Separable Multidimensional Orthogonal Matching Pursuit and its Application to Joint Localization and Communication at mmWave

Joan Palacios and Nuria González-Prelcic
Electrical and Computer Engineering Department, North Carolina State University
Email: {jbeltra, ngprelcic}@ncsu.edu

Abstract—Greedy sparse recovery has become a popular tool in many applications, although its complexity is still prohibitive when large sparsifying dictionaries or sensing matrices have to be exploited. In this paper, we start by formulating a new class of sparse recovery problems that exploit multidimensional dictionaries and the separability of the measurement matrices that appear in certain applications. Then, we develop a new algorithm, Separable Multidimensional Orthogonal Matching Pursuit (SMOMP), which can solve this class of problems with low complexity. Finally, we apply SMOMP to the problem of joint localization and communication at mmWave, and numerically show its effectiveness to provide, at a reasonable complexity, high accuracy channel and position estimations.

Index Terms—Separable sparse recovery, multidimensional orthogonal matching pursuit, joint localization and communication, millimeter wave MIMO, sparse channel estimation

I. INTRODUCTION

Greedy algorithms for sparse recovery have become a popular tool for the reconstruction of the sparse signals that appear in applications such as compressive imaging, spectrum sensing, or channel estimation for massive MIMO or millimeter wave (mmWave) MIMO [1]. In particular, orthogonal matching pursuit (OMP) [2], [3] and simultaneous orthogonal matching pursuit (SOMP) [4], [5], have been extensively exploited in many applications where low complexity is desired.

The problem of channel estimation at mmWave exploiting a hybrid MIMO architecture, and more recently the problem of joint localization and communication, have been formulated and solved exploiting OMP, SOMP or some related variations of this algorithmic approach [6]–[12]. These strategies suffer, however, from complexity limitations that come from two different aspects: 1) the exploitation of a very large sparsifying dictionary built from Kronecker products of the dictionaries for the direction of departure (DoD), the direction of arrival (DoA) and the delay; 2) the large dimensionality of the measurement matrix representing the channel sounding process, which comes from the large arrays exploited at mmWave. Note that for the joint localization and communication problem, it is especially important to operate with large dictionaries, since a high resolution of the angular and delay dictionaries are required for precise localization.

Separable sparse recovery methods proposed in prior work usually exploit Kronecker dictionaries [13]–[17] to reduce complexity. For example, tensor-OMP [13] computes a block-sparse representation of a tensor with respect to a Kronecker basis. This is an interesting solution when the observation (received signal) can be written as \( \mathbf{o} = \mathbf{Ψ}_s \), being \( \mathbf{Ψ} \) a sparsifying basis with a Kronecker structure. For the application to joint localization and communication at mmWave, a different setting arises, since the observation (received signal) depends also on a sensing matrix \( \mathbf{Φ} \), i.e. \( \mathbf{o} = \mathbf{Φ}_s \mathbf{Ψ}_s \), which destroys the Kronecker structure of the equivalent measurement matrix \( \mathbf{Φ}_s \mathbf{Ψ}_s \). In this setting, previous separable sparse recovery algorithms cannot be applied.

The Multidimensional Orthogonal Matching Pursuit (MOMP) algorithm [18] has been recently proposed to reduce the complexity of greedy approaches based on high resolution dictionaries, without requiring any Kronecker structure in the equivalent measurement matrix. MOMP splits the projection step of OMP into multiple, much simpler iterations, so it can be performed for each dictionary separately. This significantly reduces the complexity of the projection process, which is the most computationally expensive when working with large dictionaries. Nonetheless, the MOMP formulation does not tackle the problem of having a large measurement matrix when the signal representation has a high dimensionality. For example, when applying MOMP to joint localization and communication at mmWave [18], [19], increasing the number of antennas or frequency carriers heavily increases the dimension of the measurement matrix, so it is not possible to store it in an average computer memory or run the algorithm.

In this paper, we propose first an alternative algorithm to MOMP which does not require the explicit construction of the measurement matrix. Our new approach, Separable Multidimensional Orthogonal Matching Pursuit (SMOMP), builds upon the MOMP formulation, not only simplifying the computational complexity of the projection step when the number of dictionary elements increases, but also reducing complexity when the measurement matrix is separable. Then, we formulate the problem of joint localization and communication at mmWave as a sparse reconstruction problem.
with multidimensional dictionaries and multiple measurement matrices, such that it can be solved with SMOMP. Numerical results show how SMOMP can provide channel estimates that lead to high accuracy positioning in practical setups for mmWave MIMO, where complexity and memory requirements prevent the execution of other greedy approaches.

A. Notation

Throughout the paper, \(x, x, X, \mathcal{X}\) will be the styles for scalar, vector, matrix or tensor and set. For a matrix \(X\), \([X]_{a,:}\) and \([X]_{1:b}\) are respectively, the \(a\)-th row and the \(b\)-th column, this notation is extended to tensors with multi-index like \([X]_{a_1,a_2,b}\) for \(a = [a_1, a_2]\). The operator \(\|x\|, \|X\|\) to denote the Euclidean and Frobenius norms. To work with a more compact expression of the dimensions of some particular tensors, we denote \(X \in \mathbb{C}^{N_1 \times N_2 \times \ldots \times N_L}\) as \(X \in \mathbb{C}^{\otimes_{l=1}^{L}N_k}, k = 1, \ldots, L\).

II. SEPARABLE MULTIDIMENSIONAL ORTHOGONAL MATCHING PURSUIT

A. Background

OMP is a generalization of OMP when working with multiple independent dictionaries [18], [19]. To define the sparse recovery problem to be solved with MOMP, let us assume first that we have \(N_D\) dictionaries, with the \(k\)-th dictionary \(\Psi_k \in \mathbb{C}^{N_i \times N_i^n}\) consisting of \(N_i^n\) atoms in \(\mathbb{C}^{N_i^n}\). Then we define the coefficients of the sparse signal in the set of dictionaries, i.e. \(\mathcal{C} \in \mathbb{C}^{N_1 \times \ldots \times N_D \times N_D^n}, \) with \(N_m\) the number of measurements. We also need to write the measurement matrix as a tensor, i.e., \(\mathcal{F} \in \mathbb{C}^{N_0 \times N_1 \times \ldots \times N_D, N_D^n}\), with \(N_0^n\) the number of entries in each measurement. Finally, we define the set of entry coordinate combinations \(\mathcal{I} = \{i = (i_1, \ldots, i_{N_D}) \in \mathbb{N}^{N_D} \text{ s.t. } i_k \leq N_i^n, \forall k \leq N_D\}\), and the set of dictionary index combinations \(\mathcal{F} = \{j = (j_1, \ldots, j_{N_D}) \in \mathbb{N}^{N_D} \text{ s.t. } j_k \leq N_i^n, \forall k \leq N_D\}\) to cycle over each dictionary atom entry index and dictionary atom index, respectively. With all these definitions, the equivalent multidimensional matching projection problem can now be formulated as

\[
\min_{\mathcal{C}} \left\| \mathcal{O} - \sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{F}} \mathcal{F}_{i,j} \left( \prod_{k=1}^{N_D} \Psi_k^{j_k} \right) \mathcal{C}_{j,i} \right\|^2, \tag{1}
\]

where \(\mathcal{O} \in \mathbb{C}^{N_0 \times N_m}\) is the observation matrix. In this version of the problem, the sparsity condition is applied to the set \(\mathcal{C} \subset \mathcal{F}\) of index \(j \in \mathcal{F}\) such that \(\left\| \mathcal{C}_{j,i} \right\| > 0\). It is proven in [18] that the multi-dimensional matching pursuit problem is an extension of the original matching pursuit formulation with multiple dictionaries, which can be solved with MOMP.

B. Problem statement

The memory requirements and complexity of operating with the measurement tensor \(\mathcal{F}\) in (1) are high when the number of dictionaries \(N_D\), atom sizes \(N_k^n\), and observation length \(N_0^n\) increase, since the total number of elements in \(\mathcal{F}\) is given by \(N_0^n \prod_{k=1}^{N_D} N_k^n\). In some problems, it is possible, however, to split this tensor into multiple ones and create independent measurements. We will formulate now the matching pursuit problem with multidimensional dictionaries for the cases in which it is possible to separate \(\mathcal{F}\) into multiple tensors. First, we will need to define an additional index \(f \leq N_F\) to cycle over the different \(\Phi_f\) tensors. This means that for each \(\Phi_f\), we will need its own collection of \(N_D^f\) dictionaries, with the \(k\)-th dictionary \(\Phi_{f,k} \in \mathbb{C}^{N_i \times N_i^n}\) consisting of \(N_i^n\) atoms in \(\mathbb{C}^{N_i^n}\). This results in \(\Phi_f \in \mathbb{C}^{N_i \otimes_{k=1}^{N_D^f} N_i^n}\), and consequently, we need to define the observation as \(\mathcal{O} \in \mathbb{C}^{N_0^n \otimes_{f=1}^{N_F} N_F^n \times N_m}\). For simplicity, we define a set of observation index combinations \(\mathcal{O} = \{o = (o_1, \ldots, o_{N_F}) \in \mathbb{N}^{N_F} \text{ s.t. } o_f \leq N_F^n \forall f \leq N_F\}\). We keep in a single tensor the coefficients of the sparse signal in the multiple sets of dictionaries and we denote it as \(\mathcal{C} \in \mathbb{C}^{N_0^n \otimes_{f=1}^{N_F} N_F^n \times N_m}\). We define the index sets \(I_f = \{i_f = (i_{f,1}, \ldots, i_{f,N_F}) \in \otimes_{k=1}^{N_D^f} N_i^n\}, J_f = \{j_f = (j_{f,1}, \ldots, j_{f,N_F}) \in \otimes_{k=1}^{N_D^f} N_k^n\}\) and \(J = \{j = (j_1, \ldots, j_{N_F}) \in \otimes_{f=1}^{N_F} J_f\}\). The equivalent separable multidimensional matching projection problem is now defined as the minimization over \(\mathcal{C}\) of

\[
\sum_{o \in \mathcal{O}} \left\| \mathcal{O}_{o} - \sum_{i \in I_f, j \in J_f} \prod_{k=1}^{N_D^f} \Phi_{f,k}^{j_k} \left( \mathcal{C}_{j_f, o} \right) \right\|^2. \tag{2}
\]

The interaction between the different elements in this expression is illustrated with a toy example in Fig. 1.

To prove the equivalence with the MOMP formulation, we compress \(\mathcal{O}\) as a simple index \(\mathcal{O} = \sum_{f=1}^{N_F} o_f \prod_{f=1}^{f-1} N_f^n\), and group the \(f\) and \(k\) indices into the linear index \(k = k + \sum_{f=1}^{f-1} N_f^n\). The last indices to reshape are \(\bar{I} = I_f\) and \(\bar{J} = J_f\). We can also compute the parameters \(N_F^n = \prod_{f=1}^{N_F} N_f^n\), \(N_D = \sum_{f=1}^{N_F} N_D^f\), \(N_{i}^n = N_{i}^n\), and \(N_{x}^n = N_{x}^n\). This allows us to define the equivalent parameters as \(\mathcal{O}_{0,0} = \mathcal{O}_{0,0} = \prod_{f=1}^{N_F} \mathcal{O}_{0,0} \mathcal{F}_{f,k}^{j_f, k_f, k_f} \Psi_{f,k}\). Under these definitions, (2) can be rewritten as (1), proving them to be equivalent.

C. Separable multidimensional matching pursuit

The goal of this section is to describe SMOMP, a modification of MOMP to solve problem (2). To do so, we analyze the steps in MOMP and describe the necessary transformations to define the corresponding step in SMOMP.

Like every OMP algorithm, MOMP starts by initializing the residual observation to \(\mathcal{O}_{\text{res}} \leftarrow \mathcal{O}\). This translates into SMOMP initializing \(\mathcal{O}_{\text{res}} \leftarrow \mathcal{O}\). Next, we have an iterative process with two steps, namely projection step and residual update step. For MOMP, the projection step is divided into two sub-steps, projection initialization and refinement. To simplify the notation, the MOMP formulation makes use of the definition \(\Phi_{\text{res}} \in \mathbb{C}^{N_m \otimes_{f=1}^{N_F} N_F^n}\) as \(\mathcal{O}_{\text{res}} C_{f,i} = \mathcal{O}_{\text{res}} C_{f,i}^H \Phi_{f,k}\). Equivalently, SMOMP uses the definition \(\Phi_{\text{res}} C_{f,i} = \sum_{o \in \mathcal{O}} \mathcal{O}_{\text{res}} C_{f,i}^H \prod_{f=1}^{N_F} \Phi_{f,k}^{j_f, k_f, k_f}\).
For the projection initialization, MOMP iteratively solves for the different \( \vec{k} \) the expression in (3), with \( \vec{E} \) being the set of already estimated indices \( \vec{k} \), and \( \vec{E} \) being the set of indices \( \vec{k}' \) which have not been estimated yet, excluding index \( k \).

To adapt (3) to SMOMP, we define \( \vec{E} \) as the set of index pairs \((f', k')\) already estimated, \( \vec{E} \) as the set of index pair which have not been estimated yet, excluding the pair \((f, k)\) and \( \vec{E}' = \vec{E} \cup \{(f, k)\}\) and their slices \( \vec{E}_{f} = \{(f', k') \in \vec{E} \text{ s.t. } f' = f\} \) and \( \vec{E}_{f'} = \{(f', k') \in \vec{E} \text{ s.t. } f' = f\} \). To simplify the formulation, we also define, for any given set of indices \( \mathcal{A} \), the set \( \mathcal{I}_{\mathcal{A}} = \{i_{A} : i \in \mathcal{A}\} \), that is, the set of possible values of \( i_{A} \). Using this definition and the expression of the linear variables, applying the distributive property and eliminating constant terms, we get the expression of the SMOPM projection initialization as

\[
\max_{j_{f}, k} \sum_{i_{j} \in \mathcal{I}_{\mathcal{E}_{f}}} \sum_{i_{f} \in \mathcal{I}_{\mathcal{E}_{f}}} \left\| \frac{1}{\Phi_{j_{f}, i_{f}}} \sum_{(f', k') \in \vec{E}_{f}} [\mathcal{O}]_{i_{j}} \prod_{i' \neq f'} [\Psi_{j_{f}, i_{f}, k'}]_{i_{f}, i_{f'}} \right\|^2
\]

The expression in (4) has the same solution as (3), with a lower complexity due to the simplifications in the denominator.

For the projection refinement, MOMP iteratively solves the following problem for the different \( \vec{k} \), assuming all other indexes estimations to be known:

\[
\max_{j_{f}, k} \left( \sum_{i_{j} \in \mathcal{I}_{\mathcal{E}_{f}}} \sum_{i_{f} \in \mathcal{I}_{\mathcal{E}_{f}}} \left\| \frac{1}{\Phi_{j_{f}, i_{f}}} \sum_{(f', k') \in \vec{E}_{f}} [\mathcal{O}]_{i_{j}} \prod_{i' \neq f'} [\Psi_{j_{f}, i_{f}, k'}]_{i_{f}, i_{f'}} \right\|^2 \right)
\]

Following the same steps that we used to transform (3) into (4), we obtain the SMOPM projection refinement as

\[
\max_{j_{f}, k} \left( \sum_{i_{j} \in \mathcal{I}_{\mathcal{E}_{f}}} \sum_{i_{f} \in \mathcal{I}_{\mathcal{E}_{f}}} \left\| \frac{1}{\Phi_{j_{f}, i_{f}}} \sum_{(f', k') \in \vec{E}_{f}} [\mathcal{O}]_{i_{j}} \prod_{i' \neq f'} [\Psi_{j_{f}, i_{f}, k'}]_{i_{f}, i_{f'}} \right\|^2 \right)
\]

The expression in (6) is equivalent to (5), with a much lower complexity due to the huge simplification of the calculations in the denominator.

For the residual update step, MOMP updates \( \mathcal{O}_{\text{res}} \) as \( \mathcal{O} \rightarrow \mathcal{O} + \frac{\mathcal{O}_{\text{res}}}{\Phi_{\mathcal{C}}^\top \Phi_{\mathcal{C}}} \mathcal{C}_{\mathcal{C}} \), where \( \Phi_{\mathcal{C}} \) is obtained by stacking the columns \( \sum_{i_{j} \in \mathcal{I}_{\mathcal{E}}} \sum_{i_{f} \in \mathcal{I}_{\mathcal{E}}} [\Psi_{j_{f}, i_{f}, k'}]_{i_{f}, i_{f'}} \) for \( j \in \mathcal{C} \), and \( \mathcal{C}_{\mathcal{C}} \) is the solution to \( \mathcal{C}_{\mathcal{C}} = -\sum_{i_{j} \in \mathcal{I}_{\mathcal{E}}} \sum_{i_{f} \in \mathcal{I}_{\mathcal{E}}} [\mathcal{O}]_{i_{j}} \prod_{i' \neq f'} [\Psi_{j_{f}, i_{f}, k'}]_{i_{f}, i_{f'}} \). In the case of SMOPM, the columns of \( \Phi_{\mathcal{C}}^\top \mathcal{C}_{\mathcal{C}} \) can be written as \( \prod_{f=1}^{N_{f}} \sum_{i_{j} \in \mathcal{I}_{\mathcal{E}}} \sum_{i_{f} \in \mathcal{I}_{\mathcal{E}}} [\Psi_{j_{f}, i_{f}, k'}]_{i_{j}, i_{f}, k'} \). We can then retrieve \( \mathcal{O}_{\text{res}} = \mathcal{O}_{\text{res}}(\vec{C}) \) and \( \mathcal{C}_{\mathcal{C}} = \mathcal{C}_{\mathcal{C}}(\mathcal{C}) \). This concludes the description of the SMOPM algorithm. An implementation of SMOPM is available online [20].
III. SMOMP-BASED JOINT CHANNEL ESTIMATION AND LOCALIZATION

We consider a MIMO communication system operating at mmWave frequencies based on a hybrid architecture and uniform rectangular arrays (URA) at both ends, with sizes $N_T = N_T^x \times N_T^y$ and $N_R = N_R^x \times N_R^y$ and using $M_T$ and $M_R$ RF-chains at the transmitter and receiver respectively. We consider the transmission of $N_S \leq M_T$ streams. For training purposes, we choose square digital precoders and combiners, therefore $N_S = M_T$. $U_T$ and $U_R$ denote the sets of feasible analog precoder and combiner entries, given by the resolution of the phase shifting stages. During the link establishment phase, the transmitter sends $M = M_1 M_2$ sequences of $Q$ training symbol vectors to the receiver. The $m$-th training frame, with $m = m_1 M_2 + m_2$, sounds the channel with a hybrid precoder $F_{m_2} = F_{m_2}^{RF} F_{m_2}^{BB}$ and a hybrid combiner $W_{m_1} = W_{m_1}^{RF} W_{m_1}^{BB}$, where $F_{m_2} \in \mathbb{U}_T^{M_T \times M_R}$ and $W_{m_1} \in \mathbb{U}_R^{N_R \times M_R}$ are the analog counterparts of the precoder and combiner, while their digital counterparts are $F_{m_2}^{BB} \in \mathbb{C}^{M_T \times M_R}$ and $W_{m_1}^{BB} \in \mathbb{C}^{M_R \times M_R}$. $D$ is the delay spread of the channel. The training symbol matrix of length $Q$ and $D$ symbols of zero padding for the $m$-th frame is denoted as $S_{m_1} \in \mathbb{C}^{M_R \times (Q+D)}$.

The frequency selective mmWave channel is modeled using a geometric channel model with $L$ paths [9]. The $d$-th delay tap of the channel, for $d \leq D$, is represented as

$$H_d = \sum_{l=1}^{L} \alpha_l R_l(\theta_l) a_l^H(\phi_l) p((d-1)T_n + \tau_0 - \tau_l),$$

where $\alpha_l \in \mathbb{C}$, $\tau_l \in \mathbb{R}$ and $\theta_l, \phi_l \in \{ v \in \mathbb{R}^3 \text{ s.t. } \|v\| = 1 \}$ are the complex gain, delay, direction of arrival (DoA), and direction of departure (DoD) for the $l$-th path, $p(t)$ is the band limited pulse shaping filter including the contributions of the transmitter and receiver, $\tau_0$ is the clock offset, and $a_T$ and $a_R$ denote the steering vectors for the transmitter and the receiver.

Because of the multiple array configurations exploited simultaneously with a hybrid architecture, the transmission of each training frame will generate a set of received signals $Y_m \in \mathbb{C}^{M_R \times Q}$, comprised of $M_R$ combinations of the $M_T$ pilot signal streams. Considering a transmission power $P$, the expression of $Y_m$ is

$$[Y_m]_{:,q} = \sqrt{P} \sum_{d=1}^{D} [W_{m_1}^{H}]_{:,d} F_{m_2} [S_{m_2}]_{:,q} + D - d + W_{m_1}^{H} [N_m]_{:,q},$$

for a noise matrix $N_m \in \mathbb{C}^{N_R \times Q}$ with independent identically distributed entries following a distribution $N(0, \sigma^2)$, being $\sigma^2$ the noise power.

The whitened version of (8) can be obtained from the Cholesky decomposition of the noise covariance matrix $R_m = L_m L_m^{H} = W_m^{H} W_m$, as $L_m^{-1} Y_m$. As proven in [18], [19], the channel estimation problem corresponds to the MOMP problem with $N_Y = 5$, $N^m = 1$, the observation given by

$$[\bar{S}]_{mM_RQ+M_RQ+q} = [L_{m_1}^{-1} Y_m]_{mR,q},$$

the measurement matrix defined as

$$[\bar{S}]_{mM_RQ+M_RQ+q,i} = \sqrt{P} L_{m_1}^{-1} W_{m_1}^{H} [L_{m_2} S_{m_2}]_{:,i} N_{m_1} + i_{i} + D - i_1, \quad (8)$$

and five dictionaries $\Psi_1, \Psi_2, \Psi_3, \Psi_4$ and $\Psi_5$ as in [19], two of them to represent two dimensions of the AoA, two additional ones to represent two dimensions of the AoD, and a fifth dictionary for the delay domain.

Now we obtain a compact expression for the dictionaries that represent the AoA domain. To this aim, we define $\theta^l = (\theta^x_l, \theta^y_l, \theta^z_l)$. Then we exploit that for a horizontal uniform rectangular array of size $N_R^x \times N_R^y$, the array response vector $a_T(\theta) \in \mathbb{C}^{N_R^x N_R^y}$ can be decomposed into two sub-components $a_T^x(\theta^x_l) \in \mathbb{C}^{N_R^x}$ and $a_T^y(\theta^y_l) \in \mathbb{C}^{N_R^y}$ such that $a_T(\theta_l) = a_T^x(\theta^x_l) \otimes a_T^y(\theta^y_l)$. This means that we can rewrite the entries of $a_T(\theta)$ as $[a_T(\theta)]_{n_R^x n_R^y + n_y} = [a_T^x(\theta^x_l)]_{n_R^x} [a_T^y(\theta^y_l)]_{n_R^y}$, with the $(n_R^x N_R^y + n_R^y)$-th antenna element located in position $(n_R^x - 1, n_R^y - 1, 0)$. Similarly, we can define $\phi_l = (\phi^x_l, \phi^y_l, \phi^z_l)$ to obtain the expression of the dictionaries for the AoD for a uniform rectangular array of size $N_T^x \times N_T^y$ at the transmitter. This way we find the decompositions $[a_T(\phi)]_{n_T^x N_T^y + n_y} = [a_T^x(\phi^x_l)]_{n_T^x} [a_T^y(\phi^y_l)]_{n_T^y}$, respectively.

Fig. 2: Illustration of the features of first order channel paths in indoor environments exploited for localization: (a) a first order reflection on a wall travels the same vertical distance as the LoS; (b) first order reflection on the ceiling or floor, travels the same horizontal distance as the LoS.
To derive final expressions for the dictionaries, we can consider the discrete final expressions for the dictionaries, we can consider the discrete domains for $\theta_1$, $\theta_2$, $\phi_1$, $\phi_2$ and $\tau_1 - \tau_0$ with resolutions $N_\theta^2$, $N_\phi^2$, $N_\phi^2$ and $N_\tau^2$, respectively, in discrete domains of size $N_\theta^2$:  \[
abla_1 = \{a_{\theta_1}(\theta_1), \ldots, a_{\theta_1}(\theta_{N_\theta^2})\} \quad (11)
\]
\[
abla_2 = \{a_{\phi_1}(\phi_1), \ldots, a_{\phi_1}(\phi_{N_\phi^2})\} \quad (12)
\]
\[
abla_3 = \{a_{\phi_2}(\phi_{N_\phi^2}), \ldots, a_{\phi_2}(\phi_{N_\phi^2})\} \quad (13)
\]
\[
abla_4 = \{a_{\phi_2}(\phi_{N_\phi^2}), \ldots, a_{\phi_2}(\phi_{N_\phi^2})\} \quad (14)
\]
\[
abla_5 = \{a_{\tau_1}(\tau_{N_\tau^2}), \ldots, a_{\tau_1}(\tau_{N_\tau^2})\} \quad (15)
\]

Next, we transform the MOMP formulation into a SMOMP formulation. From (10), we see that this is an easy task when splitting the measurement matrix using the identity $m = m_1 M_2 + m_2$. The new sparse recovery problem can be written in terms of the following tensors

\[
[O]_{m_1 M_1 + m_2 M_2} = [L_{m_1}^{-1}, Y_{m_1}]_{m_1, q}, \quad (16)
\]
\[
[\Phi_1]_{m_1 M_1 + m_1 M_1} = \sqrt{P}[L_{m_1}^{-1} W_{m_1}]_{m_1, i_1, 1}, \quad (17)
\]
\[
[\Phi_2]_{m_2 M_2 + q, i_2} = [F_{m_2} S_{m_2}]_{i_2, 1}, \quad (18)
\]
\[
\Psi_{1, 1} = \Psi_1, \quad \Psi_{1, 2} = \Psi_2, \quad (19)
\]
\[
\Psi_{2, 1} = \Psi_3, \quad \Psi_{2, 2} = \Psi_4, \quad \Psi_{2, 3} = \Psi_5, \quad (20)
\]

and solved using the SMOMP algorithm described in the previous section.

From the estimated channel parameters after training it is possible to obtain the user location by assuming that the BS location is known and exploiting the geometric relationships between the path parameters and the position of the user and scatters [12], [18], [19], [21]–[23]. In particular, we will consider the strategy described in Section IV in [18], which is suitable for indoor environments. It exploits the concept of virtual anchors to prove that any first order reflection on a wall travels the same vertical distance as the LoS, i.e.

\[
\theta_1 \tau_1 = \theta_2 \tau_1, \quad (21)
\]
as illustrated in Fig. 2(a). Moreover, any first order reflection on the ceiling or floor, travels the same horizontal distance as the LoS, i.e.

\[
\sqrt{(\theta_1^2 + \theta_2^2)} = \sqrt{(\theta_1^2 + \theta_2^2)} \tau_1, \quad (22)
\]
as illustrated in Fig. 2(b). Leveraging these relationships, it is possible to establish an overdetermined system of linear equations to obtain $\tau_0$. From $\tau_0$ and the angles and delays provided by the channel estimation algorithm, it is possible to compute the user location as explained in detail in [18].

IV. NUMERICAL RESULTS

We consider the uplink of an indoor mmWave MIMO system with two possible definitions of array sizes and number of RF-chains. For System I, $N_\theta = N_\phi = M_T = 4$ and $N_\theta = N_\phi = M_R = 16$. For System II, $N_\theta = N_\phi = M_T = 8$ and $N_\theta = N_\phi = M_R = 8$.

We generate the channels using a ray tracing simulation of a home office scenario as described in [18]. The user has a height of 1.7m and moves along 218 different locations, connecting to the access point with highest gain. The noise power is set to $\sigma^2 = 81$dBm. The delay spread is $D = 64$ and $Q = 64$ training symbols are used. To simplify the analysis, we make use of dictionaries of size $N_{T,k} = K_{res} N_{T,k}$ for $K_{res} = 512$. We build the pilot signal as the first $M_T$ rows of a $64 \times 64$ Hadamard matrix with 64 and 32 zeros of padding before and after the pilot. $W_{m_1}$ and $F_{m_2}$ are created by dividing the matrices resulting from the Kronecker product of DFT matrices with sizes $N_\theta$ and $N_\phi$ into blocks of as many columns as RF-chains.

![Table I: Execution time and performance comparison between MOMP and SMOMP.](image-url)

|                  | MOMP | SMOMP |
|------------------|------|-------|
| Computation time | 0.5s | 0.2s  |
| Average angular error | 0.34° | 0.34° |

Table I: Execution time and performance comparison between MOMP and SMOMP.

First we summarize the comparison between MOMP and SMOMP in terms of memory requirements and execution time (comparisons between MOMP and conventional OMP can be found in [18]). The size of the measurement matrix for MOMP is $Q N_\theta N_\phi \times N_T N_R D$, which results in less than 200 million elements for System I, while for System II it requires 8 billion elements, equivalent to more 128 Gb of memory. Because of this, MOMP can only run with the parameters in System I. For System I, the average computation time is 9.5s for MOMP, while it can be reduced to 0.2s with SMOMP. Both algorithms provide similar performance, with an average angular error in the estimation of strongest path of 0.34° when the transmit power is set to 20 dBm. For System II, the average computation time is 1.4s for SMOMP. Table I summarizes the complexity comparison between MOMP and SMOMP.

Fig. 3 depicts the strongest path’s angular error provided by SMOMP when computed as $\cos^{-1}(\hat{\theta}_1^2 \hat{\theta}_1)$. By comparing the blue and green graphs, we can observe the improvement in angular accuracy achieved when increasing the size of the antenna array at the user side. The low budget scenario in System I seems to saturate its accuracy around $-10$dB, while the high budget scenario in System II only requires around $-30$dB to saturate. We believe that this saturation is caused by the limitations of the greedy nature of matching pursuit.

The localization error using the channel estimates provided by SMOMP as the input to the positioning algorithm in [18] can be observed in Fig. 4 as a function of the transmit power. For both System I and System II, over 50% of the users can be located within sub-meter accuracy, while very high accuracy (in the order of cm) can be achieved for over 5% of the users.

These numerical results illustrate the high channel estimation and localization accuracies that can be achieved by a com-
V. CONCLUSIONS

We formulated a new class of sparse recovery problems with multidimensional dictionaries and multiple measurement matrices. We also proposed an algorithm called SMOMP to solve this type of problem with a reasonable complexity, which enables operation with high resolution dictionaries and large separable measurement matrices. We applied the proposed approach to the problem of joint channel estimation and localization at mmWave for a random deployment of users in an indoor scenario simulated by ray tracing. We showed that the proposed approach can provide high accuracy results when considering practical system parameters where other greedy approaches pose unfeasible requirements in terms of memory and complexity.

REFERENCES

[1] J. Tropp, “Greed is good: algorithmic results for sparse approximation,” IEEE Transactions on Information Theory, vol. 50, no. 10, pp. 2231–2242, 2004.

[2] Y. C. Pati, R. Rezaifar, and P. S. Krishnaprasad, “Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition,” in Conference Record Of The Twenty-Seventh Asilomar Conference On Signals, Systems and Computers, 1993, pp. 1–3.

[3] G. M. Davis, S. G. Mallat, and Z. Zhang, “Adaptive time-frequency decompositions,” Optical Engineering, vol. 33, no. 7, pp. 2183–2191, 1994.

[4] J. Tropp, A. Gilbert, and M. Strauss, “Simultaneous sparse approximation via greedy pursuit,” in Proceedings. (ICASSP ‘05). IEEE International Conference on Acoustics, Speech, and Signal Processing, 2005., vol. 5, 2005, pp. v/721–v/724 Vol. 5.

[5] J. A. Tropp, A. C. Gilbert, and M. J. Strauss, “Algorithms for simultaneous sparse approximation. Part I: Greedy pursuit,” Signal Processing, vol. 86, no. 3, pp. 572–588, 2006.

[6] J. Lee, G.-T. Gil, and Y. H. Lee, “Exploiting spatio-spectral sparsity for estimating channels of hybrid MIMO systems in millimeter wave communications,” in 2014 IEEE global communications conference. IEEE, 2014, pp. 3326–3331.

[7] R. Méndez-Rial, C. Rusu, A. Alkhateeb, N. González-Prelcic, and R. W. Heath, “Channel estimation and hybrid combining for mmWave: Phase shifters or switches?” in 2015 Information Theory and Applications Workshop (ITA), 2015, pp. 90–97.

[8] Z. Gao, C. Hu, L. Dai, and Z. Wang, “Channel Estimation for Millimeter-Wave Massive MIMO With Hybrid Precoding Over Frequency-Selective Fading Channels,” IEEE Communications Letters, vol. 20, no. 6, pp. 1259–1262, 2016.

[9] K. Venugopal, A. Alkhateeb, N. Prelic-González, and R. W. Heath, “Channel estimation for hybrid architecture-based wideband millimeter wave systems,” IEEE Journal on Selected Areas in Communications, vol. 35, no. 9, pp. 1996–2009, 2017.

[10] J. Rodríguez-Fernández, N. González-Prelcic, K. Venugopal, and R. W. Heath, “Frequency-domain compressive channel estimation for frequency-selective hybrid millimeter wave MIMO systems,” IEEE Trans. on Wireless Commun., vol. 17, no. 5, pp. 2946–2960, 2018.

[11] X. Wu, G. Yang, F. Hou, and S. Ma, “Low-complexity downlink channel estimation for millimeter-wave FDD massive MIMO systems,” IEEE Wireless Communications Letters, 2019.

[12] A. Shahnannoosri, G. E. Garcia, G. Destino, G. Soco-Granados, and H. Wyneersch, “Position and orientation estimation through millimeter-wave MIMO in 5G systems,” IEEE Trans. on Wireless Commun., vol. 17, no. 3, pp. 1822–1835, 2018.

[13] C. F. Caiafa and A. Cichocki, “Block sparse representations of tensors using kronecker bases,” in 2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2012, pp. 2709–2712.

[14] M. F. Duarte and R. G. Baraniuk, “Kronecker compressive sensing,” IEEE Transactions on Image Processing, vol. 21, no. 2, pp. 494–504, 2012.

[15] C. F. Caiafa and A. Cichocki, “Computing sparse representations of multidimensional signals using kronecker bases,” Neural Computation, vol. 25, no. 1, pp. 186–220, 2013.

[16] R. Zhao, Q. Wang, and Y. Shen, “Selective locking tensor orthogonal matching pursuit algorithm based on block sparsity for multidimensional compressive sensing,” in 2016 IEEE International Instrumentation and Measurement Technology Conference Proceedings, 2016, pp. 1–6.

[17] I. Wickramasingha, A. Elrewainy, M. Sobhy, and S. S. Sherif, “Tensor least angle regression for sparse representations of multidimensional signals,” Neural Computation, vol. 32, no. 9, pp. 1697–1732, 2020.

[18] J. Palacios, N. González-Prelcic, and C. Rusu, “Multidimensional orthogonal matching pursuit: theory and application to high accuracy joint localization and communication at mmWave,” 2022. [Online]. Available: https://arxiv.org/abs/2208.11600.

[19] W. Zhao, Q. Wang, and X. Deng, “Low complexity position and channel estimation at millimeter wave based on multidimensional orthogonal matching pursuit,” in European Signal Processing Conference (EUSIPCO), 2022, pp. 1–6.

[20] “Separable multidimensional matching pursuit core code,” https://github.com/WiSeCom-Lab/SMOMP-core.

[21] H. Wymeersch, N. Garcia, H. Kim, G. Soco-Granados, S. Kim, F. Went, and M. Frohle, “5G mmWave downlink vehicular positioning,” in IEEE Global Commun. Conf. (GLOBECOM), IEEE, 2018, pp. 206–212.

[22] F. Zhu, A. Liu, and V. V. Lau, “Channel estimation and localization for mmWave systems: A sparse bayesian learning approach,” in IEEE Intl. Conf. on Commun. (ICC), 2019, pp. 1–6.

[23] F. Jiang, Y. Ge, M. Zhu, and H. Wymeersch, “High-dimensional channel estimation for simultaneous localization and communications,” in 2021 IEEE Wireless Communications and Networking Conference (WCNC), 2021, pp. 1–6.