Multi-level Bayes and MAP monotonicity testing

Yuri Golubev, Christophe Pouet

Multi-level Bayes and MAP monotonicity testing

Golubev Yu. * and Pouet C. †

September 20, 2019

Abstract

In this paper, we develop Bayes and maximum a posteriori probability (MAP) approaches to monotonicity testing. In order to simplify this problem, we consider a simple white Gaussian noise model and with the help of the Haar transform we reduce it to the equivalent problem of testing positivity of the Haar coefficients. This approach permits, in particular, to understand links between monotonicity testing and sparse vectors detection, to construct new tests, and to prove their optimality without supplementary assumptions. The main idea in our construction of multi-level tests is based on some invariance properties of specific probability distributions. Along with Bayes and MAP tests, we construct also adaptive multi-level tests that are free from the prior information about the sizes of non-monotonicity segments of the function.

Keywords: Haar transform, Bayes and MAP tests, multi-level hypothesis testing, stable distributions, type I and II error probabilities, critical signal-noise ratio.

AMS Subject Classification 2010: Primary 62C20; secondary 62J05.

1 Introduction

The literature on non-parametric monotonicity testing deals usually with the model

$$Y = f(X) + \xi,$$

*Institute for Information Transmission Problems, Moscow, Russia. e-mail: golubev.yuri@gmail.com

†Centrale Marseille, Aix Marseille Univ, CNRS, I2M, Marseille, France. e-mail: cpouet@centrale-marseille.fr

where Y is a scalar dependent random variable, X a scalar independent random variable, $f(\cdot)$ an unknown function, and ξ an unobserved scalar random variable with $\mathbf{E}\{\xi|X\} = 0$. We are interested in testing the null hypothesis, $\mathbf{H}_0$ that $f(x)$ is increasing against the alternative, $\mathbf{H}_1$ that there are x_1 and x_2 such that $x_1 < x_2$ and $f(x_1) > f(x_2)$. The decision is to be made based on the i.i.d. sample $\{X_i, Y_i\}_{1\le i\le n}$ from the distribution of (X, Y). Typical applications of monotonicity testing are related to econometric models, see, e.g., Chetverikov [4].

Usual approaches to this problem have in their core simple heuristic ideas and assumptions. So, the tests proposed in Gijbels et. al. [9] and Ghosal, Sen, and van der Vaart [8] are based on the signs of $(Y_{i+k} - Y_i)(X_{i+k} - X_i)$. Hall and Heckman [10] developed a test based on the slopes of local linear estimates of $f(\cdot)$. Along with these papers we can cite Schlee [15], Bowman, Jones, and Gijbels [2], Dümbgen and Spokoiny [6], Durot [7], Baraud, Huet, and Laurent [1], Wang and Meyer [17], and Chetverikov [4]. As to typical hypothesis about $f(\cdot)$, it is often assumed that $f(x)$ is a Lipschitz function, i.e.,

$$|f(y) - f(x)| \le L|y - x|,$$

where the constant $L < \infty$ may be known or unknown.

In this paper, we look at the problem of monotonicity testing from a little different and less intuitive viewpoint. As we will see below, our approach permits, in particular, to understand links between this problem and sparse vectors detection and to construct new powerful tests. In order to simplify technical details and to get rid of supplementary assumptions, we begin with monotonicity testing of an unknown function $f(t)$, $t \in [0, 1]$, in the so-called white noise model similar to that one considered in [6]. So, it is assumed we have at our disposal the noisy data

$$Y(t) = f(t) + \sigma n(t),\ \ t \in [0, 1], \tag{1}$$

where $n(\cdot)$ is a standard white Gaussian noise and $\sigma > 0$ is a known noise level. With the help of these observations we want to test

the null hypothesis

$\mathbf{H}_0:\ \ f'(t) \ge 0$, for all $t \in [0, 1]$,

vs. the alternative

$\mathbf{H}_1:\ \ f'(t) < 0$, for some $t \in [0, 1]$.

Our approach to this problem is based on estimating the following linear functionals:

$$\theta_{h,t}(f) \stackrel{\text{def}}{=} \frac{1}{h}\int_t^{t+h} f(u)\,du - \frac{1}{h}\int_{t-h}^{t} f(u)\,du$$

for all h, t that are admissible, i.e., such that $[t-h, t+h] \subseteq [0,1]$. It is clear that $\theta_{h,t}(f)/h$ may be interpreted as approximations of the derivative $f'(t)$ since

$$\lim_{h \to 0} \frac{\theta_{h,t}(f)}{h} = f'(t),$$

for any given $t \in (0,1)$.

With the help of (1), the functionals $\theta_{h,t}(f)$ are estimated as follows:

$$\hat{\theta}_{h,t}(Y) = \frac{1}{h} \int_t^{t+h} Y(u)\, du - \frac{1}{h} \int_{t-h}^{t} Y(u)\, du$$

and these estimates admit the obvious representation

$$\hat{\theta}_{h,t}(Y) = \theta_{h,t}(f) + \sigma_h \xi_{h,t}, \tag{2}$$

where

$$\sigma_h = \sigma \sqrt{\frac{2}{h}}, \quad \xi_{h,t} = \frac{1}{\sqrt{2h}} \left[\int_t^{t+h} n(u)\, du - \int_{t-h}^{t} n(u)\, du \right] \sim \mathcal{N}(0,1).$$

Notice that if $\mathbf{H}_0$ is true, then $\theta_{h,t}(f) \geq 0$ for all admissible h, t, otherwise ($\mathbf{H}_1$ is true) there exist h', t' such that $\theta_{h',t'}(f) < 0$. That is why in what follows we will focus on testing

$$\begin{aligned} &\textit{the null hypothesis} \\ &\mathbf{H}_0 : \ \theta_{h,t}(f) \geq 0, \ \textit{for all admissible } h, t \\ &\textit{vs. the alternative} \\ &\mathbf{H}_1 : \ \theta_{h,t}(f) < 0, \ \textit{for some admissible } h, t \end{aligned} \tag{3}$$

based on the observations (2).

Let us denote for brevity

$$\theta_{h,t} = \theta_{h,t}(f), \quad \hat{\theta}_{h,t} = \hat{\theta}_{h,t}(Y).$$

In order to explain our approach to the problem (3), we begin with the simple case assuming that h, t are given. So, we have to test two composite hypotheses

$$\mathbf{H}_0^{h,t} : \theta_{h,t} \geq 0 \ \text{ vs. } \mathbf{H}_1^{h,t} : \theta_{h,t} < 0.$$

Intuitively, the most powerful test with the type I error probability α rejects $\mathbf{H}_0^{h,t}$ if

$$\hat{\theta}_{h,t} \leq -\sigma_h t_\alpha, \tag{4}$$

where t_α is α-value of the standard Gaussian distribution, i.e., a solution to

$$\Phi(t_\alpha) = 1 - \alpha,$$

where

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp\left(-\frac{x^2}{2}\right) dx.$$

Of course, there exist a lot of motivations for this test. In this paper, we make use of the so-called improper Bayes approach assuming that $\theta_{h,t}$ in (2) is a random variable uniformly distributed on the interval $[0, A]$, $A > 0$, if $\mathbf{H}_0^{h,t}$ is true, and on $[-A, 0]$ if $\mathbf{H}_1^{h,t}$ is true. So, we observe a random variable $\hat{\theta}_{h,t}$ with the probability density

$$p_0^A(x|\mathbf{H}_0^{h,t} \text{ is true}) = \frac{1}{A} \int_0^A \exp\left[-\frac{(x-\theta)^2}{2\sigma_h^2}\right] d\theta$$

and

$$p_1^A(x|\mathbf{H}_1^{h,t} \text{ is true}) = \frac{1}{A} \int_{-A}^0 \exp\left[-\frac{(x-\theta)^2}{2\sigma_h^2}\right] d\theta.$$

Thus, we deal with the simple hypothesis testing and by the Neyman-Pearson lemma, the most powerful test at significance level α rejects $\mathbf{H}_0^{h,t}$ when

$$\frac{p_1^A(\hat{\theta}_{h,t})}{p_0^A(\hat{\theta}_{h,t})} \geq t_\alpha^A.$$

Taking the limit in this equation as $A \to \infty$, we arrive at the improper Bayes test that rejects $\mathbf{H}_0^{h,t}$ if

$$S\left(\frac{\hat{\theta}_{h,t}}{\sigma_h}\right) \geq t'_\alpha, \tag{5}$$

where

$$S(x) = \frac{\displaystyle\int_{-\infty}^{0} \exp\left[-(x-\theta)^2/2\right] d\theta}{\displaystyle\int_{0}^{\infty} \exp\left[-(x-\theta)^2/2\right] d\theta} = \frac{1}{\Phi(x)} - 1. \tag{6}$$

Since $S(x)$ is decreasing in $x \in \mathbb{R}$, the tests (4) and (5) are obviously equivalent.

In what follows, we will make use of the following asymptotic result:

$$S(x) = \left[1 + O\left(\frac{1}{x^2}\right)\right] \sqrt{2\pi}(1-x) \exp\left(\frac{x^2}{2}\right), \text{ as } x \to -\infty. \tag{7}$$

Along with this method, one can apply the maximum likelihood (ML) or minimax approaches. Finally, all these methods result in (4) but their initial forms are different. For instance, the ML test rejects $\mathbf{H}_0^{h,t}$ when

$$\frac{\max_{\theta<0} \exp\{-(\hat{\theta}_{h,t}-\theta)^2/(2\sigma_h^2)\}}{\max_{\theta>0} \exp\{-(\hat{\theta}_{h,t}-\theta)^2/(2\sigma_h^2)\}} = \exp\biggl\{-\frac{\hat{\theta}_{h,t}^2}{2\sigma_h^2}\operatorname{sign}(\hat{\theta}_{h,t})\biggr\} \geq t''_\alpha. \tag{8}$$

Emphasize that from a viewpoint of testing $\mathbf{H}_0^{h,t}$ vs. $\mathbf{H}_1^{h,t}$ there is no difference between (8) and (5), but the aggregation of these methods for testing $\mathbf{H}_0$ vs. $\mathbf{H}_1$ from (3) results in different tests. In this paper, we make use of the tests defined by (5) since their aggregation is simple.

In order to aggregate the statistical tests, we will make use of the so-called multi-resolution approach assuming that

1. *h belongs to the following set of dyadic bandwidths*

$$\mathcal{H} \stackrel{\text{def}}{=} \biggl\{\frac{1}{2}, \frac{1}{4}, \dots \frac{1}{2^k}, \dots\biggr\};$$

2. *t belongs to the family of dyadic grids $\mathcal{G}_h$, $h \in \mathcal{H}$, defined by*

$$\mathcal{G}_h \stackrel{\text{def}}{=} \{h, 3h, \dots, 1-h\},\ h \in \mathcal{H}.$$

There are simple arguments motivating these assumptions

- random variables $\xi_{h,t}$ and $\xi_{h',t'}$ in (2) are independent if $\{h,t\} \neq \{h',t'\}$. This fact simplifies significantly the statistical analysis of tests.
- $\sqrt{h/2}\,\hat{\theta}_{h,t}$ are the Haar coefficients admitting a fast computation in the discrete version of (1).

2 Testing at a given resolution level

Let us fix some bandwidth $h \in \mathcal{H}$ and denote for brevity by $n_h = 1/(2h)$. In this section, we focus on testing

the null hypothesis

$\mathbf{H}_0^h: \ \theta_{h,t} \geq 0$ for all $t \in \mathcal{G}_h$

vs. the alternative

$\mathbf{H}_1^h: \ \theta_{h,t} < 0$ for some $t \in \mathcal{G}_h$.

In order to construct Bayes and MAP tests, we assume that for given $h \in \mathcal{H}$

- the set $\{\theta_{h,t},\ t \in \mathcal{G}_h\}$ contains the only one negative entry $\theta_{h,\tau}$;
- τ is an unobservable random variable uniformly distributed on $\mathcal{G}_h$.

2.1 A Bayes test

With the arguments used in deriving (5), we get the following Bayes test: $\mathbf{H}_0^h$ is rejected if

$$\frac{1}{n_h} \sum_{t \in \mathcal{G}_h} S\biggl(\frac{\hat{\theta}_{h,t}}{\sigma_h}\biggr) \geq t_\alpha^B,$$

where $S(\cdot)$ is defined by (6). The critical level t_α^B is defined by a conservative way, i.e., as a solution to

$$\max_{\Theta \geq 0} \mathbf{P}_\Theta \biggl\{ \frac{1}{n_h} \sum_{t \in \mathcal{G}_h} S\biggl(\frac{\hat{\theta}_{h,t}}{\sigma_h}\biggr) > t_\alpha^B \biggr\} = \alpha,$$

where here $\mathbf{P}_\Theta$ stands for the measure generated by observations $\hat{\theta}_{h,t}$ defined by (2) for given $\Theta = \{\theta_{h,t}, h \in \mathcal{H}, t \in \mathcal{G}_h\}$.

It follows from Mudholkar's theorem [12], see also Theorem 6.2.1 in [16], that for any Θ with nonnegative entries $\theta_{h,t} \geq 0$

$$\mathbf{P}_\Theta \biggl\{ \frac{1}{n_h} \sum_{t \in \mathcal{G}_h} S\biggl(\frac{\hat{\theta}_{h,t}}{\sigma_h}\biggr) > x \biggr\} \leq \mathbf{P} \biggl\{ \frac{1}{n_h} \sum_{t \in \mathcal{G}_h} S(\xi_{h,t}) \geq x \biggr\} \tag{9}$$

and, thus, t_α^B may be computed as a solution to

$$\mathbf{P} \biggl\{ \frac{1}{n_h} \sum_{t \in \mathcal{G}_h} S(\xi_{h,t}) \geq t_\alpha^B \biggr\} = \alpha. \tag{10}$$

Therefore our next step is to study the following random variable:

$$B_h(\xi) \stackrel{\text{def}}{=} \frac{1}{n_h} \sum_{t \in \mathcal{G}_h} S(\xi_{h,t}).$$

2.1.1 A weak approximation of $B_h(\xi)$

We begin with computing a weak limit of $B_h(\xi)$ as $h \to 0$. Recall some standard definitions (see, e.g., [13]).

Definition. *Let X_1 and X_2 be independent copies of a random variable X. Then X is said to be stable if for any constants $a > 0$ and $b > 0$ the random variable $aX_1 + bX_2$ has the same distribution as $cX + d$ for some constants $c > 0$ and d.*

In the class of stable distributions there is an interesting sub-class of the so-called stable distributions with the index of stability $\alpha = 1$. For brevity, we will call them *1-stable distributions*. The formal definition of this class is as follows:

Definition. *A random variable X is called 1-stable if its characteristic function can be written as*

$$\mathbf{E}\exp(\mathrm{i}tX) = \exp\Bigl(\mu\mathrm{i}t - |ct| - \mathrm{i}\frac{2\beta|c|}{\pi}t\log(|t|)\Bigr). \tag{11}$$

The next theorem shows that the weak limit of $B_h(\xi) - \log(n_h)$ is a 1-stable distribution.

Theorem 1.

$$\lim_{h\to 0}\mathbf{E}\exp\bigl\{\mathrm{i}t\bigl[B_h(\xi) - \log(n_h) + \gamma\bigr]\bigr\} = \exp\Bigl\{\mathrm{i}t\log\frac{1}{|t|} - \frac{\pi|t|}{2}\Bigr\},$$

where $\gamma \approx 0.57721$ is Euler's constant.

In other words, this theorem states that

$$\lim_{h\to 0}\bigl[B_h(\xi) - \log(n_h) + \gamma\bigr] \stackrel{\mathcal{D}}{=} \zeta,$$

where ζ is a 1-stable random variable (see (11)) with

$$\mu = 0,\ c = \frac{\pi}{2},\ \beta = 1. \tag{12}$$

Apparently, ζ appeared firstly in [5]. Emphasize also that this random variable originate usually in Bayes hypothesis testing related to sparse vectors, see e.g. [3], [11].

The probability distribution of ζ has the following invariance property that plays an important role in Bayes tests aggregation.

Proposition 1. *Let ζ_k be i.i.d. copies of ζ and $\bar{\pi}$ be a probability distribution on $\mathbb{Z}^+$ with a bounded entropy. Then*

$$\sum_{k=1}^{\infty}\bar{\pi}_k\Bigl(\zeta_k - \log\frac{1}{\bar{\pi}_k}\Bigr) \stackrel{\mathcal{D}}{=} \zeta. \tag{13}$$

The proof of (13) follows immediately from (11) and (12).

2.1.2 A strong approximation of $B_h(\xi)$

Theorem 1 is not very informative about the tail behavior of the distribution of $B_h(\xi)$. However, for obtaining a good approximation of t_α^B in (10) this behavior may play a crucial role because in some applications α may be very small (of order 10^{-7}) and so, the Monte-Carlo method and Theorem 1 may not be good in this case.

Therefore our goal is to find an approximation of $B_h(\xi)$ that controls well the tail of its distribution. Fortunately, this can be easily done. It is clear that

$$\Phi(\xi_k) \stackrel{\mathcal{D}}{=} U_k,$$

where U_k are i.i.d. random variables uniformly distributed on $[0,1]$. Hence

$$B_h(\xi) \stackrel{\mathcal{D}}{=} \frac{1}{n_h}\sum_{k=1}^{n_h}\Big[\frac{1}{U_k}-1\Big] = \frac{1}{n_h}\sum_{i=1}^{n_h}\frac{1}{U_{(k)}}-1,$$

where $U_{(k)}$ is a non-decreasing permutation of $U_k,\ k=1,\dots,n_h$. The distribution of $U_{(1)},U_{(2)},\dots,U_{(n_h)}$ can be easily obtained with the help of the Pyke theorem [14]

$$U_{(k)} \stackrel{\mathcal{D}}{=} \frac{\mathcal{E}_k}{\mathcal{E}_{n_h+1}}, \tag{14}$$

where

$$\mathcal{E}_k = \sum_{l=1}^{k}\varkappa_l$$

is the cumulative sum of i.i.d. standard exponentially distributed random variables $\varkappa_l$

$$\mathbf{P}\{\varkappa_l \ge y\} = \exp(-y).$$

In other words, $\mathcal{E}_k \sim \mathrm{Gamma}(k,1)$. With this in mind, we obtain

$$\begin{aligned} B_h(\xi) &\stackrel{\mathcal{D}}{=} \Big[1+O\Big(\frac{1}{\sqrt{n_h}}\Big)\Big]\sum_{k=1}^{n_h}\frac{1}{\mathcal{E}_k}-1 \\ &= \Big[1+O\Big(\frac{1}{\sqrt{n_h}}\Big)\Big]\Big[\sum_{k=1}^{n_h}\Big(\frac{1}{\mathcal{E}_k}-\frac{1}{k}\Big)+\sum_{k=1}^{n_h}\frac{1}{k}\Big]-1. \end{aligned} \tag{15}$$

Next, we make use of the following simple equations:

$$\Big\{\mathbf{E}\Big[\sum_{k=n_h+1}^{\infty}\Big(\frac{1}{\mathcal{E}_k}-\frac{1}{k}\Big)\Big]^{2m}\Big\}^{1/(2m)} \le O\Big(\frac{1}{\sqrt{n_h}}\Big)$$

and

$$\sum_{k=1}^{n_h} \frac{1}{k} = \log(n_h) + \gamma + O\Big(\frac{1}{n_h}\Big).$$

So, substituting them in (15), we arrive at the following theorem.

Theorem 2. *Let*

$$\zeta^\circ = \sum_{k=1}^{\infty} \Big(\frac{1}{\mathcal{E}_k} - \frac{1}{k}\Big). \tag{16}$$

Then

$$B_h(\xi) - \log(n_h) + \gamma \overset{\mathcal{D}}{=} [1 + O(\varepsilon_h)]\zeta^\circ + 2\gamma - 1 + O(\varepsilon_h)\log(n_h), \tag{17}$$

where ε_h *is such that*

$$\big[\mathbf{E}(\varepsilon_h^{2m})\big]^{1/(2m)} \le \frac{C_m}{\sqrt{n_h}}. \tag{18}$$

Remark. *The random variable* ζ *in Theorem 1 admits the following representation*

$$\zeta = \sum_{k=1}^{\infty} \Big(\frac{1}{\mathcal{E}_k} - \frac{1}{k}\Big) + 2\gamma - 1.$$

Notice also that it follows immediately from (17) that convergence rate in Theorem 1 is $\log(n_h)/\sqrt{n_h}$*, i.e., as* $h \to 0$*,*

$$\mathbf{E}\exp\{\mathrm{i}t\big[B_h(\xi) - \log(n_h) + \gamma\big]\} = \exp\Big\{\mathrm{i}t\log\frac{1}{|t|} - \frac{\pi|t|}{2} + O\Big(\frac{\log(n_h)}{\sqrt{n_h}}\Big)\Big\}.$$

Figure 1 illustrates numerically Theorem 2 and the above remark showing log-tail approximation error

$$\Delta(x; n_h) = \log\Big[\mathbf{P}\{B_h\{\xi) - \log(n_h) + \gamma \ge x\}\Big] - \log\Big[\mathbf{P}\{\zeta^\circ + 2\gamma - 1 \ge x\}\Big].$$

computed with the help of the Monte-Carlo method with $0.5 \cdot 10^6$ replications. This picture shows that even for small $n_h = 4$ the approximation (17) works very good.

2.2 A MAP test

Similarly to the Bayes test, we can construct the MAP test that rejects $\mathbf{H}_0^h$ if

$$\max_{t\in\mathcal{G}_h} \frac{1}{n_h} S\Big(\frac{\hat{\theta}_{h,t}}{\sigma_h}\Big) \ge t_\alpha^M,$$

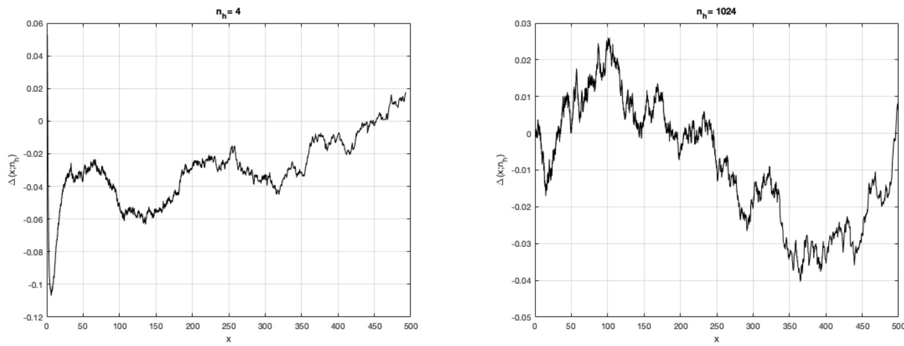

Figure 1: Log-tail approximation errors $\Delta(x; n_h)$ for $n_h = 4$ and $n_h = 1024$.

where t_α^M is defined as a solution to

$$\max_{\Theta \geq 0} \mathbf{P}_\Theta \biggl\{ \max_{t \in \mathcal{G}_h} \frac{1}{n_h} S\biggl(\frac{\hat{\theta}_{h,t}}{\sigma_h}\biggr) > t_\alpha^M \biggr\} = \alpha.$$

Similarly to (9), t_α^M may be obtained from

$$\mathbf{P}\biggl\{ \max_{t \in \mathcal{G}_h} \frac{S(\xi_{h,t})}{n_h} > t_\alpha^M \biggr\} = \alpha.$$

As to the limit distribution of $\max_{t \in \mathcal{G}_h} S(\xi_{h,t})/n_h$, as $h \to 0$, it follows immediately from (14) that

$$\max_{t \in \mathcal{G}_h} \frac{S(\xi_{h,t})}{n_h} \stackrel{\mathcal{D}}{=} \frac{1 + \varepsilon_h}{\varkappa}, \text{ as } h \to 0, \tag{19}$$

where $\varkappa$ is a standard exponential random variable and ε_h satisfies (18).

3 Multi-level testing

3.1 MAP multi-level tests

A heuristic idea behind our construction of multi-level MAP tests for (3) is related to (19) and consists in computing a positive deterministic function U_h, $h \in \mathcal{H}$, bounding from above the random process $\log(1/\varkappa_h)$, $h \in \mathcal{H}$, where $\varkappa_h$ are independent standard exponential random variables. In other words, we are looking for U_h such that

$$\zeta^U = \sup_{h \in \mathcal{H}} \biggl[\log \frac{1}{\varkappa_h} - U_h \biggr]$$

would be a non-degenerate random variable.

Let q_α^U be α-value of ζ^U, i.e., solution to

$$\mathbf{P}\{\zeta^U \geq q_\alpha^U\} = \alpha.$$

Therefore with (19), upper bounding random process $\log(1/\varkappa_h)$ by U_h, we arrive at the test that rejects $\mathbf{H}_0$ if

$$\sup_{h\in\mathcal{H}}\biggl\{\max_{t\in\mathcal{G}_h}\log\biggl[\frac{1}{n_h}S\biggl(\frac{\hat{\theta}_{h,t}}{\sigma_h}\biggr)\biggr] - U_h\biggr\} \geq q_\alpha^U. \tag{20}$$

Computing q_α^U is based on the following simple fact. Assume that

$$K^U = \log\biggl[\sum_{h\in\mathcal{H}} \mathrm{e}^{-U_h}\biggr] < \infty.$$

Then

$$\sup_{h\in\mathcal{H}}\biggl[\log\frac{1}{\varkappa_h} - U_h\biggr] - K^U \stackrel{\mathcal{D}}{=} \log\frac{1}{\varkappa}. \tag{21}$$

The proof of this identity is very simple. Indeed,

$$\begin{aligned}
&\mathbf{P}\Bigl\{\sup_{h\in\mathcal{H}}[\log\frac{1}{\varkappa_h} - U_h] - K^U > x\Bigr\} \\
&\quad= 1 - \prod_{h\in\mathcal{H}}\mathbf{P}\Bigl\{\log\frac{1}{\varkappa_h} \leq U_h + x + K^U\Bigr\} \\
&\quad= 1 - \exp\biggl\{-\sum_{h\in\mathcal{H}}\exp\Bigl[-x - U_h - K^U\Bigr]\biggr\} = 1 - \exp[-\exp(-x)].
\end{aligned}$$

Let us we denote

$$\bar{\pi}_h = \mathrm{e}^{-U_h}\Big/\sum_{h\in\mathcal{H}}\mathrm{e}^{-U_h},$$

then (21) can be rewritten in the following form:

Proposition 2. *Let $\bar{\pi}$ be a probability distribution on $\mathcal{H}$. Then*

$$\sup_{h\in\mathcal{H}}\biggl[\log\frac{1}{\varkappa_h} + \log(\bar{\pi}_h)\biggr] \stackrel{\mathcal{D}}{=} \log\frac{1}{\varkappa}. \tag{22}$$

Therefore with the help of (21) we can compute α-critical level q_α^U in (20)

$$q_\alpha^U = q_\alpha^\varkappa + K^U,$$

where

$$q_\alpha^{\varkappa} = -\log\biggl(\log\frac{1}{1-\alpha}\biggr)$$

is α-value of $\log(1/\varkappa)$.

Summarizing (see (20)), the MAP multi-level test rejects $\mathbf{H}_0$ if

$$\sup_{h\in\mathcal{H}}\bigl\{Z_h^M + \log(\bar{\pi}_h)\bigr\} \ge q_\alpha^{\varkappa}, \tag{23}$$

where

$$Z_h^M = \max_{t\in\mathcal{G}_h}\log\biggl[\frac{1}{n_h}S\biggl(\frac{\hat{\theta}_{h,t}}{\sigma_h}\biggr)\biggr]$$

and $\bar{\pi}$ is a probability distribution on $\mathcal{H}$.

In order to study the performance of this method, we analyze the type II error probability. For given $\{\rho,\tau:\rho\in\mathcal{H},\ \tau\in\mathcal{G}_g\}$ and $A\in\mathbb{R}^+$ define

$$\Theta_{\rho,\tau}(A) = \Bigl\{\theta_{h,t}:\theta_{\rho,\tau}=-A;\ \theta_{h,t}\ge 0,(h,t)\ne(\rho,\tau)\Bigr\}. \tag{24}$$

In other words, we consider the situation, where all shifts $\theta_{h,t}$ in (2) are positive except the only one. The position of the negative entry $\{\rho,\tau\}$ and its amplitude are unknown, but it is assumed that $\{\rho,\tau\}$ are random variables with the distribution defined by

- $\mathbf{P}\{\rho=h\}=\bar{\pi}_h$,
- $\mathbf{P}\{\tau=t|\rho=h\}=n_h^{-1}$,

where $\bar{\pi}$ is a probability distribution on $\mathcal{H}$ with a bounded entropy

$$H_{\bar{\pi}} = \sum_{h\in\mathcal{H}}\bar{\pi}_h\log\frac{1}{\bar{\pi}_h}.$$

In what follows, we will deal with priors $\bar{\pi}$ with large uncertainties assuming that $\bar{\pi}\to 0$, or more precisely, $\sup_{h\in\mathcal{H}}\bar{\pi}_h\to 0$, but such that

$$\lim_{\bar{\pi}\to 0}\frac{1}{\log[H_{\bar{\pi}}]}\sum_{h\in\mathcal{H}}\bar{\pi}_h\biggl|H_{\bar{\pi}}-\log\frac{1}{\bar{\pi}_h}\biggr| = 0. \tag{25}$$

In particular, we will consider the following class of prior distributions:

$$\bar{\pi}_h = \bar{\pi}_h^{\omega,\nu} = \nu\biggl[\frac{\log_2(1/h)}{\omega}\biggr]\bigg/\sum_{k=1}^{\infty}\nu\biggl(\frac{k}{\omega}\biggr) \approx \frac{1}{\omega}\nu\biggl[\frac{\log_2(1/h)}{\omega}\biggr]. \tag{26}$$

This class is characterized by the bandwidth $\omega > 1$ and the probability density $\nu(x),\ x \in \mathbb{R}^+$, which is assumed to be continuous, bounded, and with

$$H_\nu = \int_0^\infty \nu(x) \log \frac{1}{\nu(x)} dx < \infty, \qquad \int_0^\infty \nu(x) \log(x+1)\, dx < \infty. \tag{27}$$

A typical example of a such distribution is the uniform one that corresponds to $\nu(x) = 1,\ x \in [0, 1]$.

It is clear that $\bar\pi_h^{\omega,\nu} \to 0$ as $\omega \to \infty$ and that Condition (25) holds.

Let us begin with the case, where the prior distribution is known, the case of unknown $\bar\pi$ will be considered later in Section 4.

The type II error probability over $\Theta_{\rho,\tau}(A)$ of the MAP test (23) is defined as follows:

$$\beta^M_{\rho,\tau}(A) = \sup_{\Theta\in\Theta_{\rho,\tau}(A)} \mathbf{P}_\Theta \Big\{ \max_{h\in\mathcal{H}} [Z_h^M + \log(\bar\pi_h)] \le q_\alpha^\varkappa \Big\}.$$

Our goal is to study the average type II error probability

$$\bar\beta^M_{\bar\pi}(\mathbf{A}) = \sum_{h\in\mathcal{H}} \frac{\bar\pi_h}{n_h} \sum_{t\in\mathcal{G}_h} \beta^M_{h,t}(A_h),$$

where here and below $\mathbf{A} = \{A_h,\ h \in \mathcal{H}\}$.

Denote for brevity

$$R_h(q, H) = 2[q + \log(n_h) + H] - \log[4\pi(q + \log(n_h) + H)]$$

and

$$\log^*(x) = \log[\log(x)], \quad H^*_{\bar\pi} = \log(H_{\bar\pi}).$$

The next theorem shows that $R_h(q_\alpha^\varkappa, H_{\bar\pi})$ is a critical signal/noise ratio. Roughly speaking, this means that if

$$\frac{A_h}{\sigma_h} \overset{\bar\pi}{\le} \sqrt{R_h(q_\alpha^\varkappa, H_{\bar\pi})} + x$$

for any given $x > 0$, then the MAP multi-level test cannot discriminate between $\mathbf{H}_0$ and $\mathbf{H}_1$. Otherwise, if

$$\frac{A_h}{\sigma_h} \overset{\bar\pi}{\ge} \sqrt{R_h(q_\alpha^\varkappa, H_{\bar\pi})} + \sqrt{\epsilon H^*_{\bar\pi}},$$

for some $\epsilon > 0$, then reliable testing is possible.

In the next theorem, $\mathbf{E}_{\bar\pi}$ stands for the expectation w.r.t. $\bar\pi$.

Theorem 3. *Suppose* (25) *holds. If for some* $x \in \mathbb{R}$ *and* $\epsilon > 0$

$$\lim_{\bar{\pi}\to 0} \frac{1}{H_{\bar{\pi}}^*} \mathbf{E}_{\bar{\pi}} \biggl[\biggl(\frac{A_h}{\sigma_h} - x \biggr)^2 + \epsilon H_{\bar{\pi}}^* - R_h(q_\alpha^\varkappa, H_{\bar{\pi}}) \biggr]_+ = 0, \tag{28}$$

then

$$\lim_{\bar{\pi}\to 0} \bar{\beta}_{\bar{\pi}}^M(\mathbf{A}) \geq (1-\alpha)[1-\Phi(x)]. \tag{29}$$

If for some $\epsilon > 0$

$$\lim_{\bar{\pi}\to 0} \frac{1}{H_{\bar{\pi}}^*} \mathbf{E}_{\bar{\pi}} \biggl[R_h(q_\alpha^\varkappa, H_{\bar{\pi}}) + 2\sqrt{\epsilon H_{\bar{\pi}}^*} \frac{A_h}{\sigma_h} - \frac{A_h^2}{\sigma_h^2} \biggr]_+ = 0, \tag{30}$$

then

$$\lim_{\bar{\pi}\to 0} \bar{\beta}_{\bar{\pi}}^M(\mathbf{A}) = 0. \tag{31}$$

3.2 Multi-level Bayes tests

To construct these tests, let us consider the following statistics:

$$Z_h^B = \frac{1}{n_h} \sum_{t\in\mathcal{G}_h} S\biggl(\frac{\hat{\theta}_{h,t}}{\sigma_h}\biggr) - \log(n_h) - \gamma + 1, \ \ h \in \mathcal{H}.$$

When all $\theta_{h,t} = 0$, in view of Theorem 2, these random variables are approximated by the family of independent and identically distributed random variables $\zeta_h^\circ, h \in \mathcal{H}$, defined by (16). An important property of this family is provided by (13), which is used in our construction multi-level Bayes tests. More precisely, the multi-level Bayes test rejects $\mathbf{H}_0$ if

$$\sum_{h\in\mathcal{H}} \bar{\pi}_h \biggl(Z_h^B - \log\frac{1}{\bar{\pi}_h} \biggr) \geq q_\alpha^\circ,$$

where q_α° is α-value of ζ°.

The type II error probability over $\Theta_{\rho,\tau}(A)$ (see (24)) is defined by

$$\beta_{\rho,\tau}^B(A) = \sup_{\Theta\in\Theta_{\rho,\tau}(A_\rho)} \mathbf{P}_\Theta \biggl\{ \sum_{h\in\mathcal{H}} \bar{\pi}_h \biggl(Z_h^B - \log\frac{1}{\bar{\pi}_h} \biggr) \leq q_\alpha^\circ \biggr\}$$

and our goal is to analyze the average type II error probability

$$\bar{\beta}_{\bar{\pi}}^B(\mathbf{A}) = \sum_{h\in\mathcal{H}} \frac{\bar{\pi}_h}{n_h} \sum_{\tau\in\mathcal{G}_h} \beta_{h,\tau}^B(A_h).$$

Theorem 4. *Suppose* (25) *holds and for some* $x \in \mathbb{R}$ *and* $\epsilon > 0$

$$\lim_{\bar{\pi}\to 0} \frac{1}{H^*_{\bar{\pi}}} \mathbf{E}_{\bar{\pi}} \biggl[\Bigl(\frac{A_h}{\sigma_h} - x \Bigr)^2 + \epsilon H^*_{\bar{\pi}} - R_h[\log(q^\circ_\alpha), H_{\bar{\pi}}] \biggr]_+ = 0, \tag{32}$$

then

$$\lim_{\bar{\pi}\to 0} \bar{\beta}^B_{\bar{\pi}}(\mathbf{A}) \geq (1-\alpha)[1-\Phi(x)]. \tag{33}$$

If for some $\epsilon > 0$

$$\lim_{\bar{\pi}\to 0} \frac{1}{H^*_{\bar{\pi}}} \mathbf{E}_{\bar{\pi}} \biggl[R_h[\log(q^\circ_\alpha), H_{\bar{\pi}}] + 2\sqrt{\epsilon H^*_{\bar{\pi}}} \frac{A_h}{\sigma_h} - \frac{A_h^2}{\sigma_h^2} \biggr]_+ = 0, \tag{34}$$

then

$$\lim_{\bar{\pi}\to 0} \bar{\beta}^B_{\bar{\pi}}(\mathbf{A}) = 0. \tag{35}$$

Remark. *Notice that as* $\alpha \to 0$

$$\log(q^\circ_\alpha) = (1+o(1))q^{\varkappa}_\alpha = (1+o(1)) \log \frac{1}{\alpha}.$$

Therefore, since $H_{\bar{\pi}} \to \infty$ *as* $\bar{\pi} \to 0$, *conditions* (28) *and* (32) *along with* (30) *and* (34) *are almost equivalent. This means that in the considered statistical problem there is no substantial difference between MAP and Bayes tests.*

4 Adaptive multi-level tests

The main drawback of the MAP and Bayes tests is related to their dependence on the prior distribution $\bar{\pi}$ that is hardly known in practice. Therefore our next goal is to construct a test that, on the one hand, does not depend on $\bar{\pi}$, but on the other hand, has a nearly optimal critical signal-noise ratio.

In order to simplify our presentation, we will deal with the class of prior distributions $\bar{\pi}^{\omega,\nu}$ defined by (26). The entropy of $\bar{\pi}^{\omega,\nu}$ obviously satisfies

$$H_{\bar{\pi}^{\omega,\nu}} = \log(\omega) + H_\nu + o(1), \ \omega \to \infty, \tag{36}$$

and therefore denote for brevity

$$\widetilde{R}_h(q,\omega) = 2[q + \log(n_h) + \log(\omega)] - \log[4\pi(q + \log(n_h) + \log(\omega)]. \tag{37}$$

With (36), Condition (25) is checked easily and the next result follows immediately from Theorem 3.

Corollary 1. *If for some $x \in \mathbb{R}$ and $\epsilon > 0$*

$$\lim_{\omega\to\infty} \frac{1}{\log^*(\omega)} \mathbf{E}_{\bar\pi^{\omega,\nu}} \biggl[\Bigl(\frac{A_h}{\sigma_h} - x \Bigr)^2 + \epsilon \log^*(\omega) - \widetilde{R}_h(q_\alpha^{\varkappa}, \omega) \biggr]_+ = 0,$$

then

$$\lim_{\omega\to\infty} \bar\beta^M_{\bar\pi^{\omega,\nu}}(\mathbf{A}) \geq (1-\alpha)[1-\Phi(x)].$$

If for some $\epsilon > 0$

$$\lim_{\omega\to\infty} \frac{1}{\log^*(\omega)]} \mathbf{E}_{\bar\pi^{\omega,\nu}} \biggl[\widetilde{R}_h(q_\alpha^{\varkappa}, \omega) + 2\sqrt{\epsilon \log^*(\omega)} \frac{A_h}{\sigma_h} - \frac{A_h^2}{\sigma_h^2} \biggr]_+ = 0,$$

then

$$\lim_{\omega\to\infty} \bar\beta^M_{\bar\pi^{\omega,\nu}}(\mathbf{A}) = 0.$$

In order to construct an adaptive test, let us compute a nearly minimal function U_h in (21). We begin with

$$\psi_0(x) = 1 + \log(x), \quad x \in \mathbb{R}^+,$$

and then iterate this function m times

$$\psi_l(x) = \psi_0\bigl[\psi_{l-1}(x)\bigr],\ l = 1, \dots, m.$$

Finally, for given $\varepsilon \in (0,1)$, define

$$L^{m,\varepsilon}(k) = -\log\biggl\{ \frac{1}{\varepsilon[\psi_m(k)]^\varepsilon} - \frac{1}{\varepsilon[\psi_m(k+1)]^\varepsilon} \biggr\},\ k \in \mathbb{Z}^+. \tag{38}$$

Since $\psi_m(1) = 1$, it is clear that

$$\sum_{k=1}^{\infty} \exp[-L^{m,\varepsilon}(k)] = \frac{1}{\varepsilon}.$$

In what follows, we will make use of the following approximation of $L^{m,\varepsilon}(k)$ for large k. Denote (see (38))

$$\begin{aligned} \widetilde{L}^{m,\varepsilon}(k) &= -\log\biggl[-\frac{1}{\varepsilon}\frac{[d\psi_m(k)]^{-\varepsilon}}{dk} \biggr] = -\log\biggl[\frac{1}{[\psi_m(k)]^{1+\varepsilon}} \frac{d\psi_m(k)}{dk} \biggr] \\ &= \log(k) + \log[\psi_0(k)] + \dots + \log[\psi_{m-1}(k)] + (1+\varepsilon)\log[\psi_m(k)]. \end{aligned} \tag{39}$$

Since

$$\frac{d^2\psi_m(k)}{dk^2} \bigg/ \frac{d\psi_m(k)}{dk} = O\Bigl(\frac{1}{k}\Bigr),$$

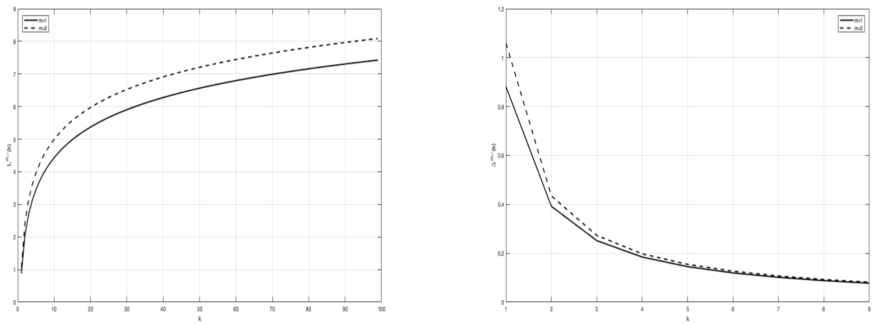

Figure 2: The functions $L^{m,\varepsilon}(\cdot)$ and the approximation errors $\Delta^{m,\varepsilon}(\cdot)$ for $m = 1, 2$ and $\varepsilon = 0.1$.

by the Taylor formula we obtain from (38)

$$\Delta^{m,\varepsilon}(k) = \widetilde{L}^{m,\varepsilon}(k) - L^{m,\varepsilon}(k) = O\Big(\frac{1}{\psi_m(k)}\Big). \tag{40}$$

Figure 2 shows that $L^{m,\varepsilon}(k)$ and approximation errors $\Delta^{m,\varepsilon}(k)$.

Since $h \in \mathcal{H} = \{2^{-1}, \dots, 2^{-k}, \dots\}$, we choose

$$U_h = L^{m,\varepsilon}[\log_2(1/h)]$$

and in view of (38) we arrive at the following prior distribution:

$$\Pi_h = \frac{1}{\{\psi_m[\log_2(1/h)]\}^{\varepsilon}} - \frac{1}{\{\psi_m[\log_2(1/h)+1]\}^{\varepsilon}}, \ h \in \mathcal{H}.$$

It is easy to check that the entropy of Π is unbounded and this is why this distribution might be viewed as an improper prior.

The MAP test associated with Π rejects $\mathbf{H}_0$ when

$$\max_{h\in\mathcal{H}}\{Z_h^M + \log(\Pi_h)\} \geq q_\alpha^{\varkappa}$$

and its type II error probability over $\Theta_{\rho,\tau}(A)$ (see (24)) is defined by

$$\beta_{\rho,\tau}^{\Pi}(A) = \sup_{\Theta\in\Theta_{\rho,\tau}(A)} \mathbf{P}_\Theta\Big\{\max_{h\in\mathcal{H}}[Z_h^M + \log(\Pi_h)] \leq q_\alpha^{\varkappa}\Big\}.$$

Denote for brevity

$$R^+(q_\alpha^{\varkappa}, \omega) = \widetilde{R}(q_\alpha^{\varkappa}, \omega) + \log^*(\omega), \tag{41}$$

where $\widetilde{R}(q_\alpha, \omega)$ is defined by (37), and let

$$\bar{\beta}^{\Pi}_{\bar{\pi}^{\omega,\nu}}(\mathbf{A}) = \sum_{h\in\mathcal{H}} \frac{\bar{\pi}_h^{\omega,\nu}}{n_h} \sum_{t\in\mathcal{G}_h} \beta^{\Pi}_{h,t}(A_h)$$

be the average type II error probability.

Theorem 5. *If for some $x \in \mathbb{R}$ and $\epsilon > 0$*

$$\lim_{\omega\to\infty} \frac{1}{\log^*(\omega)} \mathbf{E}_{\bar{\pi}^{\omega,\nu}} \left[\left(\frac{A_h}{\sigma_h} - x \right)^2 + \epsilon \log^*(\omega) - R_h^+(q_\alpha^{\varkappa}, \omega) \right]_+ = 0,$$

then

$$\lim_{\omega\to\infty} \bar{\beta}^{\Pi}_{\bar{\pi}^{\omega,\nu}}(\mathbf{A}) \geq (1-\alpha)[1-\Phi(x)].$$

If for some $\epsilon > 0$

$$\lim_{\omega\to\infty} \frac{1}{\log^*(\omega)} \mathbf{E}_{\bar{\pi}^{\omega,\nu}} \left[R_h^+(q_\alpha^{\varkappa}, \omega) + 2\sqrt{\epsilon \log^*(\omega)} \frac{A_h}{\sigma_h} - \frac{A_h^2}{\sigma_h^2} \right]_+ = 0,$$

then

$$\lim_{\omega\to\infty} \bar{\beta}^{\Pi}_{\bar{\pi}^{\omega,\nu}}(\mathbf{A}) = 0.$$

This theorem and Corollary 1 demonstrate that the critical signal-noise ratio of the adaptive test is only slightly greater, see (41), (by the additive term $\log^*(\omega)$) than the one of the MAP test that knows the prior distribution $\bar{\pi}^{\omega,\nu}$.

5 Appendix

5.1 Proof of Theorem 1

With a simple algebra we obtain

$$\begin{aligned}
\log\{\mathbf{E}\exp[\mathrm{i}tB_h(\xi)]\} &= n_h \log\biggl\{\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\cos\biggl[\frac{tS(x)}{n_h}\biggr]\mathrm{e}^{-x^2/2}\,dx \\
&\quad + \frac{\mathrm{i}}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\sin\biggl[\frac{tS(x)}{n_h}\biggr]\mathrm{e}^{-x^2/2}\,dx\biggr\} \\
&= n_h\log\biggl\{1+\int_{-\infty}^{\infty}\biggl[\cos\biggl[\frac{t}{n_h}\biggl(\frac{1}{\Phi(x)}-1\biggr)\biggr]-1\biggr]\,d\Phi(x) \\
&\quad + \mathrm{i}\int_{-\infty}^{\infty}\sin\biggl[\frac{t}{n_h}\biggl(\frac{1}{\Phi(x)}-1\biggr)\biggr]\,d\Phi(x)\biggr\} \\
&= n_h\log\biggl\{1+\int_0^1\biggl\{\cos\biggl[\frac{t}{n_h}\biggl(\frac{1}{u}-1\biggr)\biggr]-1\biggr\}\,du \qquad (42)\\
&\quad + \mathrm{i}\int_0^1\sin\biggl[\frac{t}{n_h}\biggl(\frac{1}{u}-1\biggr)\biggr]\,du\biggr\} \\
&= n_h\log\biggl\{1+\int_0^{\infty}\frac{1}{(1+u)^2}\biggl[\cos\biggl(\frac{tu}{n_h}\biggr)-1\biggr]\,du \\
&\qquad + \mathrm{i}\int_0^{\infty}\frac{1}{(1+u)^2}\sin\biggl(\frac{tu}{n_h}\biggr)\,du\biggr\} \\
&= n_h\log\biggl\{1+\frac{|t|}{n_h}\int_0^{\infty}\frac{\cos(u)-1}{(|t|/n_h+u)^2}\,du+\frac{\mathrm{i}t}{n_h}\int_0^{\infty}\frac{\sin(u)}{(|t|/n_h+u)^2}\,du\biggr\}.
\end{aligned}$$

It is clear that as $n_h \to \infty$

$$\int_0^{\infty}\frac{\cos(u)-1}{(|t|/n_h+u)^2}\,du \to \int_0^{\infty}\frac{\cos(u)-1}{u^2}\,du = -\frac{\pi}{2}. \qquad (43)$$

Let us choose $\epsilon_h < 1$ such that

$$\lim_{h\to 0}\epsilon_h = 0, \quad \lim_{h\to 0}\epsilon_h n_h = \infty.$$

Then we get by the Taylor formula

$$\begin{aligned}
\int_0^\infty \frac{\sin(u)}{(|t|/n_h+u)^2}\,du &= \int_0^{\epsilon_h} \frac{\sin(u)}{(|t|/n_h+u)^2}\,du + \int_{\epsilon_h}^\infty \frac{\sin(u)}{(|t|/n_h+u)^2}\,du \\
&= \int_0^{\epsilon_h} \frac{u}{(|t|/n_h+u)^2}\,du + O\Big[\int_0^{\epsilon_h} \frac{u^3}{(|t|/n_h+u)^2}\,du\Big] \\
&\quad + (1+o(1))\int_{\epsilon_h}^\infty \frac{\sin(u)}{u^2}\,du \qquad (44)\\
&= \log\left(1+\frac{\epsilon_h n_h}{|t|}\right) - \frac{\epsilon_h n_h}{(|t|+\epsilon_h n_h)} \\
&\quad + (1+o(1))\int_{\epsilon_h}^\infty \frac{\sin(u)}{u^2}\,du + O(\epsilon_h^2).
\end{aligned}$$

Next, integrating by parts, we obtain

$$\begin{aligned}
\int_x^\infty \frac{\sin(z)}{z^2}\,dz = &-\int_x^\infty \frac{\sin(z)}{z}\,d\Big(\log\frac{1}{z}\Big) \\
&= \frac{\sin(x)}{x}\log\frac{1}{x} + \int_x^\infty \frac{z\cos(z)-\sin(z)}{z^2}\log\frac{1}{z}\,dz.
\end{aligned}$$

Hence, as $x \to 0$

$$\begin{aligned}
\int_x^\infty \frac{\sin(z)}{z^2}\,dz &= \log\frac{1}{x} + \int_0^\infty \frac{z\cos(z)-\sin(z)}{z^2}\log\frac{1}{z}\,dz + O\Big(x^2\log\frac{1}{x}\Big) \\
&= \log\frac{1}{x} + (1-\gamma) + O\Big(x^2\log\frac{1}{x}\Big),
\end{aligned}$$

where γ is Euler's constant.

With this equation we continue (44) as follows:

$$\int_0^\infty \frac{\sin(u)}{(|t|/n_h+u)^2}\,du = \log\frac{1}{|t|} + \log(n_h) - \gamma + O\Big(\frac{|t|}{\epsilon_h n_h}\Big) + O\Big(\epsilon_h^2\log\frac{1}{\epsilon_h}\Big).$$

Substituting this equation and (43) in (42), we get

$$\log\{\mathbf{E}\exp[\mathrm{i}tB_h(\xi)]\} = -\frac{\pi|t|}{2} + \mathrm{i}t\Big(\log\frac{1}{|t|} + \log(n_h) - \gamma\Big) + o(1),$$

thus, proving the theorem.

5.2 Proof of Theorem 3

I. A lower bound. By (22) we have for any given x

$$\begin{aligned}
\beta^M_{\rho,\tau}(A) &\geq \mathbf{P}\biggl\{\biggl[\log\biggl[S\biggl(-\frac{A}{\sigma_\rho}+\xi_{\rho,\tau}\biggr)\biggr]-\log\frac{n_\rho}{\bar{\pi}_\rho}\biggr] \\
&\qquad \bigvee\biggl[\zeta_\rho-\log\frac{1}{\bar{\pi}_\rho}\biggr]\leq q_\alpha^{\varkappa}\biggr\}\mathbf{P}\biggl\{\max_{h\in\mathcal{H}}\biggl[\zeta_h-\log\frac{1}{\bar{\pi}_h}\biggr]\leq q_\alpha^{\varkappa}\biggr\} \\
&= \mathbf{P}\biggl\{S\biggl(-\frac{A}{\sigma_\rho}+\xi_{\rho,\tau}\biggr)\leq \exp\biggl[q_\alpha^{\varkappa}+\log\frac{n_\rho}{\bar{\pi}_\rho}\biggr]\biggr\} \\
&\quad \times \mathbf{P}\biggl\{\zeta_\rho\leq q_\alpha^{\varkappa}+\log\frac{1}{\bar{\pi}_\rho}\biggr\}\mathbf{P}\{\zeta\leq q_\alpha^{\varkappa}\} \\
&\geq (1-\alpha)^{1+\bar{\pi}_\rho}\mathbf{P}\biggl\{S\biggl(-\frac{A}{\sigma_\rho}+\xi_{\rho,\tau}\biggr)\leq \exp\biggl[q_\alpha^{\varkappa}+\log\frac{n_\rho}{\bar{\pi}_\rho}\biggr];\xi_{\rho,\tau}\geq x\biggr\} \\
&\geq (1-\alpha)^{1+\bar{\pi}_\rho}\mathbf{P}\{\xi_{\rho,\tau}\geq x\}\mathbf{1}\biggl\{S\biggl(-\frac{A}{\sigma_\rho}+x\biggr)\leq \exp\biggl[q_\alpha^{\varkappa}+\log\frac{n_\rho}{\bar{\pi}_\rho}\biggr]\biggr\}.
\end{aligned} \tag{45}$$

Let $R(z)\geq 0$ be a solution to

$$S\Bigl[-\sqrt{R(z)}\Bigr]=z.$$

It is easy to check with the help of (7) that as $z\to\infty$

$$R(z)=2\log(z)-\log[4\pi\log(z)]+o(1). \tag{46}$$

Denote for brevity

$$r_h(q,u)=2\biggl(q+\log\frac{n_h}{u}\biggr)-\log\biggl[4\pi\biggl(q+\log\frac{n_h}{u}\biggr)\biggr].$$

With (46) and the Markov inequality we obtain for any $\epsilon>0$

$$\begin{aligned}
&\mathbf{E}_{\bar{\pi}}\biggl[\mathbf{1}\biggl\{S\biggl(-\frac{A_h}{\sigma_h}+x\biggr)\leq \exp\biggl[q_\alpha^{\varkappa}+\log\frac{n_h}{\bar{\pi}_h}\biggr]\biggr\}\biggr] \\
&\quad = 1-\mathbf{E}_{\bar{\pi}}\biggl[\mathbf{1}\biggl\{S\biggl(-\frac{A_h}{\sigma_h}+x\biggr)> \exp\biggl[q_\alpha^{\varkappa}+\log\frac{n_h}{\bar{\pi}_h}\biggr]\biggr\}\biggr] \\
&\quad = 1-\mathbf{E}_{\bar{\pi}}\biggl[\mathbf{1}\biggl\{\biggl(\frac{A_h}{\sigma_h}-x\biggr)^2-r_h(q_\alpha^{\varkappa},\bar{\pi}_h)>o(1)\biggr\}\biggr] \\
&\quad \geq 1-\frac{1}{\epsilon H^*_{\bar{\pi}}}\mathbf{E}_{\bar{\pi}}\biggl[\biggl(\frac{A_h}{\sigma_h}-x\biggr)^2-r_h(q_\alpha^{\varkappa},\bar{\pi}_h)+\epsilon H^*_{\bar{\pi}}\biggr]_+.
\end{aligned} \tag{47}$$

With Condition (25) and simple algebra it is easy to check that

$$\lim_{\bar\pi\to 0}\frac{1}{H^*_{\bar\pi}}\mathbf{E}_{\bar\pi}\Big|r_h(q_\alpha^{\varkappa},\bar\pi)-R_h(q_\alpha^{\varkappa},H_{\bar\pi})\Big|=0.$$

Therefore (29) follows from (45), (47) and the above equation.

II. An upper bound. Since $S(x)$ is a decreasing function, we have obviously for any $x\ge 0$

$$\begin{aligned}\bar\beta^M_{\bar\pi}(\mathbf{A}) \le& \mathbf{E}_{\bar\pi}\biggl[\mathbf{P}\biggl\{S\Bigl(-\frac{A_h}{\sigma_h}+\xi\Bigr)\le \exp\Bigl[q_\alpha^{\varkappa}+\log\frac{n_h}{\bar\pi_h}\Bigr]\biggr\}\biggr]\\ \le& \mathbf{E}_{\bar\pi}\biggl[\mathbf{P}\biggl\{S\Bigl(-\frac{A_h}{\sigma_h}+\xi\Bigr)\le \exp\Bigl[q_\alpha^{\varkappa}+\log\frac{n_h}{\bar\pi_h}\Bigr];\xi\le x\biggr\}\biggr]\\ &+\mathbf{E}_{\bar\pi}\biggl[\mathbf{P}\biggl\{S\Bigl(-\frac{A_h}{\sigma_h}+\xi\Bigr)\le \exp\Bigl[q_\alpha^{\varkappa}+\log\frac{n_h}{\bar\pi_h}\Bigr];\xi> x\biggr\}\biggr]\\ \le& \mathbf{E}_{\bar\pi}\biggl[\mathbf{1}\biggl\{S\Bigl(-\frac{A_h}{\sigma_h}+x\Bigr)\le \exp\Bigl[q_\alpha^{\varkappa}+\log\frac{n_h}{\bar\pi_h}\Bigr]\biggr\}\biggr]+\mathbf{P}\{\xi>x\}.\end{aligned}$$

Next, with (46), the Markov inequality, and Condition (25) we get for any $\epsilon>0$

$$\begin{aligned}&\sum_{h\in\mathcal{H}}\bar\pi_h\mathbf{1}\biggl\{S\Bigl(-\frac{A_h}{\sigma_h}+x\Bigr)\le \exp\Bigl[q_\alpha^{\varkappa}+\log\frac{n_h}{\bar\pi_h}\Bigr]\biggr\}\\ &\quad\le \sum_{h\in\mathcal{H}}\bar\pi_h\mathbf{1}\biggl\{r_h(q_\alpha^{\varkappa},\bar\pi_h)-\Bigl(\frac{A_h}{\sigma_h}-x\Bigr)^2>o(1)\biggr\}\\ &\quad\le \frac{1}{\epsilon H^*_{\bar\pi}}\sum_{h\in\mathcal{H}}\bar\pi_h\biggl[r_h(q_\alpha^{\varkappa},\bar\pi_h)-\Bigl(\frac{A_h}{\sigma_h}-x\Bigr)^2+\epsilon H^*_{\bar\pi}\biggr]_+\\ &\quad\le \frac{1}{\epsilon H^*_{\bar\pi}}\sum_{h\in\mathcal{H}}\bar\pi_h\biggl[R_h(q_\alpha^{\varkappa},H_{\bar\pi})-\Bigl(\frac{A_h}{\sigma_h}-x\Bigr)^2+\epsilon H^*_{\bar\pi}\biggr]_+ + o(1).\end{aligned}$$

To complete the proof, let us choose $x=\sqrt{\epsilon H^*_{\bar\pi}}$.

5.3 Proof of Theorem 4

A lower bound. For given x and $\delta > 0$ by (17) and (13), we obtain

$$\begin{aligned}
\beta^B_{\rho,\tau}(A) &\geq \mathbf{P}\Big\{\frac{\bar{\pi}_\rho}{n_\rho} S\Big(-\frac{A}{\sigma_\rho} + \xi_{\rho,\tau}\Big) + \zeta^\circ \leq q^\circ_\alpha; \xi_{\rho,\tau} \geq x\Big\} \\
&\geq \mathbf{P}\Big\{\frac{\bar{\pi}_\rho}{n_\rho} S\Big(-\frac{A}{\sigma_\rho} + x\Big) + \zeta^\circ \leq q^\circ_\alpha; \xi_{\rho,\tau} \geq x\Big\} \\
&= \mathbf{P}\{\xi_{\rho,\tau} \geq x\}\mathbf{P}\Big\{\zeta^\circ \leq q^\circ_\alpha - \frac{\bar{\pi}_\rho}{n_\rho} S\Big(-\frac{A}{\sigma_\rho} + x\Big)\Big\} \\
&\geq \mathbf{P}\{\xi_{\rho,\tau} \geq x\}\mathbf{P}\{\zeta^\circ \leq (1-\delta)q^\circ_\alpha\}\mathbf{1}\Big\{\frac{\bar{\pi}_\rho}{n_\rho} S\Big(-\frac{A}{\sigma_\rho} + x\Big) \leq \delta q^\circ_\alpha\Big\}.
\end{aligned}$$

Similarly to (47), since $\bar{\pi} \to 0$, we get

$$\begin{aligned}
&\mathbf{E}_{\bar{\pi}}\Big[\mathbf{1}\Big\{\frac{\bar{\pi}_h}{n_h} S\Big(-\frac{A_h}{\sigma_h} + x\Big) \leq \delta q^\circ_\alpha\Big\}\Big] \\
&\quad \geq 1 - \frac{1}{\epsilon H^*_{\bar{\pi}}}\mathbf{E}_{\bar{\pi}}\Big[\Big(\frac{A_h}{\sigma_h} - x\Big)^2 + \epsilon H^*_{\bar{\pi}} - R_h[\log(\delta q^\circ_\alpha), H_{\bar{\pi}}]\Big]_+ \\
&\quad \geq 1 - \frac{1}{\epsilon H^*_{\bar{\pi}}}\mathbf{E}_{\bar{\pi}}\Big[\Big(\frac{A_h}{\sigma_h} - x\Big)^2 + \epsilon H^*_{\bar{\pi}} - R_h[\log(q^\circ_\alpha), H_{\bar{\pi}}]\Big]_+ \\
&\quad \geq 1 + o(1).
\end{aligned}$$

So, since δ is arbitrary, (33) follows from the above inequalities.

An upper bound. Since $S(x)$ is decreasing, we get

$$\begin{aligned}
\beta^B_{\rho,\tau}(A) \leq &\mathbf{P}\Big\{\frac{\bar{\pi}_\rho}{n_\rho} S\Big(-\frac{A}{\sigma_\rho} + \xi_{\rho,\tau}\Big) + \zeta^\circ \leq q^\circ_\alpha; \xi_{\rho,\tau} \geq x\Big\} \\
&+ \mathbf{P}\Big\{\frac{\bar{\pi}_\rho}{n_\rho} S\Big(-\frac{A}{\sigma_\rho} + \xi_{\rho,\tau}\Big) + \zeta^\circ \leq q^\circ_\alpha; \xi_{\rho,\tau} \leq x\Big\} \\
\leq &\mathbf{P}\{\xi_{\rho,\tau} \geq x\} + \mathbf{P}\Big\{\frac{\bar{\pi}_\rho}{n_\rho} S\Big(-\frac{A}{\sigma_\rho} + x\Big) + \zeta^\circ \leq q^\circ_\alpha\Big\}.
\end{aligned} \tag{48}$$

Next, for any given $x^\circ < q^\circ_\alpha$ we obtain with the help of the Markov inequality and (46)

$$
\begin{aligned}
&\mathbf{E}_{\bar\pi}\Big[\mathbf{P}\Big\{\frac{\bar\pi_h}{n_h}S\Big(-\frac{A_h}{\sigma_h}+x\Big)+\zeta^\circ\le q_\alpha^\circ\Big\}\Big] \\
&\quad\le \mathbf{E}_{\bar\pi}\Big[\mathbf{P}\Big\{\frac{\bar\pi_h}{n_h}S\Big(-\frac{A_h}{\sigma_h}+x\Big)+\zeta^\circ\le q_\alpha^\circ;\zeta^\circ<x^\circ\Big\}\Big] \\
&\qquad+\mathbf{E}_{\bar\pi}\Big[\mathbf{P}\Big\{\frac{\bar\pi_h}{n_h}S\Big(-\frac{A_h}{\sigma_h}+x\Big)+\zeta^\circ\le q_\alpha^\circ;\zeta^\circ\ge x^\circ\Big\}\Big] \\
&\quad\le \mathbf{P}\{\zeta^\circ<x^\circ\}+\mathbf{E}_{\bar\pi}\Big[\mathbf{1}\Big\{\frac{\bar\pi_h}{n_h}S\Big(-\frac{A_h}{\sigma_h}+x\Big)\le q_\alpha^\circ-x^\circ\Big\}\Big] \\
&\quad\le \mathbf{P}\{\zeta^\circ<x^\circ\}+\frac{1}{\epsilon H_{\bar\pi}^*}\mathbf{E}_{\bar\pi}\Big[R_h[\log(q_\alpha^\circ-x^\circ),H_{\bar\pi}]+\epsilon H_{\bar\pi}^*-\Big(\frac{A_h}{\sigma_h}-x\Big)^2\Big]_+ .
\end{aligned}
$$

Finally choosing $x=\sqrt{\epsilon H_{\bar\pi}^*}$ and combining this equation with (48), we complete the proof of (35).

5.4 Proof of Theorem 5

Let us consider

$$
H(\bar\pi^{\omega,\nu},\Pi)=\sum_{h\in\mathcal{H}}\bar\pi_h^{\omega,\nu}\log\frac{1}{\Pi_h}.
$$

One can check easily with (38)–(40) and (26) that as $\omega\to\infty$

$$
\begin{aligned}
H(\bar\pi^{\omega,\nu},\Pi)&=\log\frac{1}{\varepsilon}+\frac{1}{\omega}\sum_{k=1}^{\infty}\nu\Big(\frac{k}{\omega}\Big)L^{m,\varepsilon}(k)+o(1)=\log\frac{1}{\varepsilon}+o(1) \\
&+\int_0^\infty\nu(x)\Big[\log(x\omega+1)+\sum_{s=0}^{m-1}\log[\psi_s(x\omega)]+(1+\varepsilon)\log[\psi_m(x\omega)]\Big]\,dx.
\end{aligned}
$$

It is also clear in view of (27) that

$$
\int_0^\infty\nu(x)\log(x\omega+1)\,dx=\log(\omega)+O(1)
$$

and for any integer $s\ge 0$

$$
\int_0^\infty\nu(x)\log[\psi_s(x\omega)]\,dx=\log[\psi_s(\omega)]+O(1).
$$

Therefore as $\omega\to\infty$

$$
H(\bar\pi^{\omega,\nu},\Pi)=\log(\omega)+(1+o(1))\log^*(\omega). \tag{49}
$$

Next, with similar arguments we obtain

$$\begin{aligned}\sum_{h\in\mathcal{H}} \bar{\pi}_h^{\omega,\nu}\Big[H(\bar{\pi}^{\omega,\nu},\Pi) - \log\frac{1}{\Pi_h}\Big]_+ \\ &= \int_0^\infty \nu(x)\Big[H(\bar{\pi}^{\omega,\nu},\Pi) - \log(x\omega+1) \\ &\quad - (1+o(1))\log[\log(x\omega+1)+1]\Big]_+ dx \\ &= \int_0^\infty \nu(x)\Big[\log(\omega) + (1+o(1))\log^*(\omega)] - \log(x\omega+1) \\ &\quad - (1+o(1))\log[\log(x\omega+1)+1]\Big]_+ dx \\ &= \int_0^\infty \nu(x)\big[-\log(x+1) + o(1)\log^*(\omega)\big]_+ dx = o(1)\log^*(\omega)].\end{aligned} \tag{50}$$

In view of (49) and (50) the rest of the proof is similar to the one of Theorem 3 and therefore omitted.

References

[1] Baraud Y., Huet S., and Laurent B. (2005). Testing convex hypotheses on the mean of a Gaussian vector. Application to testing qualitative hypotheses on a regression function. Ann. Statist. **33**, 214–257.

[2] Bowman A. W., Jones M. C., and Gijbels I. (1998). Testing monotonicity of regression. J. Comput. Graph. Statist. **7**, 489–500.

[3] Burnashev M. and Bergamov I. (1990). On a problem of signal detection leading to stable distributions. Theory Probab. Appl. **35**(3), 556–560.

[4] Chetverikov D. (2019). Testing regression monotonicity in econometric models. Econometric Theory **35**(4), 729–776

[5] Dobrushin R. (1958). A statistical problem arising in the theory of detection of signals in the presence of noise in a multi-channel system and leading to stable distribution Laws. Theory Probab. Appl. **3**(2), 161–173.

[6] Dümbgen L., and Spokoiny V. (2001). Multiscale testing of qualitative hypotheses. Ann. Statist., **29**, 124–152.

[7] Durot C. (2003). A Kolmogorov-type test for monotonicity of regression. Statist. Probab. Lett. **63**, 425–433.

[8] Ghosal S., Sen A., and van der Vaart A. (2000). Testing monotonicity of regression. Ann. Statist. **28**, 1054–1082.

[9] Gijbels I., Hall, P., Jones M., and Koch I. (2000). Tests for monotonicity of a regression mean with guaranteed level. Biometrika **87**, 663–673.

[10] Hall P. and Heckman N. (2000). Testing for monotonicity of a regression mean by calibrating for linear functions. Ann. Statist. **28**, 20–39.

[11] Ingster Yu. I. and Suslina I. A. (2003). *Nonparametric Goodness-of-Fit Testing Under Gaussian Models. Lecture Notes in Statistics* **169**. New York: Springer.

[12] Mudholkar G. S. (1966). The integral of an invariant unimodal function over an invariant convex set – an inequality and applications. Proc. Amer. Math. Soc **17**, 1327–1333.

[13] Nolan J.P. (2016). *Stable Distributions: Models for Heavy-Tailed Data.* New York: Springer.

[14] Pyke R. (1965). Spacings (with discussion). J. Roy. Statist. Soc. Ser. B **27**(3), 395–449.

[15] Schlee W. (1982). *Nonparametric Tests of the Monotony and Convexity of Regression.* In Non-parametric Statistical Inference, Amsterdam: North-Holland.

[16] Tong Y. L. (1980). *Probability Inequalities in Multivariate Distributions. 1st edition.* . Academic Press.

[17] Wang J. and Meyer M. (2011). Testing the monotonicity or convexity of a function using regression splines. Canad. J. Statist. **39**, 89–107.