

(iv) The condition (7.3) holds, e.g., for $I_n = (\frac{1}{n}, \alpha_0)$, which corresponds to the refinement of Donoho and Jin [20] mentioned in Chapter 6. □

In the following we present the limit of the HC statistic under the null. The result was first proved by Jaeschke [34], but he did not use the term HC. Contemporary Eicker [21] determined the limit of very similar statistics. Proofs of their results can also be found in [56]. A useful tool within the proofs is the Hungarian construction. For a deeper discussion of this construction we recommend the textbook of Csörgő and Révész [15], the paper of Csörgő et al. [14] and, of course, the original work of Komlós et al. [44, 45].

Theorem 7.4. *Suppose that Assumption 7.2*(iii) *is fulfilled. Let* Λ *be the distribution function of a standard Gumbel distribution, see* (2.7)*. Then*

$$a_n\, HC^{n,U,j}_{(0,r_n)\cup(1-r_n,1)} - b_n \xrightarrow{\mathcal{P}} -\infty \;\; \text{as } n \to \infty, \tag{7.6}$$
$$a_n\, HC^{n,U,j}_{(u_n,1-u_n)} - b_n \xrightarrow{\mathcal{P}} -\infty \;\; \text{as } n \to \infty, \tag{7.7}$$
$$a_n\, HC^{n,U,j}_{(s_n,t_n)\cup(1-t_n,1-s_n)} - b_n \xrightarrow{\mathcal{P}} -\infty \;\; \text{as } n \to \infty. \tag{7.8}$$

Furthermore, if (7.3) *holds then for all* $t \in \mathbb{R}$

$$\mathcal{P}\left(a_n\, HC^{n,U,j}_{I_n} - b_n \le t\right) \to \Lambda(t)^{\frac{j}{2}} \quad \text{as } n \to \infty. \tag{7.9}$$

Otherwise, if (7.4) *holds then for all* $t \in \mathbb{R}$

$$\mathcal{P}\left(a_n\, HC^{n,U,j}_{I_n} - b_n \le t\right) \to \Lambda(t)^{j} \quad \text{as } n \to \infty. \tag{7.10}$$

Remark 7.5 (Critical value). Let the assumptions of Remark 7.3(iii) be fulfilled. First, observe that

$$\mathbb{E}_{\mathcal{P}}\left(\varphi^{HC}_{n,I_n,c_n,j}\left(U_{(n)}\right)\right) \;=\; \mathcal{P}\left(a_n HC^{n,U,j}_{I_n} - b_n > a_n c_n - b_n\right) \;\text{ for all } n \in \mathbb{N}. \tag{7.11}$$

Suppose (7.4). By (7.9) and (7.11) $\varphi^{HC}_{n,I_n,c_n,j}$ is an asymptotically exact test of level $\alpha \in [0,1]$, i.e.,

$$\lim_{n\to\infty} \mathbb{E}_{\mathcal{P}}\left(\varphi^{HC}_{n,I_n,c_n,j}\left(U_{(n)}\right)\right) \;=\; \alpha, \tag{7.12}$$

if and only if

$$\widetilde{c}_n := a_n c_n - b_n \;\to\; (\Lambda^j)^{-1}(1-\alpha) \;=\; -\log\left(-\frac{1}{j}\log(1-\alpha)\right) \quad \text{as } n\to\infty, \tag{7.13}$$

where we extend the logarithm continuously to $[0,\infty]$, see (A.19). The sufficiency of (7.13) for (7.12) is obvious. The necessity of (7.13) can be concluded from a simple proof by contradiction. Assume that the limit of $(c_{k_n})_{n\in\mathbb{N}}$ exists in $\bar{\mathbb{R}} = [-\infty,\infty]$ for some $\{k_n : n\in\mathbb{N}\}$ of $\mathbb{N}$ but is unequal to the limit in (7.13). Fix $\alpha\in(0,1)$. By setting

$$c_n := \frac{-\log\left(-\frac{1}{j}\log(1-\alpha)\right) + b_n}{a_n},\; n\in\mathbb{N}, \tag{7.14}$$

we get an asymptotically exact test of level α. Note that

$$c_n \;\sim_{\text{asy}}\; \sqrt{2\log\log(n)} \quad \text{as } n\to\infty.$$

The statement of this remark is still valid if (7.3) holds and we replace j by $\frac{j}{2}$ in (7.13) and (7.14). □

Proof. Note that as mentioned by Jaeschke [34] his Corollaries 2 and 3 also hold for the statistics $W_n, \widehat{V}_n, \widehat{W}_n$ introduced at the beginning of his Section 2. Consequently, (7.6) and (7.7) can be concluded from these corollaries. From (1) and (2) of Jaeschke's Theorem we obtain (7.8),

$$\begin{aligned} &\mathcal{P}\left(a_n\, HC^{n,U,j}_{(r_n,1-r_n)} - b_n \le t\right) \to \Lambda(t)^j \\ \text{and }\; &\mathcal{P}\left(a_n\, HC^{n,U,j}_{(r_n,\alpha_0)} - b_n \le t\right) \;\to \Lambda(t)^{\frac{j}{2}} \quad \text{as } n\to\infty \end{aligned}$$

for all $t\in\mathbb{R}$. Combining this and (7.6)-(7.8) yields (7.9) and (7.10). ■

Below we present sufficient conditions for the case that HC can completely separate $\mathcal{H}_{0,n}$

and $\mathcal{H}_{1,n}$. The proof of the corresponding theorem is a generalisation of the proof of Theorem 1.2 in [20]. Quite similar techniques and ideas were also used in [10] and [12].

Theorem 7.6. *Let* $\{k_n : n \in \mathbb{N}\} \subset \mathbb{N}$ *be a subsequence of* $\mathbb{N}$. *Suppose Assumption 7.2*(iii).

(i) Let $(v_n)_{n\in\mathbb{N}}$ *be a sequence in* $\mathbb{R}$ *such that*

$$v_n \in I_n \cap \left(\frac{\varepsilon}{n}, \alpha_0\right) \tag{7.15}$$

for all $n \in \mathbb{N}$ *and some* $\varepsilon > 0$. *If* $j = 1$ *then assume additionally that*

$$\liminf_{n\to\infty} \frac{\sum_{i=1}^{k_n} \varepsilon_{k_n,i} \left(\mu_{k_n,i}^{T_{k_n,i}}(0, v_{k_n}] - v_{k_n}\right)}{\sqrt{k_n\, v_{k_n}(1 - v_{k_n}) \log\log(k_n)}} > 2\sqrt{2}. \tag{7.16}$$

Otherwise, if $j = 2$ *then assume additionally that*

$$\liminf_{n\to\infty} \left| \frac{\sum_{i=1}^{k_n} \varepsilon_{k_n,i} \left(\mu_{k_n,i}^{T_{k_n,i}}(0, v_{k_n}] - v_{k_n}\right)}{\sqrt{k_n\, v_{k_n}(1 - v_{k_n}) \log\log(k_n)}} \right| > 2\sqrt{2}. \tag{7.17}$$

Then

$$a_{k_n}\, HC_{I_{k_n}}^{k_n,Y,j} - b_{k_n} \xrightarrow{\mathcal{P}} \infty \quad \text{as } n \to \infty. \tag{7.18}$$

(ii) Let $(v_n)_{n\in\mathbb{N}}$ *be a sequence in* $\mathbb{R}$ *such that*

$$v_n \in I_n \cap \left[\alpha_0, 1 - \frac{\varepsilon}{n}\right) \tag{7.19}$$

for all $n \in \mathbb{N}$ *and some* $\varepsilon > 0$. *If* $j = 2$ *and* (7.17) *is fulfilled then* (7.18) *holds.*

Remark 7.7. (i) (*Detectable*) For simplicity suppose Assumption 7.2(iii) and (7.18) with $k_n = n$ for all $n \in \mathbb{N}$. Then for every fixed $c \in \mathbb{R}$

$$\mathcal{P}\Big(a_n\, HC_{I_n}^{n,Y,j} - b_n > c\Big) \to 1 \text{ as } n \to \infty.$$

It is easy to show that there exists a sequence $(c'_n)_{n\in\mathbb{N}}$ in $\mathbb{R}$ such that $c'_n \to \infty$ and

$$\mathcal{P}\Big(a_{k_n}\, HC_{I_{k_n}}^{k_n,Y,j} - b_{k_n} > c'_n\Big) \to 1 \text{ as } n \to \infty. \tag{7.20}$$

Define for all $n \in \mathbb{N}$ the critical value c_n by

$$c_n := \frac{c'_n + b_n}{a_n}.$$

By (7.20) and Remark 7.5 the type I and type II error probabilities of $\varphi_n^{HC} := \varphi_{n,I_n,c_n,j}^{HC}$ tend to 0. Consequently, φ_n^{HC} can completely separate the null and the alternative. Note that at least $\varphi_{k_n}^{HC}$ does so if (7.18) holds only for some subsequence $\{k_n : n \in \mathbb{N}\}$ of $\mathbb{N}$.

(ii) If (7.16) holds then (7.17) does so as well.

(iii) Suppose that

$$\frac{1}{\log\log(k_n)} \sum_{i=1}^{k_n} \varepsilon_{k_n,i}^2 \;\to\; 0 \quad \text{as } n \to \infty.$$

By Lemma A.30

$$\begin{aligned} &\sup_{v\in(0,\frac{1}{2}]} \frac{1}{\sqrt{v(1-v)k_n \, \log\log(k_n)}} \sum_{i=1}^{k_n} \varepsilon_{k_n,i} v &&= o(1) \\ \text{and} \quad &\sup_{v\in[\frac{1}{2},1)} \frac{1}{\sqrt{v(1-v)k_n \, \log\log(k_n)}} \sum_{i=1}^{k_n} \varepsilon_{k_n,i}\,(1-v) &&= o(1) \text{ as } n \to \infty. \end{aligned}$$

Thus, (7.17) under (7.15) and under (7.19) can be simplified to

$$\begin{aligned} &\liminf_{n\to\infty} \frac{\sum_{i=1}^{k_n} \varepsilon_{k_n,i}\, \mu_{k_n,i}^{T_{k_n,i}}(0, v_{k_n}]}{\sqrt{k_n\, v_{k_n}(1-v_{k_n}) \log\log(k_n)}} > 2\sqrt{2} \\ \text{and} \quad &\liminf_{n\to\infty} \frac{\sum_{i=1}^{k_n} \varepsilon_{k_n,i}\, \mu_{k_n,i}^{T_{k_n,i}}(1-v_{k_n}, 1)}{\sqrt{k_n\, v_{k_n}(1-v_{k_n}) \log\log(k_n)}} > 2\sqrt{2}, \text{ respectively.} \end{aligned}$$

Moreover, (7.16) and (7.17) are equivalent under (7.15). □

Proof of Theorem 7.6. Let $(v_n)_{n\in\mathbb{N}}$ be given as in (i) or as in (ii), respectively. First, note that

$$a_{k_n} HC_{I_{k_n}}^{k_n,Y,j} - b_{k_n} = \sqrt{2}\,\log\log(k_n) \left(\frac{HC_{I_{k_n}}^{k_n,Y,j}}{\sqrt{\log\log(k_n)}} - \sqrt{2} + o(1) \right) \quad \text{as } n \to \infty.$$

Hence, it is sufficient for (7.18) to show that for some $\gamma > 0$

$$\mathcal{P}\left(\frac{HC_{I_{k_n}}^{k_n,Y,j}}{\sqrt{\log\log(k_n)}} \geq \sqrt{2} + \gamma \right) \to 1 \quad \text{as } n \to \infty. \tag{7.21}$$

By (6.3) it is sufficient for (7.21) to prove

$$\mathcal{P}\left(\frac{Z_{k_n,Y}\left(v_{k_n}\right)}{\sqrt{\log\log\left(k_n\right)}} \leq \sqrt{2}+\gamma\right) \to 0 \text{ as } n\to\infty \text{ if } j=1 \tag{7.22}$$

$$\text{and} \quad \mathcal{P}\left(\frac{\left|Z_{k_n,Y}\left(v_{k_n}\right)\right|}{\sqrt{\log\log\left(k_n\right)}} \leq \sqrt{2}+\gamma\right) \to 0 \text{ as } n\to\infty \text{ if } j=2. \tag{7.23}$$

By applying Chebyshev's inequality we show that (7.22) and (7.23) are fulfilled under the assumptions of (i) as well as (ii). For this purpose, we need to calculate the expectation and the variance of $Z_{k_n,Y}(v_{k_n})$ for all $n \in \mathbb{N}$:

$$\begin{aligned}
\mathbb{E}_{\mathcal{P}}\Big(Z_{k_n,Y}\left(v_{k_n}\right)\Big) &= \sqrt{k_n}\,\frac{k_n^{-1}\sum_{i=1}^{k_n}\widetilde{Q}_{k_n,i}(0,v_{k_n}]-v_{k_n}}{\sqrt{v_{k_n}(1-v_{k_n})}} \\
&= \frac{\sum_{i=1}^{k_n}\varepsilon_{k_n,i}\left(\widetilde{\mu}_{k_n,i}\left(0,v_{k_n}\right]-v_{k_n}\right)}{\sqrt{k_n\,v_{k_n}(1-v_{k_n})}}\,, \qquad (7.24)\\
\operatorname{Var}_{\mathcal{P}}\Big(Z_{k_n,Y}\left(v_{k_n}\right)\Big) &= k_n\,\frac{k_n^{-2}\sum_{i=1}^{k_n}\widetilde{Q}_{k_n,i}(0,v_{k_n}]\left(1-\widetilde{Q}_{k_n,i}(0,v_{k_n}]\right)}{v_{k_n}(1-v_{k_n})} \\
&\leq \frac{\sum_{i=1}^{k_n}\widetilde{Q}_{k_n,i}(0,v_{k_n}]}{k_n\,v_{k_n}(1-v_{k_n})} \\
&= \frac{\sum_{i=1}^{k_n}v_{k_n}+\sum_{i=1}^{k_n}\varepsilon_{k_n,i}\left(\widetilde{\mu}_{k_n,i}\left(0,v_{k_n}\right]-v_{k_n}\right)}{k_n\,v_{k_n}(1-v_{k_n})} \\
&= \frac{1}{1-v_{k_n}} + \frac{\mathbb{E}_{\mathcal{P}}\left[Z_{k_n,Y}\left(v_{k_n}\right)\right]}{\sqrt{k_n\,v_{k_n}(1-v_{k_n})}}. \qquad (7.25)
\end{aligned}$$

By Chebyshev's inequality we have for every real-valued random variable Z on $(\Omega,\mathcal{A},\mathcal{P})$ with a finite second moment and a non-zero expectation $\mu \in \mathbb{R}\setminus\{0\}$ that

$$\begin{aligned}
\mathcal{P}\left(|Z| \leq \frac{|\mu|}{2}\right) &= \mathcal{P}\left(|Z|-|\mu| \leq -\frac{|\mu|}{2}\right) \leq \mathcal{P}\left(-|Z-\mu| \leq -\frac{|\mu|}{2}\right) \\
&= \mathcal{P}\left(|Z-\mu| \geq \frac{|\mu|}{2}\right) \quad \leq 4\,\frac{\operatorname{Var}_{\mathcal{P}}(Z)}{\mu^2} \qquad (7.26)
\end{aligned}$$

and, moreover, if $\mu > 0$ then

$$\mathcal{P}\left(Z \leq \frac{|\mu|}{2}\right) \leq \mathcal{P}\left(Z-\mu \leq -\frac{\mu}{2}\right) \leq \mathcal{P}\left(|Z-\mu| \geq \frac{\mu}{2}\right) \leq 4\,\frac{\operatorname{Var}_{\mathcal{P}}(Z)}{\mu^2}. \tag{7.27}$$

Now suppose that the assumptions of (i) hold and $j = 2$. Then we can conclude from

(7.15), (7.17) and (7.24)-(7.26) that for some sufficiently small $\gamma > 0$

$$\begin{aligned}
&\mathcal{P}\left(\frac{|Z_{k_n,Y}(v_{k_n})|}{\sqrt{\log\log(k_n)}} \le \sqrt{2}+\gamma\right)\\
&\le \mathcal{P}\left(\frac{|Z_{k_n,Y}(v_{k_n})|}{\sqrt{\log\log(k_n)}} \le \frac{1}{2}\frac{|\mathbb{E}_{\mathcal{P}}(Z_{k_n,Y}(v_{k_n}))|}{\sqrt{\log\log(k_n)}}\right)\\
&\le \left|\mathbb{E}_{\mathcal{P}}\Big(Z_{k_n,Y}(v_{k_n})\Big)\right|^{-2}\frac{4}{1-v_{k_n}} + \left|\mathbb{E}_{\mathcal{P}}\Big(Z_{k_n,Y}(v_{k_n})\Big)\right|^{-1}\frac{4}{\sqrt{k_n v_{k_n}(1-v_{k_n})}}\\
&\le \left|\mathbb{E}_{\mathcal{P}}\Big(Z_{k_n,Y}(v_{k_n})\Big)\right|^{-2}\frac{4}{1-\alpha_0} + \left|\mathbb{E}_{\mathcal{P}}\Big(Z_{k_n,Y}(v_{k_n})\Big)\right|^{-1}\frac{4}{\sqrt{\varepsilon(1-\alpha_0)}} = o(1)
\end{aligned}$$

Hence, (7.23) holds. Replacing (7.26) by (7.27) in the argumentation above we obtain (7.22) if the assumptions of (i) are fulfilled and $j = 1$. Finally, (i) is verified.
The proof of (ii) corresponds almost completely to the previous one of (i). We omit the details. We only want to mention that instead of (7.25) the following inequality should be used:

$$\mathrm{Var}_{\mathcal{P}}\Big(Z_{k_n,Y}(v_{k_n})\Big) \le \frac{\sum_{i=1}^{k_n}\Big(1-\widetilde{Q}_{k_n,i}(0,v_{k_n}]\Big)}{k_n\, v_{k_n}(1-v_{k_n})} \le \frac{1}{v_{k_n}} - \frac{\mathbb{E}_{\mathcal{P}}\Big(Z_{k_n,Y}(v_{k_n})\Big)}{\sqrt{k_n\, v_{k_n}(1-v_{k_n})}}. \qquad \blacksquare$$

Now we present sufficient conditions for the case that HC is (asymptotically) useless to distinguish between the null and the alternative.

Theorem 7.8. *Let $\{k_n : n \in \mathbb{N}\} \subset \mathbb{N}$ be a subsequence of $\mathbb{N}$ and Assumption 7.2*(v) *be fulfilled. Moreover, assume $P_{(k_n)} \triangleleft\triangleright Q_{(k_n)}$.*

(i) Let for all $n \in \mathbb{N}$

$$I_{n,1} := [r_{k_n}, s_{k_n}] \cup [t_{k_n}, u_{k_n}]$$

and (7.3) *be fulfilled. If*

$$R_n := a_{k_n}\sqrt{k_n}\,\varepsilon_{k_n} \sup_{\tau\in I_{n,1}}\left\{\frac{\left|\mu_{k_n}^{T_{k_n}}(0,\tau]-\tau\right|}{\sqrt{\tau}}\right\} = o(1) \qquad (7.28)$$

as $n\to\infty$ then for all $t \in \mathbb{R}$

$$\mathcal{P}\Big(a_{k_n}\, HC_{I_n}^{k_n,Y,j} - b_{k_n} \le t\Big) \to \Lambda(t)^{\frac{j}{2}} \quad \text{as } n\to\infty. \qquad (7.29)$$

(ii) Let for all $n \in \mathbb{N}$

$$I_{n,2} := [1 - u_{k_n}, 1 - t_{k_n}] \cup [1 - s_{k_n}, 1 - r_{k_n}]$$

and (7.4) *be fulfilled. If in addition to* (7.28)

$$a_{k_n} \sqrt{k_n}\, \varepsilon_{k_n} \sup_{\tau \in I_{n,2}} \left\{ \frac{\left| \mu_{k_n}^{T_{k_n}}(0,\tau] - \tau \right|}{\sqrt{1-\tau}} \right\} = o(1) \tag{7.30}$$

as $n \to \infty$ *then for all* $t \in \mathbb{R}$

$$\mathcal{P}\left(a_n\, HC_{I_n}^{k_n,Y,j} - b_n \le t \right) \to \Lambda(t)^j \quad \textit{as } n \to \infty. \tag{7.31}$$

Remark 7.9. (i) Let the assumptions of Remark 7.3(iii) and (7.29) or (7.31), respectively, be fulfilled. To improve the readability we set

$$\varphi_n^{HC} := \varphi_{n,I_n,c_n,j}^{HC} \text{ for all } n \in \mathbb{N}.$$

From the equivalence of (7.12) and (7.13) we can deduce that the following implication holds for every subsequence $\{k_{n,1} : n \in \mathbb{N}\}$ of $\{k_n : n \in \mathbb{N}\}$:

$$\begin{aligned} & \mathbb{E}_{P_{(k_{n,1})}}\left(\varphi_{k_{n,1}}^{HC}\right) \to C \in [0,1] \quad \text{as } n \to \infty \\ \Rightarrow \quad & \mathbb{E}_{Q_{(k_{n,1})}}\left(\varphi_{k_{n,1}}^{HC}\right) \to C \quad \text{as } n \to \infty. \end{aligned}$$

Hence, independently of the choice of the critical values $(c_n)_{n \in \mathbb{N}}$ we obtain

$$\lim_{n \to \infty} \mathbb{E}_{P_{(k_n)}}\left(\varphi_{k_n}^{HC}\right) + \mathbb{E}_{Q_{(k_n)}}\left(1 - \varphi_{k_n}^{HC}\right) = 1.$$

Thus, HC yields no better results than the test $\varphi \equiv \alpha \in [0,1]$ (asymptotically). In other words, HC cannot successfully separate $\mathcal{H}_{0,n}$ and $\mathcal{H}_{1,n}$ (asymptotically).

(ii) Suppose that

$$a_{k_n} \sqrt{k_n}\, \varepsilon_{k_n} = o(1) \quad \text{as } n \to \infty.$$

Then as $n \to \infty$

$$\sup_{v \in \left(0, \frac{1}{2}\right]} \left\{ a_{k_n} \sqrt{k_n} \varepsilon_{k_n} \frac{v}{\sqrt{v}} \right\} = o(1) \text{ and } \sup_{v \in \left(0, \frac{1}{2}\right]} \left\{ a_{k_n} \sqrt{k_n} \varepsilon_{k_n} \frac{1-v}{\sqrt{1-v}} \right\} = o(1).$$

Hence, (7.28) and (7.30) can be simplified to

$$a_n\sqrt{n}\,\varepsilon_n \sup_{v\in[r_n,s_n]\cup[t_n,u_n]}\left\{\frac{\mu_n^{T_n}(0,v]}{\sqrt{v}}\right\} = o(1)$$

$$\text{and}\quad a_n\sqrt{n}\,\varepsilon_n \sup_{v\in[r_n,s_n]\cup[t_n,u_n]}\left\{\frac{\mu_n^{T_n}(1-v,1)}{\sqrt{v}}\right\} = o(1) \quad \text{as } n\to\infty,\ \text{respectively.}$$

(iii) By Corollary 4.12(i) we have $P_{(k_n)} \vartriangleleft \vartriangleright Q_{(k_n)}$ if (**A**) is fulfilled. Note that by a simple modification within the proof the statement of Theorem 7.8 even holds without assuming $P_{(k_n)} \vartriangleleft \vartriangleright Q_{(k_n)}$ if we replace $I_{n,1}$ by $I_n \cap (0,\frac{1}{2})$ and $I_{n,2}$ by $I_n \cap [\frac{1}{2},1)$. □

Proof of Theorem 7.8. Clearly, $\widetilde{P}_{(k_n)} \vartriangleleft\vartriangleright \widetilde{Q}_{(k_n)}$ follows from $P_{(k_n)} \vartriangleleft\vartriangleright Q_{(k_n)}$. Hence, we can deduce from (7.6) and (7.7) that

$$a_{k_n}\, HC^{k_n,Y,j}_{(0,1)\setminus(I_{n,1}\cup I_{n,2})} - b_{k_n} \xrightarrow{\mathcal{P}} -\infty \text{ as } n\to\infty.$$

That is why it remains to show

$$\mathcal{P}\left(a_{k_n}\, HC^{k_n,Y,j}_{I_{n,1}} - b_{k_n} \le t\right) \to \Lambda(t)^{\frac{j}{2}} \tag{7.32}$$

$$\text{and}\quad \mathcal{P}\left(a_{k_n}\, HC^{k_n,Y,j}_{I_{n,1}\cup I_{n,2}} - b_{k_n} \le t\right) \to \Lambda(t)^{j}, \text{ respectively, as } n\to\infty \tag{7.33}$$

for all $t\in\mathbb{R}$. Since $G_{k_n}(\tau)\ge u \Leftrightarrow \tau \ge G_{k_n}^{-1}(u)$ for all $u\in(0,1)$ and $\tau\in\mathbb{R}$ we obtain

$$HC_I^{k_n,Y,1} = \sup_{\tau\in I}\left\{\frac{\sum_{i=1}^{k_n}\left(\mathbf{1}_{(0,G_{k_n}(\tau)]}(U_i)-\tau\right)}{\sqrt{k_n\,\tau(1-\tau)}}\right\} \tag{7.34}$$

$$\text{and}\quad HC_I^{k_n,Y,2} = \sup_{\tau\in I}\left\{\left|\frac{\sum_{i=1}^{k_n}\left(\mathbf{1}_{(0,G_{k_n}(\tau)]}(U_i)-\tau\right)}{\sqrt{k_n\,\tau(1-\tau)}}\right|\right\} \tag{7.35}$$

for all $I\subseteq(0,1)$ and all $n\in\mathbb{N}$. If $n\in\mathbb{N}$ is sufficiently large then

$$1>\tau+\varepsilon_{k_n}(1-\tau)\ge G_{k_n}(\tau)\ge(1-\varepsilon_{k_n})\tau>0 \text{ for all } \tau\in(0,1). \tag{7.36}$$

For all sufficiently large $n\in\mathbb{N}$ and every $\tau\in(0,1)$ we can define

$$\Delta_{k_n,1,\tau} := \frac{\sum_{i=1}^{k_n}\left[\mathbf{1}_{(0,G_{k_n}(\tau)]}(U_i)-G_{k_n}(\tau)\right]}{\sqrt{k_n\,G_{k_n}(\tau)\,(1-G_{k_n}(\tau))}},\quad \Delta_{k_n,2,\tau} := \sqrt{\frac{G_{k_n}(\tau)}{\tau}},$$

$$\Delta_{k_n,3,\tau} := \sqrt{\frac{1-G_{k_n}(\tau)}{(1-\tau)}} \quad\text{and}\quad \Delta_{k_n,4,\tau} := \sqrt{k_n}\,\frac{G_{k_n}(\tau)-\tau}{\sqrt{\tau(1-\tau)}}.$$

Clearly, for all $n \in \mathbb{N}$ and every $I \subseteq (0,1)$

$$\frac{\sum_{i=1}^{k_n}\left(\mathbf{1}_{(0,G_{k_n}(\tau)]}(U_i)-\tau\right)}{\sqrt{k_n\,\tau(1-\tau)}} \;=\; \Delta_{k_n,1,\tau}\Delta_{k_n,2,\tau}\Delta_{k_n,3,\tau}+\Delta_{k_n,4,\tau}. \tag{7.37}$$

Now suppose that (7.28) holds. The proof of (7.29) falls naturally into four steps:

$$\lim_{n\to\infty}\sup_{\tau\in I_{n,1}}|\Delta_{k_n,l,\tau}-1| = 0 \text{ for } l\in\{2,3\}, \tag{7.38}$$

$$\lim_{n\to\infty} a_{k_n}\sup_{\tau\in I_{n,1}}|\Delta_{k_n,4,\tau}| = 0, \tag{7.39}$$

$$\lim_{n\to\infty}\mathcal{P}\left(a_{k_n}\sup_{\tau\in I_{n,1}}\{\Delta_{k_n,1,\tau}\}-b_{k_n}\le t\right) = \Lambda(t)^{\frac{1}{2}} \text{ for all } t\in\mathbb{R}, \tag{7.40}$$

$$\text{and}\quad \lim_{n\to\infty}\mathcal{P}\left(a_{k_n}\sup_{\tau\in I_{n,1}}\{|\Delta_{k_n,1,\tau}|\}-b_{k_n}\le t\right) = \Lambda(t) \quad \text{for all } t\in\mathbb{R}. \tag{7.41}$$

We can conclude from (7.36) that for all $n\in\mathbb{N}$ and every $\tau\in I_{n,1}\subseteq(0,\frac{1}{2})$

$$1-\varepsilon_{k_n} \;\le\; \frac{1-G_{k_n}(\tau)}{1-\tau} \;\le\; 1+\frac{\varepsilon_{k_n}\tau}{1-\tau} \;\le 1+\varepsilon_{k_n}.$$

Thus, (7.38) follows for $l=3$. Moreover, we can conclude from (7.28) that

$$\sup_{\tau\in I_{n,1}}\left|\frac{G_{k_n}(\tau)-\tau}{\tau}\right| \;=\; \sup_{\tau\in I_{n,1}}\left\{\frac{\varepsilon_{k_n}}{\tau}\left|\mu_{k_n}^{T_{k_n}}(0,\tau]-\tau\right|\right\} \;\le\; \frac{R_n}{a_{k_n}\sqrt{k_n r_{k_n}}} \;=\; o(1) \tag{7.42}$$

as $n\to\infty$. It follows immediately that

$$\sup_{\tau\in I_{n,1}}|\Delta_{k_n,2,\tau}-1| \;=\; \sup_{\tau\in I_{n,1}}\left|\sqrt{1+\frac{G_{k_n}(\tau)-\tau}{\tau}}-1\right| = o(1)$$

as $n\to\infty$. Consequently, (7.38) is proved. (7.39) follows from (7.28) since

$$a_{k_n}\sup_{\tau\in I_{n,1}}|\Delta_{k_n,4,\tau}| = \sup_{\tau\in I_{n,1}}\left\{\frac{a_{k_n}\sqrt{k_n}\varepsilon_{k_n}}{\sqrt{\tau(1-\tau)}}\left|\mu_{k_n}^{T_{k_n}}(0,\tau]-\tau\right|\right\} \;\le\; \frac{R_n}{\sqrt{1-u_{k_n}}} \;=\; o(1)$$

as $n\to\infty$. Set for all $n\in\mathbb{N}$

$$I_{n,4} := \left[G_{k_n}(r_{k_n}), G_{k_n}(s_{k_n})\right]\cup\left[G_{k_n}(t_{k_n}), G_{k_n}(u_{k_n})\right].$$

From (7.42) we see that for all sufficiently large $n \in \mathbb{N}$

$$\left[2r_{k_n}, \frac{1}{2}s_{k_n}\right] \cup \left[2t_{k_n}, \frac{1}{2}u_{k_n}\right] \subset I_{n,4} \subset \left[\frac{1}{2}r_{k_n}, 2s_{k_n}\right] \cup \left[\frac{1}{2}t_{k_n}, 2u_{k_n}\right].$$

For the proof of (7.40) and (7.41) note that

$$\sup_{\tau \in I_{n,1}} \{\Delta_{k_n,1,\tau}\} = HC_{I_{n,4}}^{k_n,U,1} \text{and} \quad \sup_{\tau \in I_{n,1}} \{|\Delta_{k_n,1,\tau}|\} = HC_{I_{n,4}}^{k_n,U,2} \tag{7.43}$$

for all $n \in \mathbb{N}$. Applying Theorem 7.4 with the (new) constants

$$r'_{k_n} := 2r_{k_n},\ s'_{k_n} := \frac{1}{2}s_{k_n},\ t'_{k_n} := 2t_{k_n},\ u'_{k_n} := \frac{1}{2}u_{k_n}$$
$$\text{and}\ \ \widetilde{r}_{k_n} := \frac{1}{2}r_{k_n},\ \widetilde{s}_{k_n} := 2s_{k_n},\ \widetilde{t}_{k_n} := \frac{1}{2}t_{k_n},\ \widetilde{u}_{k_n} := 2u_{k_n}$$

and combining this with (7.43) verifies (7.40) and (7.41). Finally, (7.32) is shown and so is (i). The proof of (7.33) and so the one of (ii) is very similar to the proof of (7.32). Thus, we omit it and leave the details to the reader. ■

8. Applications to practical detection models

In Chapter 5 we discussed the behaviour of LLRT for several examples. We determined, among others, the detection boundary for the case that the signal probability $\varepsilon_{n,i} = \varepsilon_n$ does not depend on i. Donoho and Jin [20] showed that the detectable areas of LLRT and HC coincide for the sparse heterogeneous normal mixture model. Cai et al. [10] confirmed this result for the sparse and dense heterogeneous normal mixture model. In Sections 8.2 and 8.4 we do so for our h-model and our sparse and dense exponential family model. Cai and Wu [12] showed that under the assumptions of their Theorem 1 (our extended version is Theorem 5.27) the detectable areas of HC and LLRT coincide as well, see their Theorem 4. By having done this they extended the results in [10, 20]. But it was an unsolved problem if these areas also coincide under the assumptions of their Theorem 3 (our extended version is Theorem 5.23). In Section 8.3 we show that it is valid indeed. Beside presenting models, for which the detectable areas of HC and LLRT coincide, we are also interested in answering the following question.

How does HC behave on the detection boundary asymptotically?

As far as we know, there are no results about this issue in the literature until recently. The accumulation points of $\{P_{(n)}, Q_{(n)}\}$ are non-trivial on the boundary for the models discussed in Sections 5.1 to 5.3. We show that for all these models

HC cannot successfully separate $\mathcal{H}_{0,n}$ and $\mathcal{H}_{1,n}$ on the boundary asymptotically.

In other words, the sum of type I and II error probabilities of HC tends to 1 for these models on the boundary as $n \to \infty$. Consequently, LLRT yields indeed better than HC (asymptotically), at least on the detection boundary.

8.1. Heteroscedastic normal mixtures

8.1.1. Sparse case

In this section we analyse the behaviour of HC for the sparse heteroscedastic normal mixture model given in Example 2.6. In Theorem 5.1 the detection boundary ρ^* for LLRT was presented, see also Figure 5.1. Cai et al. [10] already showed that HC can completely separate the null and the alternative if r exceeds the detection boundary $\rho^*(\beta, \tau)$. With respect to Theorem 5.1 we are also interested in the behaviour of HC on the detection boundary. We verify that HC cannot successfully separate the null and the alternative on the boundary (asymptotically). The results for the sparse heterogeneous normal mixture model concerning HC are visualised in Figure 1.2.

First, we present the results mentioned above. Before we give the proof of them we present and prove some technical lemmas.

Theorem 8.1. *Suppose that Assumption 7.2*(v) *and the assumptions of Theorem 5.1 hold simultaneously, where*

$$T_n \;=\; 1 - \Phi \textit{ for all } n \in \mathbb{N}.$$

(i) *If $r > \rho^*(\beta, \tau)$ then* (7.18) *holds. In this case HC can completely separate $\mathcal{H}_{0,n}$ and $\mathcal{H}_{1,n}$ (asymptotically), see Remark 7.7*(i).

(ii) *Consider $r = \rho^*(\beta, \tau)$. If* (5.2) *holds then replace $\varepsilon_n = n^{-\beta}$ by ε_n defined as in* (5.3). *Then* (7.29) *holds for all $t \in \mathbb{R}$ if* (7.3) *is fulfilled, and* (7.31) *holds for all $t \in \mathbb{R}$ if* (7.4) *is fulfilled. Thus, HC cannot successfully separate $\mathcal{H}_{0,n}$ and $\mathcal{H}_{1,n}$ (asymptotically), see Remark 7.9*(i).

Remark 8.2. (i) Note that we assume $r > 0$ in Theorem 5.1 and as well in Theorem 8.1. Hence, $r = \rho^*(\beta, \tau)$ implies that

$$\text{either } \tau < \sqrt{2} \text{ or } (\beta, \tau) \in \left(1 - \frac{1}{\tau^2}, 1\right) \times \left[\sqrt{2}, \infty\right).$$

(ii) We deduce Theorem 8.1 from Theorems 7.6 and 7.8. We need to analyse the measure $\widetilde{\mu}_n = N(\vartheta_n, \tau^2)^{1-\Phi}$ for every $n \in \mathbb{N}$. Fix $n \in \mathbb{N}$. Let X be a random variable on

$(\Omega, \mathcal{A}, \mathcal{P})$ such that $X \sim \mu_n = N(\vartheta_n, \tau^2)$. Then for all $t \in (0,1)$

$$\begin{aligned}\widetilde{\mu}_n(0,t] \;=\; \mu_n^{1-\Phi}(0,t] &= \mathcal{P}\left(1-\Phi(X)\le t\right) = \mathcal{P}\Big(X \ge \Phi^{-1}(1-t)\Big) \\ &= 1-\Phi\left(\frac{\Phi^{-1}(1-t)-\vartheta_n}{\tau}\right) \\ &= 1-\Phi\left(-\frac{\Phi^{-1}(t)+\vartheta_n}{\tau}\right). \qquad (8.1)\end{aligned}$$

□

To prove Theorem 8.1(ii) we apply Theorem 7.8. In the corresponding proof at the end of this section we split the interval $I_{n,1}$ from Theorem 7.8(i) into two subintervals. In each of the following two lemmas we discuss R_n from Theorem 7.8(i) restricted to one of these subintervals, respectively.

Lemma 8.3. *Suppose that the assumptions of Theorem 8.1 are fulfilled with $r = \rho^*(\beta,\tau)$. Let $\varepsilon > 0$ be sufficiently small such that*

$$\varepsilon \;\le\; \frac{1}{6}\Big(2\beta - 1 - \rho^*(\beta,\tau)\Big). \qquad (8.2)$$

Then for all sufficiently large $n \in \mathbb{N}$

$$R_{n,1}(\varepsilon) := a_n n^{\frac{1}{2}-\beta} \sup_{v\in(n^{-r-2\varepsilon},u_n]} \left\{\frac{\mu_n(0,v]}{\sqrt{v}}\right\} \le n^{-\varepsilon}.$$

Remark 8.4. Since $r > 0$ it is easy to verify that the term on the right side of (8.2) is positive. □

Proof of Lemma 8.3. We have

$$R_{n,1}(\varepsilon) \le a_n n^{\frac{1}{2}-\beta+\frac{1}{2}r+\varepsilon} = n^{o(1)-\frac{1}{2}[2\beta-1-\rho^*(\beta,\tau)]+\varepsilon} \;\le\; n^{o(1)-3\varepsilon+\varepsilon} \text{ as } n\to\infty. \qquad \blacksquare$$

Lemma 8.5. *Let the assumptions of Theorem 8.1 be fulfilled with*

$$\lambda_n \;\ge\; \frac{5\log\log(n)}{\log(n)} \;\textit{for all } n \in \mathbb{N}. \qquad (8.3)$$

Let $\varepsilon \in (0, \frac{1}{2}-\frac{1}{2}r)$. If $\tau < \sqrt{2}$ assume additionally that

$$r + 2\varepsilon < r\left(\frac{2}{2-\tau^2}\right)^2.$$

Then as $n \to \infty$

$$R_{n,2}(\varepsilon) \;:=\; \sup_{v\in[r_n,n^{-r-2\varepsilon}]} \left\{\frac{\widetilde{\mu}_n(0,v]}{\sqrt{v}}\right\} \le n^{J_1(\tau,\beta,r)} \,(\log(n))^{J_2(\tau,\beta,r)+o(1)}\,, \tag{8.4}$$

$$\textit{where } J_1(\tau,\beta,r) := \begin{cases} \frac{r}{2-\tau^2} & \textit{if } r < \frac{(2-\tau^2)^2}{4},\ \tau<\sqrt{2} \\ \frac{1}{2} - \left(\frac{\sqrt{r}-1}{\tau}\right)^2 & \textit{otherwise} \end{cases}$$

$$\textit{and} \quad J_2(\tau,\beta,r) := \begin{cases} -\frac{1}{8} & \textit{if } r < \frac{(2-\tau^2)^2}{4},\ \tau<\sqrt{2} \\ \frac{11}{2\tau^2}\left[1-\sqrt{r}\right] - 3 & \textit{otherwise} \end{cases}. \tag{8.5}$$

Moreover, there exists a sequence $(v_n^*)_{n\in\mathbb{N}}$ with $v_n^* \in [r_n,u_n]$ for all $n\in\mathbb{N}$ such that

$$\frac{\widetilde{\mu}_n(0,v_n^*]}{\sqrt{v_n^*}} \;=\; n^{J_1(\tau,\beta,r)+o(1)} \quad \textit{as } n\to\infty. \tag{8.6}$$

Remark 8.6. Note that we have

$$\rho^*(\beta,\tau) < \frac{(2-\tau^2)^2}{4} \text{ and } \tau^2 < \sqrt{2} \text{ if and only if } \beta < 1-\frac{\tau^2}{4} \text{ and } \tau^2<\sqrt{2}.$$

By a simple calculation we obtain

$$J_1(\tau,\beta,\rho^*(\beta,\tau)) = \begin{cases} \frac{1}{2}-\frac{1}{\tau^2} & \text{if } \beta \le 1-\frac{1}{\tau^2},\ \tau\ge\sqrt{2} \\ \beta-\frac{1}{2} & \text{otherwise} \end{cases}.$$

Hence,

$$J_1(\tau,\beta,\rho^*(\beta,\tau)) \;\ge\; \beta-\frac{1}{2}. \qquad \square$$

Proof of Lemma 8.5. Throughout this proof, we use the parametrisation $v = v_n = n^{-\kappa_n}$. Observe that

$$v_n = n^{-\kappa_n} \in [r_n, n^{-r-2\varepsilon}] \quad \text{if and only if} \quad \kappa_n \in [r+2\varepsilon, 1-\lambda_n]. \tag{8.7}$$

In all following argumentations we only use Landau terms when the corresponding convergence holds uniformly in $\kappa_n \in [r+2\varepsilon, 1-\lambda_n]$. From (8.7) and Lemma A.29 we obtain for all $v_n = n^{-\kappa_n} \in [r_n, n^{-r-2\varepsilon}]$ and some $\delta > 0$

$$\begin{aligned}\Phi^{-1}(v_n) + \vartheta_n = & -\sqrt{2\kappa_n \log(n)} + \frac{\log\log(n)}{2\sqrt{2\kappa_n \log(n)}}(1+o(1)) + \sqrt{2r\log(n)} \\ = & \sqrt{2\log(n)}\left[\sqrt{r} - \sqrt{\kappa_n} + \frac{\log\log(n)}{4\log(n)\sqrt{\kappa_n}}(1+o(1))\right] \qquad (8.8)\\ \leq & \sqrt{2\log(n)}\left[\sqrt{r} - \sqrt{r+2\varepsilon} + o(1)\right] \ \leq \ -\delta\sqrt{\log(n)} \qquad (8.9)\end{aligned}$$

as $n \to \infty$. Combining (8.1), (8.8), (8.9) and (A.23) yields for all $v_n = n^{-\kappa_n} \in [r_n, n^{-r-2\varepsilon}]$ and every $n \in \mathbb{N}$

$$\begin{aligned}& v_n^{-\frac{1}{2}} \widetilde{\mu}_n((0, v_n]) \\ & = n^{\frac{1}{2}\kappa_n} \frac{1}{\sqrt{2\pi}} \frac{-\tau}{\Phi^{-1}(v_n) + \vartheta_n} \exp\left[-\frac{1}{2}\left(-\frac{\Phi^{-1}(v_n) + \vartheta_n}{\tau}\right)^2\right](1+o(1)) \\ & = \frac{\tau}{2\sqrt{\pi}\left(\sqrt{\kappa_n} - \sqrt{r}\right)} (\log(n))^{-\frac{1}{2}} n^{\frac{1}{2}\kappa_n - \tau^{-2}\left[\sqrt{r} - \sqrt{\kappa_n} + \frac{\log\log(n)}{4\log(n)\sqrt{\kappa_n}}(1+o(1))\right]^2} (1+o(1)) \\ & = \frac{\tau}{2\sqrt{\pi}\left(\sqrt{\kappa_n} - \sqrt{r}\right)} n^{E_{n,1}(\kappa_n)} \ (\log(n))^{E_{n,2}(\kappa_n)},\end{aligned}$$

$$\text{where } E_{n,1}(\kappa_n) := \frac{\kappa_n}{2} - \frac{1}{\tau^2}\left[\kappa_n - 2\sqrt{r\kappa_n} + r\right] = -\frac{r}{\tau^2} + \frac{\tau^2 - 2}{2\tau^2}\kappa_n + 2\frac{\sqrt{r\kappa_n}}{\tau^2}, \qquad (8.10)$$

$$\begin{aligned}E_{n,2}(\kappa_n) := & -\frac{1}{2} - \frac{1}{\tau^2} 2\left(\sqrt{r} - \sqrt{\kappa_n}\right) \frac{1}{4\sqrt{\kappa_n}} \Big(1 + o(1)\Big) + o(1) \\ = & -\frac{1}{2}\left[1 + \frac{1}{\tau^2}\left(\sqrt{\frac{r}{\kappa_n}} - 1\right)\right] + o(1) \text{ as } n \to \infty. \qquad (8.11)\end{aligned}$$

If $\tau^2 \neq 2$ then

$$\begin{aligned}2\tau^2 E_{n,1}(\kappa_n) = & -2r + \Big(\tau^2 - 2\Big)\left[\sqrt{\kappa_n} + \frac{2\sqrt{r}}{\tau^2 - 2}\right]^2 - \frac{4r}{\tau^2 - 2} \\ = & \Big(\tau^2 - 2\Big)\left[\sqrt{\kappa_n} - \frac{2\sqrt{r}}{2 - \tau^2}\right]^2 - \frac{2r\tau^2}{\tau^2 - 2}. \qquad (8.12)\end{aligned}$$

Now we determine the (unique) global maximum point $\kappa_n^* \in [r + 2\varepsilon, 1 - \lambda_n]$ of the function $E_{n,1} : [r + 2\varepsilon, 1 - \lambda_n] \to \mathbb{R}$ defined by (8.10), $n \in \mathbb{N}$. For this purpose we discuss three cases.

First, assume that $\tau < \sqrt{2}$ and $r < \frac{1}{4}(2 - \tau^2)^2$. From (8.12) we see that

$$\kappa_n^* = \left(\frac{2\sqrt{r}}{2 - \tau^2}\right)^2 = r\frac{4}{(2 - \tau^2)^2} \quad \text{and} \quad E_{n,1}(\kappa_n^*) = 0 - \frac{r}{\tau^2 - 2} \text{ for all } n \in \mathbb{N}.$$

Hence, (8.6) holds for $v_n^* = n^{\kappa_n^*}$, $n \in \mathbb{N}$. Let $\delta > 0$ be sufficiently small such that

$$\kappa_n^*(1-\delta)^{-2} < 1 \text{ and } \delta < \frac{\tau^2}{2(2-\tau^2)} \text{ for all } n \in \mathbb{N}.$$

Note that $E_{n,2} : [r+2\varepsilon, 1-\lambda_n] \to \mathbb{R}$ defined by (8.11) is continuous and strictly increasing for all $n \in \mathbb{N}$. Finally, we can conclude (8.4) since for all sufficiently large $n \in \mathbb{N}$

$$\begin{aligned}\sup_{v\in[r_n, n^{-r-2\varepsilon}]}\left\{\frac{\widetilde{\mu}_n(0,v]}{\sqrt{v}}\right\} &= \sup_{\kappa_n\in[r+2\varepsilon,\kappa_n^*(1-\delta)^{-2}]}\left\{\frac{\tau\, n^{E_{n,1}(\kappa_n)}}{2\sqrt{\pi}\left(\sqrt{\kappa_n}-\sqrt{r}\right)}\;(\log(n))^{E_{n,2}(\kappa_n)}\right\}\\ &\le \frac{\tau}{2\sqrt{\pi}\left(\sqrt{r+2\varepsilon}-\sqrt{r}\right)} n^{-\frac{r}{\tau^2-2}}\,(\log(n))^{E_{n,2}\left(\kappa_n^*(1-\delta)^{-2}\right)},\end{aligned}$$

$$\begin{aligned}\text{where } E_{n,2}\left(\kappa_n^*(1-\delta)^{-2}\right) &= -\frac{1}{2}\left[1+\frac{1}{\tau^2}\left(\frac{(1-\delta)(2-\tau^2)}{2}-1\right)\right]+o(1)\\ &= -\frac{1}{4}+\frac{1}{4}\delta\,\frac{2-\tau^2}{\tau^2}+o(1) \;\le\; -\frac{1}{8}+o(1)\end{aligned}$$

as $n \to \infty$. The other two cases can be discussed simultaneously. Let either $\tau < \sqrt{2}$ and $r \ge \frac{1}{4}(2-\tau^2)^2$ or $\tau \ge \sqrt{2}$. It is easy to check that from (8.10) and (8.12) we have

$$\kappa_n^* = 1-\lambda_n \text{ for all } n \in \mathbb{N}$$

in both cases. Note that $2\sqrt{r} \ge 2-\tau^2$. Hence,

$$\begin{aligned}E_{n,1}(1-\lambda_n) &= -\frac{r}{\tau^2}+\frac{\tau^2-2}{2\tau^2}(1-\lambda_n)+\frac{2}{\tau^2}\sqrt{r\,(1-\lambda_n)}\\ &= -\frac{r}{\tau^2}+\frac{\tau^2-2}{2\tau^2}(1-\lambda_n)+\frac{2\sqrt{r}}{\tau^2}\left(1-\lambda_n\left(\frac{1}{2}+o(1)\right)\right)\\ &= \frac{1}{2}-\frac{1}{\tau^2}\left(r+1-2\sqrt{r}\right)+\lambda_n\left(\frac{1}{\tau^2}-\frac{1}{2}-\frac{\sqrt{r}}{\tau^2}\right)+o\left(\frac{\log\log(n)}{\log(n)}\right) \qquad (8.13)\\ &\le \frac{1}{2}-\frac{\left(\sqrt{r}-1\right)^2}{\tau^2}+\frac{5\log\log(n)}{\log(n)}\left(\frac{1}{\tau^2}-\frac{1}{2}-\frac{\sqrt{r}}{\tau^2}+o(1)\right) \text{ as } n\to\infty.\end{aligned}$$

From (8.13) we can deduce (8.6) for $v_n^* := n^{\kappa_n^*}$, $n \in \mathbb{N}$. Since $E_{n,2} : [r+2\varepsilon, 1-\lambda_n] \to \mathbb{R}$ is strictly increasing for all $n \in \mathbb{N}$ we can conclude (8.4) from the following equation:

$$E_{n,2}(1-\lambda_n)+5\left(\frac{1}{\tau^2}-\frac{1}{2}-\frac{\sqrt{r}}{\tau^2}+o(1)\right) = \frac{11}{2\tau^2}\left(1-\sqrt{r}\right)-3+o(1) \text{ as } n\to\infty. \qquad \blacksquare$$

Proof of Theorem 8.1. We can assume (8.3) without loss of generality, see Remark 7.3(i). Let $(v_n^*)_{n\in\mathbb{N}}$ be the sequence from Lemma 8.5. Then as $n \to \infty$

$$\frac{n\varepsilon_n\widetilde{\mu}_n(0,v_n^*]}{\sqrt{n\,v_n^*(1-v_n^*)\,\log\log(n)}} = n^{\frac{1}{2}-\beta+o(1)}\;(v_n^*)^{-\frac{1}{2}}\;\widetilde{\mu}_n(0,v_n^*] = n^{J_1(\tau,\beta,r)+\frac{1}{2}-\beta+o(1)} \qquad (8.14)$$

Note that $r \mapsto J_1(\tau, \beta, r)$ is strictly increasing in $(0, 1)$. If $r > \rho^*(\beta, \tau)$ then by Remark 8.6

$$J_1(\tau, \beta, r) > J_1(\tau, \beta, \rho^*(\beta, \tau)) \ \geq \ \beta - \frac{1}{2}.$$

Combining this, (8.14), Theorem 7.6(i) and Remark 7.7(iii) shows (i).

Now suppose that the assumptions of (ii) hold, in particular we have $r = \rho^*(\beta, \tau)$. From Corollary 4.12(i), Theorem 5.1(iii) and (iv) we see that $P_{(n)} \lhd\rhd Q_{(n)}$. Thus, by Theorem 7.8 and Remark 7.9(ii) it is sufficient to show that

$$\widetilde{R}_{n,1} := a_n \sqrt{n}\, \varepsilon_n \sup_{v \in [r_n, u_n]} \left\{ \frac{\widetilde{\mu}_n(0, v]}{\sqrt{v}} \right\} \quad \text{and} \quad \widetilde{R}_{n,2} := a_n \sqrt{n}\, \varepsilon_n \sup_{v \in [r_n, u_n]} \left\{ \frac{1 - \widetilde{\mu}_n(0, 1 - v]}{\sqrt{v}} \right\}$$

converge to 0 as $n \to \infty$. Define

$$E_1 \ := \begin{cases} \frac{1}{2} \left(1 - \frac{1}{\tau}\sqrt{1 - \beta}\right) & \text{if (5.2) holds} \\ 0 & \text{otherwise} \end{cases}.$$

In general we have

$$\varepsilon_n \ = \ n^{-\beta} (\log(n))^{E_1} \text{ for all } n \in \mathbb{N}.$$

Combining Lemma 8.3, Lemma 8.5 and Remark 8.6 yields for some sufficiently small $\varepsilon > 0$

$$\begin{aligned} \widetilde{R}_{n,1} &= \ a_n n^{\frac{1}{2} - \beta} (\log(n))^{E_1} \sup_{v \in [r_n, n^{-r-2\varepsilon}) \cup [n^{-r-2\varepsilon}, u_n]} \left\{ \frac{\widetilde{\mu}_n(0, v]}{\sqrt{v}} \right\} \\ &\leq \ \max \left\{ (\log(n))^{E_1} \ R_{n,1}(\varepsilon)\, ,\, n^{\frac{1}{2} - \beta} (\log(n))^{E_1 + o(1)} \ R_{n,2}(\varepsilon) \right\} \\ &\leq \ \max \left\{ n^{-\varepsilon + o(1)}\, ,\, n^{\frac{1}{2} - \beta + J_1(\tau, \beta, \rho^*(\beta, \tau))} (\log(n))^{J_2(\tau, \beta, \rho^*(\beta, \tau)) + E_1 + o(1)} \right\} \\ &= \ (\log(n))^{J_2(\tau, \beta, \rho^*(\beta, \tau)) + E_1 + o(1)} \quad \text{as } n \to \infty. \end{aligned} \tag{8.15}$$

Observe that

$$\begin{aligned} & J_2(\tau, \beta, \rho^*(\beta, \tau)) + E_1 \\ &= \begin{cases} -\frac{1}{8} & \text{if } \beta < 1 - \frac{\tau^2}{4},\, \tau^2 < 2 \\ \frac{11}{2\tau^2} \left[1 - \sqrt{\rho^*(\beta, \tau)}\right] - 3 + \frac{1}{2} - \frac{1}{2\tau}\sqrt{1 - \beta} & \text{otherwise} \end{cases} \\ &= \begin{cases} -\frac{1}{8} & \text{if } \beta < 1 - \frac{\tau^2}{4},\, \tau^2 < 2 \\ 5 \left[\frac{1}{\tau}\sqrt{1 - \beta} - \frac{1}{2}\right] & \text{otherwise} \end{cases}. \end{aligned}$$

It is easy to check that

$$J_2(\tau,\beta,\rho^*(\beta,\tau)) + E_1 \ < \ 0.$$

Combining this and (8.15) yields that $\widetilde{R}_{n,1}$ converges to 0 as $n\to\infty$. Furthermore, by (8.1), the monotonicity and the symmetry of Φ

$$\begin{aligned} 1-\widetilde{\mu}_n(0,1-v] \ &= \ \Phi\left(-\frac{\Phi^{-1}(1-v)+\vartheta_n}{\tau}\right) = 1-\Phi\left(\frac{-\Phi^{-1}(v)+\vartheta_n}{\tau}\right) \\ &\leq \ 1-\Phi\left(-\frac{\Phi^{-1}(v)+\vartheta_n}{\tau}\right) \ = \ \widetilde{\mu}_n(0,v] \end{aligned}$$

for every $v\in(0,1)$. Hence,

$$\widetilde{R}_{n,2} = a_n\sqrt{n}\,\varepsilon_n \sup_{v\in[r_n,u_n]}\left\{\frac{1-\widetilde{\mu}_n(0,1-v]}{\sqrt{v}}\right\} \ \leq \ \widetilde{R}_{n,1} = o(1) \quad \text{as } n\to\infty. \qquad \blacksquare$$

Remark 8.7. Theorem 8.1(i) can also be shown by applying Theorem 8.13. □

8.1.2. Dense case

In this section we focus, again, on the model introduced in Example 2.6 but unlike in the previous section we discuss the dense case, i.e., $\varepsilon_n = n^{-\beta}$ with $\beta\in(0,\frac{1}{2})$. Below, we prove that Theorem 8.1 holds analogously for the dense case. Similar to the sparse case, Cai et al. [10] already showed that HC can completely separate the null and the alternative underneath the detection boundary. We verify additionally that HC cannot separate the null and the alternative on the detection boundary. We want to emphasise that the dense exponential family model, which we will discuss in Section 8.4.1, includes the dense heterogeneous normal mixture model, i.e., the case $\tau=1$. We use this connection to Section 8.4.1 for the proof of the following theorem. Moreover, we refer the reader to Figure 8.2, see p. 142, for a visualisation of the results for the case $\tau=1$.

Theorem 8.8. *Suppose that Assumption 7.2*(v) *holds for the model given in Example 2.6 with* $\varepsilon_n := n^{-\beta}$ *for some* $\beta\in(0,\frac{1}{2})$, $\vartheta_n = n^{-r}$ *for some* $r\in(0,\frac{1}{2})$, $\tau>0$ *and* $T_n = 1-\Phi$ *for all* $n\in\mathbb{N}$. *Moreover, let* ρ^*_{dense} *be defined as in* (5.5).

(i) If $\tau=1$ *and* $r<\rho^*_{dense}(\beta,\tau)$ *then* (7.18) *holds. In this case HC can completely separate* $\mathcal{H}_{0,n}$ *and* $\mathcal{H}_{1,n}$ *(asymptotically), see Remark 7.7*(i).

(ii) Consider $r = \rho^*_{dense}(\beta, \tau)$. Then (7.29) *holds for all $t \in \mathbb{R}$ if* (7.3) *is fulfilled, and* (7.31) *holds for all $t \in \mathbb{R}$ if* (7.4) *is fulfilled. In this case HC cannot successfully separate $\mathcal{H}_{0,n}$ and $\mathcal{H}_{1,n}$ (asymptotically), see Remark 7.9*(i).

(iii) If $\tau > 1$ then (7.18) *holds. In this case HC can completely separate $\mathcal{H}_{0,n}$ and $\mathcal{H}_{1,n}$ (asymptotically), see Remark 7.7*(i).

(iv) If $\tau < 1$ and $j = 2$ then (7.18) *holds. In this case HC can completely separate $\mathcal{H}_{0,n}$ and $\mathcal{H}_{1,n}$ (asymptotically), see Remark 7.7*(i).

Proof. In Section 8.4.1 we discuss the dense case for exponential family models including the heterogeneous normal mixture model, i.e., the case $\tau = 1$. The results, which we prove there, can be used to show (i) and (ii), see Example 8.18 for details. It remains to verify (iii) and (iv). Let $\tau \neq 1$. We can assume without loss of generality that $u_n = o(1)$ as $n \to \infty$, see Remark 7.3(i). By Remark 8.2(ii) and Lemma A.29 we have for some $c > 0$

$$\begin{aligned}
\widetilde{\mu}_n(0, u_n] &= 1 - \Phi\left(-\frac{\Phi^{-1}(u_n) + \vartheta_n}{\tau}\right) \\
&\sim_{\text{asy}} \frac{-\tau}{\Phi^{-1}(u_n) + \vartheta_n} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{\Phi^{-1}(u_n) + \vartheta_n}{\tau}\right)^2\right) \\
&\sim_{\text{asy}} \frac{\tau}{\sqrt{2\pi}} \left(-2\log(u_n)\right)^{-\frac{1}{2}} \exp\left(-\frac{1}{2\tau^2}\left(-\sqrt{-2\log(u_n)} + \frac{\log(4\pi) + \log\log\left(u_n^{-1}\right) + o(1)}{2\sqrt{-2\log(u_n)}}\right)^2\right) \\
&= \frac{\tau}{2\sqrt{\pi}} \left(-\log(u_n)\right)^{-\frac{1}{2}} \exp\left(-\frac{1}{2\tau^2}\left(-2\log(u_n) - \log(4\pi) - \log\log\left(u_n^{-1}\right) + o(1)\right)\right) \\
&\sim_{\text{asy}} c\left(-\log(u_n)\right)^{-\frac{1}{2} + \frac{1}{2\tau^2}} u_n^{\frac{1}{\tau^2}} = n^{\lambda_n\left(-\frac{1}{\tau^2} + o(1)\right)} \quad \text{as } n \to \infty.
\end{aligned}$$

From this and $\beta < \frac{1}{2}$ we deduce that

$$\frac{n\varepsilon_n}{\sqrt{n\, u_n \log\log n}}\left(\widetilde{\mu}_n(0, u_n] - u_n\right) = n^{\frac{1}{2} - \beta + o(1)}\left(n^{-\lambda_n \tau^{-2} + o(\lambda_n)} - n^{-\lambda_n}\right) \to \begin{cases} \infty & \text{if } \tau > 1 \\ -\infty & \text{if } \tau < 1 \end{cases}$$

as $n \to \infty$. Finally, we can conclude (iii) and (iv) from Theorem 7.6(i). ∎

8.2. The h-model

In Section 5.3 we discussed the asymptotic behaviour of LLRT for the h-model introduced in Section 2.4. We determined, among others, the detection boundary for the case that

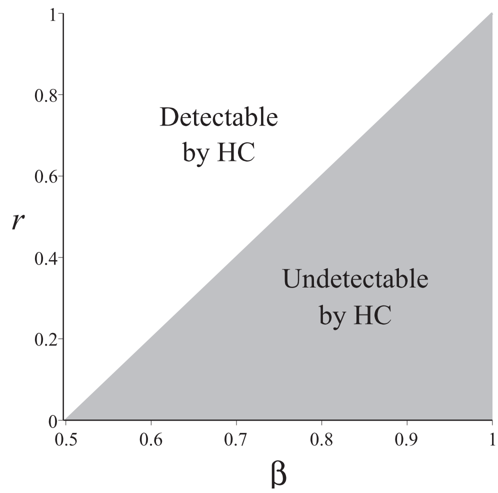

Figure 8.1.: The detectable and the undetectable area of HC are visualised for the h-model. The boundary, which splits the plane into these two areas, belongs to the undetectable area.

neither $\tau_{n,i}$ nor $\varepsilon_{n,i}$ depends on i, see Corollary 5.18 and Figure 5.3. In the following we prove that the detectable areas of LLRT and HC coincide. Under some additional assumptions we also show that HC cannot successfully separate the null and the alternative on the detection boundary (asymptotically). The results are visualised in Figure 8.1.

Theorem 8.9. *Suppose that Assumption 7.2*(v) *and the conditions of Corollary 5.18 hold simultaneously with* $T_n := T_n^{-1} := id_{(0,1)}$ *for every* $n \in \mathbb{N}$.

(i) If $\beta < \beta_h^{\#}(r)$ *then* (7.18) *holds. In this case HC can completely separate* $\mathcal{H}_{0,n}$ *and* $\mathcal{H}_{1,n}$ *(asymptotically), see Remark 7.7*(i).

(ii) Suppose $\beta = \beta_h^{\#}(r)$ *and*

$$\int_0^1 h^{2+\kappa}\,\mathrm{d}\lambda\!\!\lambda \;<\; \infty \text{ for some } \kappa > 0. \tag{8.16}$$

Then (7.29) *holds for all* $t \in \mathbb{R}$ *if* (7.3) *is fulfilled, and* (7.31) *does so if* (7.4) *holds. HC cannot successfully separate* $\mathcal{H}_{0,n}$ *and* $\mathcal{H}_{1,n}$ *(asymptotically), see Remark 7.9*(i).

Proof. Due to Remark 7.3(i) we can assume without loss of generality that

$$n^{-\lambda_n}\log\log(n) \;=\; o(1) \quad \text{as } n \to \infty. \tag{8.17}$$

Substituting $u = \tau_n y$ we deduce from (2.18) and (2.19) of Assumption 2.26 that

$$\mu_n(0, v_n] \quad = v_n - c_1 \lambda\!\!\lambda\Big((0, v_n] \cap [n^{-r}, 1)\Big) + (1 - n^{-r}) \int_{(0, \max\{1, v_n n^r\})} h \, \mathrm{d}\lambda\!\!\lambda, \tag{8.18}$$

$$\mu_n(1 - v_n, 1) = v_n - c_1 \lambda\!\!\lambda\,(1 - v_n, 1) + 0 = (1 - c_1) v_n \tag{8.19}$$

for all $v_n \in (0, \frac{1}{2})$ and all sufficiently large $n \in \mathbb{N}$.

First, we verify (i). Assume $\beta < \beta_h^{\#}(r)$. Set $v_n := n^{-r}$ for all $n \in \mathbb{N}$. By (8.18)

$$\frac{\sqrt{n}\varepsilon_n \mu_n(0, v_n]}{\sqrt{v_n(1 - v_n) \log\log n}} \geq n^{\frac{1}{2} - \beta + \frac{1}{2} r + o(1)} \int_{(0,1)} h \, \mathrm{d}\lambda\!\!\lambda = n^{\beta_h^{\#}(r) - \beta + o(1)} \, c_1$$

as $n \to \infty$. Consequently, we conclude (i) from Theorem 7.6(i) and Remark 7.7(iii). Now we verify (ii). Suppose that the corresponding conditions are fulfilled. By (8.19)

$$a_n \sqrt{n} \varepsilon_n \sup_{v \in [r_n, u_n]} \left\{ \frac{\mu_n(1 - v, 1)}{\sqrt{v}} \right\} \;=\; a_n n^{\frac{1}{2} - \beta} (1 - c_1) \frac{1}{\sqrt{u_n}} \;=\; o(1)$$

as $n \to \infty$. Combining this and Remark 7.9(ii) yields (7.30). In the following we show that (7.28) holds as well. If $r = 1 = \beta$ then by (8.17)

$$a_n \sqrt{n} \varepsilon_n \sup_{v \in [r_n, u_n]} \left\{ v^{-\frac{1}{2}} \mu_n(0, v] \right\} \;\leq\; a_n n^{\frac{1}{2} - 1} r_n^{-\frac{1}{2}} \;=\; \sqrt{2 n^{-\lambda_n} \log\log(n)} \;=\; o(1) \quad \text{as } n \to \infty.$$

We can conclude from this and Remark 7.9(ii) that (7.28) is fulfilled if $r = 1$. It remains to discuss the case $r < 1$. As mentioned in Remark 7.3(i) we can freely choose $\rho, \rho_{n,1}, \rho_{n,2}$ from Assumption 7.2(i). Suppose that

$$\rho = r, \;\; a_n n^{\left(\frac{1}{2} - \frac{1}{2+\kappa}\right) \rho_{n,2}} = o(1) \text{ and } a_n \, n^{-\frac{1}{2} \rho_{n,1}} = o(1) \quad \text{as } n \to \infty. \tag{8.20}$$

By Hölder's inequality and (8.16) there exists some $c_0 > 0$ such that

$$\int_0^x h \, \mathrm{d}\lambda\!\!\lambda \leq \left(\int_0^1 h^{2+\kappa} \, \mathrm{d}\lambda\!\!\lambda \right)^{\frac{1}{2+\kappa}} \left(\int_0^x \mathrm{d}\lambda\!\!\lambda \right)^{1 - \frac{1}{2+\kappa}} \leq c_0 \, x^{1 - \frac{1}{2+\kappa}} \;\text{ for all } x \in (0, 1]. \tag{8.21}$$

By (8.20) $\rho_{n,2}$ is negative for all sufficiently large $n \in \mathbb{N}$. Hence, $s_n = n^{-\rho + \rho_{n,2}} < n^{-r}$ for all sufficiently large $n \in \mathbb{N}$. Combining this, (8.18), (8.20) and (8.21) shows that

$$\begin{aligned}
a_n \sqrt{n} \varepsilon_n \sup_{v \in [r_n, s_n]} \left\{ v^{-\frac{1}{2}} \left| \mu_n(0, v] - v \right| \right\} \;&=\; a_n \, n^{\frac{1}{2} - \beta} (1 - n^{-r}) \sup_{v \in [r_n, s_n]} \left\{ v^{-\frac{1}{2}} \int_{(0, v n^r)} h \, \mathrm{d}\lambda\!\!\lambda \right\} \\
&\leq\; a_n \, n^{\frac{1}{2} - \beta} c_0 \sup_{v \in [r_n, s_n]} v^{\frac{1}{2} - \frac{1}{2+\kappa}} \, n^{r\left(1 - \frac{1}{2+\kappa}\right)} \\
&\leq\; c_0 \, a_n n^{\frac{1}{2} - \beta + \left(\frac{1}{2} - \frac{1}{2+\kappa}\right)(-r + \rho_{n,2}) + r\left(1 - \frac{1}{2+\kappa}\right)} \\
&\leq\; c_0 \, a_n n^{\left(\frac{1}{2} - \frac{1}{2+\kappa}\right) \rho_{n,2}} \;=\; o(1) \text{ as } n \to \infty.
\end{aligned}$$

Consequently, we obtain (7.28) since by (8.20)

$$a_n \sqrt{n} \varepsilon_n \sup_{v \in [t_n, u_n]} \left\{ v^{-\frac{1}{2}} |\mu_n(0,v] - v| \right\} \leq a_n n^{\frac{1}{2}-1} t_n^{-\frac{1}{2}} = a_n n^{-\frac{1}{2}\rho_{n,1}} = o(1) \quad \text{as } n \to \infty.$$

By Corollaries 4.12(i) and 5.18 $P_{(n)} \triangleleft \triangleright Q_{(n)}$. Hence (ii) follows from Theorem 7.8. ■

8.3. Extensions of the results of Cai and Wu

Cai and Wu [12] showed that under the assumptions of their Theorem 1, treating the case $P_{n,i} = N(0,1)$, the detectable areas of HC and LLRT coincide, see their Theorem 4. It was an unsolved problem if these areas also coincide under the assumptions of their Theorem 3, treating general $P_{n,i} = P_n$. In this section we show that it is still valid, even for our extension. At the end we also present an extension of their Theorem 4.

We begin by presenting the result concerning HC corresponding to our Theorem 5.23.

Theorem 8.10. *(i) Let Assumption 7.2*(iii), (7.4) *and the assumptions of Theorem 5.23*(ii) *be fulfilled simultaneously. Then* (7.18) *holds. Hence, HC can completely separate the null $\mathcal{H}_{0,n}$ and the alternative $\mathcal{H}_{1,n}$ (asymptotically), see Remark 7.7*(i).

(ii) Suppose that Assumption 7.2(iii), (7.3) *and the assumptions of Theorem 5.23 hold simultaneously. If* (5.54) *holds for some $\delta > 0$ and*

$$\liminf_{n\to\infty} \min_{1 \leq i \leq k_n} \lambda\!\!\lambda \left(s \in \left(\frac{\log(2)}{\log(k_n)}, 1 \right) : \beta^{\#} - \frac{1}{2} \leq \frac{h_{k_n,1,i}(s)}{\log(k_n)} - s + \frac{s \wedge 1}{2} \right) > 0 \quad (8.22)$$

then (7.18) *holds. Hence, HC can completely separate the null $\mathcal{H}_{0,n}$ and the alternative $\mathcal{H}_{1,n}$ (asymptotically), see Remark 7.7*(i).

Remark 8.11. (i) Note that there is a difference between assuming $T_{n,i} = F_{n,i}$ and assuming $T_{n,i} = 1 - F_{n,i}$ in Theorem 8.10(ii). That is one of the reason why we use a general mapping $T_{n,i}$ instead of $F_{n,i}$ or $1 - F_{n,i}$, respectively. We already mentioned this issue in Remark 5.24(i).

(ii) If (7.3) holds then the values, which are near to 1, are excluded. That is why the assumption (8.22) of (ii) only depends on $h_{k_n,1,i}$ and not on $h_{k_n,2,i}$. □

Proof. We only give the proof of (i). The one of (ii) runs analogously. Let $\tau_n \in \{\lambda_{k_n}, 1-\lambda_{k_n}\}$ for all $n \in \mathbb{N}$. By (5.54), a similar calculation as the one in (5.69), Theorem 7.6 and Remark 7.7(iii) it remains to prove that

$$\begin{aligned}
&\frac{1}{\sqrt{k_n\, k_n^{-\tau_n} \log\log(k_n)}} \sum_{i=1}^{k_n} \varepsilon_{k_n,i}\, \mu_{k_n,i}^{T_{k_n,i}}(0, k_n^{-\tau_n}] \\
&= \frac{1}{\sqrt{\log\log(k_n)}}\, k_n^{\frac{1}{2}(\tau_n - 1)} \sum_{i=1}^{k_n} \varepsilon_{k_n,i} \int_0^{k_n^{-\tau_n}} \exp\left[l_{k_n,i}\left(T_{k_n,i}^{-1}(v)\right)\right] \,\mathrm{d}v
\end{aligned} \tag{8.23}$$

or

$$\begin{aligned}
&\frac{1}{\sqrt{k_n\, k_n^{-\tau_n} \log\log(k_n)}} \sum_{i=1}^{k_n} \varepsilon_{k_n,i} \left(1 - \mu_{k_n,i}^{T_{k_n,i}}(0, 1 - k_n^{-\tau_n}]\right) \\
&= \frac{1}{\sqrt{\log\log(k_n)}} k_n^{\frac{1}{2}(\tau_n - 1)} \sum_{i=1}^{k_n} \varepsilon_{k_n,i}\, \mu_{k_n,i}^{1-T_{k_n,i}}(0, k_n^{-\tau_n}] \\
&= \frac{1}{\sqrt{\log\log(k_n)}} k_n^{\frac{1}{2}(\tau_n - 1)} \sum_{i=1}^{k_n} \varepsilon_{k_n,i} \int_0^{k_n^{-\tau_n}} \exp\left[l_{k_n,i}\left(T_{k_n,i}^{-1}(1-v)\right)\right] \,\mathrm{d}v
\end{aligned}$$

converges to infinity as $n \to \infty$. Because both terms are positive this holds if and only if the sum of both converges to infinity as $n \to \infty$. By using the substitution $v = k_n^{-s}$ the sum of both terms can be lower bounded for sufficiently large $n \in \mathbb{N}$ in the following way:

$$\begin{aligned}
S_n &:= \frac{k_n^{\frac{1}{2}(\tau_n - 1)}}{\sqrt{\log\log(k_n)}} \sum_{i=1}^{k_n} \varepsilon_{k_n,i} \int_0^{k_n^{-\tau_n}} \exp\left[l_{k_n,i}\left(T_{k_n,i}^{-1}(v)\right)\right] + \exp\left[l_{k_n,i}\left(T_{k_n,i}^{-1}(1-v)\right)\right] \,\mathrm{d}v \\
&= \frac{1}{\sqrt{\log\log(k_n)}} k_n^{\frac{1}{2}(\tau_n - 1)} \sum_{i=1}^{k_n} \varepsilon_{k_n,i} \int_{\tau_n}^{\infty} \sum_{m=1,2} \exp\Big[h_{k_n,m,i}(s)\Big] \log(k_n)\; k_n^{-s} \,\mathrm{d}s \\
&\geq \frac{\log(k_n)}{\sqrt{\log\log(k_n)}} \sum_{i=1}^{k_n} \varepsilon_{k_n,i} \int_{\tau_n}^{\infty} \exp\left[\log(k_n)\left(\frac{h_{k_n,i}(s)}{\log(k_n)} - s - \frac{1}{2} + \frac{\tau_n}{2}\right)\right] \mathrm{d}s.
\end{aligned} \tag{8.24}$$

By (5.55) there exists some $\kappa > 0$ such that for every sufficiently large $n \in \mathbb{N}$

$$\min_{1 \le i \le k_n} \left\{ \lambda\!\!\lambda\left(s \in (1, \infty) : \beta^{\#} - 1 + s \le \frac{h_{k_n,i}(s)}{\log(k_n)}\right)\right\} \geq \kappa \tag{8.25}$$

$$\text{or} \quad \min_{1 \le i \le k_n} \left\{ \lambda\!\!\lambda\left(s \in \left(\max\left\{\frac{\log(2)}{\log(k_n)}, \lambda_{k_n}\right\}, 1\right) : \beta^{\#} + \frac{s}{2} - \frac{1}{2} \le \frac{h_{k_n,i}(s)}{\log(k_n)}\right)\right\} \geq \kappa. \tag{8.26}$$

Let $\tau_n = 1 - \lambda_{k_n}$ for all $n \in \mathbb{N}$. From (5.54) and (8.24) we deduce that for all sufficiently large $n \in \mathbb{N}$, for which (8.25) holds,

$$S_n \geq \frac{\log(k_n)}{\sqrt{\log\log(k_n)}}\, k_n^{1-\beta^\# + \delta} \kappa \exp\left[\log(k_n)\left(\beta^\# - 1 - \frac{\lambda_{k_n}}{2}\right)\right] \geq k_n^{\delta - \frac{\lambda_{k_n}}{2}}.$$

Now let $\tau_n = \lambda_{k_n}$ for all $n \in \mathbb{N}$. Similar to the calculation above we conclude that for all sufficiently large $n \in \mathbb{N}$, for which (8.26) holds,

$$S_n \geq \kappa \frac{\log(k_n)}{\sqrt{\log\log(k_n)}}\, k_n^{\delta + \frac{\lambda_{k_n}}{2}}.$$

To sum up,

$$S_n \geq k_n^{\frac{\delta}{2}} \to \infty \text{ as } n \to \infty. \qquad \blacksquare$$

Combining Corollary 5.25 and Theorem 8.10 yields the following corollary.

Corollary 8.12. *Suppose that Assumption 7.2*(iii)*,* (7.4) *and the assumptions of Corollary 5.25 hold simultaneously.*

(i) If β exceeds $\beta^\#$ then LLRT and HC cannot successfully separate the null $\mathcal{H}_{0,n}$ and the alternative $\mathcal{H}_{1,n}$ (asymptotically).

(ii) If β is smaller than $\beta^\#$ then LLRT and HC can completely separate the null $\mathcal{H}_{0,n}$ and the alternative $\mathcal{H}_{1,n}$ (asymptotically).

Remember that Theorem 2 in [12] corresponds to Corollary 5.26 and is a special case of Corollary 5.25. Hence, we can now give the answer to the unanswered question mentioned in the introduction: the detection areas of HC and LLRT do coincide under the assumptions of Theorem 3 in [12]. Finally, we present the extension of Theorem 1 in [12]. The proof is very similar to the one of Theorem 8.10 and, thus, we omit it.

Theorem 8.13 (Extension of Theorem 4 in [12]). *(i) Let Assumption 7.2*(iii)*,* (7.4) *and the assumptions of Theorem 5.27*(ii) *be fulfilled simultaneously. Then* (7.18) *holds and HC can completely separate $\mathcal{H}_{0,n}$ and $\mathcal{H}_{1,n}$ (asymptotically).*

(ii) Suppose that Assumption 7.2(iii)*,* (7.3)*, the assumptions of Theorem 5.27 and* (5.65) *for some $\delta > 0$ hold simultaneously. Assume additionally that*

$$\liminf_{n\to\infty} \min_{1\leq i\leq k_n} \left\{ \lambda\!\!\lambda\left(x \in (-1,1) : \beta^\# - \frac{1}{2} \leq \frac{\widetilde{h}_{k_n,1,i}(x)}{\log(k_n)} - x^2 + \frac{x^2 \wedge 1}{2} \right) \right\} > 0.$$

Then (7.18) *holds and HC can completely separate $\mathcal{H}_{0,n}$ and $\mathcal{H}_{1,n}$ (asymptotically).*

8.4. Exponential families

8.4.1. Dense case

In this section we discuss exponential families for the dense case. In Section 5.2.2 we determined the asymptotic behaviour of LLRT for this case. We verified that it is uniquely determined by B_{k_n} as defined in (5.18). Below, we show that in the case $\varepsilon_{k_n,i} = \varepsilon_{k_n}$ the asymptotic behaviour of HC is uniquely determined by B_{k_n} as well under some additional assumptions. We will show at the end of this section that these additional assumptions are fulfilled for the heterogeneous normal mixture model, see Section 8.1.2, and for the exponential distribution mixture model introduced in Example 2.10. Due to Remark 5.9 we can also make use of the results to determine the asymptotic behaviour of HC for the Gumbel and Fréchet distribution mixture model introduced in Example 2.8 and 2.9. In all above-mentioned examples the detectable areas of LLRT and HC do coincide. We will show for these examples that HC cannot successfully separate the null and the alternative on the detection boundary (asymptotically). The results concerning the simple case $\varepsilon_{n,i} = \varepsilon_n = n^{-\beta}$, see Corollary 8.16, are visualised in the following section, see Figure 8.2.

Theorem 8.14. *Suppose that Assumption 2.23 and Assumption 7.2*(iii) *hold simultaneously, where $T_n = T_1 =: T$ and $T_n^{-1} = T_1^{-1} =: T^{-1}$ for all $n \in \mathbb{N}$.*

(i) Assume that

$$\begin{aligned} \sum_{i=1}^{k_n} \varepsilon_{k_n,i}\vartheta_{k_n} &= k_n^{\kappa+o(1)} \text{ for some } \kappa > \frac{1}{2} \text{ as } n \to \infty \\ \text{and } \vartheta_{k_n}^2 k_n^{\frac{1}{4}\lambda_{k_n}} &= o(1) \text{ as } n \to \infty. \end{aligned} \tag{8.27}$$

Moreover, suppose that

$$\limsup_{v \searrow 0} \left(\omega^{(1)}(0) + \frac{1}{v}\int_0^v h(T^{-1}(y))\,\mathrm{d}y \right) < 0. \tag{8.28}$$

Then (7.18) *holds. Hence, HC can completely separate $\mathcal{H}_{0,n}$ and $\mathcal{H}_{1,n}$ (asymptotically), see Remark 7.7*(i).

(ii) Assume that (8.27) *is fulfilled and*

$$\sum_{i=1}^{k_n} \varepsilon_{k_n,i}\vartheta_{k_n} = -k_n^{\kappa+o(1)} \text{ for some } \kappa > \frac{1}{2} \text{ as } n \to \infty.$$

Moreover, suppose that

$$\liminf_{v\searrow 0}\left(\omega^{(1)}(0)+\frac{1}{v}\int_0^v h(T^{-1}(y))\,\mathrm{d}y\right)>0.$$

Then (7.18) *holds and HC can completely separate* $\mathcal{H}_{0,n}$ *and* $\mathcal{H}_{1,n}$ *(asymptotically).*

(iii) Suppose that $j=2$*. Additionally, assume that* (8.27) *is fulfilled and*

$$\sum_{i=1}^{k_n}\varepsilon_{k_n,i}\,|\vartheta_{k_n}| \;=\; k_n^{\kappa+o(1)} \text{ for some } \kappa>\frac{1}{2} \text{ as } n\to\infty.$$

Moreover, suppose that

$$\liminf_{v\searrow 0}\left|\omega^{(1)}(0)+\frac{1}{v}\int_0^v h(T^{-1}(y))\,\mathrm{d}y\right|>0.$$

Then (7.18) *holds and HC can completely separate* $\mathcal{H}_{0,n}$ *and* $\mathcal{H}_{1,n}$ *(asymptotically).*

(iv) Suppose that $\varepsilon_{n,i}=\varepsilon_{n,1}=:\varepsilon_n$ *for all* $i\in\{1,\dots,n\}$ *and that* B_{k_n} *given by* (5.18) *converges to a positive constant as* $n\to\infty$*. Then* (7.29) *holds for all* $t\in\mathbb{R}$ *if* (7.3) *is fulfilled, and* (7.31) *holds for all* $t\in\mathbb{R}$ *if* (7.4) *is fulfilled. Thus, HC cannot successfully separate* $\mathcal{H}_{0,n}$ *and* $\mathcal{H}_{1,n}$ *(asymptotically), see Remark 7.9*(i).

Remark 8.15. If

$$\left(\frac{1}{n},\alpha_0\right)\subseteq I_n \text{ for all } n\in\mathbb{N} \text{ or } \left(\frac{1}{n},1-\frac{1}{n}\right)\subseteq I_n \text{ for all } n\in\mathbb{N}$$

then we can choose $\lambda_n=\frac{1}{\log n}$ for all $n\in\mathbb{N}$ and so (8.27) is fulfilled. Hence, (8.27) is not an actual restriction in practice. □

Before we give the proof of Theorem 8.14 we present the following immediate consequence of Theorem 8.14 in the context of the detection boundary discussed in Corollary 5.13.

Corollary 8.16. *Suppose that Assumption 7.2*(v) *and the assumptions of Corollary 5.13 hold simultaneously, where* $T_n=T_1=:T$ *and* $T_n^{-1}=T_1^{-1}=:T^{-1}$ *for all* $n\in\mathbb{N}$*. Moreover, assume* (8.28)*.*

(i) Suppose $r=\rho^*_{Exp,d}(\beta)$*. Then* (7.29) *holds for all* $t\in\mathbb{R}$ *if* (7.3) *is fulfilled, and* (7.31) *holds for all* $t\in\mathbb{R}$ *if* (7.4) *is fulfilled. Thus, HC cannot successfully separate the null and the alternative (asymptotically), see Remark 7.9*(i).

(ii) If $r>\rho^*_{Exp,d}(\beta)$ *then* (7.18) *holds. Hence, HC can completely separate* $\mathcal{H}_{0,n}$ *and* $\mathcal{H}_{1,n}$ *(asymptotically), see Remark 7.7*(i).

Proof of Theorem 8.14. Due to Remark 7.1(ii) there is no loss of generality in assuming $\mathcal{Q}_0 = \mathcal{U}_{(0,1)}$ and $T = id_{(0,1)} : (0,1) \to (0,1)$. Regarding Theorem 7.6 and Theorem 7.8 we first determine

$$\begin{aligned} \mu_{k_n,i}(0,\tau_n] - \tau_n &= \int_0^{\tau_n} C(\vartheta_{k_n}) \exp\left[-\vartheta_{k_n} h(x)\right] - 1 \,\mathrm{d}x \\ \text{and} \quad \mu_{k_n,i}(0,1-\tau_n] - (1-\tau_n) &= \tau_n - \mu_{k_n,i}\left((1-\tau_n,1]\right) \qquad (8.29) \\ &= -\int_{1-\tau_n}^{1} \left(C(\vartheta_{k_n}) \exp\left[-\vartheta_{k_n} h(x)\right] - 1\right) \,\mathrm{d}x \qquad (8.30) \end{aligned}$$

for certain $\tau_n \in [0,1]$ and all $n \in \mathbb{N}$. Note that we introduced $\varepsilon > 0$ in Assumption 2.23. Define the function $\chi_x : (-\varepsilon, \varepsilon) \to \mathbb{R}$ for all $x \in [0,1]$ by

$$\chi_x(\vartheta) := C(\vartheta) \exp\left[-\vartheta h(x)\right] - 1 \text{ for all } \vartheta \in (-\varepsilon, \varepsilon).$$

Since $\omega(\vartheta) > 0$ for all $\vartheta \in \Theta$, by Lemma 2.24 the function $\vartheta \mapsto C(\vartheta) = \omega(\vartheta)^{-1}$ is indefinitely often differentiable in $(-\varepsilon, \varepsilon)$ and so is χ_x for all $x \in [0,1]$. Note that

$$\begin{aligned} \chi_x^{(1)}(\vartheta) &= C^{(1)}(\vartheta) \exp(-\vartheta h(x)) - C(\vartheta) h(x) \exp(-\vartheta h(x)) \\ &= \left[-C(\vartheta)^2 \omega^{(1)}(\vartheta) - C(\vartheta) h(x)\right] \exp(-\vartheta h(x)) \qquad (8.31) \\ \text{and} \quad \chi_x^{(2)}(\vartheta) &= \left[C^{(2)}(\vartheta) - 2C^{(1)}(\vartheta) h(x) + C(\vartheta) h^2(x)\right] \exp(-\vartheta h(x)) \qquad (8.32) \end{aligned}$$

for all $\vartheta \in (-\varepsilon, \varepsilon)$ and every $x \in [0,1]$. By Taylor's theorem, (8.31) and (8.32) there exists some function $r_n : [0,1] \to \mathbb{R}$ for all sufficiently large $n \in \mathbb{N}$ such that for all $x \in [0,1]$

$$\begin{aligned} |r_n(x)| &\leq |\vartheta_{k_n}| \leq \frac{\varepsilon}{8} \qquad (8.33) \\ \text{and} \quad \chi_x(\vartheta_{k_n}) &= \vartheta_{k_n}\left[-\omega^{(1)}(0) - h(x)\right] + \frac{\vartheta_{k_n}^2}{2} \chi_x^{(2)}\left(r_n(x)\right). \qquad (8.34) \end{aligned}$$

Thus, by (8.29) and (8.30) we have for all $\tau \in (0,1)$ and every sufficiently large $n \in \mathbb{N}$

$$\mu_{k_n}(0,\tau] - \tau = -\vartheta_{k_n} \tau\, \omega^{(1)}(0) - \vartheta_{k_n} \int_0^{\tau} h(x) \,\mathrm{d}x + \frac{\vartheta_{k_n}^2}{2} \int_0^{\tau} \chi_x^{(2)}\left(r_n(x)\right) \,\mathrm{d}x \qquad (8.35)$$

$$\begin{aligned} \text{and} \quad &\mu_{k_n}(0,1-\tau] - (1-\tau) \\ &= \vartheta_{k_n} \tau\, \omega^{(1)}(0) + \vartheta_{k_n} \int_{1-\tau}^{1} h(x) \,\mathrm{d}x - \frac{\vartheta_{k_n}^2}{2} \int_{1-\tau}^{1} \chi_x^{(2)}\left(r_n(x)\right) \,\mathrm{d}x. \qquad (8.36) \end{aligned}$$

Now we examine the asymptotic behaviour of the second integral from (8.35) and (8.36), respectively. Since ω and C are indefinitely often differentiable in $(-\varepsilon, \varepsilon)$ there exists some

$M \geq 1$ such that

$$\left|\omega^{(k)}(\vartheta)\right| + \left|C^{(k)}(\vartheta)\right| \;\leq\; M \quad \text{for all } \vartheta \in \left[-\frac{\varepsilon}{2}, \frac{\varepsilon}{2}\right] \text{ and } k \in \{0, 1, \ldots, 8\}. \tag{8.37}$$

Let

either $B_\tau := (0, \tau)$ for all $\tau \in (0,1)$ or $B_\tau := (1-\tau, 1)$ for all $\tau \in (0,1)$.

Combining Lemma 2.24, (8.37) and Hölder's inequality shows for all $\tau \in (0,1)$:

$$\begin{aligned}\left|\int_0^1 h(x)\mathbf{1}_{B_\tau}(x)\,\mathrm{d}x\right| \;&\leq\; \left[\int_0^1 h(x)^4\,\mathrm{d}x\right]^{\frac{1}{4}} \left[\int_0^1 \mathbf{1}_{B_\tau}(x)\,\mathrm{d}x\right]^{\frac{3}{4}} \\ &=\; \tau^{\frac{3}{4}}\left|\omega^{(4)}(0)\right|^{\frac{1}{4}} \;\leq\; M\,\tau^{\frac{3}{4}}.\end{aligned} \tag{8.38}$$

We can conclude from (8.32), (8.33), (8.37) and Hölder's inequality that for all $\tau \in (0,1)$ and every sufficiently large $n \in \mathbb{N}$

$$\begin{aligned}\left|\int_{B_\tau} \chi_x^{(2)}\left(r_n(x)\right)\,\mathrm{d}x\right| \;&\leq\; M\sum_{k=0}^{2}\int_0^1 \mathbf{1}_{B_\tau}(x)\left|h(x)\right|^k e^{-r_n(x)h(x)}\,\mathrm{d}x \\ &\leq\; M\sum_{k=0}^{2}\left[\int_0^1 h^{4k}(x)e^{-4r_n(x)h(x)}\,\mathrm{d}x\right]^{\frac{1}{4}}\left[\int_0^1 \mathbf{1}_{B_\tau}(x)\,\mathrm{d}x\right]^{\frac{3}{4}} \\ &\leq\; M\,\tau^{\frac{3}{4}}\sum_{k=0}^{2}\left[\int_0^1 h^{4k}(x)\left(e^{-4\vartheta_{k_n}h(x)} + e^{4\vartheta_{k_n}h(x)}\right)\,\mathrm{d}x\right]^{\frac{1}{4}} \\ &=\; M\tau^{\frac{3}{4}}\sum_{k=0}^{2}\left[\left|\omega^{(4k)}(4\vartheta_{k_n})\right| + \left|\omega^{(4k)}(-4\vartheta_{k_n})\right|\right]^{\frac{1}{4}} \;\leq\; 6M^2\tau^{\frac{3}{4}}.\end{aligned} \tag{8.39}$$

Let $v_n := u_n = n^{-\lambda_n}$ for all $n \in \mathbb{N}$. Obviously, (7.15) holds. Furthermore, if (8.27) is fulfilled then by (8.35) and (8.39)

$$\begin{aligned}&\frac{1}{\sqrt{k_n\,v_{k_n}\log\log\left(k_n\right)}}\sum_{i=1}^{k_n}\varepsilon_{k_n,i}\Big(\mu_{k_n}\left(0, v_{k_n}\right] - v_{k_n}\Big) \\ &= -\sqrt{\frac{v_{k_n}}{k_n\,\log\log\left(k_n\right)}}\sum_{i=1}^{k_n}\varepsilon_{k_n,i}\vartheta_{k_n}\left(\omega^{(1)}(0) + \frac{1}{v_{k_n}}\int_0^{v_{k_n}} h(y)\,\mathrm{d}y + O\left(\vartheta_{k_n}^2\left(v_{k_n}\right)^{-\frac{1}{4}}\right)\right) \\ &= -k_n^{-\frac{1}{2}+o(1)}\sum_{i=1}^{k_n}\varepsilon_{k_n,i}\vartheta_{k_n}\left(\omega^{(1)}(0) + \frac{1}{v_{k_n}}\int_0^{v_{k_n}} h(y)\,\mathrm{d}y + o(1)\right) \text{ as } n \to \infty.\end{aligned}$$

Hence, applying Theorem 7.6(i) proves (i)-(iii).

Now suppose that the assumptions of (iv) hold. By Remark 7.3(i) we can assume without loss of generality that

$$a_n\, n^{-\frac{1}{4}\lambda_n} \;\to\; 0 \text{ as } n \to \infty. \tag{8.40}$$

Since $M \geq 1$ we deduce from (8.35)-(8.40) that

$$\begin{aligned}
& a_{k_n} \sqrt{k_n}\, \varepsilon_{k_n} \sup_{\tau \in [r_{k_n}, s_{k_n}] \cup [t_{k_n}, u_{k_n}]} \left\{ \frac{|\mu_{k_n}(0,\tau] - \tau|}{\sqrt{\tau}} + \frac{|\mu_{k_n}(0,1-\tau] - (1-\tau)|}{\sqrt{\tau}} \right\} \\
\leq\;& 2a_{k_n} \sqrt{k_n}\, \varepsilon_{k_n} \sup_{\tau \in [r_{k_n}, s_{k_n}] \cup [t_{k_n}, u_{k_n}]} \left\{ |\vartheta_{k_n}| \sqrt{\tau} M + |\vartheta_{k_n}|\, M \tau^{\frac{1}{4}} + \vartheta_{k_n}^2 3 M^2 \tau^{\frac{1}{4}} \right\} \\
\leq\;& 10\, a_{k_n} M^2 \sqrt{k_n}\, \varepsilon_{k_n}\, |\vartheta_{k_n}|\; u_{k_n}^{\frac{1}{4}} \\
=\;& 10\, M^2 \sqrt{B_{k_n}}\, a_{k_n}\, k_n^{-\frac{1}{4}\lambda_{k_n}} \;\to\; 0 \quad \text{as } n \to \infty.
\end{aligned}$$

Thus, (7.28) and (7.30) of Theorem 7.8 are fulfilled. Note that by Corollary 4.12(i) and Theorem 5.11(i) we have $P_{(k_n)} \lhd\rhd Q_{(k_n)}$. Hence, (iv) follows from Theorem 7.8. ■

The above-mentioned results can be applied, e.g., to the heterogeneous normal mixture model discussed in Section 8.1.2, see Example 8.18, and to the exponential distribution mixture model introduced in Example 2.10, see the following corollary.

Corollary 8.17. *Let* $r > 0$. *Suppose that Assumption 7.2*(v) *holds for one of the following models* (a)-(c).

(a) (Gumbel) The model introduced in Example 2.8 with $\vartheta_n \sim_{asy} -r \log(n)$ *as* $n \to \infty$.

(b) (Fréchet) The model introduced in Example 2.9 with $\vartheta_n \sim_{asy} n^{-\frac{r}{\alpha}}$ *as* $n \to \infty$.

(c) (Exponential) The model introduced in Example 2.10 with $\vartheta_n \sim_{asy} n^{-r}$ *as* $n \to \infty$.

Additionally, suppose that T *and* T^{-1} *are the distribution function and the left continuous quantile function of* P_0*, respectively. Furthermore, let* $\rho^*_{Exp,d}(\beta) := \beta - \frac{1}{2}$ *and* $\varepsilon_n := n^{-\beta}$ *for all* $n \in \mathbb{N}$ *and some* $\beta \in (0, \frac{1}{2})$.

(i) Suppose that $r = \rho^*_{Exp,d}(\beta)$. *Then* (7.29) *holds for all* $t \in \mathbb{R}$ *if* (7.3) *is fulfilled, and* (7.31) *holds for all* $t \in \mathbb{R}$ *if* (7.4) *is fulfilled. Thus, HC cannot successfully separate the null and the alternative (asymptotically), see Remark 7.9*(i).

(ii) If $r > \rho^*_{Exp,d}(\beta)$ *then* (7.18) *holds. Hence, HC can completely separate* $\mathcal{H}_{0,n}$ *and* $\mathcal{H}_{1,n}$ *(asymptotically), see Remark 7.7*(i).

Proof. We can assume without loss of generality that the model in (c) is given, compare to Remark 5.9. In this case $\mathcal{Q}_\vartheta = Exp(1+\vartheta)$, $\vartheta \in \Theta := (-1, \infty)$. Thus,

$$h(x) \;=\; x, \;\; \omega(\vartheta) \;=\; \frac{1}{1+\vartheta} \;\text{ and }\; T^{-1}(y) \;=\; -\log(1-y)$$

for all $x > 0$, $\vartheta \in \Theta$ and $y \in (0,1)$. Let $(v_n)_{n\in\mathbb{N}}$ be a sequence in $(0,1)$ with $v_n \to 0$ as $n \to \infty$. Then

$$\begin{aligned}
\omega^{(1)}(0) + \frac{1}{v_n}\int_0^{v_n} -\log(1-y)\;\mathrm{d}y &= -1 + \frac{1}{v_n}\Big((1-v_n)\log(1-v_n) - (1-v_n) + 1\Big) \\
&= -1 + (1-v_n)\frac{\log(1-v_n)}{v_n} + 1 \\
&= -1 + o(1) \text{ as } n \to \infty.
\end{aligned}$$

Finally, applying Corollary 8.16 completes the proof. ■

Example 8.18 (Heterogeneous normal mixtures). Let $(v_n)_{n\in\mathbb{N}}$ be a sequence in $(0,1)$ with $v_n \to 0$ as $n \to \infty$. Let $\mathcal{Q}_\vartheta := N(\vartheta, 1)$ for all $\vartheta \in \Theta := \mathbb{R}$ and $T := 1 - \Phi$. Clearly,

$$T^{-1}(u) = \Phi^{-1}(1-u), \quad h(x) = -x \quad \text{and} \quad \omega(\vartheta) = \exp\left(\frac{\vartheta^2}{2}\right)$$

for all $u \in (0,1)$, $x \in \mathbb{R}$ and every $\vartheta \in \Theta$. Substituting $1-u = \Phi(x)$ and $x^2 = 2y$ we can conclude from (A.24) of Lemma A.29 that

$$\begin{aligned}
&2\omega^{(1)}(0) + \frac{1}{v_n}\int_0^{v_n} h\left(T^{-1}(y)\right)\;\mathrm{d}y \;=\; 0 - \frac{1}{v_n}\int_0^{v_n} \Phi^{-1}(1-u)\,\mathrm{d}u \\
&= -\frac{1}{v_n}\int_{\Phi^{-1}(1-v_n)}^{\infty} x\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}\;\mathrm{d}x \;=\; -\frac{1}{v_n\sqrt{2\pi}}\int_{\frac{1}{2}[\Phi^{-1}(1-v_n)]^2}^{\infty} e^{-y}\;\mathrm{d}y \\
&= -\frac{1}{v_n\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left[\sqrt{-2\log(v_n)} - \frac{\log(4\pi) + \log\log(v_n^{-1}) + o(1)}{2\sqrt{-2\log(v_n)}}\right]^2\right) \\
&= -\frac{1}{v_n\sqrt{2\pi}}\exp\left(\log(v_n) + \frac{1}{2}\log(4\pi) + \frac{1}{2}\log\log(v_n^{-1}) + o(1)\right) \\
&\to \; -\infty \quad \text{as } n \to \infty.
\end{aligned}$$

Consequently, Theorem 8.8(i) and (ii) follow from Corollary 8.16. □

8.4.2. Sparse case

In contrast to the previous section we only present the results regarding Corollary 5.7 for the sparse exponential family model, i.e., $\varepsilon_{n,i} = \varepsilon_n$ for all $1 \leq i \leq n \in \mathbb{N}$. These are visualised in Figure 8.2 for the case $p = 0.6$. We show that the detection areas of HC and LLRT coincide under the conditions of Corollary 5.7. Note that by this they coincide in particular for the exponential, Gumbel and Fréchet distribution mixture models, respectively, i.e. under the assumptions of Corollary 5.10. Moreover, we prove that HC has no power on the detection boundary asymptotically, i.e., HC cannot successfully separate the null and the alternative there. We conclude this result from a more general theorem which can also be applied to the heterogeneous normal mixture model, i.e., the model introduced in Example 2.6 with $\tau = 1$.

Theorem 8.19. *Suppose that Assumption 7.2*(v) *and the assumptions of Corollary 5.7 hold for $T_n := T := F_0$ and $T_n^{-1} := T^{-1} := F_0^{-1}$, $n \in \mathbb{N}$, where F_0 and F_0^{-1} are the distribution function and the left continuous quantile function, see* (2.5)*, of $\mathcal{Q}_0$.*

(i) If $\beta < \beta^{\#}_{Exp}(r,p)$ then (7.18) *holds. In this case HC can completely separate $\mathcal{H}_{0,n}$ and $\mathcal{H}_{1,n}$ (asymptotically), see Remark 7.7*(i)*.*

(ii) Let $\beta < 1$ and $\beta = \beta^{\#}_{Exp}(r,p)$. Let ε_n be given by (5.11)*. Then* (7.29) *holds for all $t \in \mathbb{R}$ if* (7.3) *is fulfilled, and* (7.31) *holds for all $t \in \mathbb{R}$ if* (7.4) *is fulfilled. Thus, HC cannot successfully separate $\mathcal{H}_{0,n}$ and $\mathcal{H}_{1,n}$ (asymptotically), see Remark 7.9*(i)*.*

(iii) Suppose $p > 0$, $r = \frac{1}{p}$, $\beta = \beta^{\#}_{Exp}(r,p) = 1$ and $L \equiv K \in (0,\infty)$. Then (7.29) *holds for all $t \in \mathbb{R}$ if* (7.3) *is fulfilled, and* (7.31) *holds for all $t \in \mathbb{R}$ if* (7.4) *is fulfilled. Thus, HC cannot successfully separate $\mathcal{H}_{0,n}$ and $\mathcal{H}_{1,n}$ (asymptotically), see Remark 7.9*(i)*.*

The statement in (i) follows immediately from Corollary 8.12 and the alternative proof of Corollary 5.7(v) on pp. 93ff. The proofs of Theorem 8.19(ii) and (iii) are given at the end of this section. There we apply the following more general theorem, which gives us a sufficient condition for the case that HC cannot successfully separate the null and the alternative.

Theorem 8.20. *Let $\{k_n : n \in \mathbb{N}\}$ be a subsequence of $\mathbb{N}$. Let $(\alpha_n)_{n\in\mathbb{N}}$ and $(\beta_n)_{n\in\mathbb{N}}$ be sequences in $\mathbb{R}$ such that α_n and β_n tend to 0 as $n \to \infty$. Let Assumption 7.2*(v) *be fulfilled*

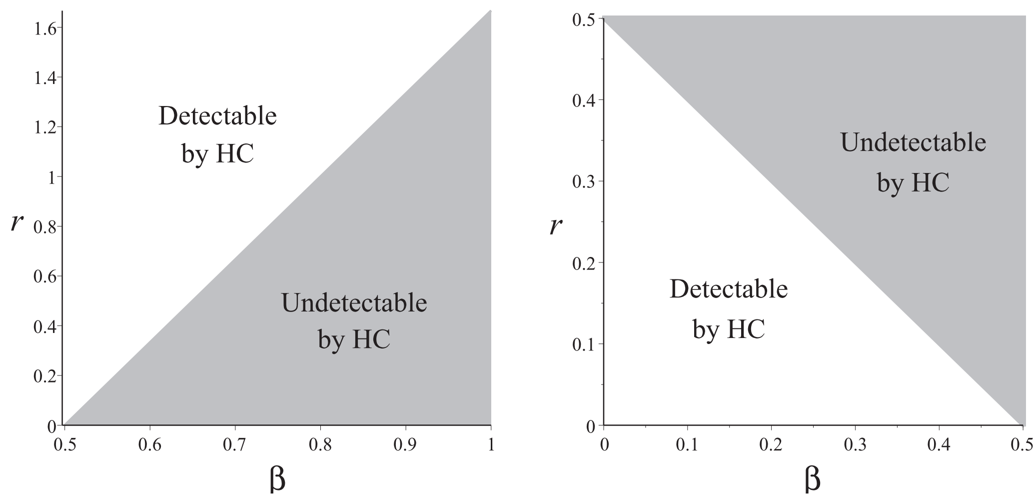

Figure 8.2.: The detectable and the undetectable area of HC are visualised for the dense exponential family mixture model (right) as well as for the sparse one in the case $p = 0.6$ (left). The boundary, which splits the plane into these two areas, belongs to the undetectable area in both models.

and $\psi : (1,3) \to \mathbb{R}$ *be a function with* $\psi(2) = \frac{1}{2}$ *such that one of the following conditions* (i)-(iii) *holds,* $P_{(k_n)} \vartriangleleft\vartriangleright Q_{(k_n)}$ *and*

$$\varepsilon_{k_n} = {k_n}^{-\beta+\beta_{k_n}} \text{ for all } n \in \mathbb{N} \text{ and some } \beta \in \left(\frac{1}{2}, 1\right]. \tag{8.41}$$

(i) We have

$$\left(\int f_{k_n}^q \, \mathrm{d}P_{k_n}\right)^{\frac{1}{q}} \leq k_n^{\psi(q)(2\beta-1)+\alpha_{k_n}} \tag{8.42}$$

for all sufficiently large $n \in \mathbb{N}$*, every* $q \in (2-\kappa, 2+\kappa) \setminus \{2\}$ *and some* $\kappa \in (0,1)$*. Moreover,* ψ *is differentiable at* $q = 2$ *with* $\psi^{(1)}(2) \in (0, \frac{1}{8\beta-4})$*.*

(ii) (8.42) *holds for all sufficiently large* $n \in \mathbb{N}$*, every* $q \in (2, 2+\kappa)$ *and some* $\kappa \in (0,1)$*. Moreover,*

$$\limsup_{q \searrow 2} \frac{\psi(q) - \frac{1}{2}}{q-2} \leq 0. \tag{8.43}$$

(iii) (8.42) *holds for every sufficiently large* $n \in \mathbb{N}$*, all* $q \in (2-\kappa, 2)$ *and some* $\kappa \in (0,1)$*. Furthermore,*

$$\liminf_{q \nearrow 2} \frac{\psi(q) - \frac{1}{2}}{q-2} \geq \frac{1}{8\beta - 4}. \tag{8.44}$$

Then (7.29) *holds for all* $t \in \mathbb{R}$ *if* (7.3) *is fulfilled, and* (7.31) *holds for all* $t \in \mathbb{R}$ *if* (7.4) *is fulfilled. Thus, HC cannot successfully separate the null and the alternative (asymptotically), see Remark 7.9*(i)*.*

Remark 8.21. Assume that Assumption 2.11 holds. Let $q \in (1,3)$ such that $\vartheta_{k_n} q \in \Theta$. Then

$$\left(\int f_{k_n,i}^q \, \mathrm{d}P_{k_n,i}\right)^{\frac{1}{q}} = \left(\int C(\vartheta_{k_n})^q \exp\left(q\vartheta_{k_n} h(x)\right) \, \mathrm{d}\mathcal{Q}_0\right)^{\frac{1}{q}} = \frac{C(\vartheta_{k_n})}{C(q\vartheta_{k_n})^{\frac{1}{q}}}.$$

Thus, the conditions (8.41) and (8.42) remind us of (4.67). □

The exponential family $(N(\vartheta,1) : \vartheta \in \mathbb{R})$ does not fulfil (iii) of Assumption 2.16. Hence, Theorem 8.20 is not applicable for this family, whereas the more general Theorem 8.20 is applicable.

Example 8.22 (Heterogeneous normal mixture model). Let $\mathcal{Q}_\vartheta := N(\vartheta,1)$ for all $\vartheta \in \Theta$. Let $\beta \in (\frac{1}{2}, \frac{3}{4})$, $r := \rho^*(\beta,1) = \beta - \frac{1}{2}$, $\vartheta_n := \sqrt{2r\log(n)}$ and $\varepsilon_n := n^{-\beta}$. Then for all $\vartheta \in \mathbb{R}$, $n \in \mathbb{N}$ and $q \in (1,3)$ we have

$$C(\vartheta) = \exp\left(-\frac{1}{2}\vartheta^2\right) \text{ and, hence, } \frac{C\left(\vartheta_{k_n}\right)}{C\left(q\vartheta_{k_n}\right)^{\frac{1}{q}}} = n^{-r(1-q)} = n^{(2\beta-1)\frac{1}{2}(q-1)}.$$

Consequently, for the case that $\tau = 1$ and $\beta \in (\frac{1}{2}, \frac{3}{4})$ Theorem 8.1(ii) can also be shown by applying Theorem 8.20 with ψ defined by $\psi(q) := \frac{1}{2}(q-1)$, $q \in (1,3)$. □

Proof of Theorem 8.20. First, we verify (i). To improve the readability we (only) give the proof for the case $\{k_n : n \in \mathbb{N}\} = \mathbb{N}$. Suppose that the assumptions of (i) are fulfilled. Let $(q_{n,1})_{n\in\mathbb{N}}$ be a sequence in $(2-\kappa, 2)$ and $(q_{n,2})_{n\in\mathbb{N}}$ be a sequence in $(2, 2+\kappa)$ such that

$$q_{n,m} = 2 + o(1) \text{ and } \left|\frac{\alpha_n}{2-q_{n,m}}\right| + \left|\frac{\beta_n}{2-q_{n,m}}\right| + \left|\frac{\log\log(n)}{(2-q_{n,m})\log(n)}\right| = o(1) \tag{8.45}$$

for $m = 1,2$ as $n \to \infty$. As mentioned in Remark 7.3(i) we can arbitrarily choose $\rho, \rho_{n,1}, \rho_{n,2}$ from Assumption 7.2(i). Define

$$\rho := 4(2\beta - 1)\psi^{(1)}(2) \in (0,1). \tag{8.46}$$

For all $n \in \mathbb{N}$ and $m \in \{1,2\}$ let

$$\rho_{n,m} := \frac{-2q_{n,m}\log\log(n)}{(q_{n,m}-2)\log(n)} + \rho - q_{n,m}(4\beta-2)\frac{\psi(q_{n,m}) - \psi(2)}{q_{n,m}-2} - \frac{2q_{n,m}\left(\alpha_n + \beta_n + \frac{a_n}{\log(n)}\right)}{q_{n,m}-2}.$$

Obviously, we can conclude from (8.45) and (8.46) that $\rho_{n,1}, \rho_{n,2} = o(1)$ as $n \to \infty$. Let

$$B_{\tau,1} := (0, \tau] \quad \text{and} \quad B_{\tau,2} := (1-\tau, 1)$$

for all $\tau \in (0,1)$. From Hölder's inequality we can deduce that for all sufficiently large $n \in \mathbb{N}$, every $\tau \in (0,1)$ and all $l, m \in \{1,2\}$:

$$\begin{aligned}
\sqrt{n}\,\varepsilon_n \,\frac{\mu_n^{T_n}(B_{\tau,l})}{\sqrt{\tau}} &= \tau^{-\frac{1}{2}}\, n^{\frac{1}{2}-\beta+\beta_n} \int f_n \,\mathbf{1}_{T_n^{-1}(B_{\tau,l})}\, \mathrm{d}P_n \\
&\leq \tau^{-\frac{1}{2}}\, n^{\frac{1}{2}-\beta+\beta_n} \left[\int f_n^{q_{n,m}}\, \mathrm{d}P_n\right]^{\frac{1}{q_{n,m}}} \left[\int \mathbf{1}_{B_{\tau,l}}\, \mathrm{d}P_n^{T_n}\right]^{1-\frac{1}{q_{n,m}}} \\
&\leq \tau^{-\frac{1}{2}}\, n^{\frac{1}{2}-\beta+\beta_n+\psi(q_{n,m})(2\beta-1)+\alpha_n} \left[\mathcal{U}_{(0,1)}\left(B_{\tau,l}\right)\right]^{1-\frac{1}{q_{n,m}}} \\
&= \tau^{\frac{1}{2}-\frac{1}{q_{n,m}}}\, n^{\frac{1}{2}-\beta+\beta_n+\psi(q_{n,m})(2\beta-1)+\alpha_n} \\
&= \tau^{\frac{q_{n,m}-2}{2q_{n,m}}}\, n^{[\psi(q_{n,m})-\psi(2)](2\beta-1)+\beta_n+\alpha_n}.
\end{aligned}$$

Combining this and the definition of $\rho_{n,m}$ shows for every $l \in \{1,2\}$

$$\begin{aligned}
a_n\sqrt{n}\,\varepsilon_n \sup_{\tau\in[r_n,s_n]} \left\{\frac{\mu_n^{T_n}(B_{\tau,l})}{\sqrt{\tau}}\right\} &\leq s_n^{\frac{q_{n,2}-2}{2q_{n,2}}}\, a_n\, n^{[\psi(q_{n,2})-\psi(2)](2\beta-1)+\alpha_n+\beta_n} \\
&= a_n\, n^{(-\rho+\rho_{n,2})\frac{q_{n,2}-2}{2q_{n,2}}+[\psi(q_{n,2})-\psi(2)](2\beta-1)+\alpha_n+\beta_n} \\
&= \exp\left(-\log\log(n)\right) \to 0 \text{ as } n \to \infty. \qquad (8.47)
\end{aligned}$$

and analogously

$$\begin{aligned}
a_n\sqrt{n}\,\varepsilon_n \sup_{\tau\in[t_n,u_n]} \left\{\frac{\mu_n^{T_n}(B_{\tau,l})}{\sqrt{\tau}}\right\} &\leq t_n^{\frac{q_{n,1}-2}{2q_{n,1}}}\, a_n\, n^{[\psi(q_{n,1})-\psi(2)](2\beta-1)+\alpha_n+\beta_n} \\
&= \exp\left(-\log\log(n)\right) \to 0 \text{ as } n \to \infty.
\end{aligned}$$

Consequently, (i) follows from Theorem 7.8 and Remark 7.9(ii).
Similar to (8.47) we conclude that for all $n \in \mathbb{N}$ and every $l \in \{1,2\}$

$$\begin{aligned}
&a_n\sqrt{n}\,\varepsilon_n \sup_{\tau\in[r_n,u_n]} \left\{\frac{\mu_n^{T_n}(B_{\tau,l})}{\sqrt{\tau}}\right\} \qquad (8.48)\\
&\leq u_n^{\frac{q_{n,2}-2}{2q_{n,2}}}\, a_n\, n^{[\psi(q_{n,2})-\psi(2)](2\beta-1)+\alpha_n+\beta_n} \\
&= \exp\left(\log(n)\frac{q_{n,2}-2}{2q_{n,2}}\left[-\lambda_n + 2q_{n,2}(2\beta-1)\frac{\psi(q_{n,2})-\psi(2)}{q_{n,2}-2} + \frac{2q_{n,2}}{q_{n,2}-2}\left(\alpha_n+\beta_n+\frac{a_n}{\log(n)}\right)\right]\right).
\end{aligned}$$

In case of strict inequality in (8.43) we can immediately deduce that (8.48) tends to 0 as $n \to \infty$. Otherwise, we can assume without loss of generality due to Remark 7.3(i) that λ_n tends to 0 sufficiently slowly as $n \to \infty$. Hence, as in the previous calculations we can verify that (8.48) also tends to 0 in case of strict inequality in (8.43). Finally, (ii) is shown. (iii) can be proven analogously. ■

Proof of Theorem 8.19(ii) *and* (iii). Supppose that the assumptions of Theorem 8.19(ii) or the ones of (iii) hold. Define $\psi : (1,3) \to \mathbb{R}$ by $\psi(q) := 1 - q^{-1}$, $q \in (1,3)$. Note that $\psi^{(1)}(2) = \frac{1}{4}$. By Lemma 2.15, (5.6) and Remark 8.21 we have for all $q \in (1,3)$

$$\left(\int f_n^q \,\mathrm{d}\mathcal{Q}_0\right)^{\frac{1}{q}} = n^{rp(1-\frac{1}{q})} \frac{L(q\vartheta_n)^{\frac{1}{q}}}{L(\vartheta_n)}(1+o(1)) = n^{(2\beta-1)\psi(q)+o(1)} \quad \text{as } n \to \infty.$$

Moreover, from Corollary 4.12(i), Corollary 5.7(ii) and (iii) we obtain $P_{(n)} \lhd \rhd Q_{(n)}$. Applying Theorem 8.20(i) completes the proof. ■

A. Appendix: additional information

A.1. Infinitely divisible distributions

In this section we present some useful results about infinitely divisible distributions from the book of Petrov [51] and modify some of his results for our purpose.

Definition A.1 (Infinitely divisible distributions, see p. 25 in [51]). *We call a probability measure P on $(\mathbb{R}, \mathcal{B})$ infinitely divisible if for every $n \in \mathbb{N}$ there exists some probability measure P_n on $(\mathbb{R}, \mathcal{B})$ such that P is equal to the n-fold convolution of P_n.*

Definition A.2 (Lévy measure). *A measure η on $(\mathbb{R} \setminus \{0\}, \mathcal{B}(\mathbb{R} \setminus \{0\}))$ is called a Lévy measure if*

$$\int_{\mathbb{R}\setminus\{0\}} \min(x^2, 1)\ \mathrm{d}\eta(x) < \infty. \tag{A.1}$$

Clearly, a Lévy measure is not necessarily finite but at least it is σ-finite on $\mathbb{R} \setminus \{0\}$.

Theorem A.3 (Lévy's formula, see Theorem II.3.5 in [51]). *Let P be a probability measure on $(\mathbb{R}, \mathcal{B})$ and let φ be its characteristic function. Then P is infinitely divisible if and only if there exists $\gamma \in \mathbb{R}$, $\sigma^2 \geq 0$ and some Lévy measure η such that*

$$\varphi(t) \ = \ \exp\left[\boldsymbol{i}\gamma t - \frac{\sigma^2 t^2}{2} + \int_{\mathbb{R}\setminus\{0\}} \left(e^{\boldsymbol{i}tx} - 1 - \frac{\boldsymbol{i}tx}{1+x^2}\right)\ \mathrm{d}\eta(x)\right]. \tag{A.2}$$

The triple (γ, σ^2, η) is called the Lévy characteristic of P and is unique.

Example A.4. Let $P = N(a, b)$ for some $(a, b) \in \mathbb{R} \times (0, \infty)$. Then P is infinitely divisible with Lévy characteristic $(a, b, 0)$, i.e., the Lévy measure η is the trivial measure, $\eta(A) = 0$ for all $A \in \mathcal{B}(\mathbb{R} \setminus \{0\})$. □

In Chapter 4 we examined the asymptotic behaviour of $\sum_{i=1}^{n} Y_{n,i}$ for a certain triangular array $(Y_{n,i})_{1 \leq i \leq n}$ with row-wise independent, real-valued random variables. In [51] the answer for the question, which limiting distribution can occur for such a sum, is given. Without additional assumptions it is easy to see that any distribution can be the limiting distribution for a suitable triangular array. Set, e.g., $Y_{n,1} := X$ for some random variable X for all $n \in \mathbb{N}$ and $Y_{n,i} := 0$ for all $2 \leq i \leq n \in \mathbb{N}$. Then the sum $\sum_{i=1}^{n} Y_{n,i}$ converges obviously in distribution to X. Hence, it is reasonable to add more restrictions for the random variables $Y_{n,i}$.

Definition A.5 (Infinite smallness, see p. 63 of [51]). *Let $(Y_{n,i})_{1\leq i\leq n}$ be a triangular array of row-wise independent, real-valued random variables on some probability space $(\Omega, \mathcal{A}, \mathcal{P})$. We say that $(Y_{n,i})_{1\leq i\leq n}$ fulfils the condition of infinite smallness if*

$$\lim_{n\to\infty} \max_{1\leq i\leq n} \left\{ \mathcal{P}\left(|Y_{n,i}| \geq \varepsilon\right) \right\} = 0 \text{ for every } \varepsilon > 0.$$

Under this additional condition all accumulation points (in the sense of convergence in distribution) of $\sum_{i=1}^{n} Y_{n,i}$ are infinitely divisible.

Theorem A.6. *Let $(Y_{n,i})_{1\leq i\leq n}$ be a triangular array of row-wise independent, real-valued random variables on some probability space $(\Omega, \mathcal{A}, \mathcal{P})$ such that the condition of infinite smallness is fulfilled. Let $\{k_n : n \in \mathbb{N}\}$ be a subsequence of $\mathbb{N}$. If*

$$\sum_{i=1}^{k_n} Y_{k_n,i} \xrightarrow{D} Y \quad \text{as } n \to \infty \tag{A.3}$$

for some real-valued random variable Y on $(\Omega, \mathcal{A}, \mathcal{P})$ then the distribution of Y is infinitely divisible. In the following we say that Y is infinite divisible.

In regard to Theorem A.3 and Theorem A.6 it is of interest to determine the Lévy characteristic (γ, σ^2, η) of Y. In order to do so we present one of the results in [51]. Since we prefer Lévy's formula we rewrite Theorem IV.2.6 in [51] using the relation between this formula and the one of Khintchine and Lévy. For more details about this relation we refer the reader to [51], in particular to Section II.2.

Theorem A.7. *Suppose that the assumptions of Theorem A.6 are fulfilled. Let $\sigma^2 \geq 0$, $\gamma \in \mathbb{R}$ and η be a Lévy measure on $(\mathbb{R} \setminus \{0\}, \mathcal{B}(\mathbb{R} \setminus \{0\}))$. Moreover, let $\tau \in \mathbb{R}$ such that $-\tau, \tau \in C(\eta)$. Then* (A.3) *holds for some infinitely divisible Y on $(\Omega, \mathcal{A}, \mathcal{P})$ with Lévy characteristic (γ, σ^2, η) if and only if the following conditions* (a)-(c) *hold.*

(a) For all $x \in C(\eta) \cap (0,\infty)$ and $y \in C(\eta) \cap (-\infty, 0)$

$$\sum_{i=1}^{k_n} \mathcal{P}\left(Y_{k_n,i} < y\right) \;\to\; \eta(-\infty, y) \quad \text{as } n \to \infty \tag{A.4}$$

$$\text{and} \quad \sum_{i=1}^{k_n} \mathcal{P}\left(Y_{k_n,i} \geq x\right) \;\to\; \eta[x, \infty) \quad \text{as } n \to \infty. \tag{A.5}$$

(b) We have

$$\lim_{\varepsilon \searrow 0} \begin{matrix}\limsup \\ \liminf \\ n\to\infty\end{matrix} \sum_{i=1}^{k_n} \left[\int Y_{k_n,i}^2 \mathbf{1}_{\{|Y_{k_n,i}|<\varepsilon\}} \, \mathrm{d}\mathcal{P} - \left(\int Y_{k_n,i} \mathbf{1}_{\{|Y_{k_n,i}|<\varepsilon\}} \, \mathrm{d}\mathcal{P} \right)^2 \right] = \sigma^2.$$

(c) We have

$$\begin{aligned} &\lim_{n\to\infty} \sum_{i=1}^{k_n} \int Y_{k_n,i} \mathbf{1}_{\{|Y_{k_n,i}|<\tau\}} \, \mathrm{d}\mathcal{P} \\ &= \gamma + \int_{(-\tau,\tau)\setminus\{0\}} \frac{x^3}{1+x^2} \, \mathrm{d}\eta(x) - \int_{\mathbb{R}\setminus[-\tau,\tau]} \frac{x}{1+x^2} \, \mathrm{d}\eta(x). \end{aligned}$$

Remark A.8. Suppose that (a) and (b) hold. Due to the equivalence above, (c) holds for some $\tau > 0$ with $-\tau, \tau \in C(\eta)$ if and only if it does for all $\tau > 0$ with $-\tau, \tau \in C(\eta)$. □

In Chapter 4 we showed that for our model certain conditions are always fulfilled. In the following we present a simplification of Theorem A.7 under this additional conditions. In order to prove this we need the following lemma.

Lemma A.9. *Let $\eta, \eta_1, \eta_2, \ldots$ be measures on $((0,\infty), \mathcal{B}((0,\infty)))$ such that*

$$\lim_{n\to\infty} \eta_n(x,\infty) \;=\; \eta(x,\infty) \in \mathbb{R} \text{ for all } x > 0 \tag{A.6}$$

$$\text{and} \quad \limsup_{n\to\infty} \int_{(0,1)} t^2 \, \mathrm{d}\eta_n(t) \;\leq\; C \text{ for some } C \in (0,\infty). \tag{A.7}$$

Then $\widetilde{\eta}$ uniquely determined by (A.8) *is a Lévy measure on $(\mathbb{R}\setminus\{0\}, \mathcal{B}(\mathbb{R}\setminus\{0\}))$:*

$$\widetilde{\eta}(-\infty, 0) := 0 \;\text{ and }\; \widetilde{\eta}(x,\infty) := \eta(x,\infty) \text{ for all } x > 0. \tag{A.8}$$

Proof. By Definition A.2 and (A.6) it is sufficient to show that

$$\int_{(0,1)} t^2 \, \mathrm{d}\eta(t) \;<\infty. \tag{A.9}$$

Note that Portmanteau's Theorem and Helly-Bray's Theorem about weak convergence of probability measures can be easily extended to finite measures. From (A.6) we obtain for every $\varepsilon \in (0,1)$

$$\eta_{n\,|(\varepsilon,\infty)} \xrightarrow{w} \eta_{|(\varepsilon,\infty)} \text{ and, hence, } \int_{(\varepsilon,1)} t^2 \,\mathrm{d}\eta_n(t) \to \int_{(\varepsilon,1)} t^2 \,\mathrm{d}\eta(t)$$

as $n \to \infty$. Combining this, the monotone convergence theorem and (A.7) yields (A.9):

$$\int_{(0,1)} t^2 \,\mathrm{d}\eta(t) = \lim_{\varepsilon \searrow 0} \int_{(\varepsilon,1)} t^2 \,\mathrm{d}\eta(t) = \lim_{\varepsilon \searrow 0} \lim_{n\to\infty} \int_{(\varepsilon,1)} t^2 \,\mathrm{d}\eta_n(t) \le C. \qquad \blacksquare$$

Theorem A.10. *Let the assumptions of Theorem A.6 be fulfilled such that for all $y < 0$*

$$\bigcup_{i=1}^{k_n} \{Y_{k_n,i} \le y\} = \emptyset \quad \text{if } n \in \mathbb{N} \text{ is sufficiently large.} \tag{A.10}$$

Assume additionally that

$$\lim_{\varepsilon \searrow 0} \lim_{n\to\infty} \sum_{i=1}^{k_n} \left(\int Y_{k_n,i} \mathbf{1}_{\{|Y_{k_n,i}|<\varepsilon\}} \,\mathrm{d}\mathcal{P} \right)^2 = 0. \tag{A.11}$$

Let $\sigma^2 \ge 0$ and $\gamma \in \mathbb{R}$. Let η be a measure on $(\mathbb{R} \setminus \{0\}, \mathcal{B}(\mathbb{R} \setminus \{0\}))$ with $\eta(-\infty,0) = 0$. Then (A.3) *holds for some infinitely divisible random variable Y with Lévy characteristic (γ, σ^2, η) if and only if the following conditions* (i)-(iii) *hold.*

(i) For every $x \in C(\eta) \cap (0,\infty)$

$$\sum_{i=1}^{k_n} \mathcal{P}\left(Y_{k_n,i} > x\right) \to \eta(x,\infty) \quad \text{as } n \to \infty.$$

(ii) We have

$$\lim_{\varepsilon \searrow 0} \begin{matrix} \limsup \\ \liminf \\ n\to\infty \end{matrix} \sum_{i=1}^{k_n} \int Y_{k_n,i}^2 \mathbf{1}_{\{|Y_{k_n,i}|\le\varepsilon\}} \,\mathrm{d}\mathcal{P} = \sigma^2. \tag{A.12}$$

(iii) We have

$$\lim_{C(\eta)\ni\tau\searrow 0} \left(\begin{matrix} \limsup \\ \liminf \\ n\to\infty \end{matrix} \left(\sum_{i=1}^{k_n} \int Y_{k_n,i} \mathbf{1}_{\{|Y_{k_n,i}|\le\tau\}} \,\mathrm{d}\mathcal{P} \right) + \int_{(\tau,\infty)} \frac{x}{1+x^2} \,\mathrm{d}\eta(x) \right) = \gamma. \tag{A.13}$$

Remark A.11. (a) The assumption that η is a Lévy measure is not needed for the statement of Theorem A.10. Applying Lemma A.9 we show in the following proof that η is a Lévy measure if (A.10) for all $y < 0$, (A.11), (i) and (ii) of Theorem A.10 are fulfilled.

(b) It is well known that the continuity points of $x \mapsto \eta(x, \infty)$ are dense in $(0, \infty)$. Thus, the term on the left side of (A.13) is well defined.

Proof of Theorem A.10. First, note that due to (A.10) the condition (A.4) holds for all $y < 0$. Consequently, (a) of Theorem A.7 and (i) of Theorem A.10 are equivalent conditions, where we replaced the relation sign $\geq$ by $>$ and the set $[\varepsilon, \infty)$ by (ε, ∞) in (i). Note that the last mentioned modifications do not affect the result since, obviously, we have under (a) of Theorem A.7 as well as under (i) of Theorem A.10

$$\lim_{n\to\infty} \sum_{i=1}^{k_n} \mathcal{P}\left(Y_{k_n,i} = x\right) = 0 \text{ and } \eta\left(\{x\}\right) = 0 \text{ for all } x \in C(\eta) \cap (0, \infty). \tag{A.14}$$

By (A.11) the conditions (b) of Theorem A.7 and (ii) of Theorem A.10 are also equivalent, where we replaced $<$ by $\leq$ in (ii). There is no loss in validity by doing this because the integrand in (A.12) is non-negative. To conclude Theorem A.10 from Theorem A.7 it remains to verify the following two statements SI and SII:

SI. If (i) and (ii) of Theorem A.10 are fulfilled then η is a Lévy measure.

SII. Condition (c) of Theorem A.7 and (iii) of Theorem A.10 are equivalent under (a) and (b) of Theorem A.7.

<u>Proof of SI:</u> Let (i) and (ii) be fulfilled and $\eta_1, \eta_2, \dots$ be measures on $((0, \infty), \mathcal{B}((0, \infty)))$ uniquely determined by

$$\eta_n(x, \infty) = \sum_{i=1}^{k_n} \mathcal{P}\left(Y_{k_n,i} > x\right) \text{ for all } x > 0 \text{ and } n \in \mathbb{N}.$$

(A.6) follows from (i) of Theorem A.10. From (i) and (ii) of Theorem A.10 we obtain (A.7). Applying Lemma A.9 yields that η is a Lévy measure.

<u>Proof of SII:</u> First, suppose that (a)-(c) of Theorem A.7 hold. As already explained, (i) and (ii) of Theorem A.10 are fulfilled. By SI the measure η is a Lévy measure and so by

Definition A.2

$$0 \leq \int_{(0,\tau)} \frac{x^3}{1+x^2}\, \mathrm{d}\eta(x) \leq \tau \int_{(0,1)} x^2\, \mathrm{d}\eta(x) \to 0 \text{ for } \tau \searrow 0. \tag{A.15}$$

Combining Remark A.8, (A.14) and (A.15) yields (iii) of Theorem A.7.
Second, suppose that (a) and (b) of Theorem A.7 as well as (iii) of Theorem A.10 hold. It is sufficient to show: for every subsequence $\{k_{n,1} : n \in \mathbb{N}\}$ of $\{k_n : n \in \mathbb{N}\}$ there exists a further subsequence $\{k_{n,2} : n \in \mathbb{N}\}$ of $\{k_{n,1} : n \in \mathbb{N}\}$ such that (c) of Theorem A.7 holds for $\{k_{n,2} : n \in \mathbb{N}\}$ and γ. Keep in mind that $\eta(-\infty, 0) = 0$. Let $\{k_{n,1} : n \in \mathbb{N}\}$ be an arbitrary subsequence of $\{k_n : n \in \mathbb{N}\}$. By (iii) and (A.14)

$$\limsup_{n\to\infty} \left| \left(\sum_{i=1}^{k_n} \int Y_{k_n,i}\, \mathbf{1}_{\{|Y_{k_n,i}|<\tau^*\}}\, \mathrm{d}\mathcal{P} \right) + \int_{(\tau^*,\infty)} \frac{x}{1+x^2}\, \mathrm{d}\eta(x) \right| < \infty \tag{A.16}$$

for some $\tau^* \in C(h) \cap (0,1)$. Hence,

$$\lim_{n\to\infty} \left(\sum_{i=1}^{k_{n,2}} \int Y_{k_{n,2},i}\, \mathbf{1}_{\{|Y_{k_{n,2},i}|<\tau^*\}}\, \mathrm{d}\mathcal{P} \right) + \int_{(\tau^*,\infty)} \frac{x}{1+x^2}\, \mathrm{d}\eta(x) \in \mathbb{R}$$

for some subsequence $\{k_{n,2} : n \in \mathbb{N}\}$ of $\{k_{n,1} : n \in \mathbb{N}\}$, and so (c) of Theorem A.7 holds for $\{k_{n,2} : n \in \mathbb{N}\}$, τ^* and some constant $\widetilde{\gamma} \in \mathbb{R}$. We deduce from the first part of the proof of SII that (iii) holds for $\{k_{n,2} : n \in \mathbb{N}\}$ and $\widetilde{\gamma}$. Obviously, (a) and (b) of Theorem A.7 as well as (iii) of Theorem A.10 hold for the subsequence $\{k_{n,2} : n \in \mathbb{N}\}$ and γ. Finally, $\gamma = \widetilde{\gamma}$. ■

A.2. Distances between probability measures

In this section we introduce different distances for probability measures and some properties of these. All definitions and results can be found in Strasser [58], Chapter 1 §2. Throughout this section, let P, Q, ν be probability measures on some measurable space $(\Omega, \mathcal{A})$ with $P, Q \ll \nu$.

Definition and Lemma A.12. *(i) We define the variational distance $||P-Q||$ between P and Q by*

$$||P-Q|| := \sup\{|P(A) - Q(A)| : A \in A\} = \frac{1}{2} \int \left| \frac{\mathrm{d}P}{\mathrm{d}\nu} - \frac{\mathrm{d}Q}{\mathrm{d}\nu} \right| \mathrm{d}\nu.$$

The variational distance is a distance with

$$0 \leq ||P - Q|| \leq 1, \text{ where } ||P - Q|| = 1 \text{ if and only if } P \perp Q,$$
$$\text{and } ||P - Q|| = 0 \text{ if and only if } P = Q.$$

(ii) The affinity of P and Q is given by

$$a(P, Q) := \int \sqrt{\frac{\mathrm{d}P}{\mathrm{d}\nu} \frac{\mathrm{d}Q}{\mathrm{d}\nu}}\, \mathrm{d}\nu.$$

The value $a(P, Q)$ does not depend on the choice of ν.

(iii) The Hellinger distance $d(P, Q)$ between P and Q is given by

$$d^2(P, Q) := \frac{1}{2} \int \left(\sqrt{\frac{\mathrm{d}P}{\mathrm{d}\nu}} - \sqrt{\frac{\mathrm{d}Q}{\mathrm{d}\nu}} \right)^2 \mathrm{d}\nu = 1 - a(P, Q). \tag{A.17}$$

The Hellinger distance is a distance with

$$0 \leq d(P, Q) \leq 1, \text{ where } d(P, Q) = 1 \text{ if and only if } P \perp Q,$$
$$\text{and } d(P, Q) = 0 \text{ if and only if } P = Q.$$

Lemma A.13 (Lemma 2.3 in [58]). *We have*

$$||P - Q|| = \sup\Big\{ \mathbb{E}_P(\varphi) - \mathbb{E}_Q(\varphi) : \ \varphi : (\Omega, \mathcal{A}) \to ([0, 1], \mathcal{B}[0, 1]) \text{ is measurable} \Big\}.$$

We can immediately conclude from Lemma 2.15 in [58]:

Lemma A.14. *We have*

$$d^2(P, Q) \leq ||P - Q|| \leq \sqrt{2}\, d(P, Q).$$

The Hellinger distance and the affinity are very useful if one deals with product measures. By their definitions it is easy to see that the following lemma holds.

Lemma A.15. *Let $P_1, Q_1, \ldots, P_n, Q_n$ be probability measures on $(\Omega, \mathcal{A})$. Then*

$$a\left(\bigotimes_{i=1}^n P_i, \bigotimes_{i=1}^n Q_i \right) = \prod_{i=1}^n a(P_i, Q_i)$$
$$\text{and, hence,} \quad d^2\left(\bigotimes_{i=1}^n P_i, \bigotimes_{i=1}^n Q_i \right) = 1 - \prod_{i=1}^n \left[1 - d^2(P_i, Q_i) \right].$$

A.3. Binary experiments

In this section we introduce binary experiments and the weak convergence of them. We present some useful results concerning these experiments and at the end we formulate the first lemma of Le Cam. To keep this section short we focus on the main results and avoid to discuss them in detail. For a fuller treatment of binary and general statistical experiments we can recommend the lecture notes of Janssen [37] and the book of Strasser [58].

Definition A.16. *Let* $(\Omega, \mathcal{A})$ *be a measurable space. Let* P *and* Q *be probability measures on* $(\Omega, \mathcal{A})$*. We call* $(\Omega, \mathcal{A}, \{P, Q\})$*, in short* $\{P, Q\}$*, a binary experiment. The likelihood-ratio of* Q *with respect of* P *is given by*

$$\frac{\mathrm{d}Q}{\mathrm{d}P} = \left(\frac{\mathrm{d}P}{d(P+Q)}\right)\left(\frac{\mathrm{d}Q}{d(P+Q)}\right)^{-1} \quad \frac{1}{2}(P+Q)\text{-}a.s., \tag{A.18}$$

where we use the convention $\frac{x}{0} = \infty$ *and* $\frac{x}{\infty} = 0$ *for* $x \in (0, \infty)$*. Let*

$$\nu_1 := \mathcal{L}\Big(\log\left(\frac{\mathrm{d}Q}{\mathrm{d}P}\right)\Big|P\Big), \; \nu_2 := \mathcal{L}\Big(\log\left(\frac{\mathrm{d}Q}{\mathrm{d}P}\right)\Big|Q\Big) \; \textit{and } \nu := \nu_1 + \nu_2,$$

where we extend the logarithm continuously to $[0, \infty]$ *by setting*

$$\log(0) := -\infty \quad \textit{and} \quad \log(\infty) := \infty. \tag{A.19}$$

We call $\{\nu_1, \nu_2\}$ *the standard form of* $\{P, Q\}$*.*

Remark A.17. (i) $\{\nu_1, \nu_2\}$ is the standard form of itself.

(ii) By 16.5 in [58] binary experiments with the same standard form are *equal informative* in the sense of 15.1 and 15.2 in [58]. Note that the following results and definitions, A.18 to A.26, only depend on the standard form and not on the special choice of the binary experiment.

(iii) Let $T : (\Omega, \mathcal{A}) \to (\widetilde{\Omega}, \widetilde{\mathcal{A}})$ be a measurable mapping. Suppose that there exists a further measurable mapping $T^{-1} : (\widetilde{\Omega}, \widetilde{\mathcal{A}}) \to (\Omega, \mathcal{A})$ such that

$$T^{-1} \circ T \;=\; id_\Omega \quad (P+Q)\text{-almost everywhere.}$$

It follows easily from Lemma A.31 that the standard form of $\{P^T, Q^T\}$ is also $\{\nu_1, \nu_2\}$.

(iv) In [58] and other references the standard form is defined without the logarithm transform. But in the case of product measures and so for the purpose of our work the logarithm transform is very useful.

(v) ν_1 and ν_2 are probability measures on $([-\infty,\infty], \mathcal{B}([-\infty,\infty]))$ with

$$\nu_1(\{\infty\}) = P\left(\frac{\mathrm{d}Q}{\mathrm{d}P} = \infty\right) = 0,\ \nu_2(\{-\infty\}) = Q\left(\frac{\mathrm{d}Q}{\mathrm{d}P} = 0\right) = 0. \qquad \square$$

Lemma A.18. *Let $\{P, Q\}$ be a binary experiment. Then*

$$\frac{\mathrm{d}\nu_1}{\mathrm{d}\nu} = \frac{1}{1+\exp}\,, \quad \frac{\mathrm{d}\nu_2}{\mathrm{d}\nu} = \frac{\exp}{1+\exp} \quad \textit{and} \quad \frac{\mathrm{d}\nu_2}{\mathrm{d}\nu_1} = \exp\,.$$

Proof. See Le Cam [48], p.24f., and (9.21) Example of Janssen et al. [39]. ■

We want to remind the reader that a sequence of probability measures $(\mathcal{P}_n)_{n\in\mathbb{N}}$ on $([-\infty,\infty], \mathcal{B}([-\infty,\infty]))$ converges weakly to a probability measure $\mathcal{P}$ on $([-\infty,\infty], \mathcal{B}([-\infty,\infty]))$ if for every continuous function $f : [-\infty,\infty] \to \mathbb{R}$

$$\int f\,\mathrm{d}\mathcal{P}_n \to \int f\,\mathrm{d}\mathcal{P} \quad \text{as } n \to \infty.$$

Definition A.19. *Let $\{P, Q\}$ and $\{P_{(n)}, Q_{(n)}\}$ be binary experiments for all $n \in \mathbb{N}$. We say that $\{P_{(n)}, Q_{(n)}\}$ converges weakly to $\{P, Q\}$ (as $n \to \infty$) if*

$$\nu_{1,n} := \mathcal{L}\left(\log\left(\frac{\mathrm{d}Q_{(n)}}{\mathrm{d}P_{(n)}}\right) \,\Big|\, P_{(n)}\right) \xrightarrow{w} \mathcal{L}\left(\log\left(\frac{\mathrm{d}Q}{\mathrm{d}P}\right) \,\Big|\, P\right) = \nu_1 \quad \textit{as } n \to \infty.$$

Corollary A.20. *Let $\{P, Q\}$ and $\{P_{(n)}, Q_{(n)}\}$ be binary experiments for all $n \in \mathbb{N}$. $\{P_{(n)}, Q_{(n)}\}$ converges weakly to $\{P, Q\}$ (as $n \to \infty$) if and only if*

$$\nu_{2,n} := \mathcal{L}\left(\log\left(\frac{\mathrm{d}Q_{(n)}}{\mathrm{d}P_{(n)}}\right) \,\Big|\, Q_{(n)}\right) \xrightarrow{w} \mathcal{L}\left(\log\left(\frac{\mathrm{d}Q}{\mathrm{d}P}\right) \,\Big|\, Q\right) = \nu_2 \quad \textit{as } n \to \infty.$$

Proof. By Theorem 16.8 in [58] it follows immediately that $\{P_{(n)}, Q_{(n)}\}$ converges weakly to $\{P, Q\}$ if and only if $\left\{Q_{(n)}, P_{(n)}\right\}$ converges weakly to $\{Q, P\}$. Furthermore,

$$\log\left(\frac{\mathrm{d}Q}{\mathrm{d}P}\right) = -\log\left(\frac{\mathrm{d}P}{\mathrm{d}Q}\right).$$

Finally, the desired equivalence follows. ■

Lemma A.21 (17.4 in [58]). *Let* $\{P, Q\}$ *and* $\{P_{(n)}, Q_{(n)}\}$ *be binary experiments for every* $n \in \mathbb{N}$*. If* $\{P_{(n)}, Q_{(n)}\}$ *converges weakly to* $\{P, Q\}$ *as* $n \to \infty$ *then*

$$d\left(P_{(n)}, Q_{(n)}\right) \to d\left(P, Q\right) \quad \text{as } n \to \infty.$$

A similar result is valid for the variational distance, see [58].

Lemma A.22. *Let* $\{P, Q\}$ *and* $\{P_{(n)}, Q_{(n)}\}$ *be binary experiments for every* $n \in \mathbb{N}$*. If* $\{P_{(n)}, Q_{(n)}\}$ *converges weakly to* $\{P, Q\}$ *as* $n \to \infty$ *then*

$$||P_{(n)} - Q_{(n)}|| \to ||P - Q|| \quad \text{as } n \to \infty.$$

Lemma A.23. *Let* $\{P_{(n)}, Q_{(n)}\}$ *be a binary experiment for all* $n \in \mathbb{N}$*. Then there exist a subsequence* $\{k_n : n \in \mathbb{N}\} \subseteq \mathbb{N}$ *and a binary experiment* $\{P, Q\}$ *such that* $\{P_{(k_n)}, Q_{(k_n)}\}$ *converges weakly to* $\{P, Q\}$ *as* $n \to \infty$ *.*

Proof. See Lemma 60.6 of Strasser [58]. ■

Definition A.24 (18.2 in [58]). *Let* $(\Omega_n, \mathcal{A}_n, \{P_{(n)}, Q_{(n)}\})$ *be a binary experiment for all* $n \in \mathbb{N}$*. The sequence* $(Q_{(n)})_{n\in\mathbb{N}}$ *is contiguous to* $(P_{(n)})_{n\in\mathbb{N}}$*, in symbols* $Q_{(n)} \lhd P_{(n)}$*, if for every sequence* $(A_n)_{n\in\mathbb{N}}$ *with* $A_n \in \mathcal{A}_n$ *for all* $n \in \mathbb{N}$

$$P_{(n)}(A_n) \to 0 \;\text{ implies }\; Q_{(n)}(A_n) \to 0 \;\text{ as } n \to \infty.$$

We call $(P_{(n)})_{n\in\mathbb{N}}$ *and* $(Q_{(n)})_{n\in\mathbb{N}}$ *mutually contiguous, in symbols* $P_{(n)} \lhd\rhd Q_{(n)}$*, if* $P_{(n)} \lhd Q_{(n)}$ *and* $Q_{(n)} \lhd P_{(n)}$*.*

Lemma A.25 (18.4 of [58]). *Let* $(\Omega_1, \mathcal{A}_1, \{P_{(1)}, Q_{(1)}\}), (\Omega_2, \mathcal{A}_2, \{P_{(2)}, Q_{(2)}\}), \ldots$ *be a sequence of binary experiments such that* $Q_{(n)} \lhd P_{(n)}$*. Let* $f_n : (\Omega_n, \mathcal{A}_n) \to (\mathbb{R}, \mathcal{B})$ *be a measurable function for all* $n \in \mathbb{N}$*. Then* $f_n \to 0$ *in* $P_{(n)}$*-probability implies* $f_n \to 0$ *in* $Q_{(n)}$*-probability (as* $n \to \infty$*).*

Lemma A.26 (First lemma of Le Cam). *Suppose that* $\{P_{(n)}, Q_{(n)}\}$ *converges weakly to* $\{P, Q\}$ *as* $n \to \infty$ *. Then the following statements (i)-(iv) are equivalent:*

(i) $Q_{(n)} \lhd P_{(n)}$. (ii) $Q \ll P$.

(iii) $\int_{\mathbb{R}} \exp \, d\nu_1 = 1$. (iv) $\nu_2\left(\{\infty\}\right) = 0$.

Moreover, the following statements (a)-(c) *are also equivalent:*

(a) $P_{(n)} \lhd Q_{(n)}$.

(b) $P \ll Q$.

(c) $\int_{\mathbb{R}} \exp(-x)\, \mathrm{d}\nu_2(x) = 1$.

(d) $\nu_1(\{-\infty\}) = 0$.

Proof. Theorem 18.11 in [58] shows the equivalence of (i) and (ii). The equivalence of (ii) and (iii) is mentioned in the introduction of §18 in [58] and is easy to verify. Finally, the equivalence of (iii) and (iv) follows immediately from Lemma A.18. The equivalence of (a)-(d) follows analogously. ■

A.4. Miscellaneous

In this section we present various results. Despite the first two ones, the results are not connected to each other.

Lemma A.27. *(i) For every* $x < 1$

$$\frac{x}{x-1} \;\le\; \log(1-x) \;\le\; -x.$$

(ii) For every $M \in (0,1)$ *there exist constants* $C_{M,1} > 1 > C_{M,2} > 0$ *such that*

$$\log(1-x) \;\ge\; -x\, C_{M,1} \quad \textit{and} \quad \log(1-y) \;\ge\; -y\, C_{M,2}.$$

for all $-M \le y \le 0 \le x \le M$. *Moreover, for* $j = 1, 2$

$$\lim_{M \searrow 0} C_{M,j} = 1. \tag{A.20}$$

Proof. (i) follows from the Mean Value Theorem. Fix $M \in (0,1)$. By (i)

$$\log(1-y) \;\ge\; \frac{-y}{1-y} \;\ge\; \frac{-y}{1+M} \text{ for all } y \in [-M, 0].$$

Using a Taylor's series expansion we obtain

$$\log(1-x) \;=\; -\sum_{k=1}^{\infty} \frac{x^k}{k} \;\ge\; -x \sum_{k=1}^{\infty} \frac{M^{k-1}}{k} \;=\; -x\, \frac{-\log(1-M)}{M}$$

for all $x \in [0, M]$. Define

$$C_{M,1} := \frac{-\log(1-M)}{M} \text{ and } C_{M,2} := \frac{1}{M+1}.$$

Obviously, (A.20) holds for $j = 2$. For $j = 1$ the left side (A.20) equals the derivative of $x \mapsto -\log(1-x)$ at 0 and so it is equal to 1. ■

The following equivalence is well known for sequences and can be easily extended to triangular arrays. For the readers, who are not familiar with the extended version, we also give the proof of it.

Lemma A.28. *Let $(h_{n,i})_{1 \le i \le n \in \mathbb{N}}$ be a triangular array of real numbers in $[0, \infty)$ such that*

$$\max_{1 \le i \le n} \{h_{n,i}\} \to 0 \text{ as } n \to \infty. \tag{A.21}$$

Then the following two conditions are equivalent:

$$(i) \quad \lim_{n\to\infty} \sum_{i=1}^{n} h_{n,i} = b \in \mathbb{R} \cup \{-\infty, \infty\}.$$

$$(ii) \quad \lim_{n\to\infty} \prod_{i=1}^{n} (1 - h_{n,i}) = \exp(-b) \in [0, \infty], \text{ where } \exp(-\infty) := 0.$$

Proof. We extend canonically the domain of log to $[0, \infty]$. Then (ii) is equivalent to

$$(iii) \quad \lim_{n\to\infty} \sum_{i=1}^{n} \log(1 - h_{n,i}) = -b.$$

By (A.21) and Lemma A.27

$$-C_{M,1} \sum_{i=1}^{n} h_{n,i} \quad \le \quad \sum_{i=1}^{n} \log(1 - h_{n,i}) \quad \le \quad -\sum_{i=1}^{n} h_{n,i}$$

for all $M \in (0, 1)$ if $n \in \mathbb{N}$ is sufficiently large. By this, (A.20) and basic calculations (i) and (iii) are equivalent, and so are (i) and (ii). ■

Lemma A.29. *Denote by φ, Φ and Φ^{-1} the density, the distribution function and the quantile function of a standard normal distributed random variable. Then*

$$\frac{x}{1+x^2} \quad \le \quad \frac{1-\Phi(x)}{\varphi(x)} \quad \le \quad \frac{1}{x} \tag{A.22}$$

for all $x > 0$. Furthermore,

$$1 - \Phi(x) \quad = \quad \frac{\varphi(x)}{x} \, \Big(1 + O(x^{-2}) \Big) \qquad \text{as } x \to \infty. \tag{A.23}$$

Moreover, for all sufficiently small $u > 0$

$$-\Phi^{-1}(u) = \Phi^{-1}(1-u) = \sqrt{2\log(u^{-1})} - \frac{\log(4\pi) + \log\log(u^{-1})}{2\sqrt{2\log(u^{-1})}} + \Psi(u), \tag{A.24}$$

$$\text{where } |\Psi(u)| \le \frac{9}{-\log(u)}. \tag{A.25}$$

Proof. (A.22) follows from inequality 7.1.13 of Abramowitz and Stegun [1], p. 298, and basic calculation. Since the proof is quite simple we give it nevertheless. From integration by parts we obtain for all $t > 0$

$$1 - \Phi(t) = \int_t^\infty \frac{1}{\sqrt{2\pi}} \frac{1}{x} \, x \exp\Big(-\frac{x^2}{2}\Big) \, \mathrm{d}x = \frac{1}{t}\varphi(t) - \int_t^\infty \frac{1}{x^2}\varphi(x) \, \mathrm{d}x.$$

Consequently, the upper bound in (A.22) follows immediately. Furthermore,

$$1 - \Phi(t) \ge \frac{1}{t}\varphi(t) - \frac{1}{t^2}(1 - \Phi(t)) \quad \text{and so} \quad \frac{t^2+1}{t^2}(1 - \Phi(t)) \ge \frac{1}{t}\varphi(t),$$

which proves the lower bound in (A.22). We deduce (A.23) from (A.22).

We can conclude from (A.22) that for every $x > 0$

$$1 - \Phi(x) \quad \le \quad \frac{\varphi(x)}{x} \quad = \quad \frac{1}{\sqrt{2\pi} x} e^{-\frac{x^2}{2}} \tag{A.26}$$

$$\text{and} \quad 1 - \Phi(x) \quad \ge \quad \frac{\varphi(x)}{x^{-1} + x} \quad = \quad \frac{1}{\sqrt{2\pi}(x^{-1} + x)} \, e^{-\frac{x^2}{2}}. \tag{A.27}$$

By Taylor's formula there exist $r_{x,1}, r_{x,2} \in [-|x|, |x|]$ for all $x \in (-1, 1)$ such that

$$\sqrt{1+x} = 1 + \frac{1}{2}x - \frac{x^2}{8}(1 + r_{x,1})^{-\frac{3}{2}} \tag{A.28}$$

$$\text{and} \quad \log(1+x) = x - \frac{x^2}{2(1 + r_{x,2})^2}. \tag{A.29}$$

In particular, $r_{x,j} \to 0$ as $x \to 0$ for $j \in \{1, 2\}$. Note that for $y > 0$

$$\log\Big(y + y^{-1}\Big) \; = \; \log(y) + \log\Big(1 + y^{-2}\Big). \tag{A.30}$$

If $|x|$ is sufficiently small and $y > 0$ is sufficiently large we have by (A.28)-(A.30) that

$$\begin{aligned} 1+\frac{1}{2}x-x^2 &\leq \sqrt{1+x} \leq 1+\frac{1}{2}x, && \text{(A.31)}\\ \frac{1}{2}\log(y) &\leq \log(y)-\frac{1}{4y^4} \leq \log\left(y+y^{-1}\right) \leq \log(y)+y^{-2} \leq \frac{4}{3}\log(y) && \text{(A.32)}\\ \text{and } \log\left(1-y^{-1}\right) &\geq -2y^{-1}. && \text{(A.33)} \end{aligned}$$

Note that Φ^{-1} is increasing and $\Phi^{-1}(1-u) \to \infty$ as $u \searrow 0$. Hence, by setting $u = 1-\Phi(x)$ for sufficiently large $x > 0$ we deduce from (A.26) and (A.31) that

$$\begin{aligned} \Phi^{-1}(1-u) &\leq \sqrt{-2\log\left(u\sqrt{2\pi}\Phi^{-1}(1-u)\right)}\\ &= \sqrt{-2\log(u)}\sqrt{1+\frac{\log(2\pi)+2\log\Phi^{-1}(1-u)}{2\log(u)}}\\ &\leq \sqrt{-2\log(u)}\left[1+\frac{1}{2}\frac{\log(2\pi)+2\log\Phi^{-1}(1-u)}{2\log(u)}\right]\\ &= \sqrt{-2\log(u)}-\frac{\log(2\pi)+2\log\Phi^{-1}(1-u)}{2\sqrt{-2\log(u)}} && \text{(A.34)}\\ \text{and in particular } \Phi^{-1}(1-u) &\leq \sqrt{-2\log(u)}. && \text{(A.35)} \end{aligned}$$

Note that for all sufficiently large $y > 0$

$$[\log(y)]^2 \leq y. \tag{A.36}$$

Similar to as in (A.34) we can conclude from (A.27), (A.31), (A.32), (A.35) and (A.36) that for all sufficiently small $u > 0$

$$\begin{aligned}
\Phi^{-1}(1-u) &\geq \sqrt{-2\log(u)}\sqrt{1+\frac{\log(2\pi)+2\log\left[\Phi^{-1}(1-u)+\Phi^{-1}(1-u)^{-1}\right]}{2\log(u)}} \\
&\geq \sqrt{-2\log(u)}\left[1+\frac{1}{2}\frac{\log(2\pi)+2\log\left[\Phi^{-1}(1-u)+\Phi^{-1}(1-u)^{-1}\right]}{2\log(u)}\right. \\
&\quad \left. -\left(\frac{3\log\left[\Phi^{-1}(1-u)+\Phi^{-1}(1-u)^{-1}\right]}{2\log(u)}\right)^2\right] \\
&\geq \sqrt{-2\log(u)}\left[1+\frac{\log(2\pi)+2\log\left[\Phi^{-1}(1-u)\right]+2\left[\Phi^{-1}(1-u)\right]^{-2}}{4\log(u)}\right. \\
&\quad \left. -4\left(\frac{\log\left[\Phi^{-1}(1-u)\right]}{\log(u)}\right)^2\right] \\
&\geq \sqrt{-2\log(u)}-\frac{\log(2\pi)+2\log\left[\Phi^{-1}(1-u)\right]}{2\sqrt{-2\log(u)}}-\frac{\left[\Phi^{-1}(1-u)\right]^{-2}}{\sqrt{-2\log(u)}}-\frac{16\,\Phi^{-1}(1-u)}{\left[-2\log(u)\right]^{\frac{3}{2}}} \\
&\geq \sqrt{-2\log(u)}-\frac{\log(4\pi)+\log\log\left(u^{-1}\right)}{2\sqrt{-2\log(u)}}-\frac{\left[\Phi^{-1}(1-u)\right]^{-2}}{\sqrt{-2\log(u)}}+\frac{8}{\log(u)} \qquad \text{(A.37)} \\
&\geq \frac{1}{2}\sqrt{-2\log(u)}. \qquad \text{(A.38)}
\end{aligned}$$

Combining (A.37) and (A.38) shows that for all sufficiently small $u > 0$

$$\begin{aligned}
\Phi^{-1}(1-u) &\geq \sqrt{-2\log(u)}-\frac{\log(4\pi)+\log\log\left(u^{-1}\right)}{2\sqrt{-2\log(u)}}-\frac{\left[\Phi^{-1}(1-u)\right]^{-2}}{\sqrt{-2\log(u)}}+\frac{8}{\log(u)} \\
&\geq \sqrt{-2\log(u)}-\frac{\log(4\pi)+\log\log\left(u^{-1}\right)}{2\sqrt{-2\log(u)}}+\frac{9}{\log(u)} \qquad \text{(A.39)} \\
&\geq \sqrt{-2\log(u)}\left[1-\left(-\log[u]\right)^{-\frac{1}{2}}\right]. \qquad \text{(A.40)}
\end{aligned}$$

By (A.33), (A.34) and (A.40)

$$\begin{aligned}
\Phi^{-1}(1-u) &\leq \sqrt{-2\log(u)}-\frac{\log(4\pi)+\log\log\left(u^{-1}\right)}{2\sqrt{-2\log(u)}}-\frac{\log\left[1-\left(-\log[u]\right)^{-\frac{1}{2}}\right]}{\sqrt{-2\log(u)}} \\
&\leq \sqrt{-2\log(u)}-\frac{\log(4\pi)+\log\log\left(u^{-1}\right)}{2\sqrt{-2\log(u)}}+\frac{2\left(-\log(u)\right)^{-\frac{1}{2}}}{\sqrt{-2\log(u)}}. \qquad \text{(A.41)}
\end{aligned}$$

Finally, (A.24) and (A.25) follow from (A.39) and (A.41). ■

The following lemma is an immediate consequence of Hölder's inequality. That is why we omit the proof and only present the result.

Lemma A.30. *Let* $x_1, \ldots, x_n \in \mathbb{R}$ *and* $p \in (1, \infty)$. *Then*

$$\left|\sum_{i=1}^{n} x_i\right| \leq \left(\sum_{i=1}^{n} |x_i|^p\right)^{\frac{1}{p}} n^{1-\frac{1}{p}} \text{ and, thus, } \frac{1}{n}\left(\sum_{i=1}^{n} x_i\right)^2 \leq \sum_{i=1}^{n} x_i^2.$$

The following lemma is well known. Nevertheless, we give a short proof of it.

Lemma A.31. *Let* $(\Omega_i, \mathcal{A}_i)$ *be a measurable space for* $i = 1, 2$. *Let* ν_1, ν_2 *be* σ*-finite measures on* $(\Omega_1, \mathcal{A}_1)$ *such that* $\nu_1 \ll \nu_2$. *Let* $T : (\Omega_1, \mathcal{A}_1) \to (\Omega_2, \mathcal{A}_2)$ *be a measurable mapping. Suppose that there is a further measurable mapping* $T^{-1} : (\Omega_2, \mathcal{A}_2) \to (\Omega_1, \mathcal{A}_1)$ *such that*

$$T^{-1} \circ T \;=\; id_{\Omega_1} \quad (\nu_2\text{-a.e.}). \tag{A.42}$$

Then

$$\frac{\mathrm{d}\nu_1^T}{\mathrm{d}\nu_2^T} = \frac{\mathrm{d}\nu_1}{\mathrm{d}\nu_2} \circ T^{-1} \quad (\nu_2^T\text{-a.e.}). \tag{A.43}$$

Proof. It is easy to see that $\nu_1^T \ll \nu_2^T$ follows from $\nu_1 \ll \nu_2$. Let $A \in \mathcal{A}_2$. From the transformation theorem for image measures and (A.42) we obtain

$$\nu_1^T(A) = \int_{T^{-1}(A)} \frac{\mathrm{d}\nu_1}{\mathrm{d}\nu_2}\,\mathrm{d}\nu_2 \;=\; \int_{T^{-1}(A)} \frac{\mathrm{d}\nu_1}{\mathrm{d}\nu_2} \circ T^{-1} \circ T \,\mathrm{d}\nu_2 = \int_A \frac{\mathrm{d}\nu_1}{\mathrm{d}\nu_2} \circ T^{-1}\,\mathrm{d}\nu_2^T.$$

Hence, (A.43) follows. ■

Lemma A.32. *Let* $F : \mathbb{R} \to \mathbb{R}$ *be some distribution function and* $F^{-1} : (0, 1) \to \mathbb{R}$ *be its left continuous quantile function, compare to* (2.5). *If*

$$F(t) \quad \sim_{asy} \; t^p L_1\left(\frac{1}{t}\right) \quad \text{as } t \searrow 0$$

for some slowly varying function $L_1 : (0, \infty) \to (0, \infty)$ *and some constant* $p > 0$ *then*

$$F^{-1}(u) \sim_{asy} \; u^{\frac{1}{p}} L_2\left(\frac{1}{t}\right) \quad \text{as } u \searrow 0$$

for some slowly varying function $L_2 : (0, \infty) \to (0, \infty)$.

Proof. Define $f : [1, \infty) \to (0, \infty)$ and $f^{\leftarrow} : [f(1), \infty) \to \mathbb{R}$ by

$$f(x) \;=\; \frac{1}{F\left(x^{-1}\right)} \text{ and } f^{\leftarrow}(s) := \inf\{y \geq 1 : f(y) > s\}$$

for all $x \geq 1$ and every $s \geq f(1)$. Note that

$$f(x) \ \sim_{\text{asy}} \ x^p L_1(x)^{-1} \ \text{ as } x \to \infty. \tag{A.44}$$

Applying Theorems 1.4.1 and 1.5.12 of Bingham et al. [8] we can conclude from (A.44)

$$f^{\leftarrow}(s) \ \sim_{\text{asy}} \ s^{\frac{1}{p}} L_3(s) \text{ as } s \to \infty$$

for some slowly varying function $L_3 : (0, \infty) \to (0, \infty)$. Moreover,

$$f^{\leftarrow}(s) \ = \ \inf\Big\{y \geq 1 : s^{-1} > F\left(y^{-1}\right)\Big\} \ = \ \inf\Big\{y \geq 1 : F^{-1}\left(s^{-1}\right) > y^{-1}\Big\} \ = \ \frac{1}{F^{-1}\left(s^{-1}\right)}$$

for all $s \geq f(1)$. Combining the last two statements completes the proof. ■

A.5. Explanatory calculation for heterogeneous normal mixtures

In this section we show how easy it is to apply our tool the heterogeneous normal mixture model by giving the proof of Theorem 5.1(iii) for $\tau = 1$. For this purpose we will apply Corollary 4.25. Let $y > 0$ be arbitrary but fixed. First, note that

$$\{x \in \mathbb{R} : \varepsilon_n f_n(x) > y\} = (\alpha_{n,y}, \infty) \ \text{ with } \ \alpha_{n,y} := \frac{(r+\beta)\log(n) + \log(y)}{\sqrt{2r\log(n)}}.$$

Since $\beta = r + \frac{1}{2}$ we obtain

$$n\varepsilon_n\mu_n\Big(x \in \mathbb{R} : \varepsilon_n f_n(x) > y\Big) = n^{1-\beta}\Big(1 - \Phi(\alpha_{n,y} - \vartheta_n)\Big) = n^{1-\beta}\Big(1 - \Phi(\gamma_{n,y})\Big),$$
$$\text{where } \gamma_{n,y} = \frac{\sqrt{\log(n)}}{2\sqrt{r}} + \frac{\log(y)}{\sqrt{2r\log(n)}}.$$

It is easy to see that $\gamma_{n,y} \to \infty$ as $n \to \infty$. Combining this and Lemma A.29 yield that for some constant $C_y > 0$

$$n\varepsilon_n\mu_n\Big(x \in \mathbb{R} : \varepsilon_n f_n(x) > y\Big) \sim_{\text{asy}} n^{\frac{1}{2}-r}\gamma_{n,y}^{-1}\varphi(\gamma_{n,y}) = \frac{C_y}{\gamma_{n,y}} n^{-\frac{1}{r}(r-\frac{1}{4})^2} \to 0 \text{ as } n \to \infty.$$

Consequently, (**B2 normal**) is fulfilled. Moreover, (**B3 normal**) holds for $y_0 = 1$:

$$\begin{aligned} n\varepsilon_n^2 \int f_n \mathbf{1}_{\{\varepsilon_n f_n \le 1\}} \, \mathrm{d}N(0,1) &= n^{1-2\beta} \int_{-\infty}^{\alpha_{n,1}} \exp(2\vartheta_n t - \vartheta_n^2) \, \mathrm{d}N(0,1)(t) \\ &= n^{1-2\beta} N(2\vartheta_n, 1)(-\infty, \alpha_{n,1}] \exp(\vartheta_n^2) \\ &= \Phi\left(\frac{\sqrt{\log(n)}}{\sqrt{2r}} \left[\frac{1}{2} - 2r \right] \right) \\ &\to \mathbf{1}_{(0,\frac{1}{4})}(r) + \frac{1}{2} \mathbf{1}_{\{\frac{1}{4}\}}(r) \text{ as } n \to \infty. \end{aligned}$$

Lists of Abbreviations and Symbols

Abbreviations

(**A**)	Condition (**A**), see p. 30.
(**A normal**)	Condition (**A normal**), see p. 41.
(**B**)	Condition (**B**), see p. 36.
(**B'**)	Condition (**B'**), see p. 38.
(**C**)	Condition (**C**), see p. 46.
a.s.	Almost surely.
f., ff.	Folio, folios following.
HC	Higher criticism test.
LLRT	Log-likelihood ratio test.
LLR_n	Test statistic of LLRT, see (3.3).
p., pp.	page, pages.
P-density of Q	Radon-Nikodym density of Q with respect to P.
i.i.d.	Independent and identically distributed.

Symbols

$\mathbf{1}_A$	Indicator function of A.
A^c	Complement of the set A.
$\|A\|, \#A$	Cardinality of the set A.
A°	Interior of the set A.
$a(P,Q)$	Affinity of P and Q, see A.12
$\mathcal{A}_1 \otimes \mathcal{A}_2$	The σ-algebra generated by the sets $A_1 \times A_2$ for all $A_1 \in \mathcal{A}_1$ and $A_2 \in \mathcal{A}_2$.
$\bigotimes \mathcal{A}_i$ $\mathcal{A}^n$	$\mathcal{A}_1 \otimes \ldots \otimes \mathcal{A}_n$ and $\bigotimes_{i=1}^n \mathcal{A}$, respectively.
$\mathcal{B}, \mathcal{B}(A), \bar{\mathcal{B}}$	Borel σ-algebra on $\mathbb{R}$, on A, on $\bar{\mathbb{R}}$, respectively.
$B(n,p)$	Binomial distribution with parameters $n \in \mathbb{N}$, $p \in [0,1]$.
$C(f), C(\eta)$	Set of continuity points of the function f and the function $x \mapsto \eta(-\infty,x)\mathbf{1}_{(-\infty,0)}(x) + \eta(x,\infty)\mathbf{1}_{(0,\infty)}(x)$ for a measure η, respectively.
$\mathbb{C}$	Complex numbers.
$d(P,Q)$	Hellinger distance between P and Q, see A.12
ϵ_x	Dirac measure centred on x.

$\operatorname{ess\,sup}_{x\in A} f(x)$	$\inf\{K \in \mathbb{R} : \lambda\!\!\lambda(x : f(x) > K) = 0\}$.
$Exp(\lambda)$	Exponential distribution with parameter $\lambda \in (0, \infty)$.
$f(t) \sim_{\text{asy}} g(t)$	The functions f and g are asymptotically equivalent, see Notation 2.18.
$f^{-1} : C \to B$	Inverse function of $f : B \to C$.
$f^{-1}(A)$	Image set of A with respect to the function $f : B \to C$, i.e., $f^{-1}(A) = \{x \in B : f(x) \in A\}$.
$f^{(k)}$	kth derivative of f.
Γ	The gamma function, see (2.11).
$\eta_1 \equiv \eta_2$	η_1 and η_2 are identically.
$\mathcal{H}_{0,n}, \mathcal{H}_{1,n}$	The general null and the general alternative, see (2.6).
$\mathbf{i}$	Imaginary unity, i.e., $\mathbf{i}^2 = -1$.
id_M	Identity function $id_M : M \to M$, i.e., $id_M(x) = x$ for all $x \in M$.
$\{k_n : n \in \mathbb{N}\}$	Subsequence of $\mathbb{N}$, see (4.1).
$\limsup\limits_{n\to\infty}$, $\liminf\limits_{n\to\infty}$	See the explanation at the end of Condition (**B**) on p. 36.
Λ	The distribution function of a standard Gumbel distribution, see (2.7).
$\mathbb{N}$	Natural numbers, $\mathbb{N} = \{1, 2, \ldots\}$.
$P \ll Q$	P is absolutely continuous with respect to Q.
$P \otimes Q$	The product measure of P and Q.
$\bigotimes_{i=1}^n P_i, P^n$	$P_1 \otimes \ldots \otimes P_n$ and $\bigotimes_{i=1}^n P$, respectively.
$P_n \lhd Q_n$	The sequence $(P_n)_{n\in\mathbb{N}}$ is contiguous to $(Q_n)_{n\in\mathbb{N}}$, see Definition A.24.
$P_n \lhd\rhd Q_n$	The sequences $(P_n)_{n\in\mathbb{N}}$ and $(Q_n)_{n\in\mathbb{N}}$ are mutually contiguous, see A.24.
$P^X = \mathcal{L}(X\|P)$	Image measure of X under P.
$P \perp Q$	The probability measures P and Q are singular.
Φ	The distribution function of a standard normal distribution.
$o(.), O(.)$	Landau symbols.
$\bar{\mathbb{R}}$, $[-\infty, \infty]$	$\mathbb{R} \cup \{-\infty, \infty\}$.
$\mathcal{U}(a, b)$	Uniform distribution on the interval (a, b).
Var_P	Variance with respect to the measure P, i.e., $\mathrm{Var}_P(X) = \mathbb{E}_P(X^2) - \mathbb{E}_P(X)^2$.
$X_{n,j}$	Projection to coordinate j, see (4.5).
$X \stackrel{D}{=} Y$	X and Y have the same distribution.
$X \sim P, F$	X is distributed according to the measure P and the distribution function F, respectively.
$\lfloor x \rfloor$	$\max\{m \in \mathbb{Z} : m \le x\}$.
$x \wedge y$	$\min(x, y)$.
$x \searrow y$, $x \nearrow y$	$y < x \to y$ and $y > x \to y$, respectively, see also Notation 2.18.
$\xrightarrow{D}, \xrightarrow{\mathrm{P}}$	Convergence in distribution and in P-probability, respectively.
$\xrightarrow{w}$	Weak convergence.
$\|\|P - Q\|\|$	Variational distance between P and Q, see A.12.

List of Figures

Bibliography

[1] Abramowitz, M. and I. A. Stegun, eds. (1970). *Handbook of mathematical functions with formulas, graphs, and mathematical tables*. Dover Publications, New York.

[2] Anderson, T. W. and D. A. Darling (1952). “Asymptotic theory of certain "goodness of fit" criteria based on stochastic processes”. *Ann. Math. Statist.* 23, pp. 193–212.

[3] Arias-Castro, E., E. J. Candès, and Y. Plan (2011). “Global testing under sparse alternatives: anova, multiple comparison and the higher criticism”. *Ann. Stat.* 39.5, pp. 2533–2556.

[4] Arias-Castro, E. and M. Wang (2013). “Distribution-Free Tests for Sparse Heterogeneous Mixtures”. *arXiv preprint arXiv:1308.0346*.

[5] Arias-Castro, E. and M. Wang (2015). “The sparse Poisson means model”. *Electronic Journal of Statistics* 9, pp. 2170–2201.

[6] Benjamini, Y. and Y. Hochberg (1995). “Controlling the false discovery rate: A practicle and powerful approach to multiple testing”. *J. Roy. Statist. Soc. Ser. B* 57, pp. 289–300.

[7] Billingsley, P. (1968). *Convergence of probability measures*. John Wiley and Sons.

[8] Bingham, N., C. Goldie, and J. Teugels (1987). *Regular Variation*. Cambridge University Press, Cambridge.

[9] Burnashev, M. V. and I. A. Begmatov (1991). “On a problem of detecting a signal leads to stable distributions”. *Theor. Probab. Appl.* 35, pp. 556–560.

[10] Cai, T., X. Jessie Jeng, and J. Jin (2011). “Optimal detection of heterogeneous and heteroscedastic mixtures”. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 73.5, pp. 629–662.

[11] Cai, T., J. Jin, and M. Low (2007). “Estimation and confidence sets for sparse normal mixtures”. *Ann. Statist.* 35.6, pp. 2421–2449.

[12] Cai, T. and Y. Wu (2014). “Optimal Detection of Sparse Mixtures Against a Given Null Distribution”. *IEEE Transactions on Information Theory* 60.4, pp. 2217–2232.

[13] Cayon, L., Jin, J., and Treaster, A. (2004). “Higher Criticism statistic: Detecting and identifying non-Gaussianity in the WMAP first year data”. *Mon. Not. Roy. Astron. Soc.* 362, pp. 826–832.

[14] Csörgő, M., S. Csörgő, L. Horváth, and D. Mason (1986). "Weighted empirical and quantile processes". *Ann. Probab.* 14.1, pp. 31–85.

[15] Csörgő, M. and P. Révész (1981). *Strong Approximations in Probability and Statistics.* Academic Press, New York, and Akadémiai Kiadó, Budapest.

[16] Dai, H., R. Charnigo, T. Srivastava, Z. Talebizadeh, and S. Qing (2012). "Integrating P-values for genetic and genomic data analysis". *J. Biom. Biostat.* 2012, pp. 3–7.

[17] Donoho, D. and J. Jin (2008). "Higher Criticism thresholding: Optimal feature selection when useful features are rare and weak". *Proc. Natl. Acad. Sci. USA* 105, pp. 14790–14795.

[18] Donoho, D. and J. Jin (2009). "Higher Criticism thresholding: Optimal phase diagram." *Philos. Trans. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.* 367, pp. 4449–4470.

[19] Donoho, D. and J. Jin (2015). "Higher Criticism for Large-Scale Inference, Especially for Rare and Weak Effects". *Statistical Science* 30.1, pp. 1–25.

[20] Donoho, D. and J. Jin (2004). "Higher criticism for detecting sparse heterogeneous mixtures". *Ann. Statist.* 32.3, pp. 962–994.

[21] Eicker, F. (1979). "The asymptotic distribution of the suprema of the standardized empirical processes". *Ann. Statist.* 7, pp. 116–138.

[22] Feller, W. (1966). *An Introduction to Probability Theory and Its Applications.* Vol. 2. John Wiley and Sons.

[23] El-Fishaway, P. (2013). "Common disease-common variant hypothesis". *Encyclopedia of Autism Spectral Disorder*, pp. 719–720.

[24] Goldstein, D. (2009). "Common genetic variation and human traits". *New England Journal of Medicine* 360, pp. 1696–1698.

[25] Gontscharuk, V., S. Landwehr, and H. Finner (2015). "The intermediates take it all. Asymptotics of higher criticism statistics and a powerful alternative based on equal local levels". *Biom. J.* 57.1, pp. 159–180.

[26] Gontscharuk, V., S. Landwehr, and H. Finner (2016). "Goodness of fit tests in terms of local levels with special emphasis on higher criticism tests". *Bernoulli* 22.3, pp. 1331–1363.

[27] Haan, L. de (1976). "Sample extremes: an elementary introduction". *Statistica Neerlandica* 30, pp. 161–172.

[28] Haan, L. de and A. Ferreira (2006). *Extreme Value Theory.* New York: Springer.

[29] Hall, P. and J. Jin (2008). "Properties of higher Criticism under strong dependence". *Ann. Statist.* 36.3, pp. 1686–1732.

[30] Hall, P. and J. Jin (2010). "Innovated higher criticism for detecting sparse signals in correleted noise". *Ann. Statist.* 38.3, pp. 1686–1732.

[31] Ingster, Y. and B. Tsybakov (2010). “Detection boundary in sparse regression”. *Electron. J. Stat.*

[32] Ingster, Y. (1997). “Some problem of hypothesis testest leading to infinitely divisible distribution”. *Math. Methods Statist.* 6, pp. 47–69.

[33] Iyengar, S. K. and R. Elston (2007). “The genetic basis of complex traits: Rare vvariant or "common gene, common disease"?” *Methods Mol. Biol.* 376, pp. 71–84.

[34] Jaeschke, D. (1979). “The asymptotic distribution of the suprema of the standardized empirical distribution function on subintervals”. *Ann. Stat.* 7.1, pp. 108–115.

[35] Jager, L. and J. Wellner (2007). “Goodness-of-fit tests via phi-divergences”. *Ann. Stat.* 35.5, pp. 2018–2053.

[36] Janssen, A. (1994). “Sums of independent triangular arrays and extreme order statistics”. *Ann. Prob.* 22.4, pp. 1766–1793.

[37] Janssen, A. (1998). *Zur Asymptotik nichtparametrischer Tests.* Münster.

[38] Janssen, A. and F. Marohn (1994). “On Statistical Information of Extreme Order Statistics, Local Extreme Value Alternatives, and Possion Point Processes”. *J. Multivariate Anal* 48.1, pp. 1–30.

[39] Janssen, A., H. Milbrodt, and H. Strasser (1985). *Infinitely Divisible Statistical Experiments.* Springer.

[40] Jin, J., Stark, J.-L., Donoho, D., Aghanim, N., and Forni, O. (2005). “Cosmological non-Gaussian signature detection: Comparing performance of different statistical tests”. *J. Appl. Signal Processing* 15, pp. 2470–2485.

[41] Jin, J. (2004). “Detecting a target in very noisy data from multiple looks.” *IMS Monography* 45, pp. 255–286.

[42] Karamata, J. (1930). “Sur un mode de croissance régulière des functions”. *Mathematica (Cluj)* 4. (french), pp. 38–53.

[43] Khmaladze, E. (1998). “Goodness of fit tests for Chimeric alternatives”. *Statistica Neerlandica* 52.1, pp. 90–111.

[44] Komlós, J., P. Major, and G. Tusnády (1975). “An approximation of partial sums of independent rv’s and the sample distribution function, I”. *Z. Wahrsch. verw. Geb.* 32, pp. 111–131.

[45] Komlós, J., P. Major, and G. Tusnády (1976). “An approximation of partial sums of independent rv’s and the sample distribution function, II”. *Z. Wahrsch. verw. Geb.* 34, pp. 33–58.

[46] Korevaar, J., T. van Aardenne Ehrenfest, and N. G. de Bruijn (1949). “A note on slowly oscillating functions”. *Nieuw Arch. Wiskunde* 23.2, pp. 77–86.

[47] Laurent, B., C. Marteau, and C. Maugis-Rabusseau (2015). "Multidimensional two component Gaussian mixtures detetection". arXiv:1509.09129.

[48] Le Cam, L. M. (1969). *Théorie asymptotique de la décision statistique*. Séminaire de Mathématiques Supérieures 33. Les Presses de l'Université de Montréal.

[49] Lóeve, M. (1977). *Probability Theory I*. 4th ed. Springer.

[50] Neill, D. and Lingwall, J. (2007). "A nonparametric scan staticistic for multivariate disease surveillance". *Advances in Disease Surveillance* 4, p. 106.

[51] Petrov, V. V. (1975). *Sums of Independent Random Variables*. Springer.

[52] Reiss, R. (1989). *Approximate Distributions of Order Statistics*. Springer.

[53] Resnick, S. (1987). *Extreme Value, Regular Variation and Point Processes*. Berlin : Springer.

[54] Resnick, S. (2007). *Heavy-Tail Phenomena*. New York: Springer.

[55] Saligrama, V. and Zhao, M. (2012). "Local anamaly detection". *JMLR W &CP* 22, pp. 969–983.

[56] Shorack, G. and J. Wellner (2009). *Empirical Processes with Applications to Statistics*. Society for Industrial and Applied Mathematics, Philadelphia.

[57] Stepanova, N. and T. Pavlenko (2016). "Goodness-of-fit tests based on sup-functionals of weighted empirical processes". arXiv:1406.0526v2.

[58] Strasser, H. (1985). *Mathematical Theory of Statistics*. Berlin/New York : De Gruyter.

[59] Tukey, J. W. (1976). "T13 N: The higher Criticism." Coures Notes. Stat 411. Princeton Univ.

[60] Tukey, J. W. (1989). "Higher Criticism for individual significances in serveral tables or parts of tables." Internal working paper, Princeton Univ.

[61] Tukey, J. W. (1994). *The Collected Works of John W. Tukey: Multiple Comparions, Volume VIII*. Ed. by H. Braun. Chapman and Hall, London.

[62] Verzelen, N. and E. Arias-Castro (2016). "Detection and Feature Selection in Sparse Mixture Models". arXiv:1405.1478.

[63] Wehrens, R. and P. Franceschi (2012). "Thresholding for biomarker selection in multivariate data using higher criticism". *Mol. Biosyst.* 8, pp. 2339–2346.

[64] Widder, D. (1946). *The Laplace transform*. Princeton university press.

[65] Witting, H. and G. Nölle (1970). *Angewandte Matematische Statistik: optimale finite und asymptotische Verfahren*. B. G. Teubner.

Erklärung

Ich versichere an Eides statt, dass die Dissertation von mir selbständig und ohne unzulässige fremde Hilfe unter Beachtung der „Grundsätze zur Sicherung guter wissenschaftlicher Praxis an der Heinrich-Heine-Universität Düsseldorf“ erstellt worden ist. Die Dissertation wurde in der vorgelegten oder ähnlicher Form noch bei keiner anderen Institution eingereicht. Ich habe bisher keine erfolglosen Promotionsversuche unternommen.

Düsseldorf, 24.November 2016

Marc Ditzhaus