Response to Reviewer Comments

Mario Coutino, Sundeep P. Chepuri, and Geert Leus

April 8, 2018

First of all, we would like to thank the reviewers for their thorough and critical review of our work, and in particular for providing useful suggestions aimed to increase the quality of the work and improve its clarity. In addition, we would like to thank the AE for handling our manuscript in such a timely and effective manner. Major changes in the manuscript have been highlighted. Following the suggestions of the reviewers and the comments of the AE we have made the following changes to the manuscript:

- Theorem 2 and its proof have been modified in order to accommodate the suggestions of Reviewer 2. That is:
  - the norm of $\boldsymbol{\theta}$ has been included in the derivation;
  - the result of Theorem 2 has been limited to that of equation (66), and the result (67) has been stated as a corollary;
  - the derivation of the lower bound has been clarified.
- the detailed version of Figure 3 has been added to the manuscript as per the suggestion from Reviewer 2.

Comments from Reviewer 1

Comment 1*: Thank you for submitting the revised version, I enjoyed reading this paper and believe it is of interest to the community. While I think adding the plots you have generated in response to my comment Q4 to the paper may be useful for some readers I leave it up to you to decide if you wish to do so.*

Response: We are glad to know that the reviewer enjoyed reading our manuscript. We are thankful for your comments in the previous round, as they have provided an interesting perspective of our work as well as helped us improve the quality of the manuscript. Although we agree with the fact that plots provide some insights in the complexity of the method, we believe that the derived theoretical description of the complexity of the method is a stronger result, and provides a more insightful contribution detached of any implementation tricks or code optimization. Therefore, we decided to not include the mentioned plots in the final version of the manuscript.

Comments from Reviewer 2

The authors have addressed most of my issues. In particular, I appreciate their improving the text after (30) to avoid another reader having the same misunderstanding as me about $s(A)$. However, there are still important issues remaining in the current version.

Response: We want to thank the reviewer for the comments provided in the previous round. Thanks to the suggestions of the reviewer we were able to clarify further our intentions within the manuscript, and provide better insights in the ideas behind our work. We really appreciate the thorough proof reading that the reviewer performed over the different versions of our manuscript, as it has improved the overall readability of the paper.

Comment 1*: Although the new statement of Theorem 2 makes it considerably clearer, the proof provided does not seem to match the result. First, the proof assume $\|\theta\| = 1$. Also, (22) does not depend on a or β, whereas the result in the proof does [see (54)]. If the result given only holds for the limiting case, then the theorem cannot claim that it holds for all $\beta \in (0,1)$. Especially since the paper will suggest the use of $\beta = 0.5$. Moreover, the proof only derives the upper bound in (52). Still, the lower bound is necessary to obtain the epsilon-Submodularity relation. It is critical that the statement and proof of Theorem 2 be reviewed so they match.*

Response: Thanks for pointing out these inconsistencies. Taking into account the reviewer comments we have made the following modifications to correct our oversights: (*i*) we have modified the proof to incorporate the norm of $\boldsymbol{\theta}$ in the derivation; (*ii*) the result in Theorem 2 has been limited to the one of (66), and the current result has been provided as a collorary; (*iii*) the lower bound, which was omitted due to an oversight is added. The modified theorem and corollary included in the manuscript are shown here:

[...]

Theorem 2. *Let $\boldsymbol{\Sigma}$ be a non-diagonal covariance matrix, with minimum eigenvalue $\lambda_{\min}\{\boldsymbol{\Sigma}\} \neq 0$, maximum eigenvalue $\lambda_{\max}\{\boldsymbol{\Sigma}\}$, condition number $\kappa := \lambda_{\max}\{\boldsymbol{\Sigma}\}/\lambda_{\min}\{\boldsymbol{\Sigma}\}$, that admits a decomposition $\boldsymbol{\Sigma} = a\boldsymbol{I}+\boldsymbol{S}$ where a is chosen as $a = \beta\lambda_{\min}$ with $\beta \in (0,1)$ to guarantee the positive definiteness of $\boldsymbol{S}$. Then the signal-to-noise ratio set function $s(\mathcal{A})$ is ϵ-approximately submodular with*

$$\epsilon \leq 4C_1\left(\frac{a}{(1-\beta)^2\lambda_{\min}^2\{\boldsymbol{\Sigma}\}} + \frac{\nu\kappa^2}{(1-\beta)^2}\right), \tag{22}$$

where $C_1 = \|\boldsymbol{\theta}\|_2^2$, with $\boldsymbol{\theta}$ being the mean vector, and $\nu = \min\{a^{-1}, (1-\beta)^{-1}\lambda_{\min}^{-1}\{\boldsymbol{\Sigma}\}\}$.

Proof. See Appendix A □

Corollary 1. *For the limiting case, $a \to 0$ or equivalently $\beta \to 0$, Theorem 2 reduces to*

$$\epsilon \leq 4C_1\kappa^2\lambda_{\min}^{-1}\{\boldsymbol{\Sigma}\}. \tag{23}$$

Proof. Follows from Theorem 2 (See Appendix A). □

[...]

The modified proof is included in the manuscript as:

[...]

Proof. First, consider the SNR set function that is defined as [cf. (20)]

$$s(\mathcal{A}) = \boldsymbol{\theta}_{\mathcal{A}}^T\boldsymbol{\Sigma}_{\mathcal{A}}^{-1}\boldsymbol{\theta}_{\mathcal{A}}. \tag{47}$$

Combining the above expression and the decomposition $\boldsymbol{\Sigma} = a\boldsymbol{I} + \boldsymbol{S}$, the signal-to-noise ratio can be rewritten as [17]

$$\begin{aligned} s(\mathcal{A}) &= \boldsymbol{\theta}_{\mathcal{A}}^T\boldsymbol{\Sigma}_{\mathcal{A}}^{-1}\boldsymbol{\theta}_{\mathcal{A}} &(48)\\ &= \boldsymbol{\theta}^T\mathbf{S}^{-1}\boldsymbol{\theta} - \boldsymbol{\theta}^T\mathbf{S}^{-1}[\mathbf{S}^{-1} + a^{-1}\mathrm{diag}(\mathbf{1}_{\mathcal{A}})]^{-1}\mathbf{S}^{-1}\boldsymbol{\theta}, &(49)\\ &= \boldsymbol{\theta}^T\boldsymbol{S}^{-1}\boldsymbol{\theta} + h(\mathcal{A}), &(50) \end{aligned}$$

where the non-zero entries of the vector $\mathbf{1}_{\mathcal{A}}$ are given by the set $\mathcal{A}$ and $h(\mathcal{A}) := -\boldsymbol{\theta}^T\boldsymbol{S}^{-1}\Big[\boldsymbol{S}^{-1} + a^{-1}\mathrm{diag}(\mathbf{1}_{\mathcal{A}})\Big]^{-1}\boldsymbol{S}^{-1}\boldsymbol{\theta}$.

For the sake of simplicity, we will rewrite the SNR as follows

$$s(\mathcal{A}) := C_1 \tilde{s}(\mathcal{A}), \tag{51}$$

where we have defined $C_1 = \|\boldsymbol{\theta}\|_2^2$, and $\tilde{s}(\mathcal{A})$ is the SNR set function with $\boldsymbol{\theta}$ being substituted by $\tilde{\boldsymbol{\theta}} = \boldsymbol{\theta}/\|\boldsymbol{\theta}\|_2$, i.e., the SNR set function is computed only considering the direction of the vector $\boldsymbol{\theta}$.

Now, let us assume that there exists a $\epsilon' \in \mathbb{R}_+$ such that

$$-\epsilon' \le \tilde{s}(\mathcal{A}) - \hat{s}(\mathcal{A}) \le \epsilon', \tag{52}$$

for any $\mathcal{A} \subseteq \mathcal{V}$ and some modular set function $\hat{s}(\mathcal{A})$. Using (52), we can obtain the following expression

$$\delta(\mathcal{A} \cup \{i\}) - \delta(\mathcal{A}) - \delta(\mathcal{A} \cup \{i, j\}) + \delta(\mathcal{A} \cup \{j\}) \ge -4\epsilon', \tag{53}$$

where we have defined $\delta(\mathcal{A}) = \tilde{s}(\mathcal{A}) - \hat{s}(\mathcal{A})$. Due to the modularity of $\hat{s}(\mathcal{A})$, i.e.,

$$\hat{s}(\mathcal{A} \cup \{i\}) - \hat{s}(\mathcal{A}) - \hat{s}(\mathcal{A} \cup \{i, j\}) + \hat{s}(\mathcal{A} \cup \{j\}) = 0, \tag{54}$$

and using expressions (53) and (54), we can show that

$$\tilde{s}(\mathcal{A} \cup \{i\}) - \tilde{s}(\mathcal{A}) - \tilde{s}(\mathcal{A} \cup \{i, j\}) + \tilde{s}(\mathcal{A} \cup \{j\}) \ge -4\epsilon'. \tag{55}$$

From this it is clear that the set function $\tilde{s}(\mathcal{A})$ is ϵ-submodular with an $\epsilon \le 4\epsilon'$ (or equivalently, $s(\mathcal{A})$ is ϵ-submodular with an $\epsilon \le 4\epsilon' C_1$). Therefore, for completing the proof we require to establish the bound in (52) for the specific ϵ' given in (22). In the following, we devote ourselves to this task.

For this proof, we select the following auxiliary set function:

$$\hat{s}(\mathcal{A}) = \tilde{\boldsymbol{\theta}}^T \boldsymbol{S}^{-1} \tilde{\boldsymbol{\theta}} + \hat{h}(\mathcal{A}), \tag{56}$$

where $\hat{h}(\mathcal{A})$ is chosen to be a *modular* [cf. (54)] set function and it is given by

$$\hat{h}(\mathcal{A}) = -\tilde{\boldsymbol{\theta}}^T \boldsymbol{S}^{-1} \left(a^{-1}\boldsymbol{I} + a^{-1}\mathrm{diag}(\mathbf{1}_{\mathcal{A}})\right)^{-1} \boldsymbol{S}^{-1} \tilde{\boldsymbol{\theta}}. \tag{57}$$

In (57), a scaled identity matrix has been introduced instead of the inverse of $\boldsymbol{S}$ [cf. (50)] to construct a modular set function. Here, it should be noticed that other set functions besides (57) could have been used for finding an upper bound on the ϵ constant. Depending on this choice, different bounds might be obtained.

To prove (52), we equivalently will establish the following inequalities

$$-\epsilon' \le \tilde{h}(\mathcal{A}) - \hat{h}(\mathcal{A}) \le \epsilon', \tag{58}$$

where we have defined $\tilde{h}(\mathcal{A}) = C_1^{-1} h(\mathcal{A})$. To obtain these inequalities, we can bound the difference of the positive definite (PD) matrices that are part of the quadratic forms in the set functions in (58). That is, we need to show that

$$\begin{aligned} -\epsilon' \boldsymbol{I} \quad &\preceq \boldsymbol{S}^{-1}\left[a^{-1}\boldsymbol{I} + a^{-1}\mathrm{diag}(\mathbf{1}_{\mathcal{A}})\right]^{-1}\boldsymbol{S}^{-1} \\ &\qquad - \boldsymbol{S}^{-1}\left[\boldsymbol{S}^{-1} + a^{-1}\mathrm{diag}(\mathbf{1}_{\mathcal{A}})\right]^{-1}\boldsymbol{S}^{-1} \preceq \epsilon' \boldsymbol{I}. \end{aligned} \tag{59}$$

Considering $\mathrm{diag}(\mathbf{1}_{\mathcal{A}}) = \boldsymbol{\Phi}_{\mathcal{A}}^T \boldsymbol{\Phi}_{\mathcal{A}}$, we can apply the matrix inversion lemma to expand the difference of the matrices in between brackets above as

$$\boldsymbol{\Delta} := a\left(\boldsymbol{I} - \frac{1}{2}\boldsymbol{\Phi}_{\mathcal{A}}^T \boldsymbol{\Phi}_{\mathcal{A}}\right) - \boldsymbol{S} + \boldsymbol{S}\boldsymbol{\Phi}_{\mathcal{A}}^T\left(a\boldsymbol{I} + \boldsymbol{\Phi}_{\mathcal{A}}\boldsymbol{S}\boldsymbol{\Phi}_{\mathcal{A}}^T\right)^{-1}\boldsymbol{\Phi}_{\mathcal{A}}\boldsymbol{S}. \tag{60}$$

Hence (59) becomes,

$$-\epsilon' \boldsymbol{I} \preceq \boldsymbol{S}^{-1} \boldsymbol{\Delta} \boldsymbol{S}^{-1} \preceq \epsilon' \boldsymbol{I}. \tag{61}$$

As all terms in (60) are PD matrices, we upper bound the matrix in (60) by removing the negative terms in (60), and lower bound it by removing all the positive terms. That is,

$$-\left(\frac{a}{2}\boldsymbol{I} + \boldsymbol{S}\right) \preceq \boldsymbol{\Delta} \preceq a\boldsymbol{I} + \boldsymbol{S}\boldsymbol{\Phi}_{\mathcal{A}}^T\left(a\boldsymbol{I} + \boldsymbol{\Phi}_{\mathcal{A}}\boldsymbol{S}\boldsymbol{\Phi}_{\mathcal{A}}^T\right)^{-1}\boldsymbol{\Phi}_{\mathcal{A}}\boldsymbol{S}. \tag{62}$$

From the definition of $\boldsymbol{S}$ and a, we can notice that a possible lower bound for the expression above is given by

$$-\lambda_{\max}\{\boldsymbol{\Sigma}\}\boldsymbol{I} \preceq \boldsymbol{\Delta}. \tag{63}$$

For the upper bound, we notice that by the maximum singular value of the second matrix, the following inequality holds

$$\begin{aligned}
\boldsymbol{\Delta} &\preceq a\boldsymbol{I} + \sigma_{\max}\left\{\boldsymbol{S}\boldsymbol{\Phi}_{\mathcal{A}}^T\left(a\boldsymbol{I} + \boldsymbol{\Phi}_{\mathcal{A}}\boldsymbol{S}\boldsymbol{\Phi}_{\mathcal{A}}^T\right)^{-1}\boldsymbol{\Phi}_{\mathcal{A}}\boldsymbol{S}\right\}\boldsymbol{I} \\
&\preceq a\boldsymbol{I} + \lambda_{\min}^{-1}\left\{a\boldsymbol{I} + \boldsymbol{\Phi}_{\mathcal{A}}\boldsymbol{S}\boldsymbol{\Phi}_{\mathcal{A}}^T\right\}\sigma_{\max}^2\left\{\boldsymbol{\Phi}_{\mathcal{A}}\boldsymbol{S}\right\}\boldsymbol{I} \\
&\preceq a\boldsymbol{I} + \min\{a^{-1}, \frac{1}{(1-\beta)\lambda_{\min}\{\boldsymbol{\Sigma}\}}\}\lambda_{\max}^2\{\boldsymbol{S}\}\boldsymbol{I}, \\
&\preceq (a + \nu\lambda_{\max}^2\{\boldsymbol{S}\})\boldsymbol{I},
\end{aligned}$$

where the submultiplicativity and subadditivity of singular values is used in the second and third inequality, respectively, and we have defined $\nu = \min\{a^{-1}, (1-\beta)^{-1}\lambda_{\min}^{-1}\{\boldsymbol{\Sigma}\}\}$. Here, the minimum of these two quantities is considered to simplify the expression of the limiting case, i.e., $a \to 0$.

Considering that

$$\lambda_{\max}\{\boldsymbol{\Sigma}\} \leq a + \nu\lambda_{\max}^2\{\boldsymbol{S}\}, \tag{64}$$

we can bound the matrix in (60) by both sides as follows

$$-(a\boldsymbol{I} + \nu\lambda_{\max}^2\{\boldsymbol{S}\}\boldsymbol{I}) \preceq \boldsymbol{\Delta} \preceq a\boldsymbol{I} + \nu\lambda_{\max}^2\{\boldsymbol{S}\}\boldsymbol{I}. \tag{65}$$

Hence, we solely continue deriving the upper bound for the expression above as the obtained ϵ' will hold for both lower and upper bound.

Now, considering that the eigenvalues of $\boldsymbol{\Sigma}$ are larger than those of $\boldsymbol{S}$ by definition, $\boldsymbol{S}^{-1} \preceq \lambda_{\min}^{-1}\{\boldsymbol{S}\}\boldsymbol{I}$, and recalling that $a = \beta\lambda_{\min}\{\boldsymbol{\Sigma}\}$ we obtain

$$\boldsymbol{S}^{-1}\boldsymbol{\Delta}\boldsymbol{S}^{-1} \preceq \frac{a}{(1-\beta)^2\lambda_{\min}^2\{\boldsymbol{\Sigma}\}}\boldsymbol{I} + \frac{\nu\kappa^2}{(1-\beta)^2}\boldsymbol{I} = \epsilon'\boldsymbol{I}, \tag{66}$$

where κ is the condition number of the matrix $\boldsymbol{\Sigma}$, proving the result of the theorem.

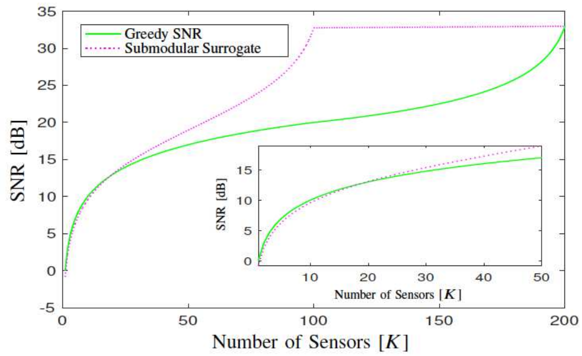
For the limiting case, $a \to 0$ or equivalently $\beta \to 0$, we obtain:

$$\epsilon' = \frac{\kappa^2}{\lambda_{\min}\{\boldsymbol{\Sigma}\}}, \tag{67}$$

which shows a relation with the typical experiment design metrics, i.e., maximization of the minimum eigenvalue and log determinant (which promotes a good matrix condition). □

[...]

Comment 2: *I do think that the detailed figure should be included in Fig. 3. What is more, since $K << M$ is the situation of interest in practice (and in this work, for that matter, given the "sparse" in the title), there seems to be no advantage in using the submodular surrogate over greedy SNR, at least in the example presented. Would there be a better example where the submodular surrogate clearly outperforms greedy SNR in the cases of interest? If not, what is the advantage of this surrogate objective?*



Response: Thanks for pointing this out. Following your suggestion in the current version of the manuscript we have added the detailed figure related to Fig. 3. While we do agree that for $K \leq 0.25M$ the submodular surrogate does not show any clear benefit over greedy SNR, we stress that in this example for the range $K \geq 0.25M$, the surrogate demonstrates its benefits. Moreover, it is true that although we focus on sparse sensing, in many instances reducing the number of sensors too much might not lead to systems meeting minimum requirements, e.g., operational SNR. In this regard, looking at Fig. 3, for a nominal system with a requirement of $SNR \geq 20dB$ only using Greedy SNR will lead to a solution containing a high number of sensors, while using the proposed surrogate leads to a solution that involves less sensors. We want to emphasize that even though throughout the manuscript we focus on the cardinality constraint formulation (P-CC) [cf. (11)], the proposed methods are also suitable for performance constraint (P-DC) formulations [cf. (12)]. Therefore, in cases that a fixed SNR is required, using the proposed surrogate provides clear advantages over Greedy SNR. The following comment about this observation has been added in the manuscript:

> [...] We want to emphasize that even though throughout the manuscript we focus on the cardinality constraint formulation (P-CC) [cf. (11)], the proposed methods are also suitable for performance constraint (P-DC) formulations [cf. (12)]. In many instances reducing the number of sensors too much might not lead to systems meeting minimum requirements, e.g., operational SNR. In this regard, observing Fig. 3, for a nominal system with a requirement of $SNR \geq 20dB$ (i.e., the performance constraint in (11)) only using Greedy SNR will lead to a solution containing a high number of sensors, while using the proposed surrogate leads to a solution that involves less sensors. Therefore, despite that for low values of K, there is no notable difference between both algorithms, in cases that a fixed SNR is required, using the proposed surrogate provides a clear advantage with respect to Greedy SNR. [...]

Comment 3: *I understand the value of Theorem 2 in explaining when greedy SNR minimization may be a good approach. However, as a guarantee, it is only valid if the bound on epsilon is small compared to $f^\star/K$. Otherwise, it gives no performance guarantee. It would be good to see numerical values of ϵ (perhaps as a function of κ and other parameters) to make sure that the result is informative.*

Response: The reviewer is right in the fact that a small ϵ implies that the set function is in an *overall* sense near submodular (with respect to all possible subsets). However, only one pair of odd behaving subsets is required to obtain an ϵ constant large compared to $f^\star/K$. In addition, any intuition that can be drawn from observing this constant for small problems (where it is possible to compute it), might not be extrapolated properly to large problem instances. It would not be proper to make any claim on the behaviour of ϵ by testing simple computationally feasible examples. As a result, the provided bound on ϵ is aimed to provide a description of the behaviour of the constant, rather than to provide a definite statement. In this answer, we include a figure that shows the average behaviour of ϵ for a series of Monte Carlo simulations for PD matrices with different condition numbers. The reported value, r_κ, is defined as the ratio between ϵ and

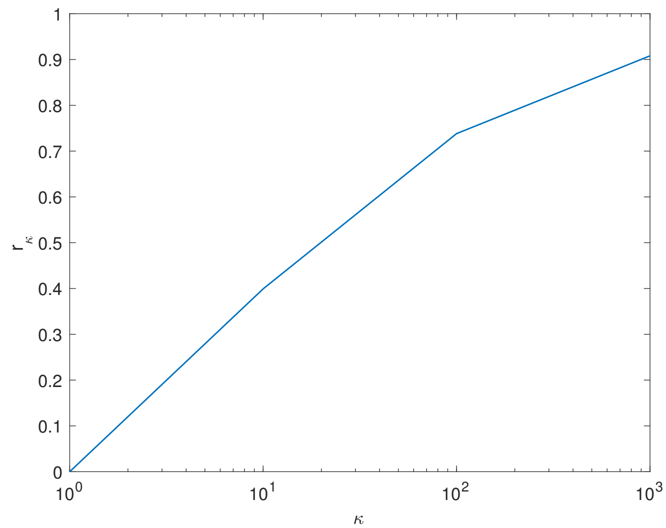


Figure 1: Average behaviour of ϵ constant as a function of condition number, κ, of $\mathbf{\Sigma}$

$f(\mathcal{V})$. From this plot, it can be seen that the ϵ constant is indeed monotone increasing with respect to the condition number of the matrices, as concluded from the bound. However, beyond the monotonicity of the ϵ constant with respect to κ, no further conclusions can be drawn from the plot. Therefore, we decided not to include these kind of plots in the manuscript. Finally, we would like to emphasize that in our work, our main intention for deriving this bound was to clarify for which instances greedy might perform well.

Comment 4: *"number of function evaluations required by the method scales linearly in the number of selected sensors": this should read "polynomially". K is a fraction of M, say rM. Therefore, greedy requires $O(rM^2)$ cost function evaluations.*

Response: In this respect, we want to emphasize that K and M are independent of each other. That is, when the number of selected sensors is increased the method scales linearly in K, as M is kept fixed. Similarly, for a fixed number of desired sensors K, for different number of available sensors the method scales linearly in M. Therefore, as each of this quantities are independent of each other, our original statement is valid.

Comments from Reviewer 3

Comment 1: *Although the authors did not implement all of my recommended revisions, I am happy with the extension to the case of unequal means and unequal variances, hence I recommend the paper for publication.*

Response: Thanks for the valuable suggestions in the previous round. Through them we were able to generalize our work and increase considerably the quality of the manuscript.