GRADIENT-BASED ALGORITHM WITH SPATIAL REGULARIZATION FOR OPTIMAL SENSOR PLACEMENT

Fateme Ghayem, Bertrand Rivet, Christian Jutten

Univ. Grenoble Alpes, CNRS, Grenoble INP*,
GIPSA-lab, 38000 Grenoble, France
*Institute of Engineering Univ. Grenoble Alpes

Rodrigo Cabral Farias

Univ. Côte d'Azur, CNRS,
I3S, 06900 Sophia-Antipolis, France

ABSTRACT

In this paper, we are interested in optimal sensor placement for signal extraction. Recently, a new criterion based on output signal to noise ratio has been proposed for sensor placement. However, to solve the optimization problem, a greedy approach is used over a grid, which is not optimal. To improve this method, we present an optimization approach to locate all the sensors at once. We further add a constraint to the problem that controls the average distances between the sensors. To solve our problem, we use an alternating optimization penalty method. As the associated cost function is non-convex, the proposed algorithm should be carefully initialized. We propose to initialize it with the result of the greedy method. Experimental results show the superiority of the proposed method over the greedy approach.

***Index Terms*—** Sensor placement, Signal extraction, Signal to noise ratio, Alternating optimization, Penalty method.

1. INTRODUCTION

Whenever a physical phenomenon is under study, it is required to collect some data using specified number of sensors. Several considerations, including reduction of expenses, having ergonomic design, and reducing computational costs, make it necessary to limit the number of sensors. As such, it is essential to place as minimum sensors as possible in those positions that maximize information gathered by the sensors. This is called optimal sensor placement, which has been studied in various domains, with different solutions depending on the application, see, *e.g.*, [1–7].

One specific application of the optimal sensor placement problem is to extract a source signal given some measurements obtained from a limited number of sensors. In the classical kriging approaches [8–10], this problem is studied in two steps. In the first step, the optimal sensor placement problem is solved such that the sensor recordings provide a good estimation of the spatial map. Afterwards, in the second step, an extraction approach such as a linear extraction model is used to extract the source signal from the sensor measurements.

Recently, by assuming a stochastic model, a robust sensor placement criterion based on the maximization of the average signal to noise ratio (SNR) has been proposed [11]. This method works by maximizing the average SNR of the desired source, and has shown a significant improvement compared to the classical kriging approaches in terms of the output SNR. To solve the maximization problem, a greedy approach is used in which the sensors are added one by one from a fixed grid of candidate sensor positions. Two limitations can be observed: first the sensor location are restricted to be on a predefined grid and secondly the greedy approach is sub-optimal. Consequently, to be precise, the grid should be fine enough leading to a high computation cost. To solve these issues, we propose a first order optimization-based approach that in contrast to the one-by-one strategy adopted by the greedy method on a grid, optimizes all the sensor positions at once and does not discretize the search space. Finally, since placing 2 sensors very close to each other may not be feasible, e.g., due to the physical size of the sensors, a regularizing term is added to avoid to choose too close sensor positions. As the cost function is non-convex, to avoid sub optimal local minima, the proposed algorithm is initialized by the solution of the greedy approach.

The rest of this paper is organized as follows: in Section 2 a review on the problem of optimal sensor placement for source extraction, as well as the criterion suggested in [11] are presented. The details of the proposed method are discussed in Section 3. Section 4 presents the performance of the proposed method and compares it with greedy approach, before the conclusion in Section 5.

2. BACKGROUND

In this section we present an overview on signal extraction for optimal sensor placement suggested in [11]. Let $y(\mathbf{x}, t)$ denote the observation recorded at location $\mathbf{x}$ and at time t which is modelled as

$$y(\mathbf{x}, t) = a(\mathbf{x})s(t) + n(\mathbf{x}, t), \tag{1}$$

where $a(\mathbf{x})$ is the spatial gain between the signal of interest $s(t)$ and the sensor at location $\mathbf{x} \in \mathbb{R}^D$, and $n(\mathbf{x}, t)$ is a spatially correlated additive noise. The noise is assumed

978-1-5090-6631-5/20/$31.00 ©2020 IEEE

to be uncorrelated with the source. Assuming that M sensors are located at positions $\mathbf{X}_M = \{\mathbf{x}_i\}_{i\in\{1,\cdots,M\}}$, linear source extraction amounts to design a vector $\mathbf{f} \in \mathbb{R}^M$ to estimate the source as $\hat{s}(t) = \mathbf{f}^T\mathbf{y}(\mathbf{X}_M, t)$, where $\mathbf{y}(\mathbf{X}_M, t) = [y(\mathbf{x}_1, t), \ldots, y(\mathbf{x}_M, t)]^T$. To find the best $\mathbf{f}$, a classical criterion is the output signal-to-noise ratio (SNR) defined by [11]

$$SNR(\mathbf{f}) = \mathbb{E}[(\mathbf{f}^T\mathbf{a}_M s(t))^2] \,/\, \mathbb{E}[(\mathbf{f}^T\mathbf{n}_M(t))^2], \qquad (2)$$

where $\mathbf{a}_M = [a(\mathbf{x}_1), \ldots, a(\mathbf{x}_M)]^T$ denotes the vector of spatial gains and $\mathbf{n}_M(t) = [n(\mathbf{x}_1, t), \ldots, n(\mathbf{x}_M, t)]^T$ the vector of noise. The spatial gain $a(\mathbf{x})$ of the source of interest is modelled as a stochastic Gaussian process to take into account the uncertain knowledge on it:

$$\hat{a}(\mathbf{x}) \sim \mathcal{GP}(m^a(\mathbf{x}), k^a(\mathbf{x}, \mathbf{x}')), \qquad (3)$$

where $m^a(\mathbf{x})$ is the mean function and $k^a(\mathbf{x}, \mathbf{x}')$ is the covariance function. Furthermore, we consider a zero mean Gaussian process with covariance matrix $k^n(\mathbf{x}, \mathbf{x}')$ to model the noise. The criterion to optimize is then obtained as [11]:

$$J(\mathbf{X}_M) = (\mathbf{m}_M^a)^T(\mathbf{R}_M^n)^{-1}\mathbf{m}_M^a + \mathrm{Tr}((\mathbf{R}_M^n)^{-1}\mathbf{R}_M^a), \quad (4)$$

where $\mathbf{R}_M^a \in \mathbb{R}^{M\times M}$ and $\mathbf{R}_M^n \in \mathbb{R}^{M\times M}$ are the covariance matrices of the spatial gain and the noise respectively, whose (i, j)th elements are $k^a(\mathbf{x}_i, \mathbf{x}_j)$ and $k^n(\mathbf{x}_i, \mathbf{x}_j)$. Also, $\mathbf{m}_M = \{m_i^a\}_{i\in\{1,\cdots,M\}}$ is the set of means at locations $\{\mathbf{x}_i\}_{i\in\{1,\cdots,M\}}$, and $\mathrm{Tr}(\cdot)$ is the trace operator. The optimal sensor locations are finally obtained as

$$\hat{\mathbf{X}}_M = \arg\max_{\mathbf{X}_M} J(\mathbf{X}_M). \qquad (5)$$

Directly maximizing (4) in a grid requires a combinatorial search, which leads to a high computational cost. Therefore, a greedy approach has been introduced in [11] that selects the M sensors by sequentially selecting $N < M$ sensors at a time. Assuming that K sensors have already been placed, to choose the locations of the next N sensors, the following criterion is optimized:

$$J(\mathbf{X}_N|\mathbf{X}_K) = \mathbb{E}[\hat{\mathbf{a}}_{K+N}^T(\mathbf{R}_{K+N}^n)^{-1}\hat{\mathbf{a}}_{K+N} \,|\, \mathbf{X}_K], \qquad (6)$$

where $K + N$ means $\{\mathbf{X}_N \bigcup \mathbf{X}_K\}$ and thus $\hat{\mathbf{a}}_{K+N} \in \mathbb{R}^{K+N}$ can be divided as $\hat{\mathbf{a}}_{K+N} = [\hat{\mathbf{a}}_K^T, \hat{\mathbf{a}}_N^T]^T$. Once the sensor locations $\hat{\mathbf{X}}_M$ are obtained, the source of interest is extracted by using the following separation vector [11]

$$\hat{\mathbf{f}}_M = (\mathbf{R}_M^n)^{-1}\mathbf{m}_M^a. \qquad (7)$$

3. PROPOSED METHOD

In this section we present our proposed framework to solve the optimization problem for sensor placement. Unlike the greedy approach, our proposed method directly provides the positions of all the required number of sensors. By considering a one dimensional situation, we want to minimize $f(\mathbf{x}_M) = -J(\mathbf{x}_M)$. In order to control the average distances between each pair of the sensors, we constrain the sum of the squared distances to be greater than a threshold. Furthermore, due to the spatial constraints of the boundaries, we consider a normalized case where $0 \le x_i \le 1$. Therefore, we study the following minimization problem:

$$\min_{\mathbf{x}_M} f(\mathbf{x}_M) \;\; \text{s.t.} \;\; \begin{cases} \|\mathbf{D}\mathbf{x}\|_2^2 \ge \epsilon \\ 0 \le x_i \le 1 \quad i \in \{1, 2, \ldots, M\} \end{cases}, \quad (8)$$

where $\mathbf{D} \in \mathbb{R}^{\frac{M(M-1)}{2}\times M}$ is a matrix that enumerates all the possible combinations of positions in pairs of size two. For instance, if the number of sensors is $M = 3$, then

$$\mathbf{D}\mathbf{x} = \begin{pmatrix} 1 & -1 & 0 \\ 1 & 0 & -1 \\ 0 & 1 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_1 - x_2 \\ x_1 - x_3 \\ x_2 - x_3 \end{pmatrix}.$$

To solve (8), we define an auxiliary variable $\mathbf{z}_M = \mathbf{D}\mathbf{x}_M$, and reformulate (8) as the following problem:

$$\min_{\mathbf{x}_M, \mathbf{z}_M} f(\mathbf{x}_M) \;\; \text{s.t.} \;\; \begin{cases} \mathbf{z}_M \in \mathcal{A}_\epsilon, \\ \mathbf{z}_M = \mathbf{D}\mathbf{x}_M, \\ 0 \le x_i \le 1 \;\; i \in \{1, \ldots, M\}, \end{cases} \qquad (9)$$

where $\mathcal{A}_\epsilon = \left\{\mathbf{z}_M \in \mathbb{R}^M \;\middle|\; \|\mathbf{z}_M\|_2^2 \ge \epsilon\right\}$. To solve (9), we use the penalty method [12], by adding the constraint $\mathbf{z}_M = \mathbf{D}\mathbf{x}_M$ as a penalty to the target function with the penalty parameter α:

$$\min_{\mathbf{x}_M, \mathbf{z}_M \in \mathcal{A}_\epsilon} \left\{ f(\mathbf{x}_M) + \frac{1}{2\alpha}\|\mathbf{z}_M - \mathbf{D}\mathbf{x}_M\|_2^2 \right\}$$
$$\text{s.t.} \;\; 0 \le x_i \le 1 \;\; i \in \{1, \ldots, M\}. \qquad (10)$$

To solve (10), we do the minimization over $\mathbf{x}_M$ and $\mathbf{z}_M$ alternately. At iteration l, first, the cost is optimized over $\mathbf{z}_M$, fixing $\mathbf{x}_M$ to its current estimate $\mathbf{x}_M^{(l)}$. That is:

$$\mathbf{z}_M^{(l)} = \operatorname*{argmin}_{\mathbf{z}_M \in \mathcal{A}_\epsilon} \frac{1}{2\alpha}\|\mathbf{z}_M - \mathbf{D}\mathbf{x}_M^{(l)}\|_2^2. \qquad (11)$$

The solution to the above minimization is a projection onto the set $\mathcal{A}_\epsilon$ as follows:

$$\mathbf{z}_M^{(l)} = \begin{cases} \mathbf{D}\mathbf{x}_M^{(l)} & \text{, if} \quad \|\mathbf{D}\mathbf{x}_M^{(l)}\|_2^2 \ge \epsilon \\ \frac{\mathbf{D}\mathbf{x}_M^{(l)}}{\|\mathbf{D}\mathbf{x}_M^{(l)}\|_2^2}\epsilon & \text{, otherwise.} \end{cases} \qquad (12)$$

For the second step, the variable $\mathbf{z}_M$ is fixed as in (12), and we do the minimization over $\mathbf{x}_M$ as follows:

$$\mathbf{x}_M^{(l+1)} = \operatorname*{argmin}_{\mathbf{x}_M} \left\{ f(\mathbf{x}_M) + \frac{1}{2\alpha}\|\mathbf{z}_M^{(l)} - \mathbf{D}\mathbf{x}_M\|_2^2 \right\}$$
$$\text{s.t.} \;\; 0 \le x_i \le 1, \;\; i \in \{1, \ldots, M\}. \qquad (13)$$

Since the constraint is a quite simple convex set, to solve (13), a projected gradient descent is used: after a gradient descent

(GD) update, the result is projected onto $[0,1]$. That is, by defining the cost function as follows:

$$g(\mathbf{x}_M) = f(\mathbf{x}_M) + \frac{1}{2\alpha}\|\mathbf{z}_M^{(l)} - \mathbf{D}\mathbf{x}_M\|_2^2, \tag{14}$$

the gradient step to optimize (13) is

$$\mathbf{x}_M^{(l+1)} = \mathbf{x}_M^{(l)} - \mu\nabla_{\mathbf{x}_M} g(\mathbf{x}_M^{(l)}), \tag{15}$$

where, $\nabla g(\mathbf{x}_M^{(l)})$ is the gradient of the smooth function $g(.)$ at the previously updated point $\mathbf{x}_M^{(l)}$, and $\mu > 0$ is a step size. To derive the gradient of $g(.)$, we can write:

$$\nabla_{\mathbf{x}_M} g(\mathbf{x}_M^{(l)}) = \nabla_{\mathbf{x}_M} f(\mathbf{x}_M^{(l)}) - \alpha^{-1}\mathbf{D}^T(\mathbf{z}_M^{(l)} - \mathbf{D}\mathbf{x}_M^{(l)}). \tag{16}$$

We use the chain rule to calculate $\nabla f(\mathbf{x}_M^{(l)})$ as follows:

$$\frac{\partial f(\mathbf{x}_M^{(l)})}{\partial x_i} = \text{Tr}\left[\left(\frac{\partial f(\mathbf{m}_M^a)}{\partial \mathbf{m}_M^a}\right)^T \frac{\partial \mathbf{m}_M^a}{\partial x_i}\right]$$
$$+\text{Tr}\left[\left(\frac{\partial f(\mathbf{R}_M^a)}{\partial \mathbf{R}_M^a}\right)^T \frac{\partial \mathbf{R}_M^a}{\partial x_i}\right] + \text{Tr}\left[\left(\frac{\partial f(\mathbf{R}_M^n)}{\partial \mathbf{R}_M^n}\right)^T \frac{\partial \mathbf{R}_M^n}{\partial x_i}\right].$$

where x_i is the i^{th} element of $\mathbf{x}_M$. The above expression is simplified to the following:

$$\frac{\partial f(\mathbf{x}_M^{(l)})}{\partial x_i} = \text{Tr}\Bigg[-2(\mathbf{m}_M^a)^T(\mathbf{R}_M^n)^{-1}\frac{\partial \mathbf{m}_M^a}{\partial x_i} - (\mathbf{R}_M^n)^{-1}\frac{\partial \mathbf{R}_M^a}{\partial x_i}$$
$$+ (\mathbf{R}_M^n)^{-1}(\mathbf{m}_M^a {\mathbf{m}_M^a}^T + \mathbf{R}_M^a)(\mathbf{R}_M^n)^{-1}\frac{\partial \mathbf{R}_M^n}{\partial x_i}\Bigg],$$

where $\frac{\partial \mathbf{m}_M^a}{\partial x_i} = [\partial m_j^a/\partial x_i]_j$, and $\frac{\partial C}{\partial x_i} = [\partial \mathbf{C}_{ij}/\partial x_i]_{(i,j)}$ in which $\mathbf{C}$ represents any covariance matrix, and $\mathbf{C}_{ij}$ corresponds to its (i,j)th entry. This way, we have computed the gradient of $f(\mathbf{x}_M^{(l)})$ over the i^{th} sensor position, providing thus the expression of the gradient vector $\nabla_{\mathbf{x}_M} f(\mathbf{x}_M^{(l)})$.

To determine μ in (15), we use a backtracking line search strategy [12]. After updating $\mathbf{x}_M$ using (15), any element of $\mathbf{x}_M^{(l+1)}$ is projected into $[0,1]$.

Finally, to solve (10) we start with an initial point and alternate between the projection step (12) and the GD step (15). As done in standard penalty methods [12], the problem (10) should be solved for a decreasing sequence of α, *e.g.*, as $\{\alpha_0, \alpha_1, \ldots\}$ where $\alpha_{k+1} = \eta\alpha_j$, with $0 < \eta < 1$. For each fixed value of α, we perform a few iterations between (12), (15) projecting $\mathbf{x}$. Moreover, iterations corresponding to α_{j+1} are initialized by the final estimate found for α_j. The final algorithm to solve (8) is summarized in Algorithm 1.

Since the problem (9) is non-convex, its initialization is important to find an appropriate minimizer. We propose to initialize the algorithm with the solution obtained by the greedy approach [11]. In this way, the algorithm is more likely to end up with a good local minimum.

Algorithm 1 Alternating minimization (AM) for solving (9)

```
Inputs: $\{\mathbf{x}^{(0)}, \mathbf{z}^{(0)}\}$, $\mu_0$, $\alpha_0$, $Q$
Initialization: Set $\mu = \mu_0$, $l = 0$
for $j = 1, 2, \cdots, Q$ do
    while stopping criterion not met do
        $t^{(l)} = \|\nabla g(\mathbf{x}^{(l)})\|_2^2$
        while $g(\mathbf{x}^{(l)} - \mu\nabla g(\mathbf{x}^{(l)})) > g(\mathbf{x}^{(l)}) - \frac{\mu}{2}t^{(l)}$ do
            $\mu \leftarrow \beta \cdot \mu$
        end while
        $\mathbf{x}^{(l+1)} = \mathbf{x}^{(l)} - \mu\nabla g(\mathbf{x}^{(l)})$
        Project $\mathbf{x}^{(l+1)}$ into $[0,1]$
        Perform projection (12) to obtain $\mathbf{z}^{(l+1)}$
        $l \leftarrow l + 1$
    end while
    $\alpha_{j+1} = \eta \cdot \alpha_j$
end for
Output: $\mathbf{x}^{(l)}$
```

4. NUMERICAL EXPERIMENTS

In this section, the numerical setup is first presented. Then the influence of the initialization, the effect of the regularization based on the sensors distances and of the smoothness of the spatial gain are presented in Subsections 4.2, 4.3 and 4.4, respectively.

4.1. Numerical setup

Synthetic data are generated in a 1D space, where the range of the sensor locations is normalized between 0 and 1. We consider a prior on the spatial gain and noise to be generated from $\mathcal{GP}(m(x), C(x,x'))$, with a square exponential covariance function $C(x,x') = \sigma^2\exp(-(x-x')^2/(2\rho^2))$. The mean of the noise is set to be 0. The mean of the gain is given by $\mathbf{m}^a(\mathbf{x}) = \sum_{i=1}^{5}\gamma_i \sin^{d_i}(w_i\pi\mathbf{x})$, where, γ_i, d_i and w_i are the i^{th} elements of the vectors $\mathcal{G} = [0.1,\, 0.2,\, \cdots,\, 0.9]$, $\mathcal{D} = [1,1,3,1,2]$, and $\mathcal{W} = [25,\, 26,\, \cdots,\, 29]$, respectively. The smoothness parameters ρ_n, and ρ_a, and the variances σ_n and σ_a as well as the size of the spatial grid for greedy initialization take different values for each experiment. Also, we set $\alpha_0 = 1$, $Q = 50$, $\eta = 0.5$, $\mu_0 = 1$, and $\beta = 0.5$.

4.2. Influence of the initialization

We set the size of the spatial grid to be 100. Two different values of the uncertainty on the spatial gain are considered, $\sigma_a = 1$ and $\sigma_a = 3$. The noise variance σ_n is accordingly set such that the SNR becomes 0.8 dB. The smoothness of the uncertainty on the spatial gain (ρ_a) is set to $\rho_a = 0.001$, which corresponds to an uncertainty with almost no spatial correlation. The spatial smoothness of the noise ρ_n is set to $\rho_n = 0.01\rho_a$. The lower bound ϵ on $\|\mathbf{D}\mathbf{x}\|_2^2$ is also considered to be $\epsilon = \frac{M(M-1)}{2} \times 10^{-3}$.

The true value of the output SNR (2) computed with estimated extraction vector $\hat{\mathbf{f}}_M$ (7) versus the number of desired

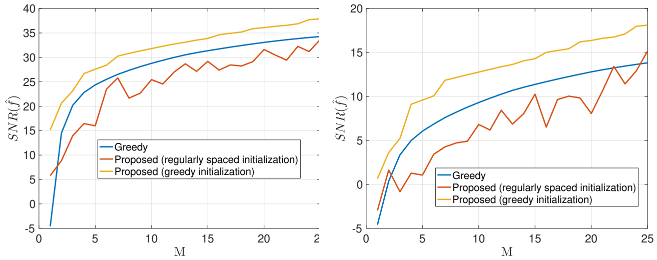

Fig. 1. Influence of the initialization. Output SNR *vs.* the number of sensors. left: $\sigma_a = 1$, right: $\sigma_a = 3$.

sensors M are depicted in Fig. 1. If no prior information is used for the initialization, one can initialize the sensor locations regularly-spaced between 0 and 1. On the contrary, one can use the previously proposed greedy approach, where each sensor are added one by one [11], to initialize the sensor locations. Firstly, one can see that the greedy initialization leads to a better extraction of the source $s(t)$ than using regularly-spaced initial locations for the sensors before applying our proposed method to adjust the sensor locations. Indeed, the difference of output SNRs varies between 10dB for a single sensor and 5dB for 25 sensors.

Moreover, the proposed method to adjust the sensor locations leads to improve the SNR of about 3dB to 5dB compared to the greedy approach. Indeed, this result is expected since the proposed method tackles the optimization of the sensor locations all at the same time instead of one after the other as in the greedy method. It is also worth noting that the output SNR is worse by applying the proposed method with a regularly-spaced initialization than by just choosing the sensor locations by the greedy method proposed in [11] with no additional adjustment.

4.3. Regularizing sensors distances

Figure 2 shows the effect of regularizing sensor distances and how it can help to control the average distances between pairs of sensors. In this part, all the parameters are set as in the previous section with $\sigma_a = 1$, except that here we consider a tighter grid of size 320. Also, the number of desired sensors are set to be $M = 15$. For the proposed method, two different values for the lower bound are considered: $\epsilon = 0.5$ and $\epsilon = 1$.

The second and the third sub-figures demonstrate the effect of ϵ in tuning the average distances between the sensors. The final SNR values for each approach, from top to bottom, are 29.22 dB (initial SNR), 32.22 dB ($\epsilon = .5$), and 31.53 dB ($\epsilon = 1$), respectively. However, increasing ϵ leads to only a slight decrease of the output SNR while increasing the average distance between the sensors.

4.4. Effect of the smoothness parameter ρ_a

In this part we study the performance for different smoothness levels of the uncertainty on the spatial gain (ρ_a). We consider an almost difficult situation with the uncertainty parameter of the spatial gain to be $\sigma_a = 5$. We also considered SNR to be 2 dB to set σ_n. The rest of the parameters are set similarly as in subsection 4.2. In Fig. 3, the output SNR versus the degree of the spatial gain smoothness is depicted. We can see that as the signal becomes more non-smooth, the performance of the greedy approach deteriorates much faster than the proposed method. This is due to the presence of highly informative sensor positions in between grid points, which cannot be chosen by the greedy approach.

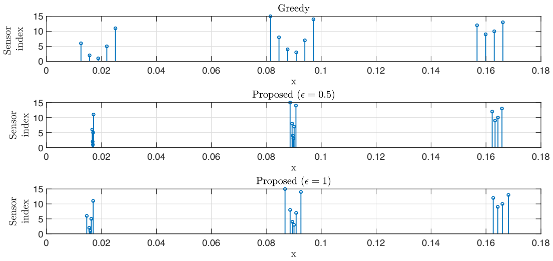

Fig. 2. Effect of the regularization parameter ϵ to control the sensor distances. Top: initial sensors localisation, middle and bottom final sensors localisation for $\epsilon = .5$ and $\epsilon = 1$, respectively.

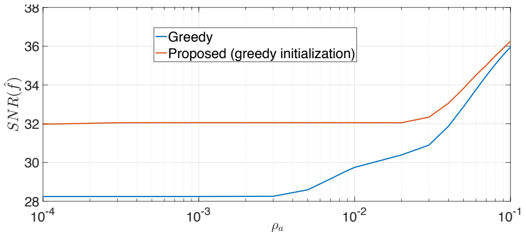

Fig. 3. Effect of the smoothness parameter of spatial gain ρ_a.

5. CONCLUSIONS AND PERSPECTIVES

In this paper, we addressed optimal sensor placement for signal extraction by maximizing output signal to noise ratio. In contrast to the greedy approach proposed in [11], the proposed method adjust all the sensors locations at once instead of choosing them one at a time. To this end, a gradient-based method is proposed to search for the sensor locations over the whole space. A constraint, controlling the average distances between sensors, is also considered to avoid to choose too close sensors (*e.g.*, depending of their size). Due to the non-convexity of the cost function, the proposed algorithm is initialized with the solution of the greedy approach. Experimental results demonstrate that the proposed method provides about 3 dB improvements over the greedy approach. Also, thanks to the new constraint, the proposed method is shown to be able to control the average distances between the sensors. In future works, an explicit constraint on each distance between pair of sensors will be studied as well as other global optimization algorithms to avoid convergence to a local optimum.

6. REFERENCES

[1] J. W. Berry, L. Fleischer, W. E. Hart, C. A. Phillips, and J. P. Watson, “Sensor placement in municipal water networks,” *Journal of Water Resources Planning and Management*, vol. 131, no. 3, May 2005.

[2] M. Meo and G. Zumpano, “On the optimal sensor placement techniques for a bridge structure,” *Engineering Structures*, vol. 27, no. 10, pp. 1488–1497, August 2005.

[3] H. J. Hermens, B. Frenriks, C. Disselhorst-Klugb, and G. Rau, “Development of recommendations for SEMG sensors and sensor placement procedures,” *Journal of Electromyography and Kinesiology*, vol. 10, no. 5, pp. 361–374, October 2000.

[4] S. Y. Chen and Y. F. Li, “Automatic sensor placement for model-based robot vision,” *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 34, no. 1, pp. 393–408, February 2004.

[5] C. K. Daniel, “Sensor placement for on-orbit modal identification and correlation of large space structures,” *Journal of Guidance, Control, and Dynamics*, vol. 14, no. 2, pp. 251–259, 1991.

[6] S. Martinez and F. Bullo, “Optimal sensor placement and motion coordination for target tracking,” *Automatica*, vol. 42, no. 4, pp. 661–668, April 2006.

[7] A. Krause, J. Leskovec, C. Guestrin, J. VanBriesen, and C. Faloutsos, “Efficient sensor placement optimization for securing large water distribution networks,” *Journal of Water Resources Planning and Management*, vol. 134, no. 6, November 2008.

[8] M. C. Shewry and H. P. Wynn, “Maximum entropy sampling,” *Journal of Applied Statistics*, vol. 14, no. 2, pp. 165–170, 1987.

[9] N. Cressie, “The origins of kriging,” *Mathematical Geology*, vol. 22, no. 3, pp. 239–252, 1990.

[10] A. Krause, A. Singh, and C. Guestrin, “Near-optimal sensor placements in gaussian processes: Theory, efficient algorithms and empirical studies,” *Journal of Machine Learning Research*, vol. 9, pp. 235–284, 2008.

[11] F. Ghayem, B. Rivet, C. Jutten, and R. Cabral Farias, “Optimal sensor placement for signal extraction,” *IEEE International Conference on Acoustics, Speech, and Signal Processing*, pp. 4978–4982, May 2019.

[12] D. P. Bertsekas, *Nonlinear Programming*, Athena Scientific, Belmont, MA, 1999.