On multi-channel signal detection

Yu. Golubev
joint work with E. Burnaev

Aix-Marseille University and IITP RAS

Luminy, 18-22 December 2017

Dobrushin, R. (1954). *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.

Galager, R. (1969). *Information Theory and Reliable Communication* (Section 8.2).

Burnashev, M. and Bergamov, I. (1990). *On a Problem of Signal Detection Leading to Stable Distributions*. Theory Probab. Appl., 35(3), 556—560.

- Ingster, Yu. (1997.) *Some problems of hypothesis testing leading to infinitely divisible distributions*. Math. Methods of Stat. 6, 47–69.
- Ingster, Yu. and Lepski, O. (2002). *On multichannel signal detection*. WIAS preprint.
- Ingster, Yu., Pouet,C., Tsybakov, A. (2009). *Classification of sparse high-dimensional vectors*. Royal Society Philosophical Transactions A, v.367, 4427-4448.

Convex combinations of i.i.d.

Let $\kappa_1, \kappa_2, \ldots$ be i.i.d. Cauchy random variables and $\bar{\pi} = (\bar{\pi}_1, \bar{\pi}_2, \ldots)$ be a probability distribution on $\mathbb{Z}^+$, i.e.,

$$\bar{\pi}_i \geq 0, \quad \sum_{i=1}^{\infty} \bar{\pi}_i = 1.$$

Then

$$\mathbf{P}\left\{ \sum_{i=1}^{\infty} \bar{\pi}_i \kappa_i \leq x \right\} = \mathbf{P}\{\kappa_1 \leq x\},$$

or, equivalently,

$$\sum_{i=1}^{\infty} \bar{\pi}_i \kappa_i \stackrel{\mathbf{P}}{=} \kappa_1.$$

In particular,

$$\frac{1}{n} \sum_{i=1}^{n} \kappa_i \stackrel{\mathbf{P}}{=} \kappa_1$$

Information entropy

$$H(\bar{\pi}) = -\sum_{i=1}^{\infty} \bar{\pi}_i \log(\bar{\pi}_i)$$

measures the average amount of information produced by a random source.

Are there i.i.d. random variables ζ_i such that

$$\sum_{i=1}^{\infty} \bar{\pi}_i \zeta_i \stackrel{\mathbf{P}}{=} H(\bar{\pi}) + \zeta_1?$$

Or, in particular,

$$\frac{1}{n} \sum_{i=1}^{n} \zeta_i \stackrel{\mathbf{P}}{=} \log(n) + \zeta_1.$$

Let

$$E_k = \sum_{s=1}^{k} e_s,$$

where e_s are i.i.d. standard exponential random variables

$$\mathsf{P}\{e_s \le x\} = 1 - \exp(-x), \quad x \in \mathbb{R}^+.$$

Then

$$\zeta = \sum_{k=1}^{\infty} \left(\frac{1}{E_k} - \frac{1}{k} \right) + \gamma,$$

where $\gamma = 0.577215\ldots$ is the Euler constant.

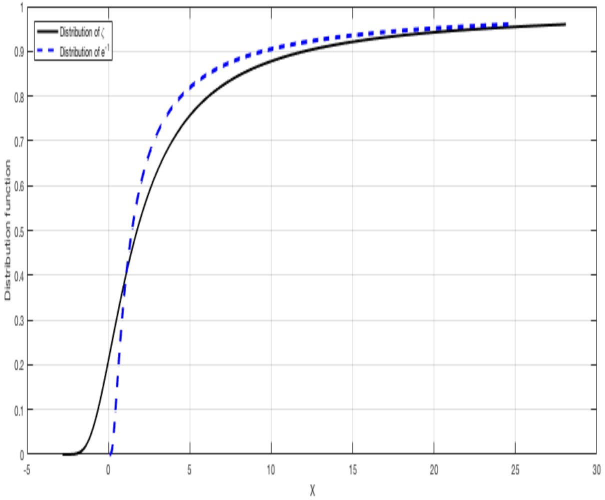

Figure: Distribution functions of ζ (solid line) and $1/e_1$ (dashed line).

- mass concentration

$$\mathbf{P}\{\zeta > 26.01\} = 0.05 \quad \text{and} \quad \mathbf{P}\{\zeta < -1.02\} = 0.05.$$

- heavy tail

$$\mathbf{P}\{\zeta \geq x\} \asymp \frac{1}{x}, \quad x \to \infty.$$

Theorem

$$\sum_{i=1}^{\infty} \bar{\pi}_i \zeta_i \stackrel{\mathbf{P}}{=} H(\bar{\pi}) + \zeta_1.$$

Suppose we observe

$$Y_i = \theta_i \times \mathbf{1}(\tau = i) + \sigma\xi_i, \quad i = 1, 2, \dots,$$

where

- ξ_i are i.i.d. $\mathcal{N}(0, 1)$;
- $\tau \in \mathbb{Z}^+$ is a random variable with a known distribution

$$\mathsf{P}\{\tau = i\} = \bar{\pi}_i;$$

- $\theta = (\theta_1, \dots)^\top \in \mathbb{R}^\infty$ is an unknown vector.

The goal is to test

the simple hypothesis

$$H_0 : \quad \theta = 0$$

vs. the compound alternative

$$H_1 : \quad \theta \neq 0.$$

Assumption. Let $\pi(x)$ be a continuous probability density on $\mathbb{R}^+$ with a bounded entropy

$$H(\pi) = -\int_0^\infty \log[\bar{\pi}(x)]\bar{\pi}(x)\, dx.$$

Then

$$\bar{\pi}_i = \bar{\pi}_i^n = \pi\left(\frac{i}{n}\right) \Big/ \sum_{k=1}^{\infty} \pi\left(\frac{k}{n}\right),$$

where n is large.

MAP test

This test accepts H_1 if

$$M(Y) \geq t_\alpha^M,$$

where

$$\begin{aligned} M(Y) &= \max_{k \in \mathbb{Z}^+} \max_{\theta_k} \left\{ \bar{\pi}_k^n \exp\left[-\frac{(Y_k - \theta_k)^2}{2\sigma^2} + \frac{Y_k^2}{2\sigma^2} \right] \right\} \\ &= \max_{k \in \mathbb{Z}^+} \bar{\pi}_k^n \exp\left(\frac{Y_k^2}{2\sigma^2} \right) \end{aligned}$$

and t_α^M is a solution to

$$\mathbf{P}\left\{ \max_{k \in \mathbb{Z}^+} \left[\bar{\pi}_k^n \exp\left(\frac{\xi_k^2}{2} \right) \right] \geq t_\alpha^M \right\} = \alpha.$$

Bayes test

This test accepts H_1 if

$$B(Y) \geq t_\alpha^B,$$

where

$$B(Y) = \sum_{k \in \mathbb{Z}^+} \bar{\pi}_k^n \max_{\theta_k} \left\{ \exp\left[-\frac{(Y_k - \theta_k)^2}{2\sigma^2} + \frac{Y_k^2}{2\sigma^2} \right] \right\}$$

$$= \sum_{k \in \mathbb{Z}^+} \bar{\pi}_k^n \exp\left(\frac{Y_k^2}{2\sigma^2} \right)$$

and t_α^B is defined by

$$\mathbf{P}\left\{ \sum_{k \in \mathbb{Z}^+} \bar{\pi}_k^n \exp\left(\frac{\xi_k^2}{2} \right) \geq t_\alpha^B \right\} = \alpha.$$

A limit theorem

Theorem

As $n \to \infty$

$$\sqrt{\pi H(\bar{\pi}^n)} \max_{k \in \mathbb{Z}^+} \left[\bar{\pi}_k^n \exp\left(\frac{\xi_k^2}{2} \right) \right] \stackrel{\mathbf{P}}{=} (1 + o(1)) \frac{1}{e_1},$$

$$\sqrt{\pi H(\bar{\pi}^n)} \left[\sum_{k \in \mathbb{Z}^+} \bar{\pi}_k^n \exp\left(\frac{\xi_k^2}{2} \right) - \mu(\bar{\pi}^n) \right] \stackrel{\mathbf{P}}{=} (1 + o(1)) \zeta.$$

where

$$\mu(\bar{\pi}^n) = \frac{2}{\pi} \sqrt{\pi H(\bar{\pi}^n)} - \frac{\log \sqrt{\pi H(\bar{\pi}^n)}}{\sqrt{\pi H(\bar{\pi}^n)}},$$

$$\zeta = \sum_{k=1}^{\infty} \left(\frac{1}{E_k} - \frac{1}{k} \right) + \gamma, \;\; E_k = \sum_{s=1}^{k} e_s.$$

Proof

Consider for simplicity the case where $\bar{\pi}_k^n = 1/n$.

I. Notice that as $nx \to \infty$

$$q_n(x) \stackrel{\text{def}}{=} \mathbf{P}\left\{ \frac{1}{n} \exp\left(\frac{\xi_i^2}{2} \right) \geq x \right\} = \frac{1 + o(1)}{nx\sqrt{\pi \log(nx)}}.$$

Hence

$$q_n^{-1}(x) = \frac{1 + o(1)}{nx\sqrt{\pi}} \left[\log\left(\frac{1}{x\sqrt{\pi}} \right) \right]^{-1/2}, \quad x \to \infty.$$

II. Let U_i, $i = 1, \dots, n$, be i.i.d. random variables uniformly distributed on $[0, 1]$ and $U_{(1)} \leq \dots \leq U_{(n)}$ be their nondecreasing permutation.
By Pyke's theorem

$$U_{(k)} \stackrel{\mathbf{P}}{=} \frac{E_k}{E_{n+1}}.$$

Therefore, as $n \to \infty$

$$U_{(n-k)} \stackrel{\mathbf{P}}{\approx} 1 - \frac{E_{k+1}}{n}$$

and thus when $k \ll n$

$$\frac{1}{n} \exp\left(\frac{\xi^2_{(n-k)}}{2}\right) \stackrel{\mathbf{P}}{\approx} q_n^{-1}\left(\frac{E_{k+1}}{n}\right) \stackrel{\mathbf{P}}{=} \frac{1 + o(1)}{\sqrt{\pi} E_{k+1}} \log^{-1/2}\left(\frac{n}{\sqrt{\pi} E_{k+1}}\right).$$

III. Next

$$\frac{1}{n}\sum_{k=1}^{n}\exp\left(\frac{\xi_k^2}{2}\right) = \frac{1}{n}\sum_{k=1}^{n}\exp\left(\frac{\xi_k^2}{2}\right)\mathbf{1}\big\{|\xi_k| > h_n\big\}$$
$$+\frac{1}{n}\sum_{k=1}^{n}\exp\left(\frac{\xi_k^2}{2}\right)\mathbf{1}\big\{|\xi_k| \le h_n\big\},$$

where h_n is defined by

$$\frac{1}{h_n}\exp\left(\frac{h_n^2}{2}\right) = \frac{n}{\sqrt{2\pi}M_n\log(n)}, \quad \lim_{n\to\infty}\frac{\log(M_n)\log[\log(n)]}{\log(n)} = 0.$$

We have (low of large numbers)

$$\frac{1}{n}\sum_{k=1}^{n}\exp\left(\frac{\xi_k^2}{2}\right)\mathbf{1}\left\{|\xi_k|\le h_n\right\} \stackrel{\mathbf{P}}{=} \sqrt{\frac{2}{\pi}}h_n + o\left(\frac{1}{\sqrt{\log(n)}}\right).$$

and

$$\frac{1}{n}\sum_{k=1}^{n}\exp\left(\frac{\xi_k^2}{2}\right)\mathbf{1}\left\{|\xi_k| > h_n\right\}$$
$$= \frac{1}{n}\sum_{k=0}^{n}\exp\left(\frac{\xi_{(n-k)}^2}{2}\right)\mathbf{1}\left\{|\xi_{(n-k)}| > h_n\right\} \stackrel{\mathbf{P}_0}{=} \sum_{k=0}^{\tau}\frac{1+o(1)}{E_{k+1}\sqrt{\pi\log(n/E_{k+1})}}$$

where

$$\tau = \min\left\{k : \frac{1}{E_{k+1}\sqrt{\pi\log(n/E_{k+1})}} \le \frac{1+o(1)}{\sqrt{\pi\log(n)}M_n}\right\}$$

Critical values

As $n \to \infty$

$$t_\alpha^M = \frac{t_\alpha^* + o(1)}{\sqrt{\pi H(\bar{\pi}^n)}},$$
$$t_\alpha^B = \mu_n + \frac{t_\alpha^\circ + o(1)}{\sqrt{\pi H(\bar{\pi}^n)}},$$

where t_α^* and t_α° are the quantiles of order α of e_1^{-1} and ζ respectively

$$t_\alpha^* = 1 - \exp(-1/\alpha) \approx 1/\alpha, \text{ as } \alpha \to 0,$$
$$\mathsf{P}\{\zeta \geq t_\alpha^\circ\} = \alpha.$$

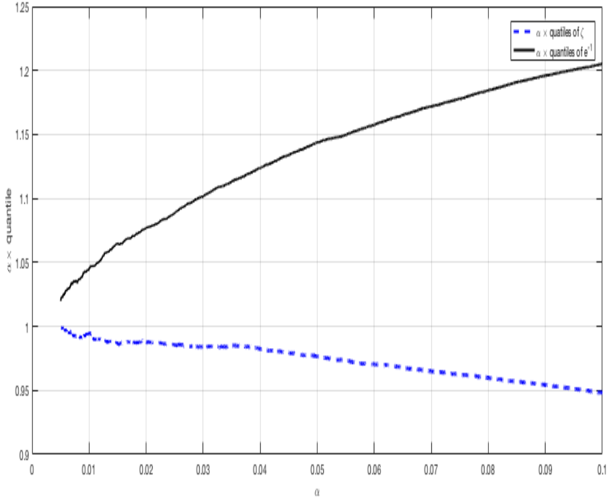

Figure: $\alpha\times$ quantiles of ζ (solid line) and $1/e_1$ (dashed line).

Small signal detection

Let

$$\rho_i^2 = \log \frac{1}{\bar{\pi}_i^n \sqrt{\pi H(\bar{\pi}^n)}}.$$

Define the critical sets:

$$\Theta^M(\tau) = \Big\{\theta : \theta_\tau^2 \le 2\sigma^2 \big[\rho_\tau^2 + \log(t_\alpha^M) - 1\big];\ \theta_k = 0, k \ne \tau\Big\},$$

$$\Theta^B(\tau) = \Big\{\theta : \theta_\tau^2 \le 2\sigma^2 \big(\rho_\tau^2 - 1\big);\ \theta_k = 0, k \ne \tau\Big\}.$$

Theorem

$$\lim_{n\to\infty} \inf_{S\in\Theta^M(\tau)} \mathbf{P}_S\Big\{M(Y) \le t_\alpha^M \Big| \tau\Big\} = \frac{1-\alpha}{2},$$

$$\lim_{n\to\infty} \inf_{S\in\Theta^B(\tau)} \mathbf{P}_S\Big\{B(Y) \le t_\alpha^B \Big| \tau\Big\} = \frac{1-\alpha}{2}.$$

Large signal detection

Define

$$R_\tau^M(\theta) = -\log \mathbf{P}_\theta\left\{M(Y) \le t_\alpha^M \middle| \tau\right\},$$
$$R_\tau^B(\theta) = -\log \mathbf{P}_\theta\left\{B(Y) \le t_\alpha^B \middle| \tau\right\}.$$

Theorem

As $n \to \infty$

$$R_\tau^M(\theta) = R_\tau^B(\theta) + O(\alpha)$$

uniformly in $\theta \notin (1+\epsilon)\Theta^M(\tau),\ \epsilon > 0,$

A general model

We observe

$$Y_i = \theta_i \times \mathbf{1}\{i \in \tau\} + \sigma\xi_i, \quad i = 1, 2, \dots,$$

where

- $\tau = \{\tau_1, \dots, \tau_S\}$ is the multi-index with i.i.d. components $\tau_i \in \mathbb{Z}^+$
$$\mathbf{P}\Big\{\tau_k = i\Big\} = \bar{\pi}_i;$$
- $\theta = (\theta_1, \dots)^\top \in \mathbb{R}^\infty$ is an unknown vector.

The MAP test

MAP test accepts H_1 if

$$M_S(Y) = \max_{\tau} \prod_{k\in\tau} \bar{\pi}_k^n \exp\left(\frac{Y_k^2}{2\sigma^2}\right) \geq t^M_{\alpha,S}.$$

Let

$$Z_k = \bar{\pi}_k^n \exp\left(\frac{Y_k^2}{2\sigma^2}\right), \quad k = 1, 2, \dots,$$

and $Z_{(1)} \geq Z_{(2)}, \dots$ be nondecreasing permutation of $Z_k,\ k = 1, 2, \dots$. Then

$$M_S(Y) = \prod_{k=1}^{S} Z_{(k)}.$$

and $t^M_{\alpha,S}$ is defined by

$$\mathbf{P}_0\left\{\prod_{k=1}^{S} Z_{(k)} \geq t^M_{\alpha,S}\right\} = \alpha.$$

Bayes test

This test accepts H_1 if

$$B_S(Y) = \sum_{\tau} \prod_{k \in \tau} \bar{\pi}_k^n \exp\left(\frac{Y_k^2}{2\sigma^2}\right) \geq t_{\alpha,S}^B,$$

where $t_{\alpha,S}^B$ is a solution to

$$\mathsf{P}\left\{\sum_{\tau} \prod_{k \in \tau} \bar{\pi}_k \exp\left(\frac{\xi_k^2}{2}\right) \geq t_{\alpha,S}^B\right\} = \alpha.$$

Bayes test complexity

Numerical complexities of the MAP and the Bayes tests are similar since, for example,

$$B_2(Y) = \frac{1}{2}\left[\sum_{k=1}^{\infty} Z_k\right]^2 - \frac{1}{2}\sum_{k=1}^{\infty} Z_k^2.$$

A limit theorem

Theorem

As $n \to \infty$

$$[\pi H(\bar{\pi}^n)]^{S/2} M_S(Y) \stackrel{\mathbf{P}_0}{=} (1 + o(1)) \prod_{k=1}^{S} \frac{1}{E_k},$$

$$\left(\frac{2}{\pi}\right)^{1-S} [\pi H(\bar{\pi}^n)]^{1-S/2} \left[B_S(Y) - \mu_S(\bar{\pi}^n)\right] \stackrel{\mathbf{P}_0}{=} (1 + o(1))\zeta,$$

where

$$\mu_S(\bar{\pi}^n) = \left(\frac{2}{\pi}\right)^S [\pi H(\bar{\pi}^n)]^{S/2} \left[\frac{1}{S} - \frac{\log \sqrt{\pi H(\bar{\pi}^n)}}{2H(\bar{\pi}^n)}\right].$$

Small signal detection with the MAP test

Let

$$\Theta^M(\tau) = \left\{\theta : \frac{\|\theta\|^2}{2\sigma^2} \le \sum_{k\in\tau} \rho_k^2 + \log(t^*_{\alpha,S}) - S;\ \theta_k = 0, k \notin \tau\right\},$$

where

$$\mathbf{P}\left\{\prod_{k=1}^{S} \frac{1}{E_k} \ge t^*_{\alpha,S}\right\} = \alpha, \quad \rho_k^2 = \log \frac{1}{\bar{\pi}_k^n \sqrt{\pi H(\bar{\pi}^n)}}.$$

Notice

$$\log(t^*_{\alpha,S}) \approx \log \frac{1}{\alpha} - S \log \frac{S}{\mathrm{e}}.$$

Theorem

$$\lim_{n\to\infty} \inf_{\theta\in\Theta^M(\tau)} \mathbf{P}_\theta\left\{M_S(Y) \le t^M_{\alpha,S} \middle| \tau\right\} = \frac{1-\alpha}{2}.$$

Small signal detection with the Bayes test

Let

$$\Theta^B(\tau) = \left\{ \theta : \frac{\theta_k^2}{2\sigma^2} \leq \rho_k^2 - 1,\ k \in \tau;\ \theta_k = 0, k \notin \tau \right\}.$$

Theorem

$$\lim_{n\to\infty} \inf_{\theta\in\Theta^B(\tau)} \mathbf{P}_\theta \left\{ B_S(Y) \leq t_{\alpha,S}^{B} \middle| \tau \right\} = \frac{1-\alpha}{2^S},$$

$$\lim_{n\to\infty} \inf_{\theta\in\Theta^B(\tau)} \mathbf{P}_\theta \left\{ B_1(Y) \leq t_{\alpha}^{B,1} \middle| \tau \right\} = \frac{1-\alpha}{2^S}.$$

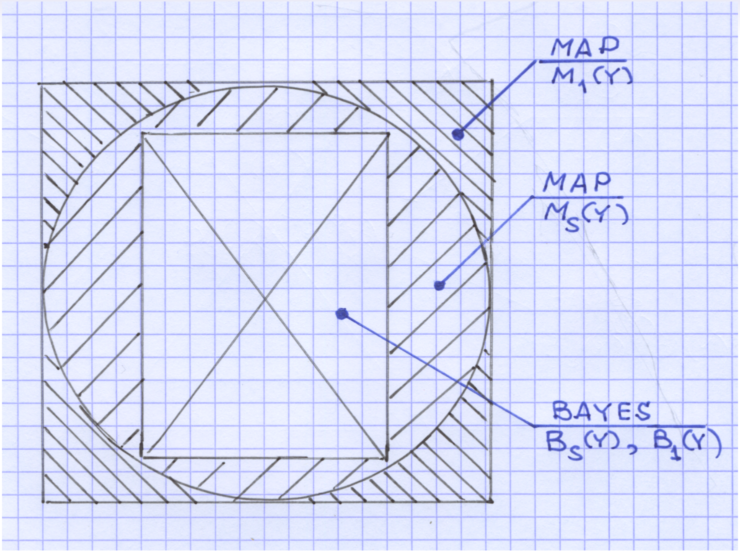

Figure: Undetectable signals.

THANK YOU FOR ATTENTION

MANY THANKS

to

OLEG and SACHA