Near-Optimal Greedy Sensor Selection for
MVDR Beamforming with Modular Budget Constraint

Mario Coutino, Sundeep Chepuri, and Geert Leus

Delft University of Technology, Delft, The Netherlands.

ABSTRACT

In this paper, we present a greedy sensor selection algorithm for minimum variance distortionless response (MVDR) beamforming under a modular budget constraint. In particular, we propose a submodular set-function that can be maximized using a linear-time greedy heuristic that is near optimal. Different from the convex formulation that is typically used to solve the sensor selection problem, the method in this paper neither involves computationally intensive semidefinite programs nor convex relaxation of the Boolean variables. While numerical experiments show a comparable performance between the convex and submodular relaxations, in terms of output signal-to-noise ratio, the latter finds a near-optimal solution with a significantly reduced computational complexity.

Index Terms— submodularity, MVDR beamforming, greedy algorithm, budget constraint, sensor selection

1. INTRODUCTION

Nowadays, advances in technology have allowed the deployment of large-scale sensor networks for distributed data sampling. In this setup, sensor nodes are deployed in different locations and are capable of autonomously processing the gathered data. As a whole, the network aims to obtain relevant information from the process it monitors. In this paper, we focus on spatial beamforming, where we are interested in extracting the signal that impinges on an antenna array from a particular direction, and filter out spatial signals which are not of interest.

Large networks usually generate prohibitively large datasets which are gathered at a central unit for further processing. Hence, methods for optimally selecting the data, i.e., removing non-informative measurements, are required to alleviate any possible bottleneck in the processing chain. As a result, sensor selection for large-scale networks is of great importance [1],[2], specially when budget constraints are enforced by the system requirements. Particularly for spatial beamforming applications, deploying new sensors in the network incurs an additional cost. For example, a sensor incurs a higher communications cost when it is deployed far from the central unit as compared to a sensor that is deployed closer to the central unit. In such cases, not only does the inference performance metric have to be optimized, but also the cost incurred due to the sensor placement has to be taken into account.

Typically, such sensor selection problems are combinatorial in nature and are NP-hard. As a result they become intractable even with a small number of candidate sensors. However, several methods are at our disposal for finding approximate solutions to the sen-

sor selection problem. Popular solutions are based on convex optimization techniques [3]-[5], which are extensively used within the signal processing community. Most of these methods are expressed as semidefinite programs [6] which can be efficiently solved with a complexity that is cubic in the number of sensors. However, great interest has arisen in sensor selection methods that optimize submodular surrogates of the cost function to optimize [7]-[11]. This is due to the near-optimality guarantee [12] of the solution provided by a greedy heuristic which has a complexity that is linear in the number of sensors. Such methods perfectly fit very large scale problems [13].

In this paper, we focus on a fast and scalable solution for the sensor selection problem related to spatial minimum variance distortionless response (MVDR) beamforming with a modular budget constraint. To achieve this, we propose a submodular relaxation of the output signal-to-noise ratio, which can be solved through a greedy heuristic and it also has a link with the common convex approach for this problem. Due to the relation between the submodular and convex relaxations, comparable results are achieved, in linear time, through the usage of the submodular machinery.

2. PRELIMINARIES

2.1. Data Model

Assume that a signal of interest, e.g., speech signal, radiation of a star, or a target reflection, impinges from a direction θ on an M-element sensor array of arbitrary geometry. The received signal vector of the antenna array can then be expressed as

$$\mathbf{x} = \mathbf{a}(\theta)s + \mathbf{n} \in \mathbb{C}^{M \times 1}, \tag{1}$$

where $\mathbf{a}(\theta) \in \mathbb{C}^{M \times 1}$ is the array manifold vector. The signal of interest $s \sim \mathcal{CN}(0, \sigma_s^2)$ and the noise vector $\mathbf{n} \sim \mathcal{CN}(\mathbf{0}, \mathbf{R}_n)$ are considered to be zero-mean complex Gaussian distributed. The covariance matrix $\mathbf{R}_x = E\{\mathbf{x}\mathbf{x}^H\}$ of the received data is given by

$$\mathbf{R}_x = \sigma_s^2 \mathbf{a}(\theta)\mathbf{a}^H(\theta) + \mathbf{R}_n \in \mathbb{C}^{M \times M}, \tag{2}$$

where $E\{\cdot\}$ denotes the expectation operation. In (2), it is implied that the signal and noise are mutually uncorrelated.

2.2. MVDR Beamforming

When the direction θ_0 of the signal of interest is known, one can apply a spatial filter $\mathbf{z}$ to remove the noise, and possibly the interference, to obtain an estimate of the desired signal s, i.e.,

$$\hat{s} = \mathbf{z}^H \mathbf{x} \in \mathbb{C}. \tag{3}$$

A widely used filter option is the so-called MVDR beamformer. It aims to minimize the average output power with the constraint that the signal at the direction of interest is undistorted. Mathematically, the MVDR beamformer is given by the solution to

$$\underset{\mathbf{z} \in \mathbb{C}^M}{\text{minimize}} \quad \mathbf{z}^H \mathbf{R}_x \mathbf{z}, \text{ s.t. } \mathbf{z}^H \mathbf{a}(\theta_0) = 1. \tag{4}$$

A closed form solution to the above problem can be given in terms of the noise covariance matrix:

$$\mathbf{z}^* = \frac{\mathbf{R}_n^{-1}\mathbf{a}(\theta_0)}{\mathbf{a}^H(\theta_0)\mathbf{R}_n^{-1}\mathbf{a}(\theta_0)}. \tag{5}$$

A sparse variation of the MVDR might be achieved by enforcing a ℓ_1-norm constraint [18][18], i.e., $\|\mathbf{z}\|_1 \leq K$, in problem (4) or by its LASSO formulation. However, these approaches neither guarantee a fixed cardinality of the selected subset nor provide a straightforward way to enforce a budget constraint. In addition, despite that the LASSO solution meets the MVDR property, after thresholding the solution's low-valued coefficients this might not be true anymore.

3. PROBLEM STATEMENT

Despite the fact that using all the M available sensors provides the best performance in terms of the problem (4), in many cases, budget constraints, e.g., cost of deployment, processing power, etc., do not allow to use them. Hence, it is of interest to obtain the subset of sensors that meets all the budget constraints as well as achieves the best performance among all possible selections. Unfortunately, this problem is in general NP-hard, i.e., no polynomial time algorithm can provide the optimal solution. As a result, one is required to settle with approximate solutions through surrogate cost functions which are much simpler to optimize.

To select the best subset of K sensors out of M candidate sensors under a budget constraint, the MVDR beamforming problem in (4) can be written in terms of a cost set-function with an additional budget constraint:

$$\begin{array}{ll} \underset{\mathcal{A}}{\text{minimize}} & \mathbf{z}_{\mathcal{A}}^H \mathbf{R}_{x,\mathcal{A}} \mathbf{z}_{\mathcal{A}} \\ \text{subject to} & \mathbf{z}_{\mathcal{A}}^H \mathbf{a}_{\mathcal{A}}(\theta_0) = 1, \; B(\mathcal{A}) \leq \beta, \; |\mathcal{A}| = K, \end{array} \tag{6}$$

where $\mathbf{z}_{\mathcal{A}}$ denotes a column vector which consists only of the entries of $\mathbf{z}$ indexed by the set $\mathcal{A}$, $\mathbf{R}_{x,\mathcal{A}}$ denotes the matrix generated by selecting only the rows and columns of $\mathbf{R}_x$ indexed by the set $\mathcal{A}$, and $B(\mathcal{A})$ is a budget function. The set $\mathcal{A}$ represents the subset of selected sensors, where $\mathcal{A} \subseteq \mathcal{V}$ with $\mathcal{V} = \{1, 2, \ldots, M\}$ being the underlying finite ground set. For example, a typical budget set-function $B(\mathcal{A})$, representing the transmission costs incurred for transmitting data from sensors to the central unit in a distributed setup, can be expressed as

$$B(\mathcal{A}) = \sum_{i \in \mathcal{A}} b_i, \tag{7}$$

where each b_i is the cost related to the ith sensor in the subset of sensors $\mathcal{A}$. In general, set-functions of the form (7) are known as *modular* set-functions. By applying the optimal solution for the filter coefficients in (5) to the problem (6) it can be shown that (6) is equivalent to

$$\begin{array}{ll} \underset{\mathcal{A}}{\text{maximize}} & \mathbf{a}_{\mathcal{A}}^H(\theta_0) \mathbf{R}_{n,\mathcal{A}}^{-1} \mathbf{a}_{\mathcal{A}}(\theta_0) \\ \text{subject to} & B(\mathcal{A}) \leq \beta, \; |\mathcal{A}| = K, \end{array} \tag{8}$$

which implies that we aim to maximize the output signal-to-noise ratio, given that the constraints are met. In this work in particular, we focus on instances of the problem (8) where it is of interest to find a subset of sensors of size $K \ll M$ when the budget set-function is modular. In (8) the set-function has been defined in terms of the noise covariance matrix, however the methods here proposed do not change when $\mathbf{R}_{x,\mathcal{A}}$ is used instead. Here, we focus on the case in which $\mathbf{R}_n$ is known beforehand e.g., known interferers covariance matrices, or it is has been estimated a priori, e.g., environmental noise capture by a microphone array.

4. CONVEX OPTIMIZATION BASED DESIGN

A classical convex relaxation for problem (8) can be constructed starting from expressing the cost function using a linear sampling scheme, i.e.,

$$f(\mathbf{w}) := \mathbf{a}^T(\theta_0) \mathbf{\Phi}^T(\mathbf{w}) [\mathbf{\Phi}(\mathbf{w}) \mathbf{R}_n \mathbf{\Phi}^T(\mathbf{w})]^{-1} \mathbf{\Phi}(\mathbf{w}) \mathbf{a}(\theta_0), \tag{9}$$

where $\mathbf{\Phi}(\mathbf{w}) \in \{0,1\}^{K \times M}$ is a selection matrix whose entries are determined by the Boolean selection vector $\mathbf{w} \in \{0,1\}^{M \times 1}$. Here, $[\mathbf{w}]_i = 1$ indicates that the ith sensor element is selected. Considering that the noise covariance matrix can be expressed as

$$\mathbf{R}_n = \mathbf{S} + a\mathbf{I}, \quad \text{for some} \;\; \mathbf{S} \succ 0, \tag{10}$$

for $a \in \mathbb{R}_+$. Using the matrix inversion lemma on (9) leads to

$$\begin{aligned} f(\mathbf{w}) = \mathbf{a}^H(\theta_0) [\mathbf{S}^{-1} - \\ \mathbf{S}^{-1} (\mathbf{S}^{-1} + a^{-1} \text{diag}(\mathbf{w}))^{-1} \mathbf{S}^{-1}] \mathbf{a}(\theta_0). \end{aligned} \tag{11}$$

As only the second term of (11) depends on the selection variable $\mathbf{w}$, the problem (8) is equivalent to

$$\begin{array}{ll} \underset{\mathbf{w}}{\text{minimize}} & \mathbf{a}^H(\theta_0) \mathbf{S}^{-1} (\mathbf{S}^{-1} + a^{-1} \text{diag}(\mathbf{w}))^{-1} \mathbf{S}^{-1} \mathbf{a}(\theta_0) \\ \text{subject to} & \mathbf{w}^T \mathbf{b} \leq \beta, \; \|\mathbf{w}\|_0 = K, \; \mathbf{w} \in \{0,1\}^{M \times 1}, \end{array} \tag{12}$$

where the vector $\mathbf{b} \in \mathbb{R}^M$ contains the costs associated to the sensor elements, i.e., $\mathbf{b} = [b_1, \ldots, b_M]^T$. Unfortunately, problem (12) is not convex due to the ℓ_0-norm constraint and the Boolean nature of the selection variable. The convex relaxation for such a problem, expressed using its epigraph form, can be stated as [1]

$$\begin{array}{ll} \underset{\mathbf{w}, t}{\text{minimize}} & t \\ \text{subject to} & \mathbf{w}^T \mathbf{b} \leq \beta, \; \|\mathbf{w}\|_1 = K, \; \mathbf{w} \in [0,1]^{M \times 1} \\ & \begin{bmatrix} \mathbf{S}^{-1} + a^{-1} \text{diag}(\mathbf{w}) & \mathbf{S}^{-1} \mathbf{a}(\theta_0) \\ \mathbf{a}^H(\theta_0) \mathbf{S}^{-1} & t \end{bmatrix} \succeq 0, \end{array} \tag{13}$$

which involves the addition of the auxiliary variable $t \in \mathbb{R}$, the substitution of the ℓ_0-norm by the convex ℓ_1-norm, and the relaxation to the box $[0, 1]$ for the elements of $\mathbf{w}$. At this point, any off-the-shelf solver can be used to compute the solution of the convex problem (13). After the non-Boolean solution is obtained, a Boolean approximate solution that satisfies the constraints can be retrieved by thresholding methods or randomization techniques [20]. Note that the obtained solution will always satisfy the minimum variance distortionless property.

5. PROPOSED SUBMODULAR OPTIMIZATION

Instead of solving the semidefinite program (13), which only provides a non-Boolean approximate of the solution to (8), we now address the sensor selection problem using a submodular set-function and a greedy heuristic.

5.1. Submodularity

Formally, a set function $f : 2^{\mathcal{V}} \mapsto \mathbb{R}$, for a finite ground set $\mathcal{V}$, is called submodular, if for all sets $\mathcal{A} \subseteq \mathcal{B} \subseteq \mathcal{V}$ and all $i \notin \mathcal{B}$ it holds that $f(\mathcal{A} \cup \{i\}) - f(\mathcal{A}) \geq f(\mathcal{B} \cup \{i\}) - f(\mathcal{B})$. This property holds for set-functions that commonly appear in operation research and machine learning, and provides a notion of *diminishing returns*. Similar to convex functions, submodular set-functions have certain properties that allow for an efficient optimization [14]. It has been shown by Nemhauser et al. [12] that for the maximization of a non-decreasing submodular set-function, f, with $f(\emptyset) = 0$, a simple *greedy* heuristic finds a solution that is at least a constant fraction of $1 - 1/e \approx 63\%$ of the optimal value. In this context, a set-function f is considered non-decreasing, if and only if, $f(\mathcal{B}) \geq f(\mathcal{A})$ holds for all sets $\mathcal{A} \subseteq \mathcal{B} \subseteq \mathcal{V}$.

Inspired by this result, Algorithm 1 presents the greedy heuristic for maximizing a non-decreasing submodular set-function, subject to modular constraints [15]. The set $\mathcal{A}$ returned from Algorithm 1 obtains the near-optimality guarantee given in [12] when the cost

---

Algorithm 1: COST-BENEFIT GREEDY ALGORITHM.

Result: $\mathcal{A} : |\mathcal{A}| = K$
1   initialization $\mathcal{A} = \emptyset$, $k = 0$;
2   **while** $k < K$ **do**
3      $a^* = \arg\max\limits_{a \in \mathcal{V}} \dfrac{f(\mathcal{A} \cup \{a\}) - f(\mathcal{A})}{b_a}$;
4   **if** $B(\mathcal{A} \cup \{a^*\}) \leq \beta$ **then**
5         $\mathcal{A} = \mathcal{A} \cup \{a^*\}$;
6         $k = k + 1$;
7   **end**
8      $\mathcal{V} = \mathcal{V} \setminus a^*$;
9   **end**

---

of all elements are equal, i.e., $b_i = b_j$, $\forall i, j$, and no budget constraint is active. However, interesting instances of problem (8) arise when each sensor has a different cost, and hence a *knapsack* constraint is active. For these instances a modified version of Algorithm 1 can be employed. First, (i) the cost-benefit solution $\mathcal{A}_{\mathrm{cb}}$ using Algorithm 1 is computed. Then, (ii) a second solution is obtained by employing Algorithm 1 with the line 3 modified as: $a^* = \arg\max_{a \in \mathcal{V}} f(\mathcal{A} \cup \{a\}) - f(\mathcal{A})$. This is known as the uniform cost solution $\mathcal{A}_{\mathrm{uc}}$. Finally, the solution with the greatest set-function value among the two available solutions is selected as the final output. This modification provides the following near-optimality guarantee [16]

$$\max\{f(\mathcal{A}_{\mathrm{uc}}), f(\mathcal{A}_{\mathrm{dc}})\} \geq \frac{1}{2}(1 - e^{-1})f(\mathcal{A}_{\mathrm{opt}}), \quad (14)$$

where $\mathcal{A}_{\mathrm{opt}}$ is the subset that attains the optimal set-function value.

5.2. Submodular Relaxation

To make use of these existing results with respect to the greedy optimization of submodular functions, in the following we derive a submodular set-function surrogate for approximating the sensor selection problem (8). First, let us recall the non-negative property of the output signal-to-noise ratio, i.e., $f(\mathbf{w}) \geq 0$. It is possible to express this condition using a linear matrix inequality (LMI) in $\mathbf{w}$ (or equivalently in $\mathcal{A}$) similar to the one found in the convex relaxed problem [cf. (13)]. That is,

$$\mathbf{M}_{\mathcal{A}} = \begin{bmatrix} \mathbf{S}^{-1} + a^{-1}\mathbf{I}_{\mathcal{A}} & \mathbf{S}^{-1}\mathbf{a}(\theta_0) \\ \mathbf{a}^H(\theta_0)\mathbf{S}^{-1} & \mathbf{a}^H(\theta_0)\mathbf{S}^{-1}\mathbf{a}(\theta_0) \end{bmatrix} \succeq 0, \quad (15)$$

where instead of the additional variable t, the first term of (11) appears in the expression. In (15), $\mathbf{I}_{\mathcal{A}}$ is a diagonal matrix with ones in the entries (i, i), $i \in \mathcal{A}$ entries. It can be easily shown that the determinant of the above matrix can be expressed as

$$\det\left(\mathbf{M}_{\mathcal{A}}\right) = \det\left(\mathbf{S}^{-1} + a^{-1}\mathbf{I}_{\mathcal{A}}\right)f(\mathcal{A}) = g(\mathcal{A})f(\mathcal{A}), \quad (16)$$

where the fact that $\mathbf{w}$ and $\mathcal{A}$ can be used interchangeably has been used for clarity. From (16) we can observe that the determinant of $\mathbf{M}_{\mathcal{A}}$ consists of the product of two terms. The second term is the output signal-to-noise ratio $f(\mathcal{A})$, while the first is a determinant that is inversely proportional to the loss in signal-to-noise ratio in (11). Although $g(\mathcal{A})$ does not depend on the array steering vector $\mathbf{a}(\theta_0)$, we can consider the following optimization problem as an alternative for approximating the solution of (8):

$$\underset{\mathcal{A}}{\text{maximize}} \quad \ln\det\left(\mathbf{M}_{\mathcal{A}}\right), \quad \text{s.t. } B(\mathcal{A}) \leq \beta, \; |\mathcal{A}| = K. \quad (17)$$

The cost set-function in problem (17) requires to satisfy the conditions of monotonicity and submodularity in order to be able to
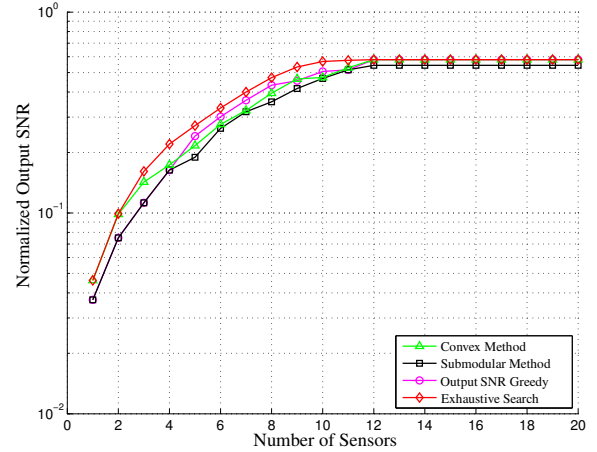
Fig. 1: Comparison of output SNR for different sensor selection schemes for $M = 20$ sensors. The budget constraint is set to 80% of the total sensor cost.

claim the near-optimality guarantees discussed before. In the following, we present two propositions that are required to provide the guarantees for near-optimality when the problem (17) is solved greedily.

Proposition 1. (Monotonicity of Cost Set-Function) *The cost function from* (17) *is a monotone non-decreasing set-function.*

   *Proof.* See Appendix A

Proposition 2. (Submodularity of Cost Set-Function) *The cost set-function from* (17) *is a submodular set-function.*

    Proof. See Appendix B

    In the following section, we provide numerical results to demonstrate the developed theory.

6. NUMERICAL EXPERIMENTS

This section presents numerical results for the MVDR sensor selection performance of three different schemes: (i) semidefinite program on the *convex* relaxation (13), (ii) greedy solution to the *submodular* relaxation (17), and (iii) the greedy solution to problem (8) (where the fact that the output SNR is not submodular is ignored). In all greedy methods, we use the modification of Algorithm 1 which obtains the sets $\{\mathcal{A}_{\mathrm{uc}}, \mathcal{A}_{\mathrm{dc}}\}$ and selects the one that obtains the greatest output SNR. The method (iii) is further on referred to as *Output SNR Greedy*. For illustration purposes, first we consider a half wavelength linear array consisting of $M = 20$ elements. Furthermore, to each array element a random non-negative cost b_i is assigned. In this example, a MVDR beamformer is desired for an angular direction $\theta_0 = -20^o$. The noise covariance matrix is assumed to consist of an interference at $\theta_i = -10^o$ and white Gaussian noise at -10dB. In Fig. 1, a comparison between the methods with respect to the exhaustive search is shown. In this example the budget constraint β is selected to be equal to $0.8B(\mathcal{V})$, i.e., 80% of the total sensor cost. The output SNR is normalized with respect to the maximum output SNR, i.e., output SNR when all the elements of the array are considered. From Fig. 1, it can be observed that the three schemes perform close to each other, and that they are not far from the exhaustive search. Notice that even without any performance guarantees the Output SNR Greedy method performs close to the exhaustive search solution. However, we should be cautious when using it as it could
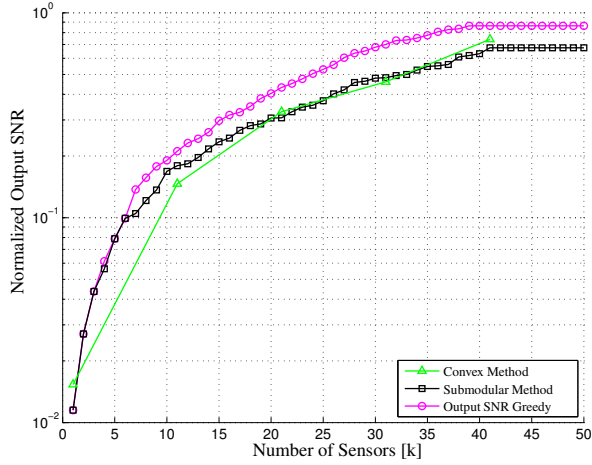
Fig. 2: Output signal-to-noise ratio comparison of the three methods for different subset sizes. The budget constraint is set to 80% of the total cost.

get stuck in some cases. For the second scenario, we consider the same setup as before, but increase the number of available sensors to $M = 50$. In this instance it is not possible to perform an exhaustive search, and therefore only a comparison among the three methods is possible. This comparison, in terms of the normalized output SNR, is shown in Fig. 2. As expected, the performance of the submodular method and the convex relaxation are close to each other. This is due to the fact that the cost functions that both methods try to optimize are related through the Schur complement. The best output SNR is obtained by the Output SNR Greedy. However, as mentioned before one should be careful when using its result due to the fact that it can perform arbitrary bad for a given problem. Despite this, when the solution of the submodular problem is obtained it could be straightforward to compare it with the one from the Output SNR Greedy method and select the best. Both solutions are computed in linear time and can be scaled easily to very large instances of the sensor selection problem. Finally, an illustration of the beam pattern for the solution based on $K = 21$ is shown in Fig. 3. It it clear from this figure that all the beamformers have a narrow main lobe centered at -20^0, and a null at -10°, however they present different sidelobe characteristics.

7. CONCLUSIONS

In this paper, we have linked a submodular set-function to the common convex relaxation approach for sensor selection. By means of the Schur complement, the log-determinant of a positive semidefinite matrix can be used to optimize the output SNR of an MVDR beamformer. It has been shown that using a modified greedy heuristic, which has linear complexity, a near-optimal solution can be obtained. Even though the theoretical bounds for submodular optimization are not tight, the experimental results show that the greedy heuristic provides a performance that is comparable to the one based on convex relaxation at a significantly reduced complexity.

A. PROOF OF PROPOSITION 1

Let us define the following:

$$\boldsymbol{T} = \begin{bmatrix} \mathbf{S}^{-1} & \mathbf{S}^{-1}\boldsymbol{a}(\theta_0) \\ \mathbf{a}^H(\theta_0)\mathbf{S}^{-1} & \mathbf{a}(\theta_0)\mathbf{S}^{-1}\mathbf{a}(\theta_0) \end{bmatrix}, \ \mathbf{L}_{\mathcal{A}} = \begin{bmatrix} a^{-1}\mathbf{I}_{\mathcal{A}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}.$$

We can express the cost set-function from (17) as $f(\mathcal{A}) = \ln \det(\mathbf{T} + \mathbf{L}_{\mathcal{A}})$, where we recall that $\mathbf{M}_{\mathcal{A}} := \mathbf{T} + \mathbf{L}_{\mathcal{A}}$. To prove that the set-
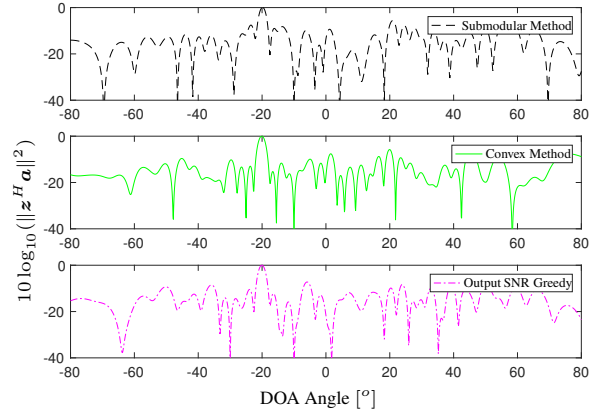
Fig. 3: Beam pattern comparison for the different sensor selection strategies when $K = 21$ sensors are selected out of $M = 50$ available elements.

function is monotone, we need to show that

$$f(\mathcal{A} \cup \{i\}) - f(\mathcal{A}) = \ln \det(\mathbf{M}_{\mathcal{A}} + \mathbf{L}_i) - \ln \det(\mathbf{M}_{\mathcal{A}}) \geq 0.$$

In other words, we should prove that $\det(\mathbf{M}_{\mathcal{A}} + \mathbf{L}_i) \geq \det(\mathbf{M}_{\mathcal{A}})$. This implies that $\mathbf{M}_{\mathcal{A}} + \mathbf{L}_i \succeq \mathbf{M}_{\mathcal{A}}$, which is always true as we choose $a \geq 0$.

B. PROOF OF PROPOSITION 2

Let us consider the previous definitions for $\mathbf{T}$ and $\mathbf{L}_{\mathcal{A}}$. Then, we can evaluate the cost set-function in the following sets:

$$f(\mathcal{A}) = \ln \det(\mathbf{M}_{\mathcal{A}}), \ \ f(\mathcal{A} \cup \{i\}) = \ln \det(\mathbf{M}_{\mathcal{A}} + \mathbf{L}_i),$$
$$f(\mathcal{A} \cup \{i,j\}) = \ln \det(\mathbf{M}_{\mathcal{A}} + \mathbf{L}_i + \mathbf{L}_j).$$

We need to prove that the following expression is always positive

$$f(\mathcal{A} \cup \{i\}) - f(\mathcal{A}) - f(\mathcal{A} \cup \{i,j\}) + f(\mathcal{A} \cup \{j\}) =$$
$$\ln \frac{\det(\mathbf{M}_{\mathcal{A}} + \mathbf{L}_i) \det(\mathbf{M}_{\mathcal{A}} + \mathbf{L}_j)}{\det(\mathbf{M}_{\mathcal{A}}) \det(\mathbf{M}_{\mathcal{A}} + \mathbf{L}_i + \mathbf{L}_j)} \geq 0.$$

The above inequality is equivalent to

$$\det(\mathbf{M}_{\mathcal{A}} + \mathbf{L}_i) \det(\mathbf{M}_{\mathcal{A}} + \mathbf{L}_j) \geq \det(\mathbf{M}_{\mathcal{A}}) \det(\mathbf{M}_{\mathcal{A}} + \mathbf{L}_i + \mathbf{L}_j).$$

Noticing that $\mathbf{L}_i = a^{-1}\mathbf{e}_i\mathbf{e}_i^T$ is a dyadic product, and that $\mathbf{M}_{\mathcal{A}}$ and $\mathbf{M}_{\mathcal{A}} + \mathbf{L}_j$ are invertible by definition, we can apply the matrix determinant lemma and rewrite the previous expression as

$$\frac{\det(\mathbf{M}_{\mathcal{A}}) \det(\mathbf{M}_{\mathcal{A}} + \mathbf{L}_j)(1 + a^{-1}\mathbf{e}_i^T\mathbf{M}_{\mathcal{A}}^{-1}\mathbf{e}_i)}{\det(\mathbf{M}_{\mathcal{A}}) \det(\mathbf{M}_{\mathcal{A}} + \mathbf{L}_j)(1 + a^{-1}\mathbf{e}_i^T(\mathbf{M}_{\mathcal{A}} + \mathbf{L}_j)^{-1}\mathbf{e}_i)} \geq 1,$$

leading to

$$\frac{1 + a^{-1}\mathbf{e}_i^T\mathbf{M}_{\mathcal{A}}^{-1}\mathbf{e}_i}{1 + a^{-1}\mathbf{e}_i(\mathbf{M}_{\mathcal{A}} + \mathbf{M}_j)^{-1}\mathbf{e}_i} \geq 1.$$

The above inequality is equivalent to $\mathbf{e}_i^T\mathbf{M}_{\mathcal{A}}^{-1}\mathbf{e}_i \geq \mathbf{e}_i(\mathbf{M}_{\mathcal{A}} + \mathbf{M}_j)^{-1}\mathbf{e}_i$. This can be proven using the following property of positive definite matrices. As $\mathbf{M}_{\mathcal{A}} + \mathbf{M}_i \succeq \mathbf{M}_{\mathcal{A}}$, is always true, we have that $\mathbf{M}_{\mathcal{A}}^{-1} \succeq (\mathbf{M}_{\mathcal{A}} + \mathbf{L}_i)^{-1}$. Therefore, $\mathbf{e}_i^T(\mathbf{M}_{\mathcal{A}}^{-1} - (\mathbf{M}_{\mathcal{A}}\mathbf{M}_i)^{-1})\mathbf{e}_i \geq 0$, and the result is proven.

C. REFERENCES

[1] Chepuri, S.P and Leus, G., "Sparse Sensing for Statistical Inference", Foundations and Trends in Signal Processing: Vol. 9 (2016): No. 34, pp 233-368.

[2] Rowaihy, H., et al. "A survey of sensor selection schemes in wireless sensor networks." Defense and Security Symposium. International Society for Optics and Photonics, 2007.

[3] Joshi, S., and Boyd, S., "Sensor selection via convex optimization." IEEE Transactions on Signal Processing 57.2 (2009): 451-462.

[4] Liu, S., et al. "Sensor selection for estimation with correlated measurement noise." IEEE Transactions on Signal Processing 64.13 (2016): 3509-3522.

[5] Chepuri, S. P., and Leus G., "Sparsity-promoting sensor selection for non-linear measurement models." IEEE Transactions on Signal Processing 63.3 (2015): 684-698.

[6] Vandenberghe, L., and Boyd S., "Semidefinite programming." SIAM review 38.1 (1996): 49-95.

[7] Guestrin, C., Krause A., and Singh, A. P., "Near-optimal sensor placements in gaussian processes." Proceedings of the 22nd international conference on Machine learning. ACM, 2005.

[8] Shamaiah, M., Banerjee S., and Vikalo H., "Greedy sensor selection: Leveraging submodularity." Decision and Control (CDC), 2010 49th IEEE Conference on. IEEE, 2010.

[9] Ranieri, J., Chebira A., and Vetterli, M., "Near-optimal sensor placement for linear inverse problems." IEEE Transactions on signal processing 62.5 (2014): 1135-1146.

[10] Krause, A., and Guestrin, C., "Near-optimal observation selection using submodular functions." AAAI. Vol. 7. 2007.

[11] Chamon, L.F., and Ribeiro, A., Near-optimality of greedy set selection in the sampling of graph signals, in 2016 IEEE Global Conf. Signal and Inform. Process. Washington, DC: IEEE, 7-9 Dec. 2016.

[12] Nemhauser, G. L., Wolsey, L. A., and Fisher, M. L., "An analysis of approximations for maximizing submodular set functions." Mathematical Programming 14.1 (1978): 265-294.

[13] Minoux, M., "Accelerated greedy algorithms for maximizing submodular set functions." Optimization Techniques. Springer Berlin Heidelberg, 1978. 234-243.

[14] Fujishige, S., "Submodular functions and optimization". Vol. 58. Elsevier, 2005. APA

[15] Khuller, S., Moss A., and Naor, J.S., "The budgeted maximum coverage problem." Information Processing Letters 70.1 (1999): 39-45.

[16] Leskovec, J., et al. "Cost-effective outbreak detection in networks." Proceedings of the 13th ACM SIGKDD international conference on Knowledge discovery and data mining. ACM, 2007.

[17] Sviridenko, M., "A note on maximizing a submodular set function subject to a knapsack constraint." Operations Research Letters 32.1 (2004): 41-43.

[18] Emura, S., "ℓ_1-constrained MVDR-based selection of nonidentical directivities in microphone array." Acoustics, Speech and Signal Processing (ICASSP), 2015 IEEE International Conference on. IEEE, 2015.

[19] Nguyen, Nam, Ingo Wiegand, and Douglas L. Jones. "Sparse beamforming for active underwater electrolocation." Acoustics, Speech and Signal Processing, 2009. ICASSP 2009. IEEE International Conference on. IEEE, 2009.

[20] Luo, Z.Q., et al. "Semidefinite relaxation of quadratic optimization problems." IEEE Signal Processing Magazine 27.3 (2010): 20-34.