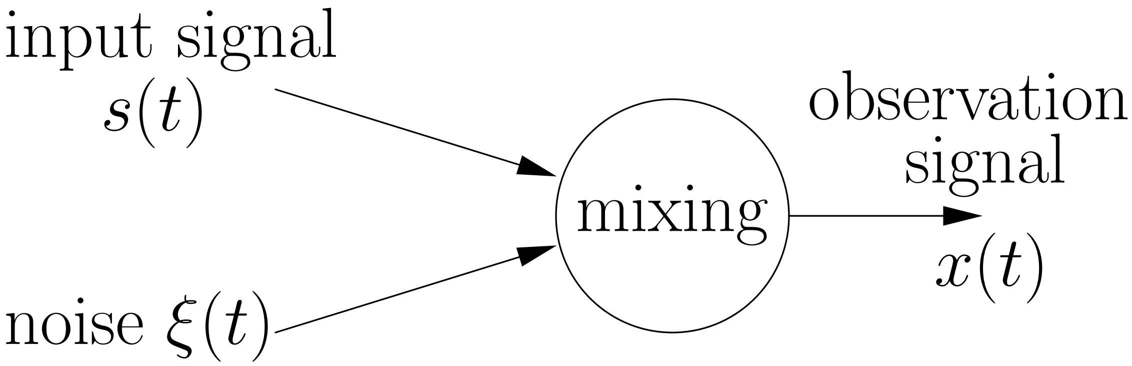
UPoN 2008, Lyon, France, 2–6 June 2008.

# Raising the noise to improve the performance of optimal processing

François Chapeau-Blondeau
David Rousseau
Laboratoire d'Ingénierie des Systèmes Automatisés (LISA),
University of Angers, France.

# General optimal-processing problem




At $N$ given times $t_k$, collect $N$ observations $x(t_k) = x_k, \quad$ for $k = 1$ to $N$.

From the $N$ observations $(x_1, \ldots x_N) = \boldsymbol{x}$,
perform some inference about $s(t)$,
optimal in the sense of a meaningful criterion of performance.

Performance of optimal processor improvable by raising the noise:

- i) Demonstration of feasibility by examples.
- ii) A general mechanism of explanation.
- iii) Open problems on noise-improved optimal processing.

# 1 Optimal detection

## 1.1 Classic theory of optimal detection

From $\boldsymbol{x} = (x_1, \ldots x_N)$, decide between two possible hypotheses:

$\mathrm{H}_0$: $s(t) \equiv s_0(t)$ known (prior probability $P_0$)

$\mathrm{H}_1$: $s(t) \equiv s_1(t)$ known (prior probability $P_1 = 1 - P_0$)

Criterion of performance: probability of detection error

$$P_{\mathrm{er}} = \Pr\{s_1 \text{ decided } |\mathrm{H}_0 \text{ true }\} \times P_0 + \Pr\{s_0 \text{ decided } |\mathrm{H}_1 \text{ true }\} \times P_1$$
$$= P_1 + \int_{\mathcal{R}_1 \subset \mathbb{R}^N} \Big[P_0 p(\boldsymbol{x}|\mathrm{H}_0) - P_1 p(\boldsymbol{x}|\mathrm{H}_1)\Big] d\boldsymbol{x} \,.$$

is minimized by optimal detector $L(\boldsymbol{x}) = \dfrac{p(\boldsymbol{x}|\mathrm{H}_1)}{p(\boldsymbol{x}|\mathrm{H}_0)} \underset{\mathrm{H}_0}{\overset{\mathrm{H}_1}{\gtrless}} \dfrac{P_0}{P_1}$,

achieving the minimum probability of error

$$P_{\mathrm{er}}^{\min} = \frac{1}{2} - \frac{1}{2} \int_{\mathbb{R}^N} \Big|P_1 p(\boldsymbol{x}|\mathrm{H}_1) - P_0 p(\boldsymbol{x}|\mathrm{H}_0)\Big| d\boldsymbol{x} \,.$$

The performance of the optimal detector

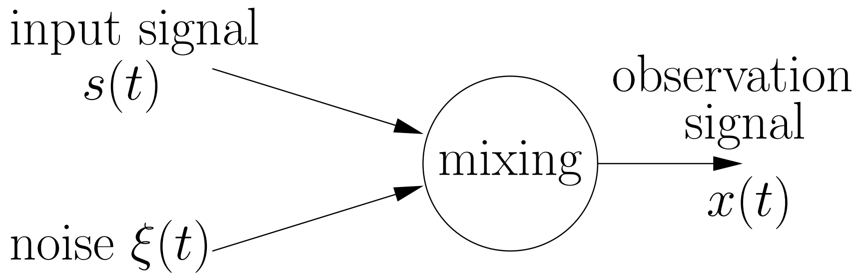
$$P_{\text{er}}^{\min} = \frac{1}{2} - \frac{1}{2}\int_{\mathbb{R}^N} \Big| P_1 p(\boldsymbol{x}|\mathrm{H}_1) - P_0 p(\boldsymbol{x}|\mathrm{H}_0) \Big| d\boldsymbol{x}$$

is not bound to degrade as the level of noise increases.

On the contrary, the optimal performance $P_{\text{er}}^{\min}$ can sometimes improve when the level of noise increases.

## 1.2 Example with additive signal-noise mixture

- Observation signal $x(t) = s(t) + \xi(t)$



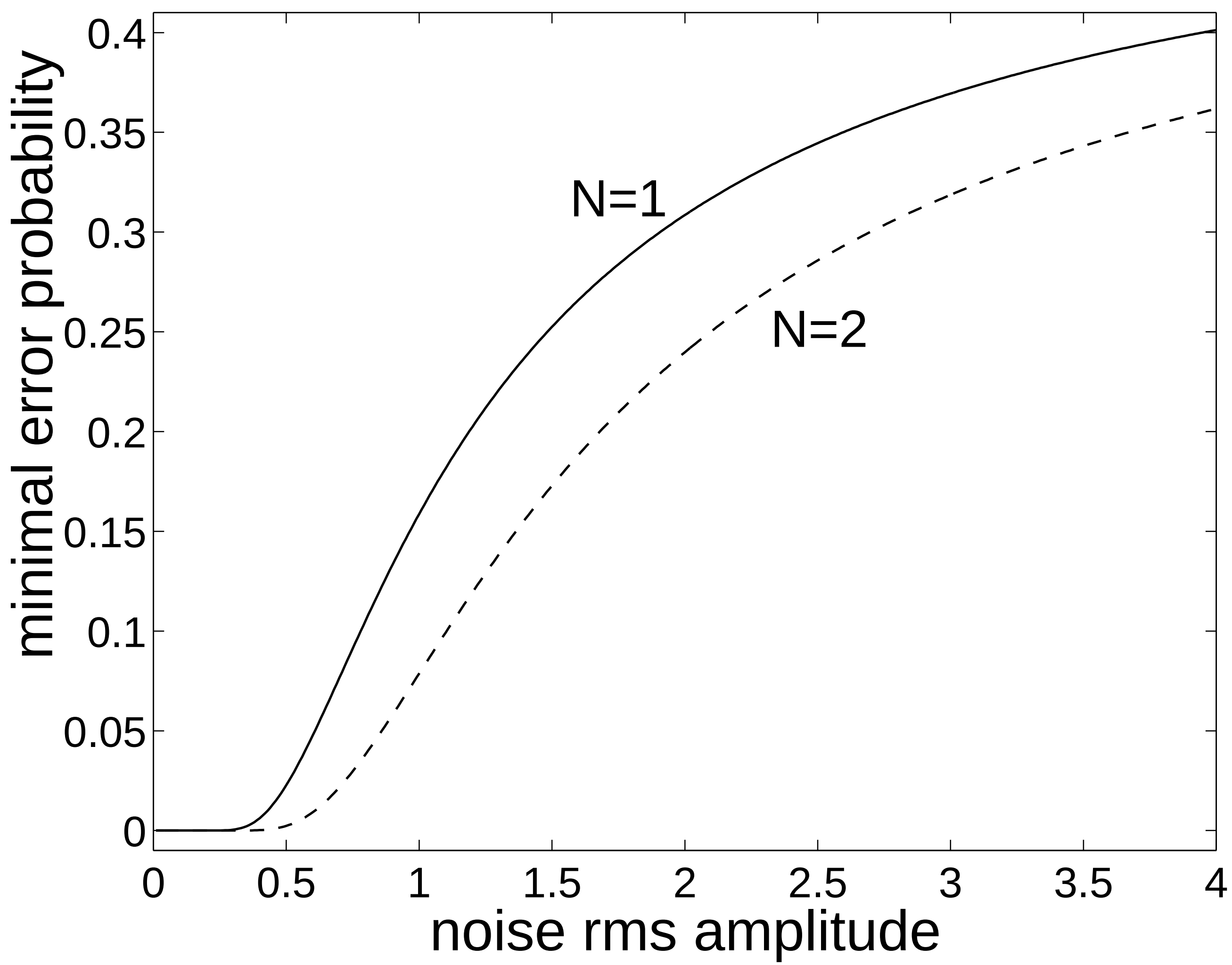

- Stationary white noise $\xi(t)$ for i.i.d. samples $\xi(t_k) \sim$ density $f_\xi(\cdot) \Longrightarrow$

density $p(\boldsymbol{x}|\mathrm{H}_j) = \prod_{k=1}^{N} p(x_k|\mathrm{H}_j)$, with $p(x_k|\mathrm{H}_j) = f_\xi[x_k - s_j(t_k)]$

- Two constant signals $s(t) = s_0$ or $s(t) = s_1$, for all $t$.

- With zero-mean Gaussian noise $\xi(t)$ :

$$P_{\mathrm{er}}^{\min} = \frac{1}{2}\left[1 + P_1\,\mathrm{erf}\left(\sqrt{N}\frac{x_T - s_1}{\sqrt{2}\sigma}\right) - P_0\,\mathrm{erf}\left(\sqrt{N}\frac{x_T - s_0}{\sqrt{2}\sigma}\right)\right] .$$




$P_{\mathrm{er}}^{\min}$ degrades (increases) as the noise rms amplitude $\sigma$ increases.

$s_0 = -1,\ s_1 = 1,\ P_0 = 1/2$ .

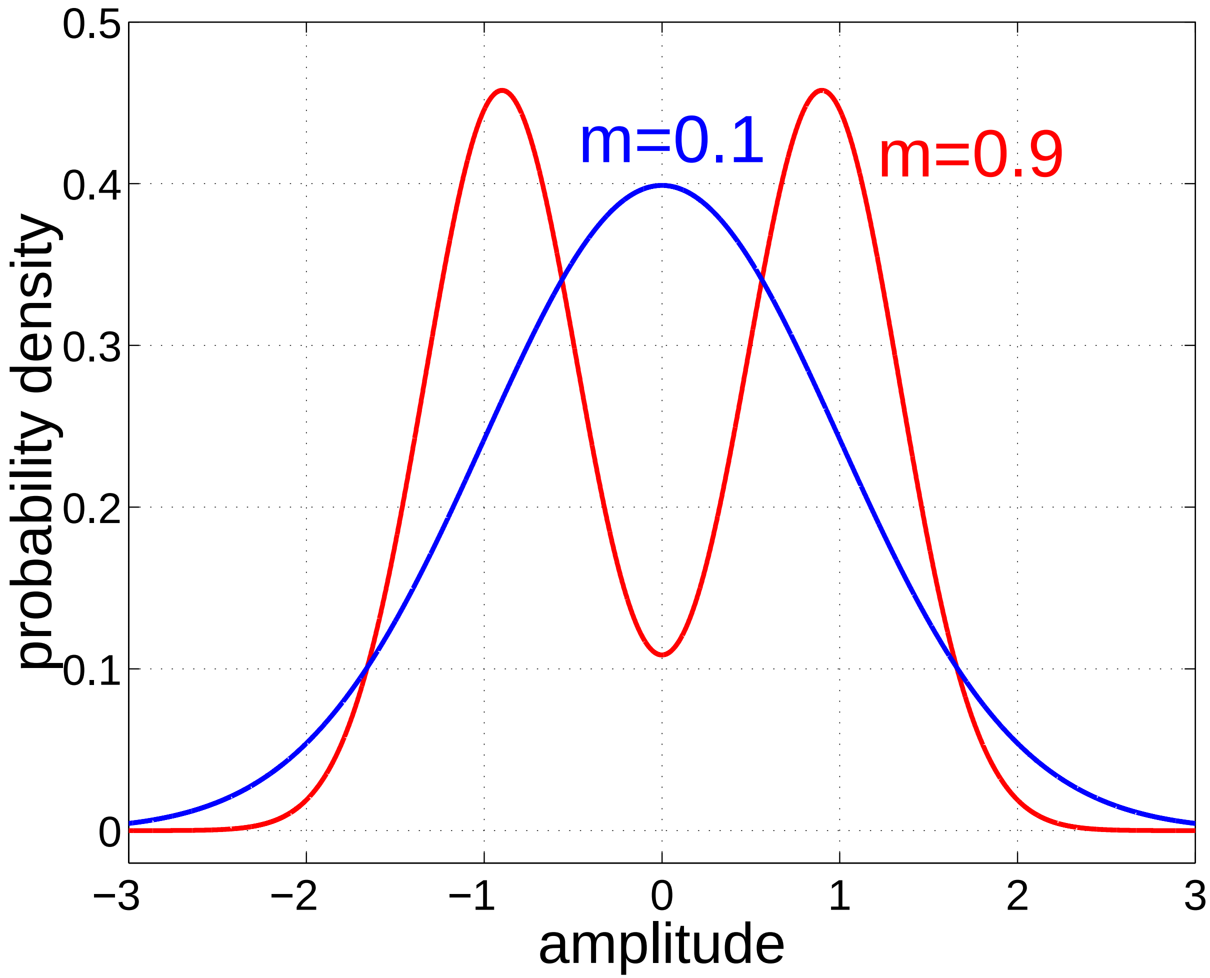
- With zero-mean non-Gaussian noise $\xi(t)$ :

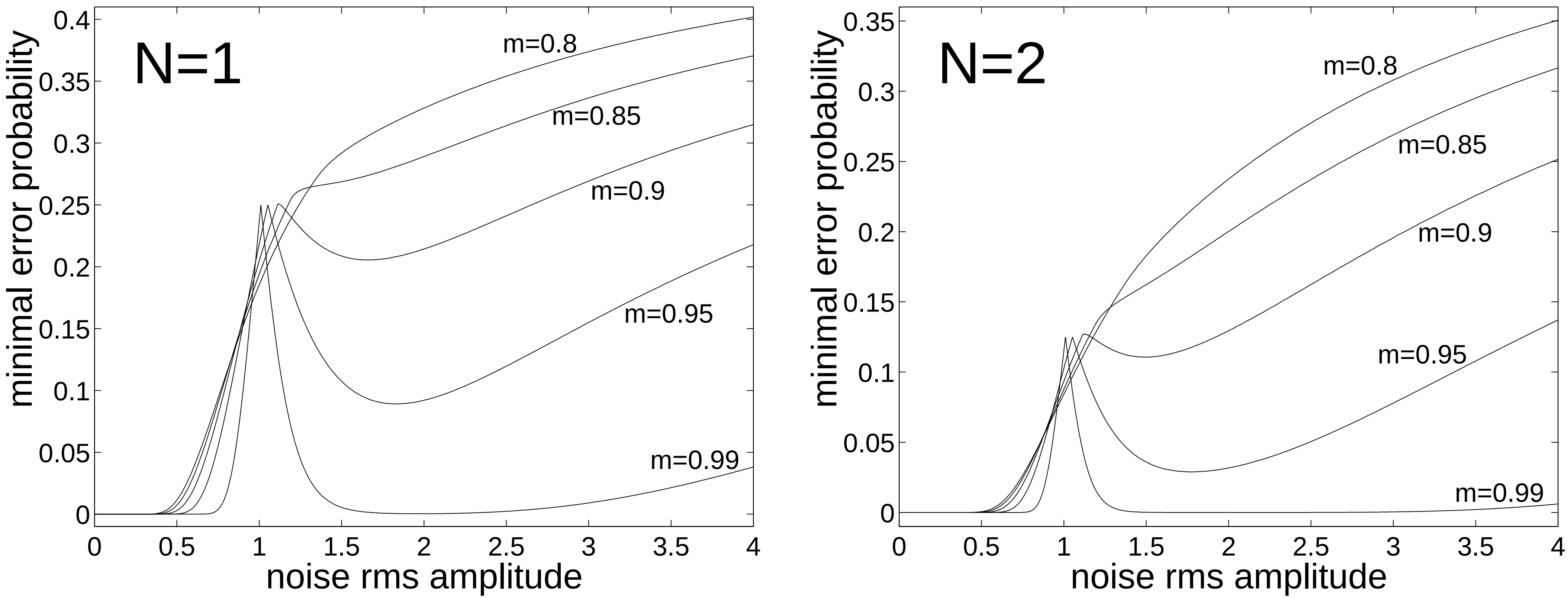
$$f_{\text{gm}}(u) = \frac{1}{2\sqrt{2\pi}\sqrt{1-m^2}} \left\{ \exp\left[-\frac{(u+m)^2}{2(1-m^2)}\right] + \exp\left[-\frac{(u-m)^2}{2(1-m^2)}\right] \right\} ,$$

Bi-Gaussian mixture $f_{\text{gm}}(u)$:

$0 \leq m < 1$

$$\xi(t) \sim \frac{1}{\sigma} f_{\text{gm}}\Big(\frac{u}{\sigma}\Big) .$$






Zero-mean bi-Gaussian mixture noise $\xi(t)$, with $s_0 = -1$, $s_1 = 1$, $P_0 = 1/2$ .

The performance $P_{\text{er}}^{\text{min}}$ of the optimal detector,

can sometimes improve,

when the noise rms amplitude $\sigma$ increases, over some ranges.

Qualitatively, the two peaks of the noise density,
make the two noisy constants $s_0$ and $s_1$ more distinguishable,
as the noise level $\sigma$ is raised, over some range.

**Improvement by noise of an optimal detector**,

with an additive signal-noise mixture with non-Gaussian noise $\xi(t)$.

$\Longrightarrow$ the level of the initial noise $\xi(t)$ cannot be increased by addition of another independent noise $\eta(t)$.

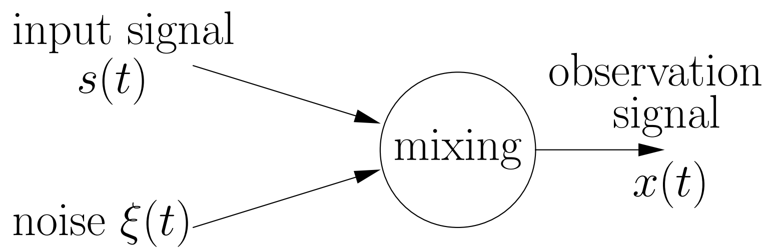
In practice: control of a more internal parameter of the physical process, like a temperature, has to be assumed to increase the level of the noise $\xi(t)$.

# 2 Optimal estimation

## 2.1 Classic theory of optimal estimation

$s(t) \equiv s_\nu(t)$ , with $\nu$ an unknown deterministic parameter.

From data $\boldsymbol{x} = (x_1, \dots x_N)$, estimate $\nu$.




Estimator $\widehat{\nu}(\boldsymbol{x})$, with rms error $\mathcal{E} = \sqrt{\int_{\mathbb{R}^N} [\widehat{\nu}(\boldsymbol{x}) - \nu]^2 \; p(\boldsymbol{x};\nu) d\boldsymbol{x}}$ .

Maximum likelihood estimator $\widehat{\nu}_{\text{ML}}(\boldsymbol{x}) = \arg\max_{\nu} \; p(\boldsymbol{x};\nu)$ .

Asymptotically at large $N$, estimator $\widehat{\nu}_{\text{ML}}(\boldsymbol{x})$ minimizes $\mathcal{E}$ ,

achieving the minimum rms error $\mathcal{E}_{\min} = \sqrt{1/J(\boldsymbol{x})}$ ,

with Fisher information $J(\boldsymbol{x}) = \int_{\mathbb{R}^N} \frac{1}{p(\boldsymbol{x};\nu)} \left[ \frac{\partial}{\partial \nu} p(\boldsymbol{x};\nu) \right]^2 d\boldsymbol{x}$ .

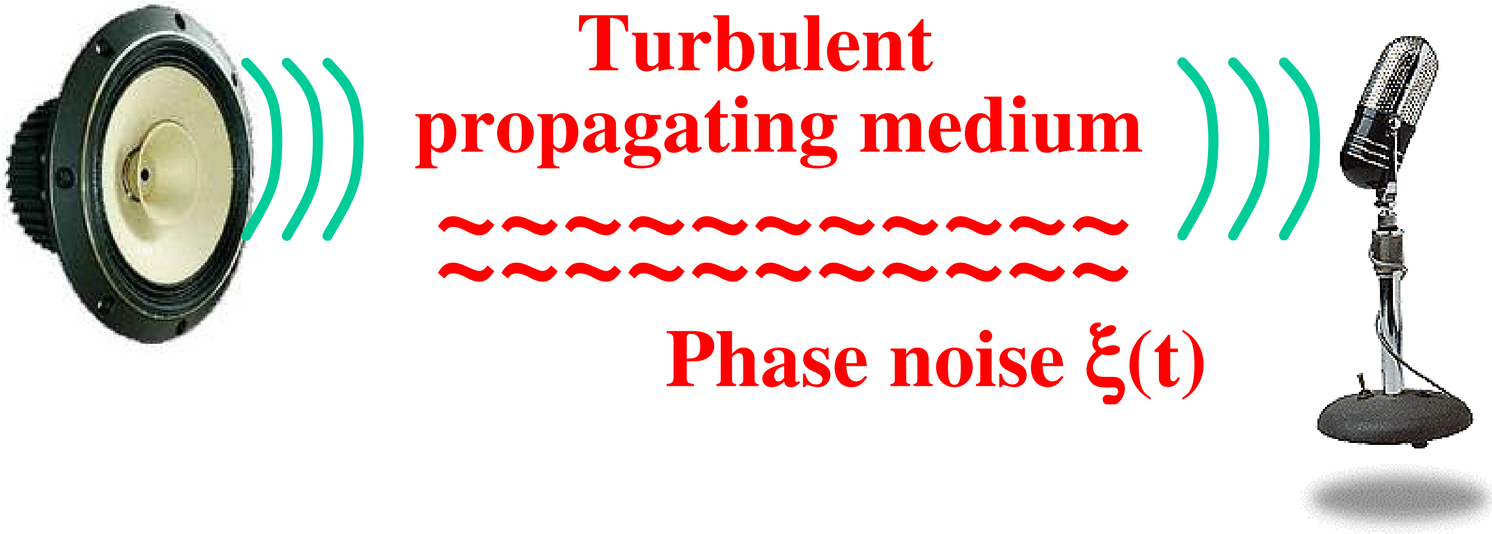
The optimal performance $\mathcal{E}_{\min} = \sqrt{1/J(\boldsymbol{x})}$ is not bound to degrade as the level of noise increases.

On the contrary, the optimal performance $\mathcal{E}_{\min} = \sqrt{1/J(\boldsymbol{x})}$ can sometimes improve when the level of noise increases.

## 2.2 Example with phase noise

- Observation signal $x(t) = w[\nu t + \xi(t)]$, of unknown frequency $\nu$,

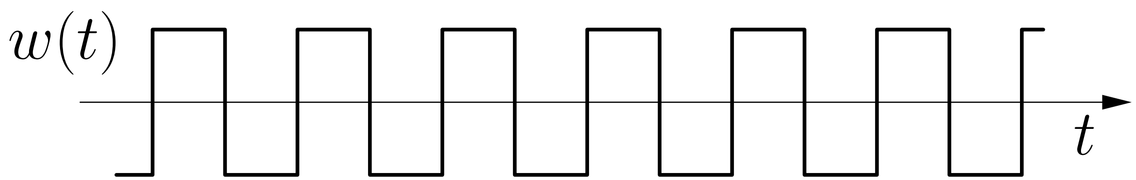
with $w(t)$ a known "mother" waveform of period 1.




- Stationary white noise $\xi(t)$ for i.i.d. samples $\xi(t_k) \sim f_\xi(\cdot) \Longrightarrow$

density $p(\boldsymbol{x};\nu) = \prod_{k=1}^{N} p(x_k;\nu)$.

- Square wave $w(t) = \pm 1$.



$\Longrightarrow$ theoretical expression for the minimal rms error $\mathcal{E}_{\min} = \sqrt{1/J(\boldsymbol{x})}$ ,

with Fisher information $J(\boldsymbol{x}) = \sum_{k=1}^{N} J(x_k)$, and

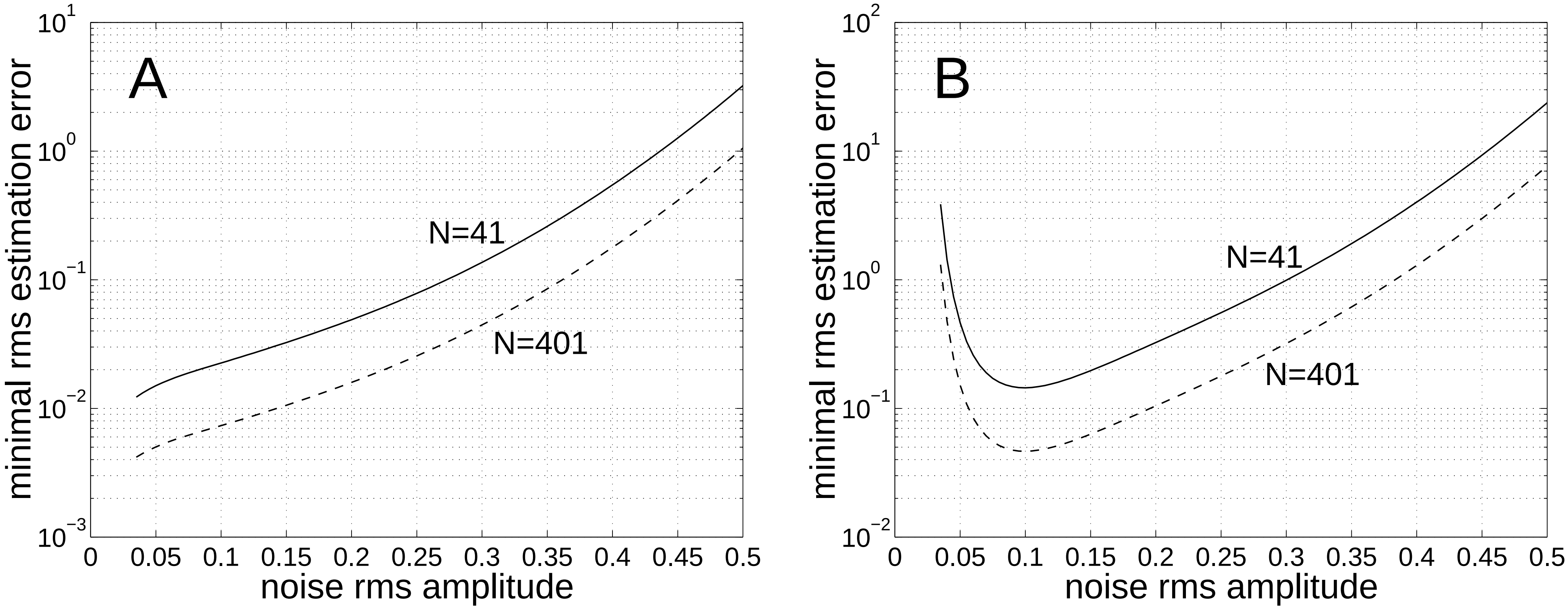
$$J(x_k) = \sum_{x_k=-1,1} \frac{1}{\Pr\{x_k;\nu\}} \left[ \frac{\partial}{\partial \nu} \Pr\{x_k;\nu\} \right]^2 ,$$

with the probabilities

$$\Pr\{x_k = 1;\nu\} = \sum_{\ell=-\infty}^{+\infty} \left[ F_\xi(\ell - \nu t_k + 1/2) - F_\xi(\ell - \nu t_k) \right]$$

and $\qquad \Pr\{x_k = -1;\nu\} = 1 - \Pr\{x_k = 1;\nu\}$ ,

and $F_\xi(\cdot)$ cumulative distribution function of the noise $\xi(t)$.




Gaussian noise $\xi(t)$, with $N$ observations at $\boldsymbol{t} = (t_1, t_2, \ldots t_N)$ in units of $1/\nu$:

Panel A: $\boldsymbol{t} = [t_1 = 0 : \Delta t = 5 \times 10^{-2} : t_N = 2]$ for $N = 41$ (solid line),
$\boldsymbol{t} = [t_1 = 0 : \Delta t = 5 \times 10^{-3} : t_N = 2]$ for $N = 401$ (dashed line).

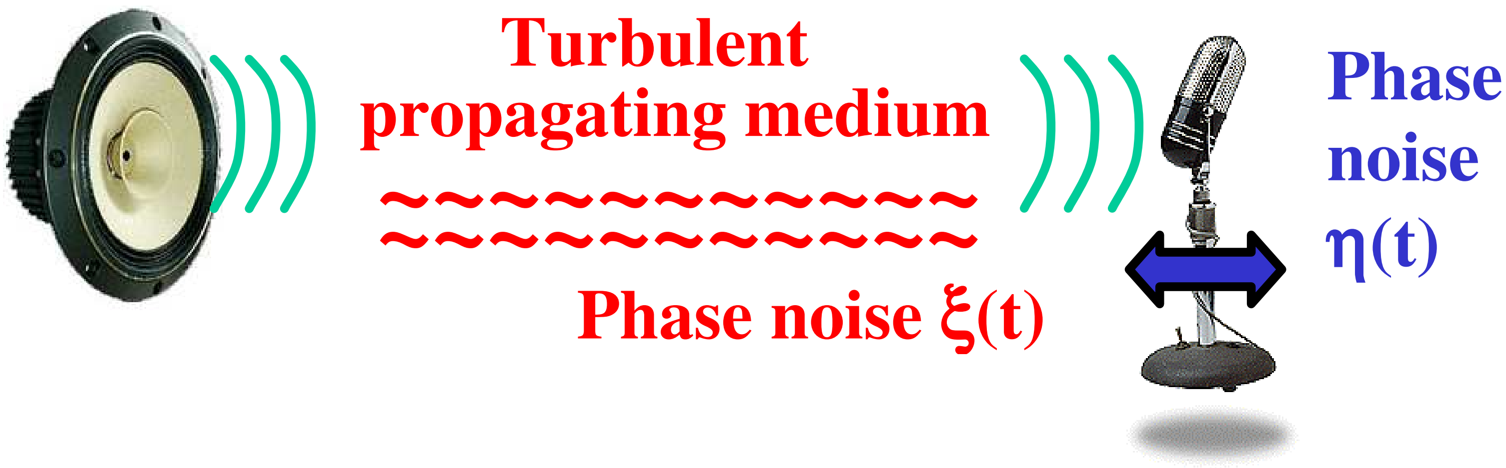
Panel B: $\boldsymbol{t} = [t_1 = 0.25 : \Delta t = 2.5 \times 10^{-3} : t_N = 0.35]$ for $N = 41$ (solid line),
$\boldsymbol{t} = [t_1 = 0.25 : \Delta t = 2.5 \times 10^{-4} : t_N = 0.35]$ for $N = 401$ (dashed).

**Improvement by noise of an optimal estimator,**

with a non-additive signal-noise mixture with Gaussian noise $\xi(t)$.

$\Longrightarrow$ the level of the initial noise $\xi(t)$ can be increased by addition of another independent Gaussian noise $\eta(t)$,

for instance, by randomly shaking the receiver.




$\Longrightarrow$ observed signal $x(t) = w[\nu t + \xi(t) + \eta(t)]$ .

# 3 Mechanism of improvement by noise

How is it possible to improve the performance of an optimal processor?

And to improve it by increasing the noise?

The measure of performance

$$P_{\text{er}} = P_1 + \int_{\mathcal{R}_1 \subset \mathbb{R}^N} \Big[ P_0 p(\boldsymbol{x}|\text{H}_0) - P_1 p(\boldsymbol{x}|\text{H}_1) \Big] d\boldsymbol{x} \quad \text{in detection,}$$

$$\mathcal{E} = \sqrt{\int_{\mathbb{R}^N} [\widehat{\nu}(\boldsymbol{x}) - \nu]^2 \, p(\boldsymbol{x};\nu) d\boldsymbol{x}} \quad \text{in estimation,}$$

is optimized by determining the best deterministic processing of data $\boldsymbol{x}$,

in presence of a fixed probabilization of the problem (fixed functions $p(\boldsymbol{x}|\text{H}_j)$ and $p(\boldsymbol{x};\nu)$ established by the statitiscal properties of the initial noise $\xi(t)$ ).

When more noise is injected, the probabilization (functions $p(\boldsymbol{x}|\text{H}_j)$ or $p(\boldsymbol{x};\nu)$) is changed.

Nothing a priori prohibits the optimal processor according to the new probabilization, to perform better than the optimal processor according to the initial probabilization (as verified by the examples).

Improvement by noise of optimal processing is made possible by a change of probabilization of the problem.

At the root, the processing problem remains the same (Which signal $s(t)$ hidden in the noise?).

It is only the measure of performance,
viewed as a functional of the probabilization established by the noise,
which changes its functional form,
when more noise is injected.

# 4 Open problems of noise

• Which optimal processing problems can benefit from a change of probabilization by injection of noise?

• How important are the characters additive / non-additive and Gaussian / non-Gaussian of the noise?

• For an *additive* signal-noise mixture with *Gaussian* white noise, is it possible to improve optimal detection or optimal estimation?

• Which changes of probabilization by injection of noise, are compatible with (authorized by) the underlying physics of the process?