Signal Processing 200 (2022) 108647

Contents lists available at ScienceDirect

Signal Processing

journal homepage: www.elsevier.com/locate/sigpro

Collaborative one-shot beamforming under localization errors: A discrete optimization approach[☆]

Yagiz Savas[a,*], Erfaun Noorani[b], Alec Koppel[c], John Baras[b], Ufuk Topcu[a], Brian M. Sadler[c]

[a] Department of Aerospace Engineering, University of Texas at Austin, TX, USA
[b] Department of Electrical and Computer Engineering, University of Maryland, MD, USA
[c] U.S. Army Research Laboratory, MD, USA

ARTICLE INFO

Article history:
Received 23 July 2021
Revised 14 May 2022
Accepted 5 June 2022
Available online 6 June 2022

Keywords:
Beamforming
Localization error
Discrete optimization

ABSTRACT

We consider a mobile multi-agent network in which the agents locate themselves in an environment through imperfect measurements and aim to transmit a message signal to a far-field base station via collaborative beamforming. The agents imperfect measurements yield localization errors that degrade the quality of service at the base station due to unknown phase offsets in the channels. Assuming that the localization errors follow Gaussian distributions, we study the design of a one-shot (non-iterative) beamforming strategy that ensures reliable communication between the agents and the base station despite the localization errors. We formulate a risk-sensitive discrete optimization problem to choose an agent subset for transmission so that the desired signal-to-interference-plus-noise ratio (SINR) at the base station is attained with minimum variance. We show that, when the localization errors have small variances characterized in terms of the carrier frequency, greedy algorithms globally minimize the variance of the received SINR. Moreover, when the localization errors have large variances, we show that the variance of the received SINR can be locally minimized by exploiting the supermodularity of the mean and variance of the received SINR. Simulations demonstrate that the proposed algorithms synthesize beamformers orders of magnitude faster than convex optimization-based approaches while achieving comparable performance with fewer agents.

1. Introduction

A mobile multi-agent network is composed of a (possibly large) number of agents, each of which has sensing, computation, communication, and mobility capabilities [1,2]. With the decreasing size and cost of available hardware, e.g., drones [3], batteries [4], and antennas [5], the deployment of mobile networked agents is an attractive option for various applications such as environmental monitoring, tracking, and surveillance.

For example, in mobile sensor networks, agents continuously traverse the environment to collect information, share the collected information with each other through low-cost short distance communications, and transmit a common information signal to a far-field base station through collaborative beamforming (CB) [6]. CB is a wireless communication technique in which the agents participating in beamforming adjust the phase and amplitude of the message signal so that the transmitted signals add coherently at the base station. Compared to single agent transmission, CB has the potential to significantly increase the range and rate of communication, to improve the directivity of the beam pattern, and to decrease the agents' individual power consumption [7,8].

In many mobile applications, agents locate themselves through imperfect sensor measurements as they travel across the environment [9]. These imperfect measurements cause localization errors that translate into unknown phase offsets in the agents' communication channels, degrading the potential coherent gain [10]. To remedy the negative effects of localization errors, in this paper, we study the problem of collaborative beamforming under localization errors and develop algorithms to establish a reliable communication link with the base station despite the agents' localization errors.

We consider a setting in which the agents' localization errors follow Gaussian distributions and the channel between the agents

☆ Y. Savas (corresponding author) and E. Noorani contributed equally to this work. This work is supported by the collaborative agreement ARL DCIST CRA W911NF-17-2-0181. E. Noorani is a Clark Doctoral Fellow at the Clark School of Engineering.

* Corresponding author.
E-mail addresses: yagiz.savas@utexas.edu (Y. Savas), enoorani@umd.edu (E. Noorani), alec.e.koppel.civ@mail.mil (A. Koppel), baras@umd.edu (J. Baras), utopcu@utexas.edu (U. Topcu), brian.m.sadler6.civ@mail.mil (B.M. Sadler).

https://doi.org/10.1016/j.sigpro.2022.108647
0165-1684/

and the base station has a strong direct-path component. In many applications, the mean and variance of the location estimates can be measured through sampling techniques, and Gaussian distributions can be used to approximate the localization errors [11,12]. The strong direct path channel occurs in free-space, and with strong Rician channels such as when one link is elevated. An additional case of interest occurs at lower frequencies, such as the lower VHF and upper HF bands, where long wavelengths result in strong barrier penetration and low levels of scattering. We note that VHF propagation studies in urban environments [13,14] show the applicability of a Rician model, even when there is no line-of-sight path.

Localization errors are associated with the topology of the network and translate to phasing errors in the transmission [6]. In the literature, the beam pattern characteristics for randomly generated network topologies are analyzed using random array theory [6,15,16]. Specifically, the authors in [15] consider a setting in which each agent's location in the environment is sampled from the same Gaussian distribution. They prove that, in this setting, the signal-to-noise ratio (SNR) received by the base station decays exponentially with a rate proportional to the variance of the Gaussian distribution. In [16], the authors show that, when the agents have fixed transmission powers and their phasing errors are identically distributed, the expected SNR increases quadratically with the number of agents so long as the expected cosine of the phasing errors is close to one. However, in multi-agent self-localization scenarios, the agents' position estimates follow non-identical distributions, in which case the aforementioned results may not be applicable. Accordingly, in this paper, we extend the existing results on the analysis of localization errors in beamforming by considering a setting in which the agents' localization errors follow Gaussian distributions with potentially different mean and covariance.

From the algorithmic perspective, different approaches are proposed to synthesize beamformers that mitigate the undesired effects of phasing errors in the transmission. The work [10] considers a case in which the agents have no statistical information regarding their localization errors and propose an iterative feedback algorithm that maximizes the SNR at the base station. Although feedback-based approaches successfully improve the quality-of-service (QoS) at the base station, their convergence to desired QoS levels may, in general, require a considerable number of two-way transmission iterations depending on network topology. When statistical channel information is available, algorithms based on semi-definite programs (SDPs) are proposed to ensure that the received SNR is above a threshold with a desired probability [17]. Similar conic optimization-based formulations are also common in the robust beamforming literature [18,19]. While SDP formulations provide a powerful method to improve the QoS at the base station, they are computationally expensive and do not scale well with the number of agents.

In this paper, we approach the beamformer design problem from a discrete optimization perspective. Specifically, given a network of agents with associated non-identical Gaussian localization errors, we propose a one-shot (non-iterative) approach that seeks a subset of agents to form a beam that achieves the desired QoS requirements without base station feedback. By employing only a subset of agents in beamforming, the proposed discrete optimization-based approach not only mitigates the undesired effects of localization errors but also increases the operational time of the agent network by consuming less energy. Moreover, our approach can also be seen as providing a one-shot initialization for feedback-based approaches by utilizing the uncertainty information and may potentially help these approaches improve the QoS at the base station faster.

The main contributions of this paper are as follows:

- First, using the variance of the SINR at the base station as a risk measure, we formulate a risk-sensitive optimization problem to form a reliable communication link between the agents and the base station despite the agents' non-identical Gaussian localization errors. Specifically, we aim to find a subset of agents to transmit the message signal such that the desired SINR level at the base station is achieved with minimum variability.
- Second, we propose an efficient sorting-based algorithm, Greedy, to solve the formulated discrete optimization problem and prove sufficient conditions for its optimality. In particular, we show that the proposed algorithm globally minimizes the variance of the received SINR when the agents' localization errors are below a certain threshold which is a function of the carrier frequency. This result characterizes a fundamental trade-off between the localization accuracy and the QoS at the base station as a function of the carrier frequency. Hence, it provides practitioners a guideline about the required localization accuracy for achieving desired QoS requirements at different frequencies. We also provide an extension of this algorithm, Double-Loop-Greedy (DLG), which improves the empirical performance.
- Third, we develop an algorithm, Difference-of-Submodular (DoS), which exploits the supermodularity of the expected value and variance of the received SINR and always returns a subset that is locally optimal for a certain relaxation of the formulated optimization problem. The DoS algorithm ensures that the agents included in beamforming locally minimizes the variance of the SINR at the base station even when their localization errors have high variances. This means that, for scenarios in which the localization accuracy required to globally minimize the variance of the received SINR cannot be achieved, the DoS algorithm can be used to provide local optimality guarantees on the QoS.

We compare the performance of the proposed algorithms with an SDP-based beamformer and demonstrate that all three algorithms (Greedy, DLG, and DoS), exhibit similar performance to the SDP-based beamformer while using significantly fewer beamforming agents. Moreover, for problem instances with a large number of agents, the Greedy and DLG algorithms compute the agent subset orders of magnitude faster than the SDP-based beamformer.

Related work: A preliminary version of this paper appeared in [20], where we present the Greedy algorithm to solve the optimization problem formulated in this paper. In this paper we extend this as follows. First, we present the DLG algorithm which improves the empirical performance over the Greedy algorithm on instances that violate the sufficient condition for optimality. Second, we prove the supermodularity of the mean and variance of the received SINR as a function of the selected agent subsets, and present the DoS algorithm to locally minimize the variance of the received SINR. Third, we provide numerical simulations to compare the performance of the proposed algorithms. Finally, we provide proof sketches for all technical results. Due to space restrictions, full proofs are given in the supplementary material and the online version [21].

In addition to the aforementioned references, the subject of this paper is also related to beamformer design when the agents have only local position information. Specifically, in [22,23], the authors consider a setting in which the global location information is not available at the agents and design an antenna array that approximates the performance of a linear antenna array using only the information of exact inter-agent distances. Here, we consider a setting in which the statistics of the global location information are available at the agents, so the problem formulation and solution are significantly different.

The idea of beamforming using only a subset of available agents has been investigated for various purposes. In [15,24,25], the authors choose a subset of sensor nodes to control the maximum sidelobe level. The work [26] develops a discete-optimization based algorithm to design a sensor array for spatial sensing applications. Finally, the reference Mehanna et al. [27] studies the antenna selection problem in multicast beamforming. Unlike these works, we consider the problem of achieving the desired SINR level at the base station with minimum variability despite localization errors and design discrete optimization-based algorithms that have provable performance guarantees.

2. System model

We consider a group of $N \in \mathbb{N}$ agents that are distributed in an environment. Each agent is equipped with a single ideal isotropic antenna with a constant transmit power $P > 0$. The agents' objective is to transmit a common message signal $m(t)$ to a base station equipped with a single antenna.The message $m(t)$ may represent raw measurement data or a waveform encoded with digital data.

2.1. Communication channel

We assume that the base station is located in the far-field region, each agent $i \in [N]$ transmits the signal $m(t)$ over a narrow-band wireless channel $h_i \in \mathbb{C}$, and the channel between the agents and the base station has a strong direct-path component.

This arises in free-space and with strong Rician channels such as when one link is elevated, and also at lower frequencies (long wavelengths). The lower frequency case is particularly of interest because the tolerable localization error scales with the wavelength. For example, as experimentally validated in [28], significant signal penetration through obstacles is possible at low-VHF frequencies. Moreover, in dense urban scenarios, low-power low-VHF communication yields significantly improved penetration and reduced multi-path [29]. Therefore, with sufficiently long wavelength the channel may become direct-path dominated even in dense clutter.

We also assume that the distances between the agents and the base station is much larger than the inter-agent distances, and local oscillators of all agents are time- and frequency-synchronized. Distributed time and frequency synchronization in wireless ad-hoc networks have been extensively studied in the literature [30,31], e.g., synchronization may be achieved using a short-range radio protocol [16,32]. There are also available methods for addressing potential implementation challenges [33]. Small synchronization errors can also result in phase errors, so these could be potentially folded into the approaches presented here, although we do not consider this further in the paper.

2.2. Collaborative transmission model

We consider a subset $\mathcal{S} \subseteq [N]$ of agents that collectively transmit the message signal $m(t)$ to the base station. All agents modulate $m(t)$ with the carrier signal $\text{Re}\{e^{j2\pi f_c t}\}$, where f_c is the carrier frequency. Our results are applicable for any f_c, while noting that longer wavelengths demand less localization error, and we will characterize this relationship explicitly.

Each agent $i \in \mathcal{S}$ adjusts the phase of the transmission with the complex gain $w_i \in \mathbb{C}$ where $|w_i| = \sqrt{P}$, where P is the transmit power. Then, the signal received by the base station is

$$y_{\mathcal{S}}(t) := \text{Re}\left\{e^{j2\pi f_c t} m(t) \sum_{i\in\mathcal{S}} w_i h_i\right\} + n(t)$$

where $n(t)$ denotes the interference-plus-noise. Without loss of generality, we let $w_i = \sqrt{P}e^{j\delta_i}$ and $h_i = a_i e^{j\eta_i}$ for each $i \in [N]$. The angle $\delta_i \in [0, 2\pi)$ denotes the phase of the gain w_i, and it is a design parameter. The magnitude $a_i > 0$ and the phase $\eta_i \in [0, 2\pi)$ characterize the channel h_i between the base station and the agent $i \in [N]$. Then, the phase offset η_i at the base station relative to a signal transmitted by an agent located at $\vec{r}_i \in \mathbb{R}^3$ (in Cartesian coordinates) is [34]

$$\eta_i = -\frac{2\pi f_c}{C}\langle \vec{r}_i, \vec{r}_c\rangle. \quad (1)$$

In (1), $\vec{r}_c \in \mathbb{R}^3$ is the unit vector pointing in the known direction of the base station, C is the speed of light, and $\langle \cdot, \cdot \rangle$ is the vector inner product.

We assume that the local position $\vec{r}_i$ of each agent $i \in [N]$ satisfies $\vec{r}_i \sim \mathcal{N}(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i)$ where $\boldsymbol{\mu}_i \in \mathbb{R}^3$ and $\boldsymbol{\Sigma}_i \in \mathbb{R}^{3\times 3}$ are, respectively, the known mean and covariance of the Gaussian distribution. The first and second order statistics of position estimates are typically easy to obtain in practice [35,36]. Moreover, the agents can utilize variations of Kalman filtering approaches to maintain Gaussian distributions for localization errors throughout their motion [37].

We also assume that $\vec{r}_i$ and $\vec{r}_j$ are independent for $i, j \in [N]$ such that $i \neq j$. Note that the independence assumption holds in practice, e.g., if the agents locate themselves in the environment based only on their own sensor measurements and do not share information with each other to improve localization.

For a given subset $\mathcal{S} \subseteq [N]$ and the corresponding phase parameters δ_i for each $i \in \mathcal{S}$, let the *array factor* be

$$F(\mathcal{S}, \boldsymbol{\delta}) := \left|\sum_{i\in\mathcal{S}} e^{j(\delta_i+\eta_i)}\right|$$

where $\boldsymbol{\delta} := [\delta_i | i \in \mathcal{S}]$ is the vector of phase parameters. Assuming that $|h_i| = |h_j|$ for all $i, j \in [N]$, the magnitude of the array factor is proportional to the square root of the SINR received by the base station [10]. We note that the assumption $|h_i| = |h_j|$ is introduced just to simplify the notation, and the results of this paper can be easily extended to cases in which $|h_i| \neq |h_j|$. In practice, the assumption $|h_i| = |h_j|$ may hold when the distance between the agents and the base station is significantly larger than the inter-agent distances. Let the *total phase* be $\Phi_i := \delta_i + \eta_i$. The square of the array factor yields the *beamforming gain* $G(\mathcal{S}, \boldsymbol{\delta})$ that is proportional to the received SINR and given by

$$G(\mathcal{S}, \boldsymbol{\delta}) := F^2(\mathcal{S}, \boldsymbol{\delta}) = \sum_{i\in\mathcal{S}}\sum_{j\in\mathcal{S}} \cos\left(\Phi_i - \Phi_j\right). \quad (2)$$

3. Problem statement

The beamforming gain $G(\mathcal{S}, \boldsymbol{\delta})$ is a fundamental quantifier of the quality of a communication link with the base station as it is proportional to the received SINR. Hence, to establish a reliable communication link, we want the beamforming gain to be high with minimum variability.

In this paper, we focus on a scenario in which the agents' exact local positions $\{\vec{r}_i : i \in [N]\}$ are not known. Each agent's potential local positions are expressed with a Gaussian distribution $\vec{r}_i \sim \mathcal{N}(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i)$. As a result, $G(\mathcal{S}, \boldsymbol{\delta})$ is a random variable. In such a scenario, a reasonable objective might be to minimize the *outage probability*, i.e., the probability that the beamforming gain falls below a certain threshold. However, there are two major difficulties involved in the optimization of the outage probability. First, due to the nonlinear structure of the beamforming gain $G(\mathcal{S}, \boldsymbol{\delta})$, deriving a closed form expression for the outage probability is not trivial. Second, as the agent subset $\mathcal{S}$ is a discrete variable and the vector $\boldsymbol{\delta}$ of phase parameters is a continuous variable, joint optimization of $G(\mathcal{S}, \boldsymbol{\delta})$ in the pair $(\mathcal{S}, \boldsymbol{\delta})$ is computationally challenging.

When the closed form probability distribution is hard to obtain or has a multimodal structure, a common approach for optimizing performance is to consider alternative risk-sensitive formulations [38,39]. Accordingly, in this paper, we consider a mean-variance optimization framework which is a widely used approach in a number of domains, ranging from cognitive radio networks [40] to finance [41].

We remedy the computational challenges involved in the joint optimization of $G(\mathcal{S}, \boldsymbol{\delta})$ in the pair $(\mathcal{S}, \boldsymbol{\delta})$ as follows. In Section 4, we show that, for any given agent subset $\mathcal{S}$, one can maximize the expected beamforming gain $\mathbb{E}[G(\mathcal{S}, \boldsymbol{\delta})]$ by selecting the vector $\boldsymbol{\delta}$ of phase parameters as $\boldsymbol{\delta} = \hat{\boldsymbol{\delta}}$ where $\hat{\delta}_i := -\mathbb{E}[\eta_i]$ for all $i \in [N]$. This result implies that, for any given agent subset $\mathcal{S}$, one can maximize the expected SINR by aligning the agents' total phases Φ_i in expectation. Therefore, we fix the phase paramater δ_i of each agent $i \in [N]$ to $\delta_i = \hat{\delta}_i$ and focus on selecting an agent subset that attains a desired level of SINR with minimum variability.

Finally, we provide the formal problem statement as follows.

Problem 1: (Subset selection) For a constant $\Gamma > 0$, and the fixed vector of phase parameters $\boldsymbol{\delta} = \hat{\boldsymbol{\delta}}$, find $\mathcal{S}^\star \subseteq [N]$ such that

$$\mathcal{S}^\star \in \arg\min_{\mathcal{S}\subseteq[N]} \quad \mathrm{Var}\Big(G(\mathcal{S}, \hat{\boldsymbol{\delta}})\Big) \tag{3a}$$

$$\text{subject to:} \quad \mathbb{E}\Big[G(\mathcal{S}, \hat{\boldsymbol{\delta}})\Big] \geq \Gamma. \tag{3b}$$

4. Statistical properties of the beamforming gain

In this section, we first derive the explicit form of the expected beamforming gain $\mathbb{E}[G(\mathcal{S}, \boldsymbol{\delta})]$ and show that, for any given subset $\overline{\mathcal{S}} \subseteq [N]$, we can maximize $\mathbb{E}[G(\overline{\mathcal{S}}, \boldsymbol{\delta})]$ by setting $\boldsymbol{\delta} = \hat{\boldsymbol{\delta}}$. We then derive the explicit form of $\mathrm{Var}(G(\mathcal{S}, \hat{\boldsymbol{\delta}}))$.

Consider the definition of $G(\mathcal{S}, \boldsymbol{\delta})$, given in (2), and recall that, for all $i \in [N]$, $\Phi_i = \eta_i + \delta_i$ where $\eta_i = 2\pi f_c \langle \vec{r}_i, \vec{r}_c\rangle / C$ and $\vec{r}_i \sim \mathcal{N}(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i)$. Then, for a given vector $\boldsymbol{\delta} \in [0, 2\pi)^N$ of phase parameters, we have $\Phi_i \sim \mathcal{N}(\theta_i, \gamma_i)$ where

$$\theta_i := \frac{2\pi f_c}{C}\Big\langle \boldsymbol{\mu}_i, \vec{r}_c\Big\rangle + \delta_i \quad \text{and} \quad \gamma_i := \frac{4\pi^2 f_c^2}{C^2}\Big\langle \vec{r}_c, \boldsymbol{\Sigma}_i \vec{r}_c\Big\rangle. \tag{4}$$

We refer to γ_i as *the effective error variance* in the localization of the ith agent.

Proposition 1. *Let $\nu_i := \exp(-\gamma_i)$. We have*

$$\mathbb{E}\Big[G(\mathcal{S}, \boldsymbol{\delta})\Big] = \Big|\mathcal{S}\Big| + \sum_{i\in\mathcal{S}} \sum_{\substack{j\in\mathcal{S} \\ j\neq i}} \sqrt{\nu_i \nu_j} \cos(\theta_i - \theta_j). \tag{5}$$

Proof (sketch): We obtain the result by taking the expectation of both sides in (2) and using the linearity of expectation, independence of $\vec{r}_i$ and $\vec{r}_j$ for all $i \neq j$, and the fact that $\mathbb{E}[\cos(X)] = e^{-\sigma^2/2}\cos(\mu)$ where $X \sim \mathcal{N}(\mu, \sigma^2)$. □

Proposition 2. *For any given $\overline{\mathcal{S}} \subseteq [N]$, $\overline{\boldsymbol{\delta}} \in \max_{\boldsymbol{\delta}\in[0,2\pi)^N} \mathbb{E}[G(\overline{\mathcal{S}}, \boldsymbol{\delta})]$ if and only if, for all $i, j \in \overline{\mathcal{S}}$, we have*

$$\Big((\overline{\delta}_i + \mathbb{E}[\eta_i]) - (\overline{\delta}_i + \mathbb{E}[\eta_i])\Big) \mod 2\pi = 0. \tag{6}$$

Proof. For any $\overline{\mathcal{S}} \subseteq [N]$, $\mathbb{E}[G(\overline{\mathcal{S}}, \boldsymbol{\delta})]$, given in (5), is maximized if and only if $(\theta_i - \theta_j) \mod 2\pi = 0$ because (i) $\cos(x) \leq 1$ for any $x \in \mathbb{R}$ and (ii) $\cos(x) = 1$ if and only if $x \mod 2\pi = 0$. Recalling that $\theta_i = \mathbb{E}[\eta_i] + \delta_i$, we conclude the result. □

The condition in (6) implies that, in order to maximize the expected beamforming gain, the agents' total phases should be aligned in expectation. Note that the vector $\hat{\boldsymbol{\delta}}$, where $\hat{\delta}_i = -\mathbb{E}[\eta_i]$ for all $i \in [N]$, satisfies the condition in (6). In the subset selection problem, we set $\boldsymbol{\delta} = \hat{\boldsymbol{\delta}}$ and aim to find a subset $\mathcal{S} \subseteq [N]$ that solves the risk-sensitive optimization problem given in (3a)-(3b).

When $\boldsymbol{\delta} = \hat{\boldsymbol{\delta}}$, we have $\theta_i = \theta_j$ for all $i, j \in [N]$, implying that

$$\mathbb{E}\Big[G(\mathcal{S}, \hat{\boldsymbol{\delta}})\Big] = \Big|\mathcal{S}\Big| + \sum_{i\in\mathcal{S}} \sum_{\substack{j\in\mathcal{S} \\ j\neq i}} \sqrt{\nu_i \nu_j}. \tag{7}$$

Next, we derive the variance of $G(\mathcal{S}, \hat{\boldsymbol{\delta}})$ as follows.

Proposition 3. *Let $\nu_i := \exp(-\gamma_i)$. We have*

$$\mathrm{Var}\Big(G(\mathcal{S}, \hat{\boldsymbol{\delta}})\Big) = \sum_{i\in\mathcal{S}} \sum_{\substack{j\in\mathcal{S} \\ j\neq i}} \Big(1 - \nu_i \nu_j\Big)^2 + 2\sum_{i\in\mathcal{S}} \sum_{\substack{j\in\mathcal{S} \\ j\neq i}} \sum_{\substack{k\in\mathcal{S} \\ k\neq i \\ k\neq j}} \Big(1 - \nu_i\Big)^2 \sqrt{\nu_j \nu_k}.$$

Proof (sketch): Note that $\mathrm{Var}(G(\mathcal{S}, \hat{\boldsymbol{\delta}})) = \mathbb{E}[G(\mathcal{S}, \hat{\boldsymbol{\delta}})^2] - \mathbb{E}[G(\mathcal{S}, \hat{\boldsymbol{\delta}})]^2$. We prove the result by utilizing the equivalence in (7) and deriving the explicit form of $\mathbb{E}[G(\mathcal{S}, \hat{\boldsymbol{\delta}})^2]$ using the identity $\cos(2x) = 2\cos(x)^2 - 1$ for any $x \in \mathbb{R}$, and the fact that $\mathbb{E}[\cos(tX)] = e^{-t^2\sigma^2/2}$ where $X \sim \mathcal{N}(0, \sigma^2)$. □

5. Agent selection under localization errors

In this section, we propose three algorithms to solve the subset selection problem and analyze their optimality guarantees. Throughout this section, we assume that the problem in (3a)-(3b) has a feasible solution. For a given problem instance, the validity of this assumption can be easily verified by checking whether $\mathbb{E}[G([N], \hat{\boldsymbol{\delta}})] \geq \Gamma$ due to the following result.

Proposition 4. *For any $\mathcal{S} \subseteq \mathcal{S}' \subseteq [N]$, $\mathbb{E}[G(\mathcal{S}, \hat{\boldsymbol{\delta}})] \leq \mathbb{E}[G(\mathcal{S}', \hat{\boldsymbol{\delta}})]$.*

Proof. The result follows from the fact that $\mathbb{E}[G(\mathcal{S}, \hat{\boldsymbol{\delta}})]$ is a sum of nonnegative terms; hence, adding an element to the subset can only increase the sum. □

5.1. Greedy algorithm

In this section, we consider a simple greedy algorithm to solve the subset selection problem and provide sufficient conditions for its optimality. The Greedy algorithm, shown in Algorithm 1, first sorts the agents' effective error variances γ_i, defined in (4), in ascending order. We note that the sorting operation can be performed in $\mathcal{O}(N\log(N))$ for an array of length N. Initializing the output set $\mathcal{S}$ to the empty set, the algorithm then iteratively adds the agent with the next lowest effective error variance to the output set until the constraint $\mathbb{E}[G(\mathcal{S}, \hat{\boldsymbol{\delta}})] \geq \Gamma$ is satisfied.

Algorithm 1 Greedy.

1: **Input:** γ_i for all $i \in [N]$, $\Gamma \in \mathbb{R}$.
2: Sort γ_i such that $\gamma_{i_1} \leq \gamma_{i_2} \leq \ldots \leq \gamma_{i_N}$.
3: $\mathcal{S} := \emptyset$, $k := 1$
4: **while** $\mathbb{E}[G(\mathcal{S}, \hat{\boldsymbol{\delta}})] < \Gamma$ **do**, $\mathcal{S} := \mathcal{S} \cup \{i_k\}$, $k := k + 1$
5: **end while**
6: **return** $\mathcal{S}$.

We now present sufficient conditions on the set $\{\gamma_i : i \in [N]\}$ for which the Greedy algorithm returns an optimal solution to the problem in (3a)-(3b). Let *the total effective error variance of a subset* $\mathcal{S} \subseteq [N]$ be measured by the function $V : 2^{[N]} \to \mathbb{R}$ where $V(\mathcal{S}) := \sum_{i\in\mathcal{S}} \gamma_i$. Consider the problem of choosing a subset $\mathcal{S}' \subseteq [N]$ that satisfies the constraint in (3b) and has the minimum total effective error variance, i.e.,

$$\mathcal{S}' \in \arg\min_{\mathcal{S}\subseteq[N]} \quad V(\mathcal{S}) \tag{8a}$$

$$\text{subject to:} \quad \mathbb{E}\Big[G(\mathcal{S}, \hat{\boldsymbol{\delta}})\Big] \geq \Gamma. \tag{8b}$$

The next result, together with Proposition 4, implies that the Greedy algorithm yields an optimal solution to the problem in (8a)-(8b).

Proposition 5. *For any $K \in \mathbb{N}$ such that $K \leq N$, we have*

$$\arg\min_{\substack{\mathcal{S}\subseteq[N]:\\ |\mathcal{S}|=K}} V(\mathcal{S}) = \arg\max_{\substack{\mathcal{S}\subseteq[N]:\\ |\mathcal{S}|=K}} \mathbb{E}\Big[G(\mathcal{S}, \hat{\boldsymbol{\delta}})\Big].$$

Proof (sketch): We prove the result by showing that the derivative of $\mathbb{E}[G(\mathcal{S}, \hat{\boldsymbol{\delta}})]$ with respect to γ_i, where $i \in \mathcal{S}$, is always negative. Therefore, a subset $\mathcal{S}$ of fixed size K maximizes the expected beamforming gain if and only if the subset has the minimum total effective error variance $V(\mathcal{S})$ among all subsets of size K. □

It can be shown that the problems in (3a)-(3b) and (8a)-(8b) are not equivalent in general. Hence, the greedy approach is, in general, not optimal to solve the subset selection problem. However, there are certain sufficient conditions, which are formalized below, under which such an approach becomes optimal.

Theorem 1. *For a given set $\{\gamma_i : i \in [N]\}$ of effective error variances, let $\gamma_{i_1} \leq \gamma_{i_2} \leq \ldots \leq \gamma_{i_N}$ where $i_k \in [N]$. A solution to the problem in (8a)-(8b) is also a solution to the problem in (3a)-(3b) if either one of the following conditions hold:*

(C1) $\mathbb{E}[G(\mathcal{S}, \hat{\boldsymbol{\delta}})] \geq \Gamma$ *where* $\mathcal{S} = \{i_1, i_2\}$,
(C2) $\gamma_{i_N} \leq 0.83$.

Proof (sketch): The main idea in the proof is to show that the derivative of $\text{Var}(G(\mathcal{S}, \hat{\boldsymbol{\delta}}))$ with respect to $\max_{i\in\mathcal{S}} \gamma_i$ is positive. Condition (C1) follows from the fact that, when $|\mathcal{S}| \leq 2$, the derivative is always positive. Condition (C2) follows from the fact that, when $\gamma_{i_N} \leq 0.83$, the derivative is positive regardless of the size of the set $\mathcal{S}$. For such γ_{i_N}, the subset with minimum $V(\mathcal{S})$ is the one that minimizes $\text{Var}(G(\mathcal{S}, \hat{\boldsymbol{\delta}}))$; hence, the problems in (8a)-(8b) and (3a)-(3b) become equivalent when (C1) or (C2) holds. □

Theorem 1 states that if all the agents have "small" effective error variances, then the Greedy algorithm returns an optimal solution to the subset selection problem. In particular, it follows from Theorem 1 that a sufficient condition for optimality characterized by the carrier frequency is

$$\max_{i\in[N]} \Big\langle \vec{r}_c, \boldsymbol{\Sigma}_i \vec{r}_c \Big\rangle \leq \frac{0.83C^2}{4\pi^2 f_c^2}.$$

For example, suppose that $\boldsymbol{\Sigma}_i = \sigma_i^2 I_{3\times3}$, where $I_{3\times3}$ is the identity matrix, and let $\sigma_{\max}^2 := \max_i \sigma_i^2$. Then, we have $\sigma_{\max}^2 \leq \frac{0.83C^2}{4\pi^2 f_c^2}$ as the sufficient condition (C2). In Fig. 1, we plot the trade-off between the carrier frequency f_c and the maximum variance $\sigma_{\max}^2$ under which the Greedy algorithm is optimal. As f_c decreases (resulting in longer wavelength), condition (C2) allows larger position error variance. For example, below 50 MHz in the lower VHF, agents are allowed to have localization error variance approaching one square meter or more. This can be relatively easily achieved with, for example, global navigation sensors [42] and employing existing localization algorithms [43,44].

Although the Greedy algorithm provides an optimal solution to the subset selection problem under the sufficient conditions given in Theorem 1, we may have scenarios where the localization error variance ranges from small to large and (C2) does not hold for all agents. In the Appendix, we present a simple extension of the Greedy algorithm, which we refer to as the Double-Loop-Greedy (DLG) algorithm. Although the DLG algorithm has the same theoretical guarantees with the Greedy algorithm, it improves the empirical performance when the conditions given in Theorem 1 are violated, as shown in numerical examples.

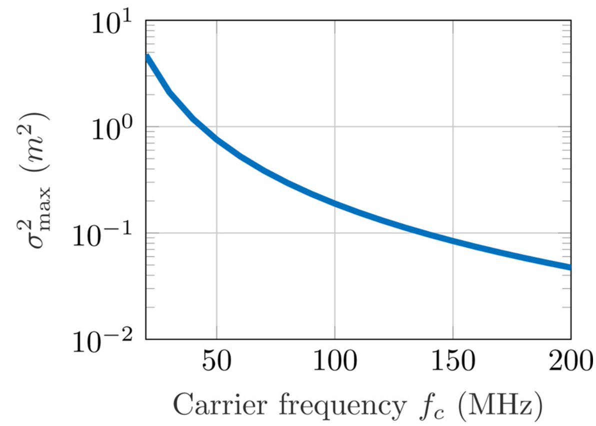

Fig. 1. Maximum localization error variance $\sigma_{\max}^2$ allowed for the optimality of the greedy algorithm as a function of the carrier frequency f_c. Note that the localization error tolerance is relaxed at lower frequencies (longer wavelengths).

5.2. Difference-of-Submodular (DoS) algorithm

The Greedy algorithm is guaranteed to return optimal solutions to the subset selection problem under the sufficient conditions stated in Theorem 1. In this section, we propose a second approach to solve the subset selection problem, which *always* returns a locally optimal solution to a certain relaxation of the subset selection problem. Although the proposed approach is computationally more demanding, its local optimality guarantee is independent of the carrier frequency unlike the Greedy algorithm.

Before presenting the Difference-of-Submodular (DoS) algorithm, we first provide a definition of submodularity and show that both the expected value and the variance of the beamforming gain are supermodular set functions.

Definition 1. A set function $f : 2^\Omega \to \mathbb{R}$ is submodular if for every $X, Y \subseteq \Omega$ with $X \subseteq Y$ and every $e \in \Omega \backslash Y$, we have $f(X \cup \{e\}) - f(X) \geq f(Y \cup \{e\}) - f(Y)$.

A set function $f : 2^\Omega \to \mathbb{R}$ is *supermodular* if the set function $-f$ is submodular.

Theorem 2. *Both $\mathbb{E}[G(\mathcal{S}, \hat{\boldsymbol{\delta}})]$ and $\text{Var}(G(\mathcal{S}, \hat{\boldsymbol{\delta}}))$ are supermodular set functions.*

Proof (sketch): For notational simplicity, let $\overline{G}(\mathcal{S}) := G(\mathcal{S}, \hat{\boldsymbol{\delta}})$. For $X, Y \subseteq [N]$ such that $X \subseteq Y$, let $X' = X \cup \{e\}$ and $Y' = Y \cup \{e\}$ where $e \in [N]\backslash Y$. We have,

$$X_{\text{diff}} := \mathbb{E}[\overline{G}(X')] - \mathbb{E}[\overline{G}(X)] = 1 + 2\sqrt{\upsilon_e} \sum_{i\in X} \sqrt{\upsilon_i},$$

$$Y_{\text{diff}} := \mathbb{E}[\overline{G}(Y')] - \mathbb{E}[\overline{G}(Y)] = 1 + 2\sqrt{\upsilon_e} \sum_{i\in Y} \sqrt{\upsilon_i}.$$

Using $\upsilon_i \geq 0$ and $X \subseteq Y$, we obtain $X_{\text{diff}} - Y_{\text{diff}} = -2\sqrt{\upsilon_e}\sum_{i\in Y\backslash X} \sqrt{\upsilon_i} \leq 0$. Hence, we conclude that $\mathbb{E}[G(\mathcal{S}, \hat{\boldsymbol{\delta}})]$ is supermodular.

To show the supermodularity of $\text{Var}(G(\mathcal{S}, \hat{\boldsymbol{\delta}}))$, we define $\overline{X}_{\text{diff}} := \text{Var}(\overline{G}(X')) - \text{Var}(\overline{G}(X))$ and $\overline{Y}_{\text{diff}} := \text{Var}(\overline{G}(Y')) - \text{Var}(\overline{G}(Y))$. Then, we show that $X_{\text{diff}} - Y_{\text{diff}} \leq 0$ by using the fact that $\upsilon_i \geq 0$ and $X \subseteq Y$. □

Next, we formalize the notion of local optimality for discrete optimization problems and introduce the DoS algorithm which utilizes the results of [45].

Definition 2. [45] For a set function $\phi : 2^{\Omega} \to \mathbb{R}$, a sequence $\{S_t \subseteq \Omega : t \in \mathbb{N}\}$ is said to converge to a local minimum if there exists a constant $M \in \mathbb{N}$ such that $\phi(S_m) = \phi(S_n)$ for all $m, n \geq M$, and for any $k \in \mathbb{N}$, $\phi(S_k) \leq \phi(S_l)$ for all $l \leq k$.

Let $f : 2^{\Omega} \to \mathbb{R}$ and $g : 2^{\Omega} \to \mathbb{R}$ be submodular set functions. In [45], the authors present an algorithm, called Submodular-Supermodular-Procedure (SSP), that returns a locally optimal solution to the following problem

$$\min_{S \subseteq \Omega} \quad f(S) - g(S). \tag{9}$$

The DoS algorithm, shown in Algorithm 2, utilizes the SSP as a

Algorithm 2 Difference-of-Submodular (DoS).

1: **Input:** γ_i for all $i \in [N]$, $\Gamma \in \mathbb{R}$, $\lambda_0 > 0$, $\alpha > 1$.
2: $\mathcal{S} := \emptyset$, $k := 0$.
3: **while** $\mathbb{E}[G(\mathcal{S}, \hat{\boldsymbol{\delta}})] < \Gamma$ **do**
4: $\quad f(\cdot) := -\lambda_k \mathbb{E}[G(\cdot, \hat{\boldsymbol{\delta}})]$, $g(\cdot) := -\mathrm{Var}(G(\cdot, \hat{\boldsymbol{\delta}}))$
5: $\quad \mathcal{S} := \mathrm{SSP}(f(\cdot), g(\cdot))$
6: $\quad k := k+1$, $\lambda_k := \alpha\lambda_{k-1}$.
7: **end while**
8: **return** $\mathcal{S}$.

subprocedure to return a locally optimal solution to a certain relaxation of the subset selection problem. In particular, it takes two parameters $\lambda_0 > 0$ and $\alpha > 1$ as inputs as well as the agents' effective localization error variances γ_i and the expected gain threshold Γ. At the kth iteration, where $k \in \mathbb{N}$, using the SSP as a subprocedure, the DoS algorithm finds a locally optimal solution to the following problem

$$\min_{\mathcal{S} \subseteq [N]} \quad \mathrm{Var}\Big(G(\mathcal{S}, \hat{\boldsymbol{\delta}})\Big) - \lambda_k \mathbb{E}\Big[G(\mathcal{S}, \hat{\boldsymbol{\delta}})\Big] \tag{10}$$

where λ_k is iteratively defined as $\lambda_k = \alpha\lambda_{k-1}$. The algorithm terminates when the solution returned by the SSP satisfies $\mathbb{E}[G(\mathcal{S}, \hat{\boldsymbol{\delta}})] \geq \Gamma$.

Convergence of the DoS algorithm: For the DoS algorithm to terminate, the subprocedure SSP should output a subset $\mathcal{S} \subseteq [N]$ such that $\mathbb{E}[G(\mathcal{S}, \hat{\boldsymbol{\delta}})] \geq \Gamma$. At the kth iteration, the SSP finds a locally optimal solution to the problem in (10) by computing successive modular approximations of the function $\mathrm{Var}(G(\mathcal{S}, \hat{\boldsymbol{\delta}}))$ and finding a globally optimal solution to each of the resulting approximation problems. Since $\lambda_0 > 0$ and $\alpha > 1$, the parameter λ_k increases at each iteration. Hence, in terms of the objective value, the globally optimal solution of the approximate problems become closer to the globally optimal solution of $\max_{\mathcal{S}} \mathbb{E}[G(\mathcal{S}, \hat{\boldsymbol{\delta}})]$, which is $\mathcal{S} = [N]$. Since we assumed at the beginning that there exists a feasible solution to the subset selection problem, the DoS algorithm is guaranteed to terminate for some finite $k \in \mathbb{N}$.

Optimality of the DoS algorithm: As mentioned earlier, at each iteration, the DoS algorithm computes a locally optimal solution to the problem in (10). Hence, the subset returned by the DoS algorithm is a locally optimal solution to the following relaxation of the subset selection problem

$$\min_{\mathcal{S} \subseteq [N]} \quad \mathrm{Var}\Big(G(\mathcal{S}, \hat{\boldsymbol{\delta}})\Big) - \lambda_{k^\star} \mathbb{E}\Big[G(\mathcal{S}, \hat{\boldsymbol{\delta}})\Big] \tag{11}$$

where $k^\star$ is the number of iterations until the convergence of the DoS algorithm. We also note that the above problem formulation is sometimes referred to as a "regularized version" of the original constrained optimization problem [46].

6. Numerical experiments

In this section, we present numerical simulation results that demonstrate the performance of the proposed algorithms. All computations are run on a 3.1-GHz desktop with 32 GB RAM using the toolbox [47] for the implementation of the SSP (step 5 in the DoS algorithm).

6.1. Suboptimality ratio on small-scale instances

A typical measure to assess the empirical performance of an optimization algorithm is its suboptimality ratio on randomly generated instances. For the purposes of this paper, the suboptimality ratio demonstrates how much the variance of the received SINR is larger than the minimum achievable one when the agents employ the proposed algorithms to transmit the message.

We compare the suboptimality of the proposed algorithms as a function of three problem parameters: the total number N of agents, the maximum localization error $\gamma_{\max} := \min\{\gamma : \gamma \geq \gamma_i, \text{for all } i \in [N]\}$, and the expected beamforming gain threshold $\Gamma = \beta\Gamma_{\max}$ where $0 < \beta \leq 1$ and $\Gamma_{\max} := \mathbb{E}[G([N], \hat{\boldsymbol{\delta}})]$ is the maximum expected beamforming gain that can be achieved by the agents.

For a given problem instance, we measure the performance of an algorithm by the suboptimality ratio (SR) of its output. Specifically, let $\mathcal{S}^\star$ be an optimal solution to the given problem instance (3a)-(3b), which, for small N, can be computed by considering all subsets $\mathcal{S} \subseteq [N]$. Moreover, let $\overline{\mathcal{S}}$ be the (possibly suboptimal) output of a given algorithm. We define the SR of the algorithm on the given instance as

$$\mathrm{SR} := \frac{\mathrm{Var}\Big(G(\overline{\mathcal{S}}, \hat{\boldsymbol{\delta}})\Big)}{\mathrm{Var}\Big(G(\mathcal{S}^\star, \hat{\boldsymbol{\delta}})\Big)}.$$

All proposed algorithms, i.e., Greedy, DLG, and DoS, have $\mathrm{SR} \geq 1$ since their output $\overline{\mathcal{S}}$ satisfies $\mathbb{E}[G(\overline{\mathcal{S}}, \hat{\boldsymbol{\delta}})] \geq \Gamma$.

In the first set of experiments, we investigate the relationship between the algorithms' SR, the total number N of agents, and the bound $\gamma_{\max}$ on the agents' effective localization error variances. For a given N and $\gamma_{\max}$, a problem instance consists of $\{\gamma_i : i \in [N]\}$ where each γ_i is uniformly randomly selected from the interval $(0, \gamma_{\max})$. We set the expected beamforming gain threshold as $\Gamma = 0.6\Gamma_{\max}$ to allow the algorithms to output subsets of different sizes if it is optimal to do so. Furthermore, we set $\lambda_0 = 1$ and $\alpha = 2$ for the DoS algorithm. Recall that the DoS algorithm has only local optimality guarantees. Hence, the SR of the algorithm's output depends on the initialization of the SSP. Accordingly, for each problem instance, we run the DoS algorithm 10 times using different initializations and report the performance of the best output.

For each $N \in \{6, 8, 10\}$ and each $\gamma_{\max} \in \{1, 2, \ldots, 20\}$, we generate 100 problem instances and illustrate the average SRs of all algorithms in Fig. 2 (top). As can be seen from the figure, all algorithms show near-optimal performance ($\mathrm{SR} \leq 1.3$) for all $(N, \gamma_{\max})$ pairs. Recall from Theorem 1 that, when $\gamma_{\max} \leq 0.83$, both the Greedy and DLG algorithms are guaranteed to have SR=1. Moreover, the DLG algorithm is always guaranteed to have smaller SR than the Greedy algorithm. The results shown in Fig. 2 (top) empirically witness these theoretical guarantees. Moreover, as can be seen from the figure, both the Greedy and DLG algorithms perform well ($\mathrm{SR} \leq 1.1$) even when the sufficient optimality condition, $\gamma_{\max} \leq 0.83$ is violated. The DoS algorithm shows comparable performance to that of the Greedy and DLG algorithms when $\gamma_{\max} \geq 10$. However, for small effective localization error variances, the Greedy and DLG algorithms perform significantly better than

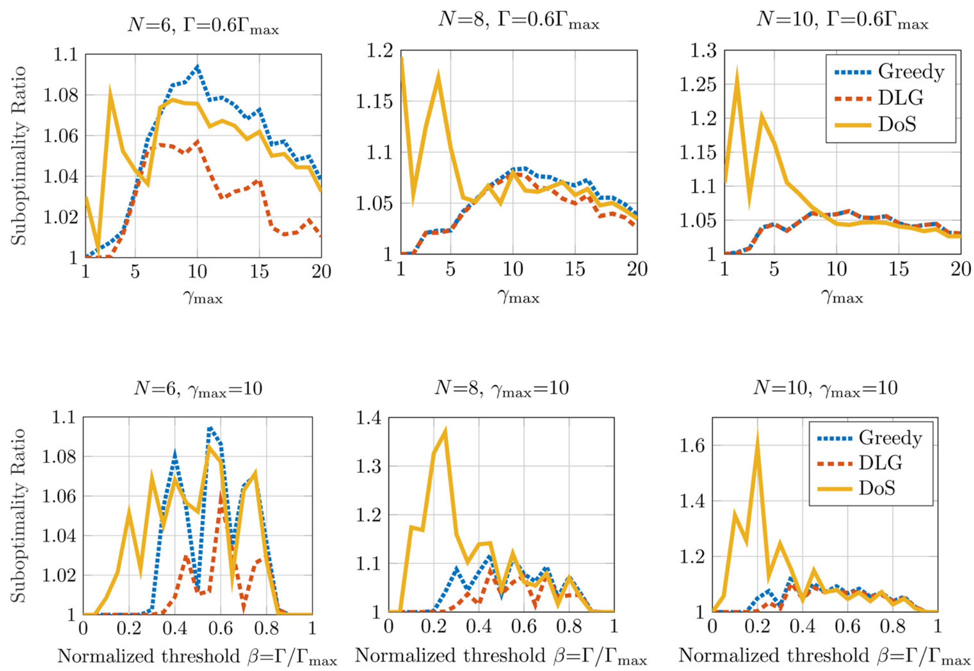

Fig. 2. Suboptimality ratios (SRs) of the proposed algorithms averaged over 100 randomly generated subset selection problem instances. (Top) SRs when the total number of agents is $N \in \{6, 8, 10\}$ and the effective localization error variances $\{\gamma_i : i \in [N]\}$ are generated randomly from the interval $(0, \gamma_{\max})$. (Bottom) SRs when the total number of agents is $N \in \{6, 8, 10\}$ and the expected beamforming gain threshold is $\Gamma = \beta\Gamma_{\max}$.

the DoS algorithm. Finally, note that the SR of the DoS algorithm increases with increasing total number N of agents in general. On the other hand, the SR of the Greedy and DLG algorithms, in general, remain at the same level despite the increasing total number of agents.

In the second set of experiments, we investigate the relationship between the algorithms' SRs, the total number N of agents, and the normalized threshold $\beta = \Gamma/\Gamma_{\max}$. For given N and β, a problem instance consists of $\{\gamma_i : i \in [N]\}$ where each γ_i is selected uniformly randomly from the interval (0,10), i.e., $\gamma_{\max} = 10$. Finally, we set $\lambda_0 = 1$ and $\alpha = 2$, and run the DoS algorithm with 10 random initializations.

For each $N \in \{4, 6, 8\}$ and each $\beta \in \{0.1, 0.2, \ldots, 1\}$, we generate 100 problem instances. The average SRs of the algorithms over the generated instances are shown in Fig. 2 (bottom). As can be seen from the figure, all algorithms have average SR less than 1.6 for each (N, β) pair. Recall from Theorem 1 that, for instances in which the threshold Γ can be attained using two agents, the Greedy and DLG algorithms have $SR = 1$. For small β values, we observe that the Greedy and DLG algorithms achieve $SR = 1$ since, in most problem instances, the threshold is attained by using two agents. Similar to the first set of experiments (Fig. 2 (top)), we observe that the SR of the DoS algorithm increases with increasing N in general. Moreover, the performance of the DoS algorithm, in general, improves with increasing β values.

The empirical performance evaluation of the proposed algorithms on small-scale instances show that all three algorithms, Greedy, DLG, and DoS, achieve near-optimal performance (SR$\leq$ 1.6) for a range of N, β, and $\gamma_{\max}$ values. Although the Greedy and DLG algorithms have theoretical optimality guarantees only for small $\gamma_{\max}$ and β values, they perform well (SR$\leq$ 1.1) even for large $\gamma_{\max}$ and β values. On the other hand, although the local optimality guarantee of the DoS algorithm is independent of the problem parameters, the performance of the algorithm is, in general, comparable (SR $\leq$ 1.1) to that of the Greedy and DLG algorithms only for large $\gamma_{\max}$ and β values.

6.2. Performance comparison with an SDP-based beamformer

We compare the performance of the proposed algorithms with a semi-definite programming-based (SDP-based) beamforming algorithm. SDP-based methods are widely used in robust beamforming to mitigate the degrading effects of uncertain parameters on the beam pattern [18,48]. Accordingly, for comparison, we synthesize a beamforming vector $\mathbf{w}^\star \in \mathbb{C}^N$ where

$$\mathbf{w}^\star \in \arg\min_{\mathbf{w}\in\mathbb{C}^N} \quad \|\mathbf{w}\|_2^2 \tag{12a}$$

$$\text{subject to:} \quad \mathbb{E}[\mathbf{w}^H\mathbf{H}\mathbf{w}] \geq \Gamma \tag{12b}$$

$$\forall i \in [N], \quad |w_i|^2 \leq 1. \tag{12c}$$

Here, $\mathbf{H} \in \mathbb{C}^{N\times N}$ is $\mathbf{H} = \mathbf{h}\mathbf{h}^H$ where $\mathbf{h}^H = [h_1, h_2, \ldots, h_N]$, and $\mathbf{w}^H = [w_1, w_2, \ldots, w_N]$. The constraint in (12c) ensures that $w_i = \sqrt{P}e^{j\delta_i}$ for some $P \leq 1$.

A solution to the problem in (12a)–(12c) is a beamformer $\mathbf{w}^\star$ that attains the desired threshold Γ with minimum total power while respecting the individual power constraints in (12c). It can be shown that a solution to the problem in (12a)–(12c) can be computed *exactly* by solving an SDP [19,48]. To synthesize the beamformer $\mathbf{w}^\star$, we utilized the SDP solver of the CVX toolbox [49] with its nominal parameters. Note that the beamformer $\mathbf{w}^\star$ minimizes the total transmit power of the antenna array while ensuring that the expected beamforming gain exceeds the desired threshold Γ. Therefore, it represents a solution to a convex relaxation of the problem

$$\min_{\mathcal{S}\subseteq[N],\boldsymbol{\delta}\in\mathbb{C}^N} \quad |\mathcal{S}|$$

$$\text{subject to} \quad \mathbb{E}[G(\mathcal{S}, \boldsymbol{\delta})] \geq \Gamma$$

which is a risk-neutral version of the subset selection problem. For given $\mathbf{w}^\star = [w_1^\star, w_2^\star, \ldots, w_N^\star]$, we let the corresponding optimal subset be $\mathcal{S}^\star = \{i \in [N] : |w_i^\star| > \epsilon\}$ where $\epsilon = 10^{-1}$.

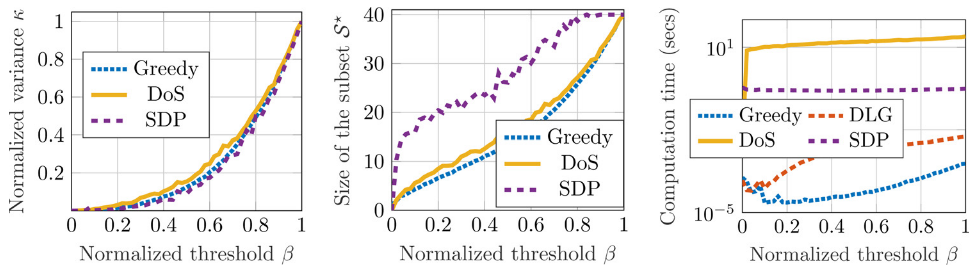

Fig. 3. Performance comparison of the proposed algorithms with an SDP-based beamformer. (Left) For a given threshold $\Gamma = \beta\Gamma_{\max}$, the normalized variances of the beamforming gains are similar for all approaches. (Middle) Proposed algorithms achieve the same performance by employing strict subsets of the agent group when possible. (Right) Greedy and DLG algorithms synthesize beamformers orders of magnitudes faster than the SDP-based approach.

We generate 100 subset selection problem instances by setting $N = 40$ and selecting the error variances $\{\gamma_i : i \in [N]\}$ uniformly randomly from the interval (0,10), i.e., $\gamma_{\max} = 10$. For the DoS algorithm, we set $\lambda_0 = 1$ and $\alpha = 2$, and run the algorithm with 10 random initializations. We compare the performance with respect to three metrics: the average variance of the beamforming gain, the average size of the selected subset, and the average computation time.

Fig. 3 (left) shows the normalized variance of the beamforming gain, i.e., $\kappa = \text{Var}(G(\mathcal{S}^{\star}, \boldsymbol{\delta}))/\text{Var}(G([N], \hat{\boldsymbol{\delta}}))$, versus the normalized threshold $\beta = \Gamma/\Gamma_{\max}$. In the figure, we do not plot the results of the DLG algorithm since it is almost exactly the same as the Greedy algorithm. As can be seen from the figure, the proposed discrete-optimization-based algorithms achieve similar performance to that of the SDP-based beamformer. The variance of the SDP-based beamformer is, in general, smaller than the variance of the proposed algorithms since the problem in (12a)–(12c) is a convex relaxation of the subset selection problem.

Fig. 3 (middle) demonstrates the trade-off between the normalized threshold β and the average size of the optimal subset $\mathcal{S}^{\star}$. The plot for the DLG algorithm is omitted since the average size of the subsets selected by the DLG algorithm is almost exactly the same as the Greedy algorithm. As can be seen from the figure, for $\beta < 1$, the proposed algorithms employ strict subsets of the agent network $[N]$ where $N = 40$. On the other hand, the SDP-based beamformer includes all the agents to beamforming for all $\beta > 0$. Combined with the results shown in Fig. 3 (left), this result suggests that the proposed algorithms achieve similar performance to that of the SDP-based approach using fewer agents. This saves resources overall and allows unallocated agents to take on other tasks.

Finally, Fig. 3 (right) shows the computation times for all algorithms. The Greedy and DLG algorithms run orders of magnitude faster than the SDP-based beamformer. On the other hand, the DoS algorithm takes longer than the SDP-based beamformer to select a subset in general. The long computation time is partially due to the fact that we run the DoS algorithm with 10 random initializations to improve its performance. We observe in our experiments that the variance of the beamforming gain for the subset selected by the DoS algorithm decreases considerably as the number of random initializations used in the DoS algorithm increases. Therefore, there is a trade-off between the computation time of the DoS algorithm and the quality of the beamformer.

The empirical evaluations presented above suggests that the proposed discrete optimization-based approaches have the potential to synthesize beamformers with similar performance to that of the convex optimization-based beamformers using significantly less number of agents. Furthermore, when the Greedy and DLG algorithms are employed to synthesize beamformers, the required computation time for the synthesis can be significantly reduced with respect to SDP-based approaches.

7. Conclusions

We considered a mobile multi-agent network in which the sensor nodes locate themselves in an environment through imperfect measurements and aim to transmit a message signal to a base station. Under the assumption that the agents have Gaussian localization errors, we developed three one-shot (non-iterative) algorithms, Greedy, Double-Loop-Greedy (DLG), and Difference-of-Submodular (DoS), each of which chooses a subset of agents to optimize the quality-of-service without requiring feedback from the base station.

When the localization errors for all agents are below a certain threshold, the Greedy algorithm globally minimizes the variance of the SINR received by the base station while guaranteeing that the expected SINR is above a desired threshold. The DLG algorithm improves the empirical performance over the Greedy algorithm. Finally, the DoS algorithm enables the agents to locally optimize the reliability of the communication link even when the localization errors are large. We empirically showed that the proposed algorithms achieve similar performance with a convex optimization-based algorithm while using significantly fewer agents. Moreover, the Greedy and DLG algorithms run orders of magnitude faster than the convex optimization-based approach.

Although the DoS algorithm achieves comparable performances to that of the convex optimization-based algorithm with fewer agents, its computational requirements may hinder its applicability to scenarios in which the size of the agent network is large. Interesting future directions include developing algorithms for large scale systems that are both fast and have performance guarantees, as well as utilizing our algorithms to initialize some further beamforming refinement such as using base station feedback.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

CRediT authorship contribution statement

Yagiz Savas: Conceptualization, Methodology, Software, Writing – original draft, Writing – review & editing. **Erfaun Noorani:** Conceptualization, Methodology, Software, Writing – original draft, Writing – review & editing. **Alec Koppel:** Conceptualization, Methodology, Supervision, Writing – original draft, Writing – review & editing. **John Baras:** Conceptualization, Methodology, Supervision, Funding acquisition, Writing – review & editing. **Ufuk Topcu:** Conceptualization, Methodology, Supervision, Funding acquisition, Writing – review & editing. **Brian M. Sadler:** Conceptualization, Methodology, Supervision, Funding acquisition, Writing – review & editing.

Appendix A

In this appendix, we present the Double-Loop-Greedy (DLG) algorithm, shown in Algorithm 3, which is an extension of the

Algorithm 3 Double-Loop-Greedy (DLG).

1: **Input:** γ_i for all $i \in [N]$, $\Gamma \in \mathbb{R}$.
2: Sort γ_i such that $\gamma_{i_1} \leq \gamma_{i_2} \leq \ldots \leq \gamma_{i_N}$.
3: $\mathcal{S}_1 := \emptyset$, $\mathcal{S}_2 := \emptyset$, $k := 1$, $l := N$
4: **while** $\mathbb{E}[G(\mathcal{S}_1, \hat{\boldsymbol{\delta}})] < \Gamma$ **do**, $\mathcal{S}_1 := \mathcal{S}_1 \cup \{i_k\}$, $k := k+1$
5: **end while**
6: **while** $\mathbb{E}[G(\mathcal{S}_2, \hat{\boldsymbol{\delta}})] < \Gamma$ **do**, $\mathcal{S}_2 := \mathcal{S}_2 \cup \{i_l\}$, $l := l-1$
7: **end while**
8: **if** $\mathrm{Var}(G(\mathcal{S}_1, \hat{\boldsymbol{\delta}})) < \mathrm{Var}(G(\mathcal{S}_2, \hat{\boldsymbol{\delta}}))$ **then** $\mathcal{S} := \mathcal{S}_1$
9: **else** $\mathcal{S} := \mathcal{S}_2$
10: **end if**
11: **return** $\mathcal{S}$.

Greedy algorithm. The idea is to form two solutions and then choose the best one. Set $\mathcal{S}_1$ is developed as the Greedy algorithm, while set $\mathcal{S}_2$ is formed in a similar way but starting from the worst case error. First, the DLG algorithm sorts the agents' effective error variances γ_i in ascending order. Sets $\mathcal{S}_1$ and $\mathcal{S}_2$ are initially empty. Starting from the agent with the lowest effective error variance, the agent with the next lowest effective error variance is iteratively added to the set $\mathcal{S}_1$ until the constraint $\mathbb{E}[G(\mathcal{S}_1, \hat{\boldsymbol{\delta}})] > \Gamma$ is satisfied. This is the same procedure as the Greedy algorithm.

To form $\mathcal{S}_2$ we proceed as follows. Starting from the agent with the highest effective error variance, the agent with the next highest effective error variance is iteratively added to the set $\mathcal{S}_2$ until the constraint $\mathbb{E}[G(\mathcal{S}_2, \hat{\boldsymbol{\delta}})] > \Gamma$ is satisfied. Finally, the DLG algorithm compares the variance of the beamforming gain for $\mathcal{S}_1$ and $\mathcal{S}_2$, and outputs the one with smaller value. We note that the time complexity of the DLG algorithm is the same as the Greedy algorithm.

For a given problem instance, the subset $\mathcal{S} \subseteq [N]$ returned by the DLG algorithm satisfies $\mathrm{Var}(G(\mathcal{S}, \hat{\boldsymbol{\delta}})) \leq \mathrm{Var}(G(\mathcal{S}', \hat{\boldsymbol{\delta}}))$ where $\mathcal{S}' \subseteq [N]$ is the subset returned by the Greedy algorithm. Hence, DLG provides an optimal solution to the problem in (8a)-(8b) under the sufficient conditions stated in Theorem 1.

Supplementary material

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.sigpro.2022.108647.

References

[1] K. Dantu, M. Rahimi, H. Shah, S. Babel, A. Dhariwal, G.S. Sukhatme, Robomote: enabling mobility in sensor networks, in: International Symposium on Information Processing in Sensor Networks, 2005, pp. 404–409.
[2] A. Howard, M.J. Matarić, G.S. Sukhatme, Mobile sensor network deployment using potential fields: a distributed, scalable solution to the area coverage problem, in: Distributed Autonomous Robotic Systems 5, Springer, 2002, pp. 299–308.
[3] D. Floreano, R.J. Wood, Science, technology and the future of small autonomous drones, Nature 521 (7553) (2015) 460–466.
[4] W. van Schalkwijk, B. Scrosati, Advances in Lithium-ion Batteries, Springer Science & Business Media, 2007.
[5] J. Oh, J. Choi, F.T. Dagefu, K. Sarabandi, Extremely small two-element monopole antenna for HF band applications, IEEE Trans. Antennas Propag. 61 (6) (2013) 2991–2999.
[6] H. Ochiai, P. Mitran, H.V. Poor, V. Tarokh, Collaborative beamforming for distributed wireless ad hoc sensor networks, IEEE Trans. Signal Process. 53 (11) (2005) 4110–4124.
[7] G. Barriac, R. Mudumbai, U. Madhow, Distributed beamforming for information transfer in sensor networks, in: Third International Symposium on Information Processing in Sensor Networks, 2004, pp. 81–88.
[8] S. Jayaprakasam, S.K.A. Rahim, C.Y. Leow, Distributed and collaborative beamforming in wireless sensor networks: classifications, trends, and research directions, IEEE Commun. Surv. Tutor. 19 (4) (2017) 2092–2116.
[9] I. Amundson, X.D. Koutsoukos, A survey on localization for mobile wireless sensor networks, in: International Workshop on Mobile Entity Localization and Tracking in GPS-less Environments, 2009, pp. 235–254.
[10] R. Mudumbai, J. Hespanha, U. Madhow, G. Barriac, Distributed transmit beamforming using feedback control, IEEE Trans. Inf. Theory 56 (1) (2009) 411–426.
[11] J. Yin, D. Wang, Y. Wu, T. Tang, Single-step localization using multiple moving arrays in the presence of observer location errors, Signal Process. 152 (2018) 392–410.
[12] N. Patwari, A.O. Hero, M. Perkins, N.S. Correal, R.J. O'dea, Relative location estimation in wireless sensor networks, IEEE Trans. Signal Process. 51 (8) (2003) 2137–2148.
[13] J.A. Pugh, R.J.C. Bultitude, P.J. Vigneron, Mobile propagation measurements with low antennas for segmented wideband communications at VHF, in: 2006 12th International Symposium on Antenna Technology and Applied Electromagnetics and Canadian Radio Sciences Conference, IEEE, 2006, pp. 1–4.
[14] P.J. Vigneron, J.A. Pugh, Propagation models for mobile terrestrial VHF communications, in: MILCOM 2008-2008 IEEE Military Communications Conference, IEEE, 2008, pp. 1–7.
[15] M.F.A. Ahmed, S.A. Vorobyov, Collaborative beamforming for wireless sensor networks with gaussian distributed sensor nodes, IEEE Trans. Wireless Commun. 8 (2) (2009) 638–643.
[16] R. Mudumbai, G. Barriac, U. Madhow, On the feasibility of distributed beamforming in wireless networks, IEEE Trans. Wireless Commun. 6 (5) (2007) 1754–1763.
[17] K. Wang, A.M. So, T. Chang, W. Ma, C. Chi, Outage constrained robust transmit optimization for multiuser MISO downlinks: tractable approximations by conic optimization, IEEE Trans. Signal Process. 62 (21) (2014) 5690–5705.
[18] R.G. Lorenz, S.P. Boyd, Robust minimum variance beamforming, IEEE Trans. Signal Process. 53 (5) (2005) 1684–1696.
[19] Z.-Q. Luo, W.-K. Ma, A.M.-C. So, Y. Ye, S. Zhang, Semidefinite relaxation of quadratic optimization problems, IEEE Signal Process. Mag. 27 (3) (2010) 20–34.
[20] E. Noorani, Y. Savas, A. Koppel, J. Baras, U. Topcu, B.M. Sadler, Collaborative beamforming for agents with localization errors, in: Asilomar Conference on Signals, Systems, and Computers, 2021.
[21] E. Noorani, Y. Savas, A. Koppel, J. Baras, U. Topcu, B.M. Sadler, Collaborative beamforming under localization errors: a discrete optimization approach, arXiv:2003.12637(2020).
[22] P.J. Vincent, M. Tummala, J. McEachen, A beamforming approach for distributed wireless sensor networks, in: International Conference on System of Systems Engineering, 2007, pp. 1–6.
[23] N. Papalexidis, T.O. Walker, C. Gkionis, M. Tummala, J. McEachen, A distributed approach to beamforming in a wireless sensor network, in: Conference on Signals, Systems and Computers, 2007, pp. 606–610.
[24] J. Chen, C. Wen, K. Wong, An efficient sensor-Node selection algorithm for sidelobe control in collaborative beamforming, IEEE Trans. Veh. Technol. 65 (8) (2016) 5984–5994.
[25] G. Sun, Y. Liu, J. Zhang, A. Wang, X. Zhou, Node selection optimization for collaborative beamforming in wireless sensor networks, Ad Hoc Netw. 37 (2016) 389–403.
[26] G. Shulkind, S. Jegelka, G.W. Wornell, Sensor array design through submodular optimization, IEEE Trans. Inf. Theory 65 (1) (2019) 664–675.
[27] O. Mehanna, N.D. Sidiropoulos, G.B. Giannakis, Joint multicast beamforming and antenna selection, IEEE Trans. Signal Process. 61 (10) (2013) 2660–2674.
[28] F.T. Dagefu, J. Choi, M. Sheikhsofla, B.M. Sadler, K. Sarabandi, Performance assessment of lower VHF band for short-range communication and geolocation applications, Radio Sci. 50 (5) (2015) 443–452.
[29] J. Choi, F.T. Dagefu, B.M. Sadler, K. Sarabandi, Low-power low-VHF ad-Hoc networking in complex environments, IEEE Access 5 (2017) 24120–24127.
[30] K. Römer, Time synchronization in ad hoc networks, in: Proceedings of the 2nd ACM International Symposium on Mobile ad hoc Networking & Computing, 2001, pp. 173–182.
[31] L.-A. Phan, T. Kim, T. Kim, J. Lee, J.-H. Ham, Performance analysis of time synchronization protocols in wireless sensor networks, Sensors 19 (13) (2019) 3020.
[32] R. Mudumbai, D.R.B. Iii, U. Madhow, H.V. Poor, Distributed transmit beamforming: challenges and recent progress, IEEE Commun. Mag. 47 (2) (2009) 102–110.
[33] D. Djenouri, M. Bagaa, Synchronization protocols and implementation issues in wireless sensor networks: a review, IEEE Syst. J. 10 (2) (2014) 617–627.
[34] R.J. Mailloux, Phased array theory and technology, Proc. IEEE 70 (3) (1982) 246–291.
[35] S. Thrun, D. Fox, W. Burgard, F. Dellaert, Robust Monte Carlo localization for mobile robots, Artif. Intell. 128 (1–2) (2001) 99–141.
[36] Z. Chen, et al., Bayesian filtering: from Kalman filters to particle filters, and beyond, Statistics 182 (1) (2003) 1–69.
[37] C. Hu, W. Chen, Y. Chen, D. Liu, et al., Adaptive Kalman filtering for vehicle navigation, J. Glob. Position. Syst. (2003) 42–47.
[38] M. Bennis, M. Debbah, H.V. Poor, Ultra reliable and low-latency wireless communication: tail, risk, and scale, Proc. IEEE 106 (10) (2018) 1834–1853.

[39] H. Levy, H.M. Markowitz, Approximating expected utility by a function of mean and variance, Am Econ Rev 69 (3) (1979) 308–317.
[40] T. Wysocki, A. Jamalipour, Mean-variance based QoS management in cognitive radio, IEEE Trans. Wireless Commun. 9 (10) (2010) 3281–3289.
[41] H.M. Markowitz, G.P. Todd, Mean-variance Analysis in Portfolio Choice and Capital Markets, 66, John Wiley & Sons, 2000.
[42] Emlid, Reach RS2, 2021, (https://emlid.com/reachrs2/) Retrieved June 28, 2021.
[43] D. Fox, S. Thrun, W. Burgard, F. Dellaert, Particle filters for mobile robot localization, in: Sequential Monte Carlo Methods in Practice, Springer, 2001, pp. 401–428.
[44] S.M. LaValle, Planning Algorithms, Cambridge university press, 2006.
[45] M. Narasimhan, J. Bilmes, A submodular-supermodular procedure with applications to discriminative structure learning, in: Conference on Uncertainty in Artificial Intelligence, 2005, pp. 404–412.
[46] S. Boyd, S.P. Boyd, L. Vandenberghe, Convex Optimization, Cambridge university press, 2004.
[47] A. Krause, Submodular function optimization, (https://www.mathworks.com/matlabcentral/fileexchange/20504-submodular-function-optimization) Retrieved June 28, 2021.
[48] A.B. Gershman, N.D. Sidiropoulos, S. Shahbazpanahi, M. Bengtsson, B. Ottersten, Convex optimization-based beamforming, IEEE Signal Process. Mag. 27 (3) (2010) 62–75.
[49] M. Grant, S. Boyd, CVX: Matlab software for disciplined convex programming, version 2.1, 2014, (http://cvxr.com/cvx).