

Low complexity joint position and channel estimation at millimeter wave based on multidimensional orthogonal matching pursuit

Joan Palacios, Nuria González-Prelcic

Cristian Rusu

North Carolina State University, USA
Email:{jbeltra, ngprelcic}@ncsu.edu

University of Bucharest, Romania
Email:cristian.rusu@unibuc.ro

Abstract—Compressive approaches provide a means of obtaining high resolution channel estimates in millimeter wave MIMO systems, despite the use of analog and hybrid architectures. Such estimates can also be used as part of a joint channel estimation and localization solution. Achieving good localization performance, though, requires high resolution channel estimates and better methods to exploit those channels. In this paper, we propose a low complexity multidimensional orthogonal matching pursuit strategy for compressive channel estimation based by operating with a product of independent dictionaries for the angular and delay domains, instead of a global large dictionary. This leads to higher quality channel estimates but with lower complexity than generalizations of conventional solutions. We couple this new algorithm with a novel localization formulation that does not rely on the absolute time of arrival of the LoS path and exploits the structure of reflections in indoor channels. We show how the new approach is able to operate in realistic 3D scenarios to estimate the communication channel and locate devices in an indoor simulation setting.

I. INTRODUCTION

Millimeter wave (mmWave) communication and MIMO technology offer additional benefits beyond high data rate communications. The large arrays at high frequencies provide the angle and delay resolvability that enables accurate localization of users and objects in the environment as a byproduct of communication [1]. Future 5G releases and 6G envision VR/AR and automated cars as main use cases, expecting an indoor accuracy < 1 cm and an outdoor accuracy < 10 cm. New research work is needed, from an algorithmic perspective, to achieve this performance.

One direction for developing high accuracy localization in a mmWave MIMO system is to take a model-driven two-stage approach [1]–[4]. The first stage performs compressive channel estimation, which has proven successful in mmWave and massive MIMO channels explicitly, since the large bandwidth or number of antennas prevents the use of conventional channel estimation approaches [5]–[8]. The second stage exploits the geometric relationships between some of the parameters of a sparse MIMO channel (angles and delays) and the position and orientation of the transmitter and the receiver to complete the localization. Despite the success of compressive approaches, further work is needed to overcome obstacles

to achieving higher localization accuracy along with relaxing key assumptions that make the algorithms more practical. To improve accuracy, more precise channel estimators are needed that do not suffer from high complexity encountered with larger dictionaries. To improve realizability, synchronization assumptions need to be relaxed. For example, prior work assumed that the absolute delay of the line of sight (LoS) component can be directly obtained by the receiver as part of the channel estimate, which requires an unrealistic system where the transmitter and receiver are triggered at the same time. To take advantage of larger arrays, more work is needed on hybrid architectures versus the fully digital architectures with high resolution converters, which is not feasible at mmWave frequencies due to its high power consumption [9]. Finally, bandlimited models are needed to avoid an equivalent filtering effect based on a Dirac delta function, which leads to an artificial enhancement of the channel sparsity.

In this paper, we present a practical approach to joint channel estimation and localization at mmWave that overcomes previous limitations. To drastically reduce complexity, we propose the use of a multidimensional orthogonal matching pursuit algorithm (MOMP) [10] that operates with a dictionary in multiple dimensions instead of a large dictionary as conventional OMP [11] would do. To enable localization with a realistic transceiver, the position estimate is obtained as a function of the time difference of arrival and the angular parameters of the different multipath components, without relying in an absolute time of arrival of the LoS path provided by the channel estimation algorithm. In addition, we also propose a novel approach to map channel parameters into position information by exploiting the special structure of reflections on walls, ceiling, and floor that appears in indoor channels. Numerical results in an indoor setting simulated by ray tracing show the effectiveness of the proposed strategy when the transmitter and receiver employ a practical hybrid MIMO architecture.

Notation: We use the following notation throughout the paper. x , \mathbf{x} , \mathbf{X} and \mathcal{X} will be the styles for scalar, vector, matrix or tensor and set. Regarding sub/supper-indices, x and \mathbf{x} are used to denote scalar and categorical values respectively.

$[\mathbf{x}]_n$ denotes the n -th entry of \mathbf{x} . For a 2D matrix \mathbf{X} , $[\mathbf{X}]_{a,b}$, $[\mathbf{X}]_{a,:}$ and $[\mathbf{X}]_{:,b}$ are respectively, the element in the a -th row and b -th column, the a -th row and the b -th column, this notation is extended to the case of tensors with multi-indices acting like multiple indices, such as $[\mathbf{X}]_{\mathbf{a},b} = [\mathbf{X}]_{a_1,a_2,b}$ for $\mathbf{a} = [a_1, a_2]$. The set $\mathcal{U}^{n \times m}$ denotes the set of unit magnitude complex matrices of size n by m . We use the operator $\|\mathbf{x}\|$, $\|\mathbf{X}\|$ to denote the Euclidean and Frobenius norms of \mathbf{x} and \mathbf{X} respectively. In this paper, indexing of tensors starts from 1 and we consider all vectors are column vectors.

II. SYSTEM AND SIGNAL MODEL

We consider a MIMO communication system with a hybrid architecture at both ends to enable operation at millimeter wave bands []. The number of antennas at the transmit array is N_T , while N_R denotes the number of antenna elements at the receive array. The number of RF-chains is denoted as M_T and M_R for the transmitter and receiver respectively. We consider the transmission of N_s streams, with $N_s \leq M_T$ in general. During the link establishment phase, the transmitter sends a sequence of M training frames to the receiver, so this can estimate the communication channel and the location of the transmitter. The training symbols are transmitted by a sequence of training hybrid precoders designed to sound the channel, and received through a set of training hybrid combiners. For training purposes, we choose $N_s = M_T$, what leads to a square digital precoder. The digital combiner is also chosen to be square during training, to avoid compressing the information captured by the receive RF chains. With these assumptions in mind, the digital counterpart of the training precoders/combiners for the m -th training frame are $\mathbf{F}_m^{\text{BB}} \in \mathbb{C}^{M_T \times M_T}$ and $\mathbf{W}_m^{\text{BB}} \in \mathbb{C}^{M_R \times M_R}$. The analog training precoder and the analog training combiner for the m -th training frame are denoted as $\mathbf{F}_m^{\text{RF}} \in \mathcal{U}^{N_T \times M_T}$ and $\mathbf{W}_m^{\text{RF}} \in \mathcal{U}^{N_R \times M_R}$. This way, the hybrid precoders/combiners are $\mathbf{F}_m = \mathbf{F}_m^{\text{RF}} \mathbf{F}_m^{\text{BB}} \in \mathbb{C}^{N_T \times M_T}$ and $\mathbf{W}_m = \mathbf{W}_m^{\text{RF}} \mathbf{W}_m^{\text{BB}} \in \mathbb{C}^{N_R \times M_R}$. Regarding the construction of the training sequence itself, we consider the transmission of M_T streams during training, with a length D zero padding and Q symbols per stream and frame. D is taken as the delay tap length of the channel. Under these assumptions and definitions, the training symbol matrix for the m -th frame is denoted as $\mathbf{S}_m \in \mathbb{C}^{M_T \times (Q+D)}$.

The frequency selective mmWave channel is modeled using a geometric channel model [] with L paths. The d -th delay tap of the channel, for $d = 1, \dots, D$, is represented as

$$\mathbf{H}_d = \sum_{l=1}^L \alpha_l \mathbf{a}_R(\boldsymbol{\theta}_l) \mathbf{a}_T^H(\boldsymbol{\phi}_l) p((d-1)T_s + \tau_0 - \tau_l), \quad (1)$$

where $\alpha_l \in \mathbb{C}$, $\tau_l \in \mathbb{R}$, $\boldsymbol{\theta}_l$, and $\boldsymbol{\phi}_l$ are the complex gain, delay, angle of arrival (AoA), and angle of departure (AoD) for the l -th path, $p(t)$ is a band limited pulse shaping function including filtering effects at the transmitter and receiver, τ_0 is the delay between the beginning of the transmission and the beginning of the reception (clock offset), and $\mathbf{a}_T(\boldsymbol{\phi}_l)$ and $\mathbf{a}_R(\boldsymbol{\theta}_l)$ denote

the array response vectors for the transmitter and the receiver. The angular directions are represented as unitary vector directions, i.e. $\boldsymbol{\theta}_l, \boldsymbol{\phi}_l \in \{\mathbf{v} \in \mathbb{R}^3 \text{ such that } \|\mathbf{v}\| = 1\}$.

Because of the multiple array configurations exploited simultaneously with a hybrid architecture, the transmission of each training frame will generate a set of receive signals that we denote as the block matrix $\mathbf{Y}_m \in \mathbb{C}^{M_R \times Q}$, comprised of M_R combinations of the M_T pilot signal streams. Considering a transmission power P , the expression of this received block matrix is

$$[\mathbf{Y}_m]_{:,q} = \sqrt{P} \sum_{d=1}^D \mathbf{W}_m^H \mathbf{H}_d \mathbf{F}_m [\mathbf{S}_m]_{:,q+D-d} + \mathbf{W}_m^H [\mathbf{N}_m]_{:,q}, \quad (2)$$

The noise block matrix $\mathbf{N}_m \in \mathbb{C}^{N_R \times Q}$ has independent identically distributed entries following a distribution $\mathcal{NC}(0, \sigma^2)$, being σ^2 the noise power. The final observation matrix \mathbf{Y} is constructed by concatenating the received signals \mathbf{Y}_m for all the training frames, i.e. $\mathbf{Y} = [\mathbf{Y}_1, \mathbf{Y}_2, \dots, \mathbf{Y}_M] \in \mathbb{C}^{M_R \times QM}$. The problem to be solved consists of the estimation of the channel matrices \mathbf{H}_d and the position of the transmitter assuming that the receiver position is fixed and known.

III. MOMP-BASED MMWAVE CHANNEL ESTIMATION

Prior work has developed a significant number of sparse recovery solutions to solve the problem of estimating the mmWave channel []. The proposed strategies incur, however, in high computational complexity. This is mainly due to exploiting a formulation, for both the sensing matrix and the sparsifying dictionary, based on Kronecker products of large matrices. The Kronecker construction leads to an equivalent measurement matrix of significant dimension for common sizes of mmWave arrays, what makes matching pursuit solutions impractical in real-world 3D scenarios. In this section, we derive a new formulation based on MOMP [10], which leverages independent dictionaries in the angular and delay domains, without building a global dictionary of larger size.

A. Background in MOMP

The multidimensional matching pursuit problem [10] consists of reconstructing a multidimensional sparse signal to best fit the available observations, assuming that it can be represented by projections on a given set of sparsifying dictionaries. Let us first consider a whitened version of the observation matrix $\mathbf{Y} \in \mathbb{C}^{M_R \times QM}$ as previously defined, which contains a set of QM observations, each one of dimension M_R . We denote the vectorized whitened measurements as $\mathbf{y} \in \mathbb{C}^{MM_R Q}$. To reconstruct a sparse tensor from this observation using MOMP, we consider N_D dictionaries, with the k -th dictionary $\boldsymbol{\Psi}_k \in \mathbb{C}^{N_k^s \times N_k^a}$ consisting of N_k^a atoms in $\mathbb{C}^{N_k^s}$. The coefficients of the sparse signal in the set of dictionaries are represented by $\mathbf{C} \in \mathbb{C}^{N_1^a \times \dots \times N_{N_D}^a}$. Our goal is to reconstruct the tensor containing these coefficients under the assumption that only a few elements are non-zero. To cycle through the multiple indices of the dictionaries, we define the set of entry coordinate combinations $\mathcal{I} = \{\mathbf{i} = (i_1, \dots, i_{N_D}) \in \mathbb{N}^{N_D} \text{ s.t. } i_k \leq N_k^s \ \forall d \leq N_D\}$, and the set of dictionary index combinations

$\mathcal{J} = \{\mathbf{j} = (j_1, \dots, j_{N_D}) \in \mathbb{N}^{N_D} \text{ s.t. } j_k \leq N_k^a, \forall d \leq N_D\}$. Finally, the sensing tensor $\Phi \in \mathbb{C}^{M M_R Q \times N_1^a \times \dots \times N_{N_D}^a}$ defines how the measurement of the sparse signal is performed. To recover the coefficients tensor we can use MOMP to solve:

$$\min_{\mathbf{C}} \|\mathbf{y} - \sum_{\mathbf{i} \in \mathcal{I}} [\Phi]_{:, \mathbf{i}} \left[\prod_{k=1}^{N_D} \Psi_{i_k, j_k^1}, \dots, \prod_{k=1}^{N_D} \Psi_{i_k, j_k^{|\mathcal{J}|}} \right] \mathbf{C}_{\mathcal{J}}\|^2. \quad (3)$$

Additionally, the support $\mathcal{C} \subset \mathcal{J}$ is defined as the set of indices such that any element of $\mathbf{C}_{\mathcal{C}}$ is non zero.

B. MOMP-based formulation

To formulate the mmWave channel estimation as a multi-dimensional orthogonal matching pursuit problem, we need to rewrite (2) to match the structure of the problem in (3). The next paragraphs derive the appropriate expressions for the dictionaries and the sensing matrix so that MOMP can be used to solve (3) given the observation matrix \mathbf{Y} .

The channel tensor $\mathbf{H} \in \mathbb{C}^{N_R \times M_T \times D}$ is defined as $[\mathbf{H}]_{:, :, d} = \mathbf{H}_d$. To write the channel \mathbf{H} as a sparse combination of products of dictionaries, it is natural to think of independent dictionaries for the angle of arrival, the angle of departure, and the delay, since these are the independent parameters of the paths that compose the channel. Moreover, two different dictionaries are used to represent the two independent dimensions of both the AoA and AoD. With this in mind, we define five sparsifying dictionaries to represent the sparse mmWave channel.

To obtain the compact expression of the dictionaries for the AoA, we consider that for a horizontal uniform rectangular array of size $N_R^x \times N_R^y$, the array response vector $\mathbf{a}_R(\theta_l) \in \mathbb{C}^{N_R^x N_R^y}$ can be decomposed into two sub-components $\mathbf{a}_R^x(\theta^x) \in \mathbb{C}^{N_R^x}$ and $\mathbf{a}_R^y(\theta^y) \in \mathbb{C}^{N_R^y}$ such that $\mathbf{a}_R(\theta_l) = \mathbf{a}_R^x(\theta^x) \otimes \mathbf{a}_R^y(\theta^y)$. This means that we can rewrite the entries of $\mathbf{a}_R(\theta_l)$ as $[\mathbf{a}_R(\theta_l)]_{n_R^x N_R^y + n_R^y} = [\mathbf{a}_R^x(\theta_l^x)]_{n_R^x} [\mathbf{a}_R^y(\theta_l^y)]_{n_R^y}$. Analogously, 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, we can find the decomposition $[\mathbf{a}_T(\phi_l)]_{n_T^x N_T^y + n_T^y} = [\mathbf{a}_T^x(\phi_l^x)]_{n_T^x} [\mathbf{a}_T^y(\phi_l^y)]_{n_T^y}$. Finally, we will define the dictionary for the delay domain from the evaluation of the pulse shaping functions $[\mathbf{a}_D(\tau)]_d = p((d-1)T_s - \tau)$, $d = 1, \dots, D$.

Now, to obtain the final expressions for the dictionaries, we can consider the discrete domains for θ^x , θ^y , ϕ^x , ϕ^y and $\tau - \tau_0$ with resolutions N_1^a , N_2^a , N_3^a , N_4^a and N_5^a , namely $\{\bar{\theta}_1^x, \dots, \bar{\theta}_{N_1^a}^x\}$, $\{\bar{\theta}_1^y, \dots, \bar{\theta}_{N_2^a}^y\}$, $\{\bar{\phi}_1^x, \dots, \bar{\phi}_{N_3^a}^x\}$, $\{\bar{\phi}_1^y, \dots, \bar{\phi}_{N_4^a}^y\}$ and $\{\bar{\tau}_1, \dots, \bar{\tau}_{N_5^a}\}$. This way, we define the dictionaries as

$$\begin{aligned} \Psi_1 &= [\mathbf{a}_R^x(\bar{\theta}_1^x), \dots, \mathbf{a}_R^x(\bar{\theta}_{N_1^a}^x)], \Psi_2 = [\mathbf{a}_R^y(\bar{\theta}_1^y), \dots, \mathbf{a}_R^y(\bar{\theta}_{N_2^a}^y)], \\ \Psi_3 &= [\mathbf{a}_T^x(\bar{\theta}_1^x)^*, \dots, \mathbf{a}_T^x(\bar{\theta}_{N_3^a}^x)^*], \Psi_4 = [\mathbf{a}_T^y(\bar{\theta}_1^y)^*, \dots, \mathbf{a}_T^y(\bar{\theta}_{N_4^a}^y)^*] \\ \Psi_5 &= [\mathbf{a}_D(\bar{\tau}_1), \dots, \mathbf{a}_D(\bar{\tau}_{N_5^a})]. \end{aligned} \quad (4)$$

To use these dictionaries, we define the sets of multi-indices $\mathcal{I} = \{\mathbf{i} = [i_1, i_2, i_3, i_4, i_5] \in \mathbb{N} \text{ such that } i_1 \leq N_R^x, i_2 \leq N_R^y, i_3 \leq N_T^x, i_4 \leq N_T^y, i_5 \leq D\}$, and $\mathcal{J} = \{\mathbf{j} = [j_1, j_2, j_3, j_4, j_5] \in \mathbb{N} \text{ such that } j_1 \leq N_1^a, j_2 \leq N_2^a, j_3 \leq$

$N_3^a, j_4 \leq N_4^a, j_5 \leq N_5^a\}$ and $N_R^x \times N_R^y \times N_T^x \times N_T^y \times D$ can be expressed as $\otimes_{k=1}^5 N_k^a$. Finally, ignoring quantization effects caused by the finite resolution of the dictionaries, we can define $\mathbf{C}_{\mathcal{J}} \in \mathbb{C}^{\otimes_{d=1}^5 N_d^a \times 1}$ as the sparse tensor

$$\mathbf{C}_{\mathcal{J}} = \begin{cases} \alpha_l, & \text{if } \theta_l^x = \bar{\theta}_{j_1}^x \text{ or } \theta_l^y = \bar{\theta}_{j_2}^y \text{ or} \\ & \phi_l^x = \bar{\phi}_{j_3}^x \text{ or } \phi_l^y = \bar{\phi}_{j_4}^y \text{ or } \tau_l = \bar{\tau}_{j_5} \\ 0, & \text{otherwise.} \end{cases} \quad (5)$$

Under these definitions, the channel entries can now be expressed as

$$[\mathbf{H}]_{i_1 N_R^y + i_2, i_3 N_T^y + i_4, i_5} = \left[\prod_{k=1}^{N_D} \Psi_{i_k, j_k^1}, \dots, \prod_{k=1}^{N_D} \Psi_{i_k, j_k^{|\mathcal{J}|}} \right] \mathbf{C}_{\mathcal{J}}. \quad (6)$$

Regarding the construction of the observation, we assume that each training frame is independent. Next, the received signal is whitened to compensate for the correlation effect introduced by the analog combiner. To this aim, we find a Cholesky decomposition of the noise correlation matrix $\mathbf{W}_m^H \mathbf{W}_m$ i.e. \mathbf{L}_m such that $\mathbf{L}_m \mathbf{L}_m^H = \mathbf{W}_m^H \mathbf{W}_m$ and multiply (2) by \mathbf{L}_m^{-1} to obtain

$$\begin{aligned} [\mathbf{L}_m^{-1} \mathbf{Y}_m]_{:, q} &= \sqrt{P} \sum_{d=1}^D \mathbf{L}_m^{-1} \mathbf{W}_m^H [\mathbf{H}]_{:, :, d} \mathbf{F}_m [\mathbf{S}_m]_{:, q+D-d} \\ &\quad + \mathbf{L}_m^{-1} \mathbf{W}_m^H [\mathbf{N}_m]_{:, q}. \end{aligned} \quad (7)$$

After the whitening stage, the expressions for the observation \mathbf{y} , and noise $\mathbf{n} \in \mathbb{C}^{M M_R Q}$ are

$$[\mathbf{y}]_{m M_R Q + m_R Q + q} = [\mathbf{L}_m^{-1} \mathbf{Y}_m]_{m_R Q, q} \quad (8)$$

$$[\mathbf{n}]_{m M_R Q + m_R Q + q} = [\mathbf{L}_m^{-1} \mathbf{W}_m^H \mathbf{N}_m]_{m_R Q, q}. \quad (9)$$

To complete the MOMP formulation, we still need to define the sensing matrix $\Phi \in \mathbb{C}^{M M_R Q \times N_R^x \times N_R^y \times N_T^x \times N_T^y \times D}$. Note that we can also write the observations in (8) as $[\mathbf{y}]_{m M_R Q + m_R Q + q} = [\Phi]_{m M_R Q + m_R Q + q, \mathbf{i}} [\mathbf{H}]_{i_1 N_R^y + i_2, i_3 N_T^y + i_4, i_5}$. Mapping the terms in this observation to (7)

$$\begin{aligned} [\Phi]_{m M_R Q + m_R Q + q, \mathbf{i}} &= \\ &= \sqrt{P} [\mathbf{L}_m^{-1} \mathbf{W}_m^H]_{m_R, i_1 N_R^y + i_2} [\mathbf{F}_m \mathbf{S}_m]_{i_3 N_T^y + i_4, q+D-i_5}. \end{aligned} \quad (10)$$

With these definitions, the multi-dimensional matching pursuit problem is completed, since we can re-write equation (2) like

$$\mathbf{y} = \sum_{\mathbf{i} \in \mathcal{I}} [\Phi]_{:, \mathbf{i}} \left[\prod_{k=1}^{N_D} \Psi_{i_k, j_k^1}, \dots, \prod_{k=1}^{N_D} \Psi_{i_k, j_k^{|\mathcal{J}|}} \right] \mathbf{C}_{\mathcal{J}} + \mathbf{n}. \quad (11)$$

Since \mathbf{n} is white noise, the maximum likelihood estimator is the solution to (3), and we can apply the MOMP algorithm to approximately and sparsely solve it.

The parameters for the estimated path corresponding to the index $\mathbf{j} \in \mathcal{C}$ can be directly extracted from the support and the sparse reconstructed matrix as $\bar{\theta}_{j_1}^x$, $\bar{\theta}_{j_2}^y$, $\bar{\phi}_{j_3}^x$, $\bar{\phi}_{j_4}^y$ and $\bar{\tau}_{j_5}$ for the angular information and relative time of arrival, and as $\mathbf{C}_{\mathcal{J}}$ for the complex gain. To fully retrieve the angular information, the z components of θ and ϕ can be computed

using the fact that these are unitary vectors, and that the z component is positive (otherwise the path would be coming from the antenna array substrate which blocks the signal), i.e. $\bar{\theta}^z = \sqrt{(\bar{\theta}^x)^2 + (\bar{\theta}^y)^2}$ and $\bar{\phi}^z = \sqrt{(\bar{\phi}^x)^2 + (\bar{\phi}^y)^2}$.

IV. LOCALIZATION

Once the channel parameters have been estimated, the geometric relationships between those parameters and the scatterers in the environment can be used to obtain an estimate of the transmitter position. These geometric relationships between the channel parameters and the position to be estimated depend on the type of path. Therefore, a method to classify the different estimated paths has to be proposed. To this aim, we will make use of two properties: a) specular reflections on the horizontal plane (floor/ceiling) do not change the x - y components of a point, and b) vertical reflections (walls) do not change the z component. With this in mind, we consider 4 possible types of paths: line of sight (LoS), wall reflection, floor/ceiling reflection, or any other path that will be labeled as spurious and will not be exploited for localization.

The input to our proposed path classification algorithm are the spherical coordinates of the different angles, computed as $\theta^{az} = \arg(\theta^x + j\theta^y)$, $\theta^{el} = \arcsin(\theta^z)$, $\phi^{az} = \arg(\phi^x + j\phi^y)$ and $\phi^{el} = \arcsin(\phi^z)$. Wall reflections and LoS paths satisfy $\theta_l^{el} + \phi_l^{el} = 0$ while floor/ceiling reflections satisfy $\theta_l^{el} = \phi_l^{el}$. Additionally, floor/ceiling reflections and LoS paths arrival and departure azimuth angles are opposite, i.e. they are separated by 180° . By defining the threshold values r_{az}, r_{el} we can obtain the conditions

$$|\sin(\theta^{el} - \phi_l^{el})| < r_{el}, \quad (12)$$

$$|\sin(\theta^{el} + \phi_l^{el})| < r_{el}, \quad (13)$$

$$\cos(\theta^{az} - \phi_l^{az}) < r_{az} - 1. \quad (14)$$

A LoS path satisfies conditions (12) and (14), floor/ceiling reflections satisfy (13) and (14), while wall reflections only satisfy (12). Any other path will be classified as spurious.

Let us define now the algorithm that provides the position estimate from the path parameters and path classification. We consider the access point (receiver) to be in the coordinate origin, and $\mathbf{u} \in \mathbb{R}^3$ to be the device (transmitter) location. To ease the formulation of the position estimation algorithm we introduce the projection matrices $\chi_{LOS} = \mathbf{I}_3$, $\chi_h = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$, and $\chi_v = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$. With these definitions, the geometric relationships created by a LoS path, a floor/ceiling reflection or a wall reflection can be expressed as

$$\chi_{LOS}\mathbf{u} = \chi_{LOS}\boldsymbol{\theta}_l\tau_l, \quad \chi_h\mathbf{u} = \chi_h\boldsymbol{\theta}_l\tau_l, \quad \chi_v\mathbf{u} = \chi_v\boldsymbol{\theta}_l\tau_l, \quad (15)$$

respectively. In other words, if we define χ_l , as the projection matrix corresponding to the l -th path type, we can rewrite the geometric relationships for that path as

$$\chi_l(\mathbf{u} - \boldsymbol{\theta}_l\tau_l) = 0. \quad (16)$$

Note, however, that the channel estimation algorithm only provides the estimation of the relative delays $\Delta\tau_l = \tau_l - \tau_0$,

with τ_0 unknown, instead of τ_l . Therefore, the equations for the geometric relationships need to be established in terms of the relative delays, i.e.

$$\chi_l(\mathbf{u} - \boldsymbol{\theta}_l(\Delta\tau_l + \tau_0)) = 0. \quad (17)$$

Now, we define the matrix $\mathbf{U}_{\boldsymbol{\theta}_l} = [\mathbf{I}, -\boldsymbol{\theta}_l]$ and the variable $\mathbf{z} = [\mathbf{u}^T, \tau_0]^T$, which contains all the unknown variables, to rewrite (17) as

$$\chi_p(\mathbf{U}_{\boldsymbol{\theta}_l}\mathbf{z} - \boldsymbol{\theta}_l\Delta\tau_l) = 0. \quad (18)$$

This equation can be alternatively solved by

$$\min_{\mathbf{z}} \mathbf{z}^T \mathbf{U}_{\boldsymbol{\theta}_l}^T \chi_l^T \chi_l \mathbf{U}_{\boldsymbol{\theta}_l} \mathbf{z} - 2\tau_l \phi_l^T \chi_l^T \chi_l \mathbf{U}_{\boldsymbol{\theta}_l} \mathbf{z} + \tau_l^2 \phi_l^T \chi_l^T \chi_l \phi_l. \quad (19)$$

We define now a weighted version of the previous equation to account for the different impact of the errors associated to different types of paths. Thus, we consider a weight value for each path w_l that we define as its estimated gain $w_l = |\alpha_l|^2$. With this in mind we can weight average (19) for all the different paths as

$$\min_{\mathbf{z}} \mathbf{z}^T \mathbf{A} \mathbf{z} - 2\mathbf{b} \mathbf{z} + c, \quad (20)$$

for $\mathbf{A} = \sum_{l=1}^L w_l \mathbf{U}_{\boldsymbol{\theta}_l}^T \chi_l^T \chi_l \mathbf{U}_{\boldsymbol{\theta}_l}$, $\mathbf{b} = \sum_{l=1}^L w_l \tau_l \phi_l^T \chi_l^T \chi_l \mathbf{U}_{\boldsymbol{\theta}_l}$ and $c = \sum_{l=1}^L w_l \tau_l^2 \phi_l^T \chi_l^T \chi_l \phi_l$. The widely known solution to (20) for a symmetric \mathbf{A} matrix is $\hat{\mathbf{z}} = \mathbf{A}^{-1}\mathbf{b}$. The location estimation $\hat{\mathbf{u}}$ can then be extracted from the first 3 entries of $\hat{\mathbf{z}}$, while the last entry provides the unknown offset between transmitter and receiver.

V. SIMULATION

To generate the evaluation data set with channels and user positions we simulate by ray tracing (using Wireless InSite software) a home office with two rooms. There is an access point in each one of the rooms. We generate 218 user locations following a path connecting both rooms. The user antenna is placed horizontally, while the access points are mounted on the walls vertically, facing the interior of their respective rooms. The ray tracing software generates the optical geometrical channel between the user positions and the access points. We associate each user location to the single access point with the highest received signal strength. For each user location, we input the parameters corresponding to the channel with its associated access point in a measurement model to generate the received signal. The transmit power is set to $P = 20\text{dBm}$, and the noise level is the thermal noise corresponding to a room at 15°C when using a 1GHz bandwidth, i.e. $\sigma^2 = -84\text{dBm}$. The pulse shaping filter is defined as a sinc function. We set the delay tap length to $D = 64$, and the training signal to be the first 6 rows of a 64 element Hadamard matrix padded with 64 zeros to the left and 32 zeros to the right, i.e. $\mathbf{S} = [0, \dots, 0, \mathbf{Had}_{64}, 0, \dots, 0]$. The training precoders/combiners are taken as columns extracted from the Kronecker product of the DFT matrices with sizes N_T^x and N_T^y/N_R^x and N_R^y .

To simplify the sparse reconstruction problem, we define the ratio between the number of atoms and atoms size $K_{\text{res}} =$

| Method | OMP | | | MOMP | | |
|--------------------------------------|-----|-----|-----|------|-----|------|
| K_{res} | 1 | 1.3 | 1.6 | 16 | 128 | 1024 |
| $\angle(\theta_1, \theta_1)[^\circ]$ | 12 | 10 | 8 | 1.1 | 0.9 | 0.9 |
| $\angle(\phi_1, \phi_1)[^\circ]$ | 44 | 30 | 25 | 2.1 | 1.8 | 1.7 |
| Run-time [s] | 1.3 | 2.2 | 8.2 | 3.2 | 3.3 | 3.5 |

TABLE I: Median main path angular error and run time as a function of K_{res} .

N_k^a/N_k^s , setting it to the same constant value independent of k . By how the dictionaries are defined we have $N_1^s = N_R^x$, $N_2^s = N_R^y$, $N_3^s = N_T^x$, $N_4^s = N_R^y$, and $N_5^s = D$.

The first metric used to evaluate the accuracy of the channel estimation algorithm is the error in the estimation of the AoA of the main path. We can see its evolution over K_{res} for both algorithms, OMP and MOMP, in Table I, when considering a 3×3 antenna array at the transmitter and a 6×6 antenna at the receiver. The number of RF chains is set to 3 for the device and to 6 for the access point. It is straight forwards to see how MOMP outperforms OMP estimations by an order of magnitude because of its ability to increase the resolution of the dictionaries without incurring in exorbitant computational times or impossible memory requirements.

We analyze next the localization error $\|\mathbf{u} - \hat{\mathbf{u}}\|$ as a function of the number of antennas and complexity. The value used for the weight is each path's estimated power $w_l = |\hat{\alpha}_l|^2$ and the threshold values for the path classification are $r_{\text{az}} = r_{\text{el}} = 0.12$. We fix the antenna size for the device to 3×3 , while three different sizes are considered for the access points: 6×6 , 8×8 , and 10×10 . The number of receive RF chains also varies accordingly as 6, 8 and 10. The localization results can be visualized in Fig. 1. Even with these small antennas sizes, the access point is able to localize most of the devices keeping the error under reasonable values when an unknown clock offset is considered in the simulation. Note that even channels with very bad SNR conditions, where the link cannot be established have been considered in these averaged results. These results clearly show the benefits of applying MOMP to the problem of joint channel estimation and localization at mmWave. Higher estimation accuracies are expected with larger antenna arrays at the transmitter, which is constrained to be only 3×3 in this simulation scenario so OMP can run.

VI. CONCLUSIONS

We developed a joint compressive channel estimation and localization strategy for a realistic mmWave MIMO communication systems. The proposed approach is based on a multidimensional matching pursuit sparse recovery algorithm, which enables operation with high resolution dictionaries, for both the time and angular domains, at a reduced complexity. Using mmWave channels generated by ray tracing we obtained the channel estimation error for the angular parameters and the localization errors for a random deployment of users, as a function of the complexity and the system parameters. We showed how the MOMP-based approach is able to provide reasonable

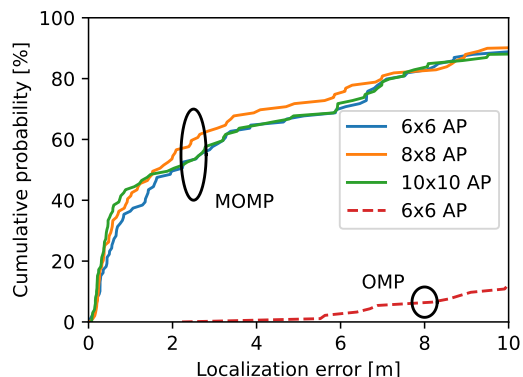


Fig. 1: Localization error as a function of the receive array size, number of RF chains and sparse recovery strategy. K_{res} has been set to 128 for MOMP and to 1.6 for OMP.

estimation accuracies in scenarios where conventional OMP is not feasible.

REFERENCES

- [1] A. Shahmansoori, G. E. Garcia, G. Destino, G. Seco-Granados, and H. Wymeersch, "Position and orientation estimation through millimeter-wave MIMO in 5G systems," *IEEE Trans. on Wireless Commun.*, vol. 17, no. 3, pp. 1822–1835, 2018.
- [2] H. Wymeersch, N. Garcia, H. Kim, G. Seco-Granados, S. Kim, F. Went, and M. Fröhle, "5G mmWave downlink vehicular positioning," in *IEEE Global Commun. Conf. (GLOBECOM)*. IEEE, 2018, pp. 206–212.
- [3] J. Talvitie, M. Koivisto, T. Levanen, M. Valkama, G. Destino, and H. Wymeersch, "High-accuracy joint position and orientation estimation in sparse 5G mmWave channel," in *IEEE Intl. Conf. on Commun. (ICC)*, 2019, pp. 1–7.
- [4] 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.
- [5] K. Venugopal, A. Alkhateeb, N. Prelicic-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.
- [6] 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.
- [7] 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.
- [8] F. Zhu, A. Liu, and 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.
- [9] R. W. Heath, N. González-Prelcic, S. Rangan, W. Roh, and A. M. Sayeed, "An overview of signal processing techniques for millimeter wave MIMO systems," *IEEE Journal of Selected Topics in Signal Processing*, vol. 10, no. 3, pp. 436–453, April 2016.
- [10] J. Palacios, N. González-Prelcic, and C. Rusu, "Multidimensional orthogonal matching pursuit: theory and application to joint channel estimation and localization at mmWave," arXiv preprint, 2022.
- [11] J. A. Tropp and A. C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Transactions on Information Theory*, vol. 53, no. 12, pp. 4655–4666, 2007.