# Joint Position and Beamforming Control via Alternating Nonlinear Least-Squares with a Hierarchical Gamma Prior

Anjaly Parayil, Amrit Singh Bedi, and Alec Koppel

*Abstract*— **We consider the problem of controlling antennae gains and positions among a set collection of mobile beamforming agents. Existing approaches predominately fall into two categories: solvers based upon convex relaxations of subset selection, and Monte Carlo sampling approaches that seek close-to-exact solutions, whose consistency requires the number of samples to approach infinity. In this work, we adopt an approach that improves upon the accuracy of prevailing convex relaxation approaches, motivated by their relative computational efficiency. Specifically, for fixed pose, we develop a modified hierarchical prior which is well-known within Bayesian inference to promote sparsity more effectively than the conventional Gaussian-gamma prior. Then, with this specification, we develop a variant of Expectation Maximization (EM) whose updates can be evaluated in closed form to obtain the beamforming gains and set of active agents. Then, when the signal phase and and amplitude are fixed, we propose a projected block descent approach, i.e., alternating nonlinear least-squares, for efficient relocation of the pruned set of agents. The inter-weaved iterative approach presented here better synthesizes the desired beam pattern with the minimum set of active agents and demands less computational load compared to the dense grid search implementations. Preliminary results indicate the proposed approach attains a superior tradeoff of sparsification and accuracy as compared to existing approaches.**

## I. Introduction

Beamforming is a fundamental capability for mobile autonomous agents equipped with directional antennae. The objective is to form a beam by selecting the antenna phase and amplitude to match a desired pattern. In of itself, this is a comple-valued non-convex optimization problem that arises in satellite communications, radars, biomedical imaging, acoustics, and remote sensing [1], [2]. Often, one would like to form a desired beam pattern with as few agents as possible, which is referred to as *sparse beamforming* [3]. In this work, we put forth a new technique for sparse beamforming and position control among a group of mobile agents that attains superior tradeoffs in computational effort and the parsimony of the number of transmitting agents. Our approach is based upon (i) a new regularizer derived from a hierarchical Bayesian hyper-parameter specification fused with Expectation Maximization (EM) to solve for the beamforming gains and cardinality of the set of active agents, which alternates with (ii) a position estimation step constructed from alternating nonlinear least squares, matrix projections, and logarithmic variable transformations.

Authors are with the U.S. Army Research Laboratory, Adelphi, MD, USA (e-mail: panjaly05@gmail.com, amrit0714@gmail.com, alec.e.koppel.civ@mail.mil).

Predominate existing approaches may be categorized into matrix pencil methods [4]–[6], global optimization schemes [7]–[9], solvers based upon convex relaxation [10]–[15], and Bayesian/Monte Carlo approaches [16]–[18]. Matrix pencil methods (MPM) offer a non-iterative scheme for the synthesis of sparse beamforming arrays with reduced computation time [4]–[6]. The approach organizes the associated data from the desired pattern in the form of the Hankel matrix and finds its lower rank approximate with singular value decomposition. Further, the matrix pencil method of [19] is applied to reconstruct the excitation and the position of the reduced set of antenna arrays. The approach is further extended to a forward-backward matrix pencil method (FBMPM) which partially mitigates imaginary solutions in the element locations [5]. The work in [6] proposes a multiple pattern synthesis version of [5] using enhanced unitary matrix pencils. Despite its effectiveness, MPM methods are offline approaches that offer less flexibility for any possible adaptation and are limited in terms of imposing user-defined geometric features on the generated beam.

Beamforming based on global optimization schemes are proposed in [7]–[9]. The works, [7]–[9] present iterative schemes based on modified real genetic algorithm, biogeography based optimization, and the population-based stochastic solve constrained optimization problem for optimal element positions. Though global optimization methods are generally efficient, unknown agents in sparse solutions add to the computational complexity and the time required.

Another class of work in this area solves the problem as a constrained convex optimization with an equivalent convex cost function. Though the problem of sparsity is in general NP-hard, [10]–[15] propose equivalent convex norms for the non-convex $\ell_0$ auxiliary cost functions. For instance, an iterative approach based on $\ell_1$ norm is presented in [10]. The approach therein eliminates the non-convex nature of the constraints, however, reduces half of the design variables and tends to converge to suboptimal solutions [11]. This is again extended to address different geometric arrays using weighted $\ell_1$ optimization problem in [11]. Along similar lines, a mixed $\ell_{1,\infty}$-norm squared for group-sparsity and semidefinite relaxation is proposed in [12], [13]. Approaches in the literature such as [14], [15] focuses on mimicking behavior of $\ell_0$ norms with re-weighted $\ell_1$ approaches to get as close as possible to the $\ell_0$-norm.

Another line of research treats the problem within the Bayesian framework and better mimics $\ell_0$ norm [16]–[18], [20]–[22]. Although, the approaches discussed above can also be treated as an optimization in Bayesian framework with a fixed factorial prior, associated cost functions fail to be sufficiently sparse [22]. Auxiliary cost functions based on $\ell_1$ norms are uni-modal function which however may not be sufficiently sparse. On the other hand, sparse enough $\ell_2$ cost functions suffer from numerous local minima, and closeness to the global solution is heavily dependent on proper initialization [21], [22].

Inspired from fast relevance vector machine algorithm introduced in [23], [24], [16] presents pattern synthesis using maximally-sparse array with the highest a-posteriori probability and solves for agent weights. Herein, a hierarchical prior is introduced on the agent weights, weights corresponding to agents are integrated out and the hyperparameters are solved by maximizing the marginal likelihood. The approach inherently forces weights corresponding to most of the agents to zero while generating the desired pattern. Different choices of priors on weights are considered in [20]. An extension to complex non-Hermitian layouts with multi-task Bayesian compressive sensing theory in [17], [18].

Though, Bayesian learning guarantees maximally sparse solutions [25], the reconstruction error is bounded by the position of available beamforming arrays [26]. In other words, *sparse recovery* is accurate if the considered agent positions are near to the fictitious agents that were used to create the beam and the error deteriorated with the mismatch. To this effect, joint optimization of both position and weights are addressed with compressed sensing based approaches [26], [27] wherein, the candidate sparse element positions are constrained onto a set of discrete grid points for higher degrees of freedom. Such a process, however is sensitive to the initial setting of the grid and increases the computational load with dense grids. Further, approaches based on Taylor's first order approximations with less computational complexity would be erratic for higher mismatch in agent positions and results in slow convergence [28].

**Contributions** We propose an interweaved iterative approach to control the cardinality and position of beamforming agents with a better beam matching. We adopt the Bayesian framework for agent selection that offers maximal sparsity for a given agent layout whose performance, however, depends on the positioning error of available active agents. First, we introduce a modified prior as opposed to the conventional prior and forces high probability mass near to the null of beamforming weights, in turn, offers better shrinkage of the active agent set (Lemma 1). Iterative expressions for beamforming weights and hyperparameters are updated based on the proposed regularizer using evidence maximization. Another challenge for sparse beamforming is the efficient relocation of the pruned agent set for better beam matching without a high computational complexity. For a given set of beamforming weights, we exploit the underlying convex structure and propose an iterative projected block descent for efficient agent positioning. The approach results in superior beam matching with less computational complexity and efficient sparse recovery in comparison with the commonly used algorithms in the literature.

The paper is organized as follows: Section II introduces the problem of sparse beamforming. The proposed approach is introduced in Section III along with an introduction to the agent selection problem and the associated cost functions. Section IV discusses its performance and compares it with available algorithms. Concluding remarks are given in Section VI with additional discussion on derived expressions in Section VII.

## II. Problem Formulation

In this work, we focus on the physical communications problem of synthesizing a beam among a set of $n$ mobile agents. The location of agent $i$ is a planar quantity $[x_i,\ y_i] \in \mathbb{R}^2$. The agents are equipped with directional antennae, and the *array factor*, which determines their ability to communicate, is a complex-valued far-field radiation pattern whose closed form is given by [29]

$$\mathrm{AF}(\theta) = \sum_{i=0}^{n-1} a_i\, e^{j(\alpha_i + kx_i\cos(\theta) + ky_i\sin(\theta))} \tag{1}$$

where $a_i$ denotes the signal amplitude and $\alpha_i$ is its phase. The excitation of each element $(a_i \exp(j\alpha_i))$ can be represented by complex weights, $w_i \in \mathbb{C},\ \forall i = 1,\ldots,n,\ k$ is the wave number and $\theta,\ \alpha \in [0, 2\pi)$. The generalized array factor can be further expressed as

$$\mathrm{AF}(\theta) = \sum_{i=0}^{n-1} w_i\, e^{j(kx_i\cos(\theta) + ky_i\sin(\theta))} \tag{2}$$

With this physical entity defined, the main technical problem we consider in this work is how, given samples from an a priori unknown desired beam pattern ($\mathrm{AF}_d$) obtained along $\theta_i,\ \forall i = 1,\ldots,N$ directions, a minimum number of mobile agents are selected to match the desired beam at the given directions accurately as possible. More specifically, *sparse beamforming* may be encapsulated as the joint optimization of position $\mathbf{r}_i = [x_i,\ y_i]$ and complex weights, $w_i$ such that the generated beam is matched with the desired pattern at $N$ directions with minimum set of $n$ agents:

$$\min_{n} \min_{\{\mathbf{r}_i, w_i\}_{i=1}^{n}} \|\mathbf{AF} - \mathbf{AF}_d\|_2^2 . \tag{3}$$

Here, $\mathbf{AF}_d \in \mathbb{C}^N$ represents $N$ samples of the desired pattern and $\mathbf{AF} \in \mathbb{C}^N$ represents corresponding values of the beam created by $n$ agents at $\theta_i$. Moreover, the desired pattern $\mathbf{AF}_d \in \mathbb{C}^N$ is created by a set of $n_d$ "fictitious" agents, which is unknown.

Next, we elaborate how this optimization problem admits a formulation as a nonlinear least-squares over real-valued range when the agent positions are fixed. Specifically, we propose lifting the complex-valued entities, i.e., the array factor $\mathbf{AF} \in \mathbb{C}^N$ [cf. (2)] by defining the matrix $\mathbf{H}(\mathbf{r})$ and coefficient vector $\mathbf{w}$ as

$$\mathbf{AF} = \mathbf{H}(\mathbf{r})\mathbf{w} \tag{4}$$

$$[\mathbf{H}(\mathbf{r})]_{ml} = e^{jk(x_l\cos\theta_m + y_l\sin\theta_m)} \tag{5}$$

where $[\mathbf{H}(\mathbf{r})]_{ml}$ denotes the $(m,l)^{\text{th}}$ entry of the matrix $\mathbf{H}(\mathbf{r})$ with $m, l = 1,\cdots,N$. Next, to define a convex formulation with respect to the excitation of agents, by utilizing the definitions in (5), we consider matching the real and imaginary parts of the patterns, i.e., reformulating the objective of (3) as

$$\min_{n} \min_{\{\mathbf{r_i}, w_i\}_{i=1}^{n}} \left\| \mathbf{\Phi}(\mathbf{r})\tilde{\mathbf{w}} - \widetilde{\mathbf{AF}}_d \right\|_2^2 \tag{6}$$

where,

$$\mathbf{\Phi}(\mathbf{r}) = \begin{bmatrix} \mathcal{R}(\mathbf{H}(\mathbf{r})) & -\mathcal{I}(\mathbf{H}(\mathbf{r})) \\ \mathcal{I}(\mathbf{H}(\mathbf{r})) & \mathcal{R}(\mathbf{H}(\mathbf{r})) \end{bmatrix}, \ \widetilde{\mathbf{AF}}_d = \begin{bmatrix} \Re(\mathbf{AF}_d) \\ \Im(\mathbf{AF}_d) \end{bmatrix} \tag{7}$$

and $\tilde{\mathbf{w}} = [\Re(\mathbf{w}), \Im(\mathbf{w})]^\top$. Observe that the problem in (6) is a nonlinear least-squares problem when we ignores the subset selection aspect of (3) and fix the agent poses $[x_i, y_i]$. In this work, however, we consider the unified problem of selecting agent antennae gains (signal amplitude and phase), positions, and the *number of agents*. This makes the problem under consideration both non-convex and an integer program. In the next section, we develop an approach based upon Bayesian relaxation of the subset selection problem and alternating nonlinear least squares.

## III. Minimum Agent Selection and Existing Approaches

The problem of reducing cardinality of set of active agents for a given $\mathbf{\Phi}$ in (6) is equivalent to minimizing agents with non-zero $\tilde{\mathbf{w}}$ given by

$$\min_{\tilde{\mathbf{w}}} \|\tilde{\mathbf{w}}\|_0, \text{ s.t. } \widetilde{\mathbf{AF}}_d = \mathbf{\Phi}\tilde{\mathbf{w}}. \tag{8}$$

Here, $\ell_0$ norm represents *indicator function*, $\|\tilde{\mathbf{w}}\|_0 = \sum \mathcal{I}[|\tilde{w}_i| > 0]$. We note that in (8), the exact sparse recovery is NP-hard and different approximation methods have been proposed in literature to mimic the solution [10]–[15]. One of the most popular among them is the $\ell_1$ regularizer (LASSO) which can be conveniently solved for the global minimum with standard optimization methods [14]. However, the obtained global minimum does not necessarily coincide with the sparsest solutions (except in the special cases, [30]). On the other hand, approaches based on multi-modal $\ell_p$ norms converge to suboptimal local minima [31]. Due to these limitations, we resort to *Sparse Bayesian Learning* (SBL) based formulations for $\ell_0$ approximation which is shown to settle to the sparsest solution for a given $\mathbf{\Phi}$ (at least in the absence of noise) and possess fewer local minima compared to $\ell_p$ approaches [31]. A sharper $\tilde{\mathbf{w}}$ distribution along with a flat tail further reduces the set of active agents compared to conventional SBL formulations. However, this leads to undesired pruning of the active agents along with a deteriorated pattern matching which can be properly corrected with the adequate position control. The following subsection proposes a solution to this joint optimization problem with an inter-weaved iterative scheme based on SBL and second-order nonlinear least square method.

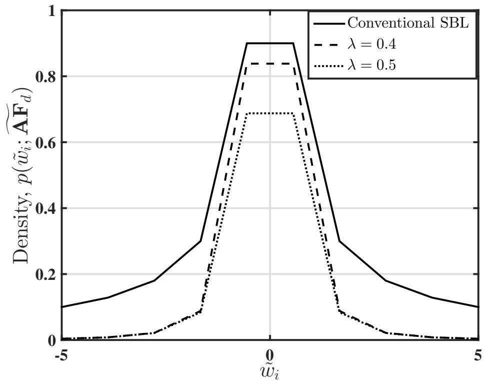



Fig. 1: Prior distribution of $\tilde{w}_i$ for the proposed approach, $p(\tilde{w}_i; \widetilde{\mathbf{AF}}_d) \propto (\frac{\lambda}{2} + \frac{\tilde{w}_i^2}{2})^{-3/2}$ for different $\lambda$ along with conventional SBL

### A. Bayesian Hyper-parameter Specification

We begin by introducing the probabilistic assumptions that give rise to the functional form of our augmented regularization. More specifically, SBL approaches to agent selection hinge upon a *Gaussian likelihood model* for beamforming with $N$ samples from the desired beam given by

$$p(\widetilde{\mathbf{AF}}_d)|\tilde{\mathbf{w}}) = (2\pi\sigma^2)^{-N/2} e^{-\frac{\|\widetilde{\mathbf{AF}}_d - \mathbf{\Phi}\tilde{\mathbf{w}}\|_2^2}{2\sigma^2}}. \tag{9}$$

The above expression is equivalent to beamforming error, $\epsilon = \|\widetilde{\mathrm{AF}}_d - \widetilde{\mathrm{AF}}\|_2$ being modelled as $\mathcal{N}(\epsilon; 0, \sigma^2\mathbf{I})$ [22]. The hierarchial approach to sparse recovery hinges upon specifying the prior distribution of $\tilde{\mathbf{w}}$ as $p(\tilde{\mathbf{w}}; \gamma) \sim \mathcal{N}(\tilde{\mathbf{w}}; 0, \mathrm{diag}[\gamma])$ with $\gamma \in \mathbb{R}^{2n}$ being the hyperparameter for the model. The *hierarchical prior* on $\tilde{\mathbf{w}}$ and $\gamma$ given by

$$\begin{aligned} p(\tilde{\mathbf{w}}|\gamma) &= \prod_{i=1}^{n} p(\tilde{w}_i|\gamma_i), \\ \text{where} \quad p(\tilde{w}_i|\gamma_i) &= \mathcal{N}(0, \gamma_i), \\ \text{and} \quad p(\gamma_i^{-1}) &= \gamma_i^{1-a} e^{-b/\gamma_i}. \end{aligned} \tag{10}$$

The prior on $\tilde{\mathbf{w}}$ and hyperparameter $\gamma$ operates as follows. When $\gamma_i$ is null, then the associated coefficient $\tilde{w}_i$ is null, which eliminates the corresponding agent, i.e., kicks it out of the set of active beamforming agents. We hypothesize that the noise prior $\sigma^2$ is either known or user-specified.

In contrast to standard SBL where an *uninformative prior* is used, i.e., $a$ and $b$ are zero in (10), we select $a = 1$ and $b = \lambda/2$. These selections result in a sharper distribution for $\tilde{\mathbf{w}}$ with flat tails as shown in Fig. 1 as opposed to the conventional case (solid curve in the figure).

Now, let's consider the *marginal log-likelihood* for $\gamma$, which may be obtained by marginalizing over $\tilde{\mathbf{w}}$

$$\begin{aligned} \mathcal{L}(\gamma) &= -2\log \int p(\widetilde{\mathrm{AF}}_d|\tilde{\mathbf{w}}) p(\tilde{\mathbf{w}}|\gamma) d\tilde{\mathbf{w}} p(\gamma) \\ &= \log|\mathbf{\Sigma}_{\widetilde{\mathrm{AF}}_d}| + \widetilde{\mathrm{AF}}_d^\top \mathbf{\Sigma}_{\widetilde{\mathrm{AF}}_d}^{-1} \widetilde{\mathrm{AF}}_d + \lambda \sum_i^n \gamma_j \end{aligned} \tag{11}$$

where, $\mathbf{\Sigma}_{\widetilde{\mathrm{AF}}_d} = \sigma^2\mathbf{I} + \mathbf{\Phi}\Gamma\mathbf{\Phi}^\top$, $\Gamma \triangleq \mathrm{diag}[\gamma]$. Maximizing the above expression, specifically, evidence maximization (*type-II maximum likelihood*) [23] allows us to estimate the hyperparameters, $\gamma$. In contrast to the conventional SBL, the last term in (11) is the additional summand which further incentivizes the reduction in the number of active beamforming agents. Doing so then yields iterative updates for the hyper-parameters which can be evaluated in closed form, as we formalize next.

*Lemma 1:* Expectation-maximization (EM) applied to the likelihood (11) can be evaluated with closed-form iterative updates for the hyperparameter $\gamma$ and agent weights $\tilde{\mathbf{w}}$ as

$$\mu = \tilde{\mathbf{w}} = \mathbb{E}\left[\tilde{\mathbf{w}}|\widetilde{\mathrm{AF}}_d, \gamma_*\right] = \Gamma\mathbf{\Phi}^\top\mathbf{\Sigma}^{-1}_{\widetilde{\mathrm{AF}}_d}\widetilde{\mathrm{AF}}_d \tag{12}$$

$$\mathbf{\Sigma} = \Gamma - \Gamma\mathbf{\Phi}^\top\mathbf{\Sigma}^{-1}_{\widetilde{\mathrm{AF}}_d}\mathbf{\Phi}\Gamma \tag{13}$$

$$\gamma_i = \frac{2\left(\mu_i^2 + \mathbf{\Sigma}_{ii}\right)}{1+\sqrt{1+4\lambda(\mu_i^2+\mathbf{\Sigma}_{ii})}}, \ \forall i = 1,\ldots,n. \tag{14}$$

*Proof:* See the appendix in Sec. VI-A. ■

Lemma 1 may be employed to iteratively find the optimal beamforming gain $\tilde{\mathbf{w}}$ and sparsity-inducing hyperparameter $\gamma$ when the agent positions are fixed. In the following subsection, we expand upon how one may improve agent positioning for beaming through a numerical search routine based upon nonlinear least-squares.

### B. Position Control

We consider an approximation of the constrained $\ell_0$ problem in (8) as

$$\begin{aligned} &\arg\min_{\mathbf{r},\mathbf{w}} \|\mathbf{w}\|_0, \\ &\text{s.t. } \|\mathbf{AF}_d - \mathbf{H}(\mathbf{r})\mathbf{w}\|_2^2 \leq \tilde{\epsilon} \end{aligned} \tag{15}$$

where we use the lifting of the complex-valued quantities via (4) and $\tilde{\epsilon}$ is the approximation introduced for the desired beam pattern. We note that for a fixed $\mathbf{w}$, the problem in (15) defines a quadratic program given by

$$\min_{\mathbf{H}_j} \|\mathbf{AF}_d - \underbrace{\begin{bmatrix}\mathbf{H}_1 & \ldots & \mathbf{H}_n\end{bmatrix}}_{\text{unknowns}, \ \mathbf{H}_j \in \mathbb{C}^N}\mathbf{w}\|_2^2 \tag{16}$$

which is convex in $\mathbf{H}_j$ for all $j = 1,\ldots,n$ and the problem in (16) is constraint satisfaction problem. Let us define $\mathbf{H} = [\mathbf{H}_1 \ldots, \mathbf{H}_n]$ and a convex set $\mathcal{C}_1$ as

$$\mathcal{C}_1 \triangleq \left\{\mathbf{H} \in \mathbb{R}^{N\times n} \text{s.t.} \forall\ j = 1,\ldots,n, \mathbf{H}_j^\top\mathbf{H}_j \leq N\right\} \tag{17}$$

which removes scaling ambiguity and constraints the feasible set of matrices $\mathbf{H}$ onto a convex set with $\ell_2$ norm of column matrices bounded by $N$. A nonlinear least-squares estimate of the above problem with a constrained solution space and logarithmic transformation yields the refined agent positions in an iterative manner. With the convex optimization problem of (16), we obtain a closed-form expression for the optimality condition by computing the gradient with respect to $\mathbf{H}$ given by

$$\nabla_{\mathbf{H}}\frac{1}{2}(\mathbf{Hw}-\mathbf{AF}_d)^\top(\mathbf{Hw}-\mathbf{AF}_d) \tag{18}$$

$$\begin{aligned} =&\frac{1}{2}\nabla_{\mathbf{H}}\mathrm{Tr}\left(\mathbf{w}^\top\mathbf{H}^\top\mathbf{Hw} - \mathbf{w}^\top\mathbf{H}^\top\mathbf{AF}_d\right. \\ &\left.-\mathbf{AF}_d^\top\mathbf{Hw} - \mathbf{AF}_d^\top\mathbf{AF}_d\right) \end{aligned} \tag{19}$$

$$=(\mathbf{Hww}^\top - \mathbf{AF}_d\mathbf{w}^\top). \tag{20}$$

Note that the Hessian of (16) with respect to $\mathbf{H}$ is written as $\mathbf{ww}^\top$. A sequential least square update of $\mathbf{H}_j$ at $i+1$-th iteration can be expressed as

$$\begin{aligned} &\mathbf{H}_j(i+1) = \mathbf{H}_j(i) + \mathrm{diag}(\mathbf{w}(j)\mathbf{w}(j)^\top)^{-1} \\ &(\mathbf{AF}_d\mathbf{w}(j)^\top - \mathbf{H}(i)\mathbf{ww}(j)^\top), \ \forall j = 1,\ldots,n \end{aligned} \tag{21}$$

$$\mathbf{H}_j(i+1) \leftarrow \frac{\mathbf{H}_j(i+1)}{\max(\|\mathbf{H}_j\|_2, N)} \tag{22}$$

The update in (22) defines the projection of estimated $\mathbf{H}_j$ onto $\mathcal{C}_1$ and removes scaling ambiguity. The position of $j$-th mobile agent $[x_j^*,\ y_j^*]$ based on the estimated $\mathbf{H}_j$ is given by

$$\begin{bmatrix}x_j^*\\ y_j^*\end{bmatrix} = \frac{1}{k}\left[(\mathbf{R}_\theta^\top\mathbf{R}_\theta)^{-1}\mathbf{R}_\theta^\top\right](\mathbf{H}_R) \tag{23}$$

where,

$$\begin{aligned} [\mathbf{R}_\theta]_{i,:} =& [\cos\theta_i \quad \sin\theta_i] \\ \mathbf{H}_R =& \frac{1}{k}\left[\mathcal{I}(\log\mathbf{H}_j)\right] = \mathbf{R}_\theta\begin{bmatrix}x_j^*\\ y_j^*\end{bmatrix}. \end{aligned} \tag{24}$$

In (24), $[\mathbf{R}_\theta]_{i,:}$ denotes the $i^{\text{th}}$ row with $i = 1,\cdots,N$ and all columns. The elements of column matrix, $\mathbf{H}_j$ should lie on a unit circle with a phase shift depending on the corresponding sampling direction, $\theta_i$ and the estimated $(\mathbf{x}_j^*, \mathbf{y}_j^*)$:

$$e^{jk*\left\langle \mathbf{r}_j^*, \left(\sum_{i=1}^N\begin{bmatrix}\cos\theta_i\\ \sin\theta_i\end{bmatrix}\right)\right\rangle} = \prod_{i=1}^N \mathbf{H}(i,j), \ \forall j = 1,\ldots,n, \tag{25}$$

$\mathcal{C}_2$ defines the corresponding space

$$\begin{aligned} \mathcal{C}_2 \triangleq \Big\{&\mathbf{H} \in \mathbb{R}^{N\times n}\text{s.t.}\forall\ j = 1,\ldots,n, \\ &\exists\mathbf{r}_j^*, e^{jk*\langle\mathbf{r}_j^*, (\mathbf{d}_{\theta_1}+\ldots+\mathbf{d}_{\theta_N})\rangle}\prod_{i=1}^N\mathbf{H}(i,j)\Big\}. \end{aligned} \tag{26}$$

Here, $\mathbf{d}_{\theta_j}$ represents unit vector, $\left(\begin{bmatrix}\cos\theta_j\\ \sin\theta_j\end{bmatrix}\right)$ along the direction of $j$-th sample. The above constraint forces each elements of the column vector, $\mathbf{H}_j$ onto a unit circle consistent with $x_j^*$ and $y_j^*$ and is given by

$$\mathbf{H}_j(i+1) \leftarrow \Pi_{\mathcal{C}_2}\left[\mathbf{H}_j(i+1)\right], \ \forall j = 1,\ldots,n, \tag{27}$$

where, $\Pi_{\mathcal{C}_2}$ represents orthogonal projection on the set $\mathcal{C}_2$ corresponding to $[x_i^*,\ y_i^*]$.

We summarize the idea for the joint positioning and beamforming control in Algorithm 1. We start with a set of agents at a given location and solve for the number of agents

**Algorithm 1** Joint positioning and beamforming control

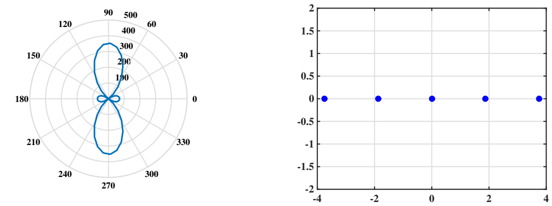
**procedure** REQUIRE( $N$ samples, $\widetilde{\mathbf{AF}}_d$, set of $n$ ($n > N+1$) agents at $\mathbf{r}$)
  **Initialize** : $\lambda = 1$, $\Gamma = \mathbf{I}_{n\times n}$, and $\tilde{\mathbf{w}}$
  Obtain $\mathbf{H}$ and $\mathbf{AF}$ from (4) - (5)
  **while** $\|\mathbf{AF}_d - \mathbf{AF}\|_2 \geq \tilde{\epsilon}$ **do**
    Obtain $\mathbf{\Phi}$ and $\widetilde{\mathbf{AF}}$ from (7)
    **while** $\epsilon$ has not converged **do** ▷ sparse recovery for a given agent layout
      compute $\mathbf{\Sigma}$, $\tilde{\mathbf{w}}$ using (12),
      update $\gamma_i$, $\forall i = 1, \ldots, n$ using (14)
    $\tilde{\mathbf{w}} \leftarrow \mathbb{E}\left[\tilde{\mathbf{w}}|\widetilde{\mathrm{AF}}_d; \gamma_*\right] = \Gamma_* \Phi^\top \mathbf{\Sigma}^{-1}_{\widetilde{\mathrm{AF}}_d} \widetilde{\mathrm{AF}}_d$
    Remove agents with $\tilde{\mathbf{w}} = 0$, update agent set and cardinality, $n$
    **while** $\mathbf{H}_j$ not converged **do** ▷ Position control for given sparse set
      **for** j=1 to n **do**
        Compute constrained least square estimate, $\mathbf{H}_j$ using (21) - (22)
        Calculate ($x_j^*$, $y_j^*$) using (23)
        Update $\mathbf{H}_j$ using (27)
    Update $\mathbf{r}$
**Return:** $\tilde{\mathbf{w}}$ **and** $\mathbf{r}$

and corresponding excitation with evidence maximization of the proposed marginal log-likelihood. The iterative updates of (12)-(14) modify the excitation of agents and reduce the search space by removing inactive agents. Further, we refine the positioning of the pruned agent set for a better beam matching. With the underlying convex structure of the problem along with combined nonlinear least-square estimates, matrix set projections, and logarithmic transformations, (21)-(27) refine the location of the active beamforming agents.

## IV. Simulations and Discussions

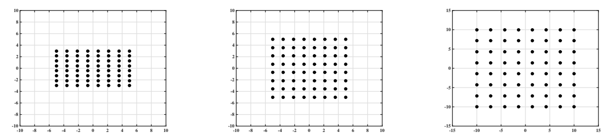
In this section, we evaluate the performance of Algorithm 1 and compare its merits against other available techniques in the literature to solve the similar problem.

For the purpose of simulations, we consider a set of 50 samples from the desired pattern, $\mathbf{AF}_d \in \mathbb{C}^N$ created by a group of fictitious agents as given in Fig. 2(a). The fictitious agents assumed to be in an equally spaced beamforming array centered around the origin as shown in Fig. 2(b), transmits at 40 MHz with $\alpha_m = \pi/4$, $a_m = 100$ for all $m = 1, \ldots, 5$. Numerical simulations consider three different cases, ranging from available sets of agents in the proximity of fictitious agents to the case with distant positioning from fictitious agents. The three different layouts with available set of 64 agents distributed along rectangular grids are shown in Figs. 3(a), 3(b), and 3(c). The performance of the algorithm for sparse beamforming is compared in terms of its accuracy and sparsification with existing approaches of Sparse Bayesian Learning (SBL) [16], reweighted-$\ell_2$ [32], reweighted-$\ell_2$ based on sparse Bayesian learning [33], $\ell_1$ with FOCal Underdetermined System Solver (Focuss) [34]




(a) (b)

Fig. 2: a) Desired Beam b) Set of fictitious agents used for desired beam



(a) (b) (c)

Fig. 3: Initial agent layouts considered for experiments

and Perturbed Compressed Sensing (PCS) [26]. Number of agents selected by the algorithms are given in Fig. 4(a) with corresponding errors, $\epsilon = \|\widetilde{\mathrm{AF}}_d - \widetilde{\mathrm{AF}}\|_2$ for all the cases are depicted in Fig. 4(b). Modified SBL of (12) - (14) , i.e., the proposed approach without including position refinement obtain a better parsimonious collection of active beamforming agents compared to conventional SBL, however, the additional shrinkage in $\tilde{\mathbf{w}}$ imposed by (11) is at the cost of deteriorated beam matching. On the other hand, the proposed approach of Algorithm 1 results in better sparse recovery and accuracy as shown in Figs. 4(a) and 4(b). In short, the approach provides more parsimonious collections of mobile beamforming agents for a fixed accuracy as described in Fig. 4(a). Similarly, for a given number of agents (excluding the Monte-Carlo approach with higher computational complexity [35]), the proposed approach provides better accuracy as shown in Fig. 4(b).

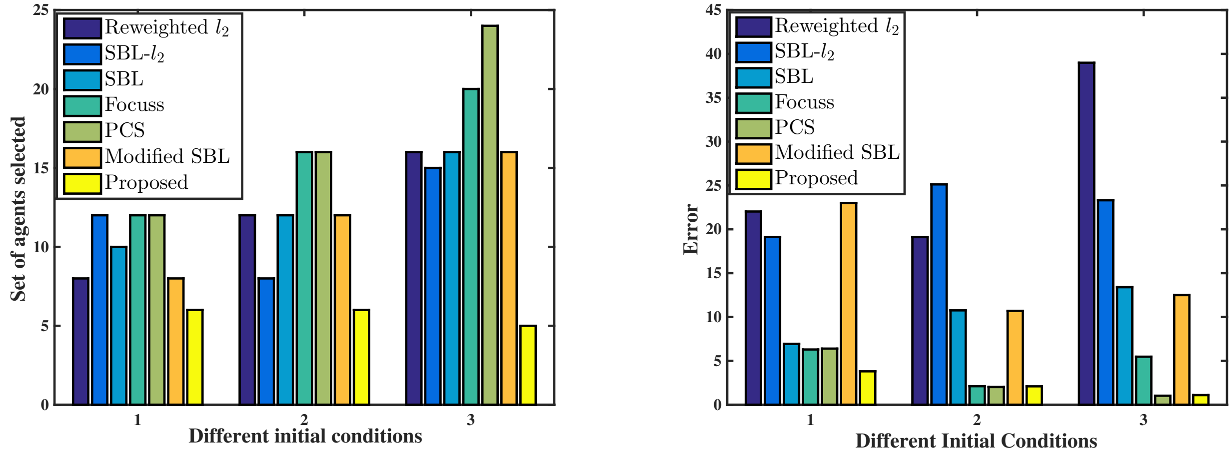
## V. Conclusion

In this work, we proposed a joint optimization approach based on modified hierarchical prior and alternating block coordinate descent to control the weights and efficient relocation of beamforming agents for accurate beam matching in the given directions. Experiment results demonstrated superior performance with better sparse recovery and accuracy as compared to the existing approaches. Future research directions involve the development of adaptive sparsification logic with selective pruning based on the agent weights, group sparsity, and inherent mutual coupling among beamforming agents.

## VI. Appendix

### A. *Proof of Lemma 1*

Consider the likelihood for the array factor (9) and apply Bayes' Rule to obtain the posterior distribution over the




(a) (b)

Fig. 4: Sparse Beamforming with different algorithms: a) Cardinality of active beamforming sets b) Achieved accuracy. The proposed technique provides the best tradeoff between accuracy and the sparsity of the solution.

weights, $\tilde{\mathbf{w}}$ can be expressed as

$$p(\tilde{\mathbf{w}}|\widetilde{\mathrm{AF}}_d,\gamma)=\frac{p(\widetilde{\mathrm{AF}}_d|\tilde{\mathbf{w}})p(\tilde{\mathbf{w}}|\gamma)p(\gamma)}{p(\widetilde{\mathrm{AF}}_d|\gamma)} \tag{28}$$

Observe that the denominator in the preceding expression is a *convolution* of Gaussians, and hence reduces to a Gaussian distribution:

$$p(\tilde{\mathbf{w}}|\widetilde{\mathrm{AF}}_d,\gamma)=\mathcal{N}(\mu,\mathbf{\Sigma}) \tag{29}$$

where the mean and covariance parameters may be evaluated using standard rules for transformation of Gaussian distributions:

$$\mu=\Gamma\mathbf{\Phi}^\top\left(\lambda_2\mathbf{I}+\mathbf{\Phi}\Gamma\mathbf{\Phi}^\top\right)^{-1}\widetilde{\mathrm{AF}}_d \tag{30}$$

$$\mathbf{\Sigma}=\Gamma-\Gamma\Phi^\top\left(\lambda_2\mathbf{I}+\mathbf{\Phi}\Gamma\mathbf{\Phi}^\top\right)^{-1}\mathbf{\Phi}\Gamma \tag{31}$$

These expressions for the mean and covariance are the as stated in Lemma 1, specifically, (12) - (13).

Now, we proceed to derive *Expected Maximization* updates for $\gamma$ (14), inspired by [23]. To do so, we first require, the *marginal log-likelihood* for $\gamma$. To obtain this likelihood, proceed then by integrating out the weights $\tilde{\mathbf{w}}$ from the *log-likelihood* [cf. (29)] as

$$\mathcal{L}=\mathbb{E}_{\tilde{\mathbf{w}}|\widetilde{\mathrm{AF}}_d,\gamma}\left[\log p(\widetilde{\mathrm{AF}}_d|\tilde{\mathbf{w}})+p(\tilde{\mathbf{w}}|\gamma)+\log p(\gamma)\right] \tag{32}$$

Using second moment of *Gaussian*, $\mathbb{E}_{\tilde{\mathbf{w}}|\widetilde{\mathrm{AF}}_d,\gamma}\left[\tilde{w}_i^2\right]=\mathbf{\Sigma}_{ii}+\mu_i^2$ and removing the terms independent of $\gamma$, the above expression simplifies to

$$\mathcal{L}=-\frac{1}{2}\left(\sum\log\gamma_i+\sum\frac{\mathbf{\Sigma}_{ii}+\mu_i^2}{\gamma_i}+\sum\frac{\lambda\gamma_i}{2}\right), \tag{33}$$

where we have substituted in choices $a=1$ and $b=\lambda/2$ for the prior in (14) to write (33) Now the maximization step, in other words, derivative of the above expression with respect to $\gamma_i$ yields

$$\gamma_i=\frac{2\left(\mu_i^2+\mathbf{\Sigma}_{ii}\right)}{1+\sqrt{1+4\lambda\left(\mu_i^2+\mathbf{\Sigma}_{ii}\right)}} \tag{34}$$

■

## REFERENCES

[1] G. Krieger, N. Gebert, and A. Moreira, "Multidimensional waveform encoding: A new digital beamforming technique for synthetic aperture radar remote sensing," *IIEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 1, pp. 31–46, 2007.

[2] M. Hawkes and A. Nehorai, "Acoustic vector-sensor beamforming and capon direction estimation," *IEEE Trans. Signal Process.*, vol. 46, no. 9, pp. 2291–2304, 1998.

[3] M. Richard and D. Brian, "Design of maximally sparse beamforming arrays," *IEEE Trans. Antennas Propag.*, vol. 39, no. 8, 1991.

[4] Y. Liu, Z. Nie, and Q. H. Liu, "Reducing the number of elements in a linear antenna array by the matrix pencil method," *IEEE Trans. Antennas Propag.*, vol. 56, no. 9, pp. 2955–2962, 2008.

[5] Y. Liu, Q. H. Liu, and Z. Nie, "Reducing the number of elements in the synthesis of shaped-beam patterns by the forward-backward matrix pencil method," *IEEE Trans. Antennas Propag.*, vol. 58, no. 2, pp. 604–608, 2010.

[6] H. Shen and B. Wang, "An effective method for synthesizing multiple-pattern linear arrays with a reduced number of antenna elements," *IEEE Trans. Antennas Propag.*, vol. 65, no. 5, pp. 2358–2366, 2017.

[7] K. Chen, Z. He, and C. Han, "A modified real GA for the sparse linear array synthesis with multiple constraints," *IEEE Trans. Antennas Propag.*, vol. 54, no. 7, pp. 2169–2173, 2006.

[8] S. K. Goudos, K. Siakavara, T. Samaras, E. E. Vafiadis, and J. N. Sahalos, "Sparse linear array synthesis with multiple constraints using differential evolution with strategy adaptation," *IEEE Antennas Wireless Propag. Lett.*, vol. 10, pp. 670–673, 2011.

[9] S. Goudos, K. Baltzis, K. Siakavara, T. Samaras, E. Vafiadis, and J. Sahalos, "Reducing the number of elements in linear arrays using biogeography-based optimization," in *IEEE 6th EUCAP*, 2012, pp. 1615–1618.

[10] S. E. Nai, W. Ser, Z. L. Yu, and H. Chen, "Beampattern synthesis for linear and planar arrays with antenna selection by convex optimization," *IEEE Trans. Antennas Propag.*, vol. 58, no. 12, pp. 3923–3930, 2010.

[11] B. Fuchs, "Synthesis of sparse arrays with focused or shaped beampattern via sequential convex optimizations," *IEEE Trans. Antennas Propag.*, vol. 60, no. 7, pp. 3499–3503, 2012.

[12] O. Mehanna, N. D. Sidiropoulos, and G. B. Giannakis, "Multicast beamforming with antenna selection," in *IEEE 13th SPAWC*, 2012, pp. 70–74.

[13] ——, "Joint multicast beamforming and antenna selection," *IEEE Trans. Signal Process.*, vol. 61, no. 10, pp. 2660–2674, 2013.

[14] B. Fuchs, A. Skrivervik, and J. R. Mosig, "Shaped beam synthesis of arrays via sequential convex optimizations," *IEEE Antennas Wireless Propag. Lett.*, vol. 12, pp. 1049–1052, 2013.

[15] X. Liu, F. Hu, W. Xiong, Y. Du, and G. Guo, "Cuboid sparse array synthesis for sensor selection by convex optimization with constrained beam pattern based on wban," in *IEEE WCSP*, 2015, pp. 1–5.

[16] G. Oliveri and A. Massa, "Bayesian compressive sampling for pattern synthesis with maximally sparse non-uniform linear arrays," *IEEE Trans. Antennas Propag.*, vol. 59, no. 2, pp. 467–481, 2010.

[17] G. Oliveri, M. Carlin, and A. Massa, "Complex-weight sparse linear array synthesis by bayesian compressive sampling," *IEEE Trans. Antennas Propag.*, vol. 60, no. 5, pp. 2309–2326, 2012.

[18] F. Viani, G. Oliveri, and A. Massa, "Compressive sensing pattern matching techniques for synthesizing planar sparse arrays," *IEEE Trans. Antennas Propag.*, vol. 61, no. 9, pp. 4577–4587, 2013.

[19] Y. Hua and T. K. Sarkar, "Matrix pencil method for estimating parameters of exponentially damped/undamped sinusoids in noise," *IEEE Trans. Acoustics Speech Signal Process.*, vol. 38, no. 5, pp. 814–824, 1990.

[20] W. Zhang, L. Li, and F. Li, "Reducing the number of elements in linear and planar antenna arrays with sparseness constrained optimization," *IEEE Trans. Antennas Propag.*, vol. 59, no. 8, pp. 3106–3111, 2011.

[21] D. P. Wipf and B. D. Rao, "An empirical bayesian strategy for solving the simultaneous sparse approximation problem," *IEEE Trans. Signal Process.*, vol. 55, no. 7, pp. 3704–3716, 2007.

[22] D. P. Wipf, "Bayesian methods for finding sparse representations," Ph.D. dissertation, UC San Diego, 2006.

[23] M. E. Tipping, "Sparse bayesian learning and the relevance vector machine," *J. Mach. Learn Res.*, vol. 1, no. June, pp. 211–244, 2001.

[24] M. E. Tipping, A. C. Faul *et al.*, "Fast marginal likelihood maximisation for sparse bayesian models." in *AISTATS*, 2003.

[25] D. P. Wipf and S. S. Nagarajan, "A new view of automatic relevance determination," in *Advances in neural information processing systems*, 2008, pp. 1625–1632.

[26] F. Yan, P. Yang, F. Yang, and T. Dong, "Synthesis of planar sparse arrays by perturbed compressive sampling framework," *IET Microwaves, Antennas & Propagation*, vol. 10, no. 11, pp. 1146–1153, 2016.

[27] G. Prisco and M. D$'$ Urso, "Maximally sparse arrays via sequential convex optimizations," *IEEE Antennas Wireless Propag. Lett.*, vol. 11, pp. 192–195, 2012.

[28] F. Yan, P. Yang, F. Yang, L. Zhou, and M. Gao, "An alternating iterative algorithm for the synthesis of complex-excitation and pattern reconfigurable planar sparse array," *Signal Processing*, vol. 135, pp. 179–187, 2017.

[29] C. A. Balanis, *Antenna theory: analysis and design*. John wiley & sons, 2016.

[30] D. L. Donoho and M. Elad, "Optimally sparse representation in general (nonorthogonal) dictionaries via $\ell_2$ minimization," *Proc. of the National Academy of Sciences*, vol. 100, no. 5, pp. 2197–2202, 2003.

[31] Z. Tan, P. Yang, and A. Nehorai, "Joint sparse recovery method for compressed sensing with structured dictionary mismatches," *IEEE Trans. Signal Process.*, vol. 62, no. 19, pp. 4997–5008, 2014.

[32] R. Chartrand and W. Yin, "Iteratively reweighted algorithms for compressive sensing," in *2008 IEEE ICASSP*, 2008, pp. 3869–3872.

[33] D. Wipf and S. Nagarajan, "Iterative reweighted $\ell_1$ and $\ell_2$ methods for finding sparse solutions," *IEEE J. Sel. Topics Signal Process.*, vol. 4, no. 2, pp. 317–329, 2010.

[34] S. F. Cotter, B. D. Rao, K. Engan, and K. Kreutz-Delgado, "Sparse solutions to linear inverse problems with multiple measurement vectors," *IEEE Trans. Signal Process.*, vol. 53, no. 7, pp. 2477–2488, 2005.

[35] T. Park and G. Casella, "The bayesian lasso," *J. American Statistical Association*, vol. 103, no. 482, pp. 681–686, 2008.