ICASSP 2020 Barcelona

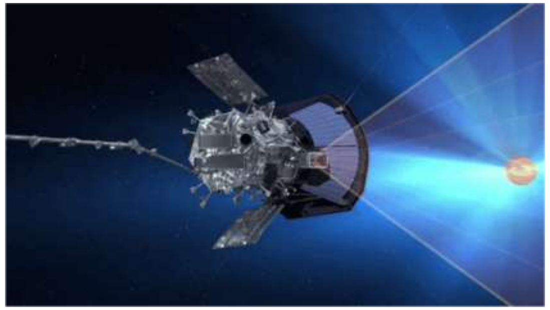
Gradient-based algorithm with spatial regularization for optimal sensor placement

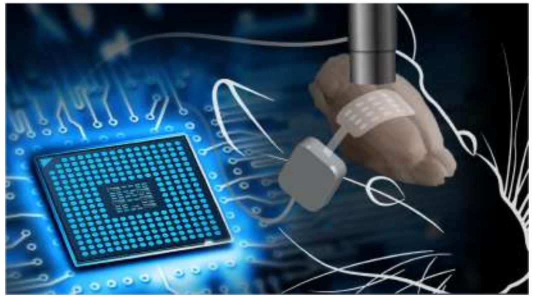
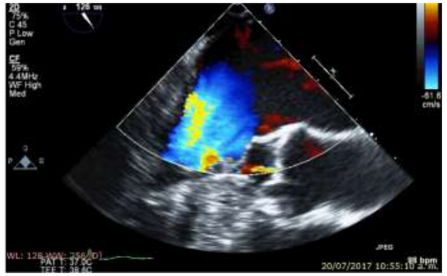
Fateme Ghayem, Bertrand Rivet, Christian Jutten, Rodrigo Cabral Farias

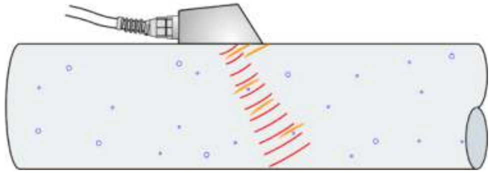
UNIVERSITÉ Grenoble Alpes
CNRS
Grenoble INP
UNIVERSITÉ CÔTE D'AZUR
gipsa-lab
Grenoble | images | parole | signal | automatique | laboratoire

Sensors are being used in a variety of domains

- Industry
- Medicine
- Wireless communications
- Aerospace engineering
- Biomedical engineering
- Civil engineering
- Environmental study
- Robotics
- …



Optimal Sensor Placement

? Why optimal sensor placement is important?

✓ Limited number of sensors:

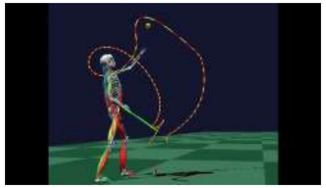
- **Economical interest**: reducing the price

- **Energy**: reducing the required energy for the power supply

- **Weight**: making the products as light as possible

- **Computational complexity**: reducing computational cost

- **Ergonomic design and arrangement** e.g. motion capture



...

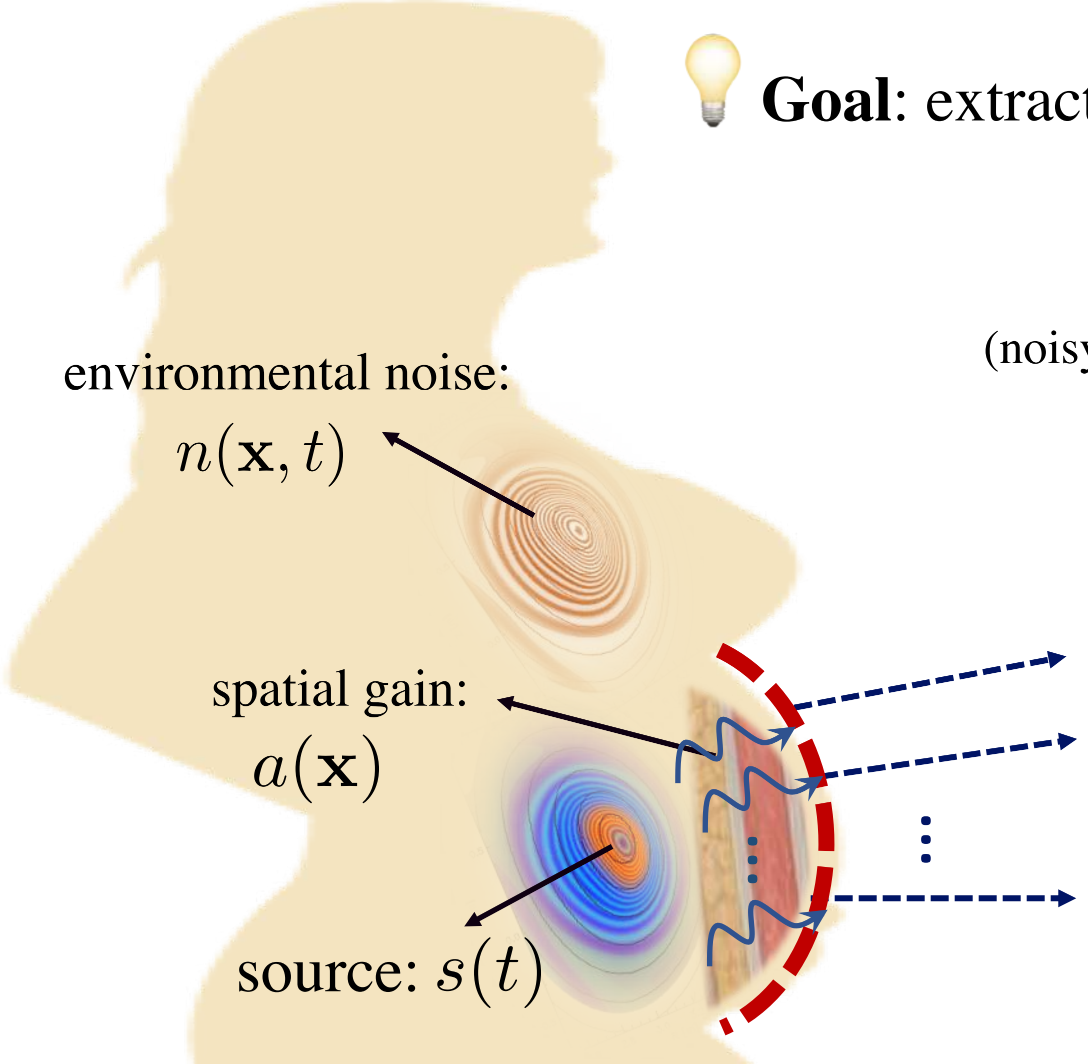
Optimal sensor placement for source extraction

Optimal Sensor Placement for Source Extraction

Goal: extract the source s(t) from a set of noisy measurements

$$\mathbf{y}(\mathbf{X}_M, t) \;\Rightarrow ?\; \hat{s}(t)$$

(noisy measurements) (source)



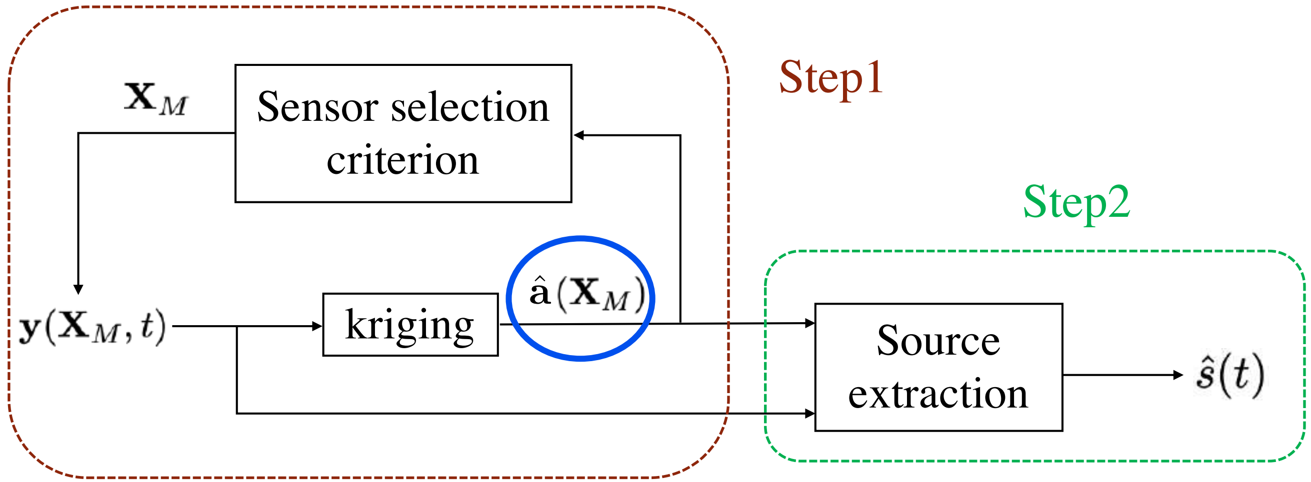
Location → Time

$$\begin{bmatrix} y(\mathbf{x}_1, t) = a(\mathbf{x}_1)s(t) + n(\mathbf{x}_1, t) \\ y(\mathbf{x}_2, t) = a(\mathbf{x}_2)s(t) + n(\mathbf{x}_2, t) \\ \vdots \\ y(\mathbf{x}_M, t) = a(\mathbf{x}_M)s(t) + n(\mathbf{x}_M, t) \end{bmatrix} = \mathbf{y}(\mathbf{X}_M, t)$$

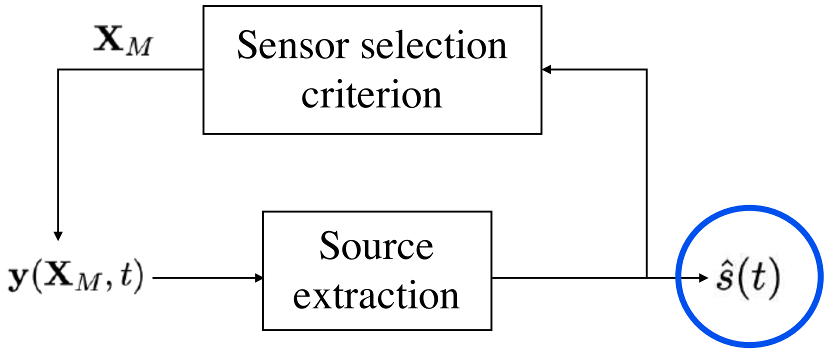
Question: Where to put the set of sensors to have the best source extraction?

Our approach *v.s*. classical kriging approaches

$$\mathbf{y}(\mathbf{X}_M, t) = \mathbf{a}(\mathbf{X}_M)s(t) + \mathbf{n}(\mathbf{X}_M, t)$$



(a) Kriging approach*



(b) Our approach**

* Maximum entropy sampling
M. C. Shewry et H. P. Wynn. Journal of Applied Statistics , (1987).

* The origins of kriging"
N. Cressie. Mathematical Geology, (1990).

** Optimal Sensor Placement for Signal Extraction
Fateme Ghayem, et.al
ICASSP2019, Brighton, UK
DOI: 10.1109/ICASSP.2019.8682432

Optimal sensor placement for Signal Extraction

- ✓ Linear source extraction: $\hat{s}(t) = \mathbf{f}^T \mathbf{y}(\mathbf{X}_M, t)$
- ✓ Targeting the output signal to noise ratio (SNR)

Max{SNR} ⟹

$$\mathbf{f}^* = (R_M^n)^{-1} \mathbf{a}(\mathbf{X}_M)$$

(Fundamentals of Statistical Signal Processing: Estimation Theory -- Steven M. Kay)

$$SNR(\mathbf{f}^*, \mathbf{X}_M) = \sigma_S^2 \, \mathbf{a}^T(\mathbf{X}_M)[\mathbf{R}^n(\mathbf{X}_M)]^{-1}\mathbf{a}(\mathbf{X}_M) = J(\mathbf{X}_M)$$

$\sigma_S^2 = \mathbb{E}[s(t)^2]$: Variance of the source

$R_M^n = \mathbb{E}[\mathbf{n}_M(t)\mathbf{n}_M^T(t)]$: Variance of the noise

$$J(\mathbf{X}_M) = \sigma_S^2 \, \mathbf{a}^T(\mathbf{X}_M)[\mathbf{R}^n(\mathbf{X}_M)]^{-1}\mathbf{a}(\mathbf{X}_M) \Longrightarrow \mathbf{X}_M^* = \arg\max_{\mathbf{X}_M} J(\mathbf{X}_M)$$

Greedy approach for optimal sensor placement

Greedy approach for optimal sensor placement

✓ Greedy approach: sequentially selecting N < M sensors at a time

Assumption: ▪ $\mathbf{X}_M$ is the union of $\mathbf{X}_K$ and $\mathbf{X}_N$: $\mathbf{X}_M = \left\{ \begin{array}{c} \mathbf{X}_K \\ \mathbf{X}_N \end{array} \right\}$

- K sensors are already allocated
- Choosing other N sensor locations

Prior knowledge

Uncertainty

Modeling assumption: stochastic Gaussian process ------> $\hat{a}(\mathbf{x}) \sim \mathcal{GP}(m^a(\mathbf{x}), R^a(\mathbf{x}, \mathbf{x}'))$

$$\hat{J}(\mathbf{X}_N|\mathbf{X}_K) = \mathbb{E}[\hat{\mathbf{a}}_{K+N}^T(\mathbf{R}_{K+N}^n)^{-1}\hat{\mathbf{a}}_{K+N}\,|\mathbf{X}_K] = (\mathbf{m}_{K+N}^a)^T(\mathbf{R}_{K+N}^n)^{-1}\mathbf{m}_{K+N}^a + \mathrm{Tr}[(\mathbf{R}_{K+N}^n)^{-1}\mathbf{R}_{K+N}^a]$$

$$f_P = f(\mathbf{X}_P)$$

$$\hat{\mathbf{X}}_N = \arg\max_{\mathbf{X}_N} \hat{J}(\mathbf{X}_N|\mathbf{X}_K)$$

$\hat{\mathbf{X}}_M$ ------> $\hat{\mathbf{f}}_M = (R_M^n)^{-1}\mathbf{m}_M^a$ ------> $\hat{s}(t)$

Greedy approach for optimal sensor placement

- Two limitations of greedy approach:

  1. Restricting sensor location on a predefined grid
  2. Suboptimal solution

- In order to be precise:

Fine grid ⇒ High computation cost

Proposed method:

Gradient-based algorithm with spatial regularization

Gradient-based algorithm with spatial regularization

$$J(\mathbf{X}_M) = (\mathbf{m}_M^a)^T(\mathbf{R}_M^n)^{-1}\mathbf{m}_M^a + \mathrm{Tr}((\mathbf{R}_M^n)^{-1}\mathbf{R}_M^a)$$

Spatial regularization

$$\min_{\mathbf{x}_M} -J(\mathbf{x}_M) \quad \text{s.t.} \quad \begin{cases} \|\mathbf{D}\mathbf{x}\|_2^2 \geq \epsilon \\ 0 \leq x_i \leq 1 \quad i \in \{1, 2, \ldots, M\} \end{cases}$$

$$\mathbf{D} \in \mathbb{R}^{\frac{M(M-1)}{2} \times M} \qquad e.g. \qquad M = 3 \implies \mathbf{D}\mathbf{x} = \begin{pmatrix} 1 & -1 & 0 \\ 1 & 0 & -1 \\ 0 & 1 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_1 - x_2 \\ x_1 - x_3 \\ x_2 - x_3 \end{pmatrix}$$

(1-D case)

Gradient-based algorithm with spatial regularization

Auxiliary variable: $\mathbf{z}_M = \mathbf{D}\mathbf{x}_M$

$$\min_{\mathbf{x}_M, \mathbf{z}_M} -J(\mathbf{x}_M) \text{s.t.} \begin{cases} \mathbf{z}_M \in \mathcal{A}_\epsilon, \\ \mathbf{z}_M = \mathbf{D}\mathbf{x}_M, \\ 0 \leq x_i \leq 1 \;\; i \in \{1, \ldots, M\} \end{cases}$$

$$\mathcal{A}_\epsilon = \left\{ \mathbf{z}_M \in \mathbb{R}^M \;\middle|\; \|\mathbf{z}_M\|_2^2 \geq \epsilon \right\}$$

Penalty method:

$$\min_{\mathbf{x}_M, \mathbf{z}_M \in \mathcal{A}_\epsilon} \left\{ -J(\mathbf{X}_M) + \frac{1}{2\alpha} \|\mathbf{z}_M - \mathbf{D}\mathbf{x}_M\|_2^2 \right\}$$

$$\text{s.t.} \;\; 0 \leq x_i \leq 1 \;\; i \in \{1, \ldots, M\}.$$

Gradient-based algorithm with spatial regularization

Alternating minimization:
$$\min_{\mathbf{x}_M, \mathbf{z}_M \in \mathcal{A}_\epsilon} \left\{ -J(\mathbf{X}_M) + \frac{1}{2\alpha} \|\mathbf{z}_M - \mathbf{D}\mathbf{x}_M\|_2^2 \right\}$$
$$\text{s.t.} \quad 0 \leq x_i \leq 1 \;\; i \in \{1, \ldots, M\}.$$

- Step1: fixing $\mathbf{x}_M$

$$\mathbf{z}_M^{(l)} = \operatorname*{argmin}_{\mathbf{z}_M \in \mathcal{A}_\epsilon} \frac{1}{2\alpha} \|\mathbf{z}_M - \mathbf{D}\mathbf{x}_M^{(l)}\|_2^2.$$

✓ Solution: projection on to the set $\mathcal{A}_\epsilon$:

$$\mathbf{z}_M^{(l)} = \begin{cases} \mathbf{D}\mathbf{x}_M^{(l)} & \text{, if} \quad \|\mathbf{D}\mathbf{x}_M^{(l)}\|_2^2 \geq \epsilon \\ \frac{\mathbf{D}\mathbf{x}_M^{(l)}}{\|\mathbf{D}\mathbf{x}_M^{(l)}\|_2^2}\epsilon & \text{, otherwise.} \end{cases}$$

- Step2: fixing $\mathbf{z}_M$

$g(\mathbf{x}_M^{(l)}, \mathbf{z}_M^{(l)})$ (smooth function)

$$\mathbf{x}_M^{(l+1)} = \operatorname*{argmin}_{\mathbf{x}_M} \left\{ -J(\mathbf{X}_M) + \frac{1}{2\alpha} \|\mathbf{z}_M^{(l)} - \mathbf{D}\mathbf{x}_M\|_2^2 \right\}$$
$$\text{s.t.} \quad 0 \leq x_i \leq 1, \;\; i \in \{1, \ldots, M\}.$$

✓ Solution: projected gradient descent:

$$\mathbf{x}_M^{(l+1)} = \mathbf{x}_M^{(l)} - \mu \nabla_{\mathbf{x}_M} g(\mathbf{x}_M^{(l)}, \mathbf{z}_M^{(l)})$$

!

✓ Chain rule

Gradient-based algorithm with spatial regularization

- Projection step:

$$\mathbf{z}_M^{(l)} = \begin{cases} \mathbf{D}\mathbf{x}_M^{(l)} & \text{, if} \quad \|\mathbf{D}\mathbf{x}_M^{(l)}\|_2^2 \geq \epsilon \\ \frac{\mathbf{D}\mathbf{x}_M^{(l)}}{\|\mathbf{D}\mathbf{x}_M^{(l)}\|_2^2}\epsilon & \text{, otherwise.} \end{cases}$$

- Gradient step:

$$\mathbf{x}_M^{(l+1)} = \mathbf{x}_M^{(l)} - \mu \nabla_{\mathbf{x}_M} g(\mathbf{x}_M^{(l)}, \mathbf{z}_M^{(l)})$$

❑ Penalty method:

$$\min_{\mathbf{x}_M, \mathbf{z}_M \in \mathcal{A}_\epsilon} \left\{ -J(\mathbf{X}_M) + \frac{1}{2\alpha} \|\mathbf{z}_M - \mathbf{D}\mathbf{x}_M\|_2^2 \right\}$$

$$\text{s.t.} \quad 0 \leq x_i \leq 1 \;\; i \in \{1, \ldots, M\}.$$

$\{\alpha_0, \alpha_1, \ldots\}$

$\alpha_{k+1} = \eta \alpha_j$, with $0 < \eta < 1$

❑ Non-convex problem

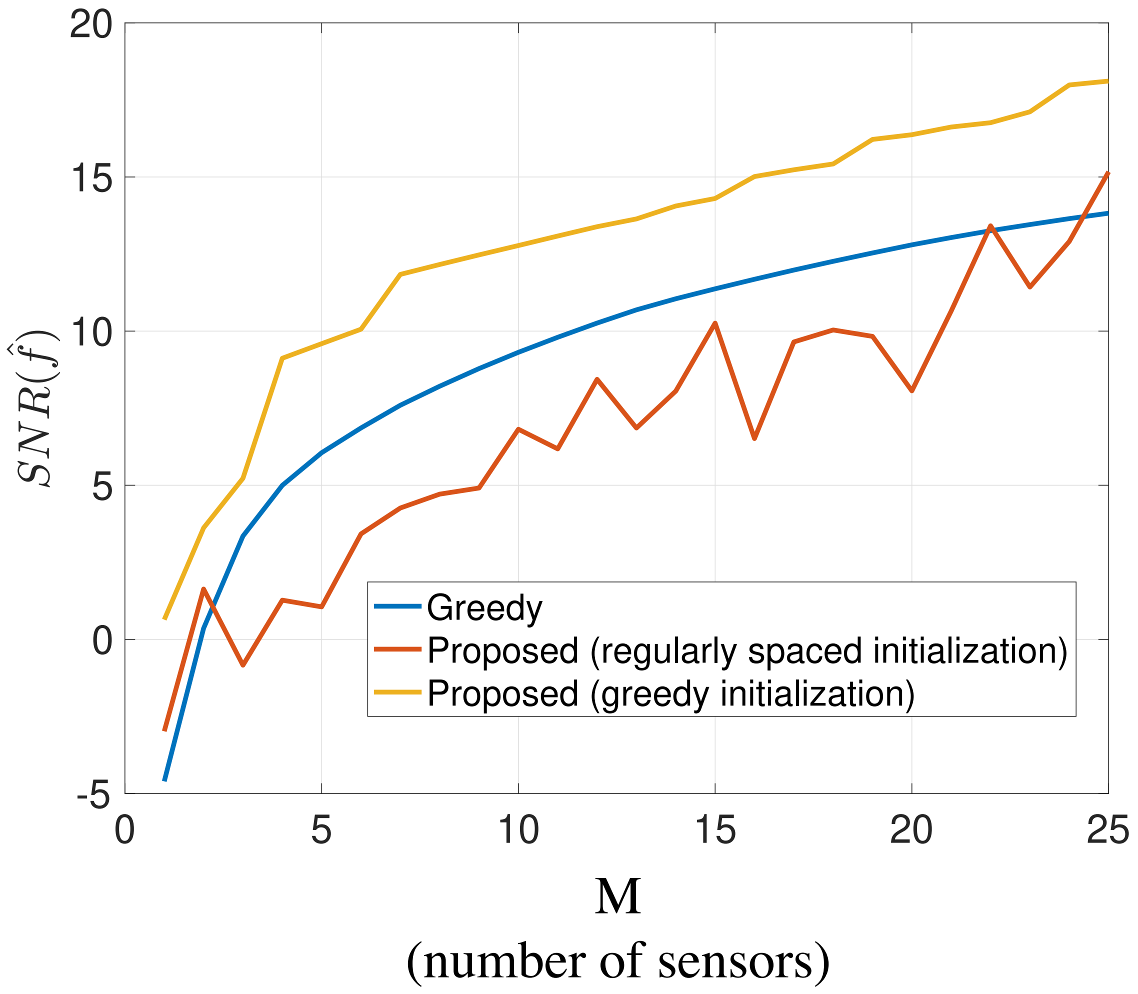
✓ Initialization with the solution obtained by the greedy approach

Numerical results

Numerical Setup

- Synthetically generated data
- D (space dimension): 1
- x : normalized in the range $x \in [0, 1]$
  - ➢ Size of the spatial grid for greedy initialization: takes different values
- $a(x)$ and $n(x)$: produced from Gaussian processes $\mathcal{GP}(m(x), C(x, x'))$
  - ➢ square exponential covariance function $C(x, x') = \sigma^2 \exp(-(x - x')^2/(2\rho^2))$
  - ➢ $m^n(x) = 0$
  - ➢ $\mathbf{m}^a(\mathbf{x}) = \sum_{i=1}^{5} \gamma_i \sin^{d_i}(w_i \pi \mathbf{x})$
  - ➢ ρ and σ : taking different values
- $\alpha_0 = 1$, $Q = 50$, $\eta = 0.5$, $\mu_0 = 1$, and $\beta = 0.5$

Influence of the initialization



Conclusion:

- Greedy initialization leads to a better extraction of the source compared with using regularly-spaced initialization.
- Proposed method improves the SNR compared to the greedy approach.
- Proposed method with regularly-spaced initialization is worse than greedy approach.

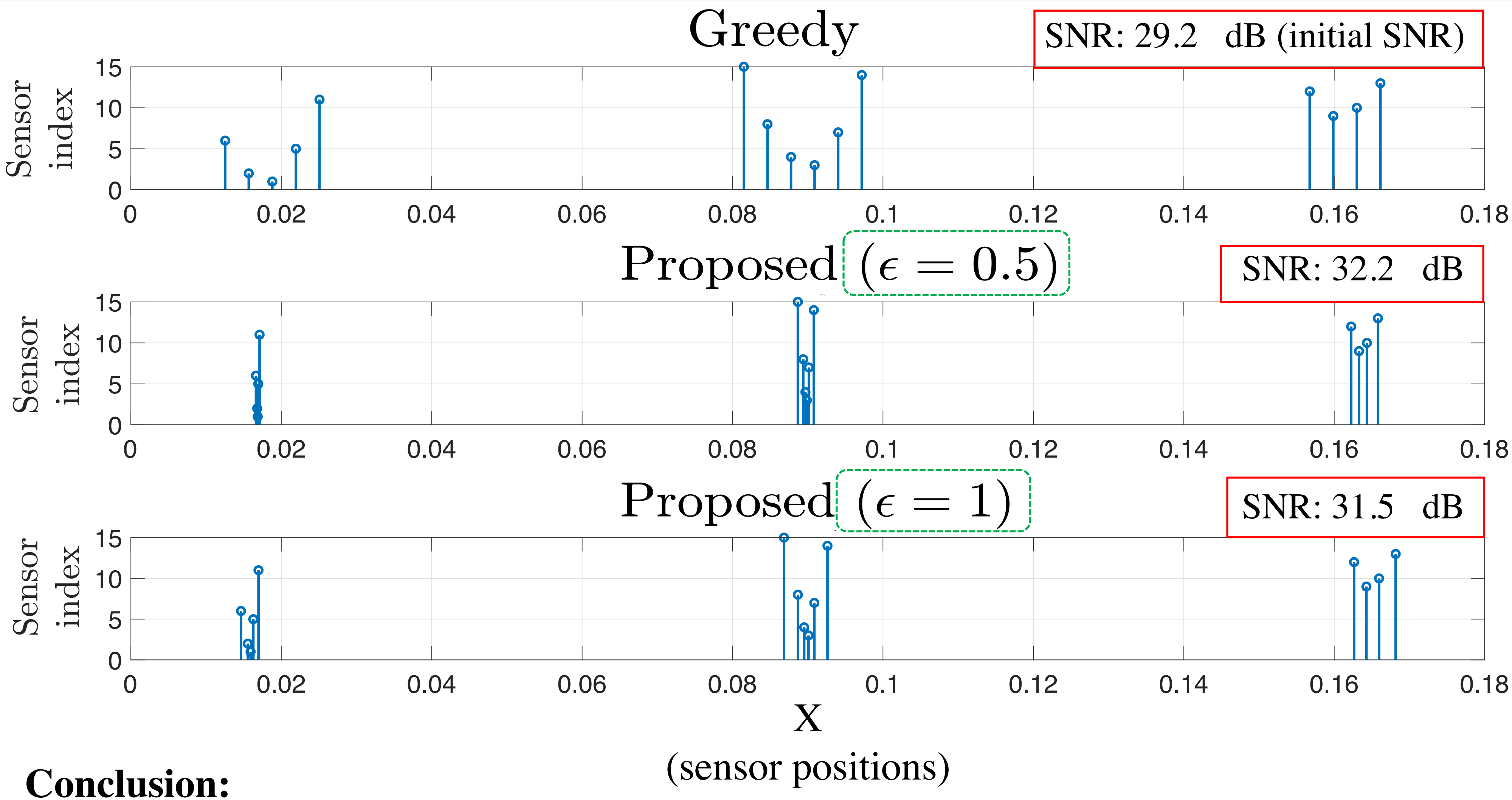
Grid size : 100

$\sigma_a = 3$ $\quad$ $\sigma_n \longrightarrow$ SNR: 0.8 dB

$\rho_a = 0.001$ $\quad$ $\rho_n = 0.01\rho_a$

lower bound ϵ on $\|\mathbf{Dx}\|_2^2$ $\quad$ $\epsilon = \frac{M(M-1)}{2} \times 10^{-3}$

Regularizing sensors distances



Grid size : 320

$\sigma_a = 1$

σ_n → SNR: 0.8 dB

$\rho_a = 0.001$

$\rho_n = 0.01\rho_a$

Number of desired sensors: M=15

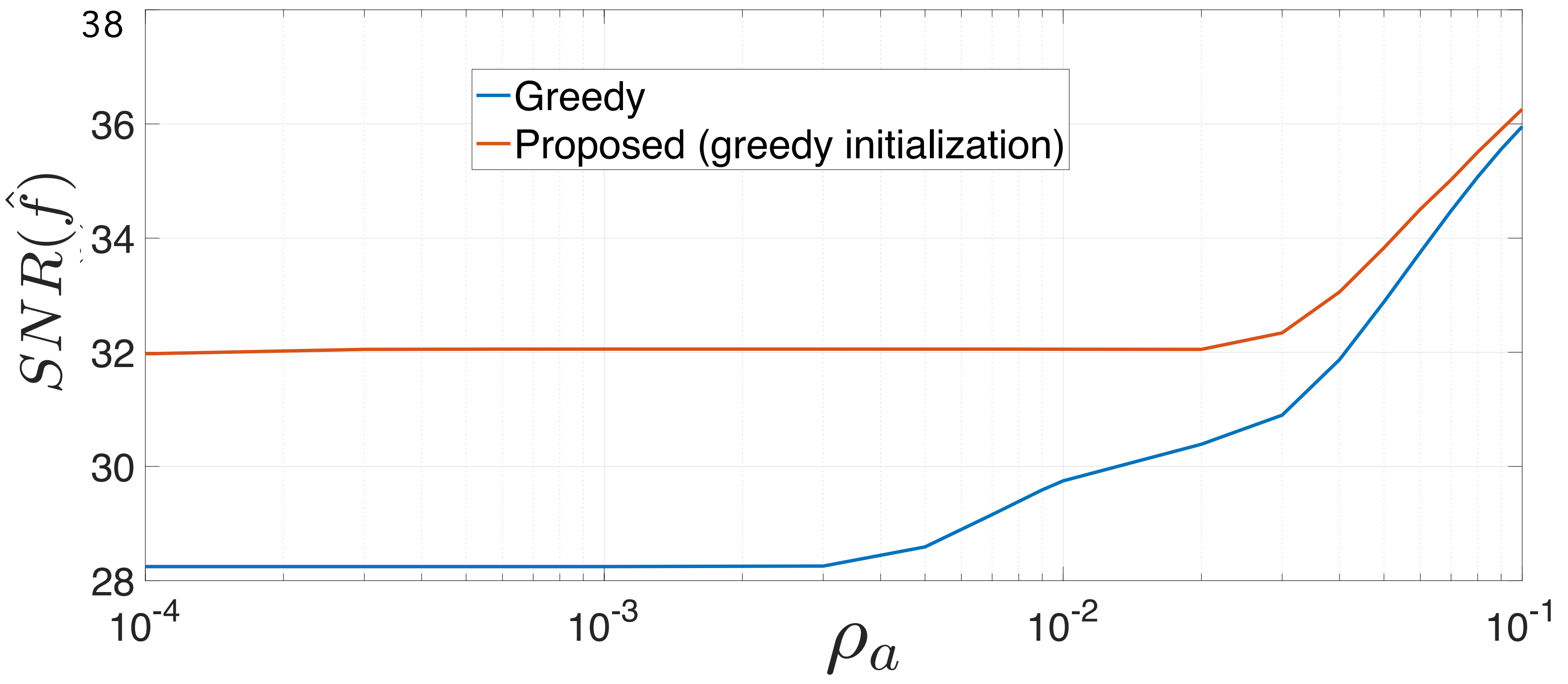
lower bound ϵ on $\|\mathbf{Dx}\|_2^2$:

$\epsilon = \frac{M(M-1)}{2} \times 10^{-3}$

Conclusion:

- Increasing ε leads to increasing the average distance between the sensors with a slightly decrease of the output SNR .

Effect of the smoothness parameter



(smoothness parameter)

Very non-smooth

Very smooth

Grid size : 100

$\sigma_a = 5$

$\sigma_n \longrightarrow$ SNR: 2 dB

$\rho_a = 0.001$

$\rho_n = 0.01\rho_a$

lower bound ϵ on $\|\mathbf{D}\mathbf{x}\|_2^2$:

$$\epsilon = \frac{M(M-1)}{2} \times 10^{-3}$$

Conclusion:

- In non-smooth cases, the performance of the greedy method deteriorates much faster than the proposed method.

Conclusions & Perspectives

Conclusions

- The problem of optimal sensor placement for signal extraction: Maximizing the output SNR
- A new gradient-based method: searching for the sensor locations over the whole space and adjusting the sensors locations at once.
- Using a spatial regularization constraint.
- Initializing with the solution of the greedy approach
- Numerical simulations:
  - Improvement of the output SNR compared to the greedy approach
  - Being able to control the average distances between the sensors

Perspectives

- An explicit constraint on each distance between pair of sensors
- Other global optimization algorithms to avoid convergence to a local optimum

Thank you!