A Computationally Efficient 2D MUSIC Approach for 5G and 6G Sensing Networks

Marcus Henninger*†, Silvio Mandelli*, Maximilian Arnold*, and Stephan ten Brink†
*Nokia Bell Labs, 70435 Stuttgart, Germany,
†Institute of Telecommunications, University of Stuttgart, 70659 Stuttgart, Germany
E-mail: marcus.henninger@nokia.com, {silvio.mandelli, maximilian.arnold}@nokia-bell-labs.com

***Abstract*—Future cellular networks are intended to have the ability to sense the environment by utilizing reflections of transmitted signals. Multi-dimensional sensing brings along the crucial advantage of being able to resort to multiple domains to resolve targets, enhancing detection capabilities compared to 1D estimation. However, estimating parameters jointly in 5G New Radio (NR) systems poses the challenge of limiting the computational complexity while preserving a high resolution. To that end, we make us of channel state information (CSI) decimation for MUltiple SIgnal Classification (MUSIC)-based joint range-angle of arrival (AoA) estimation. We further introduce multi-peak search routines to achieve additional detection capability improvements. Simulation results with orthogonal frequency-division multiplexing (OFDM) signals show that we attain higher detection probabilities for closely spaced targets than with 1D range-only estimation. Moreover, we demonstrate that for our considered 5G setup, we are able to significantly reduce the required number of computations due to CSI decimation.**

***Index Terms*—OFDM radar, MUSIC, Sensing, Localization**

I. Introduction

Joint Communication and Sensing (JCAS) is expected to be one of the main features of Beyond 5G (B5G) and 6G cellular networks and is currently attracting a lot of attention in the research community [1]. Future networks should be able to extract information about the physical world from the channel, performing de-facto sensing operations. Road user protection [2] and positioning in indoor scenarios [3] are two examples of the various possible sensing applications. In communication systems, the channel state information (CSI) is estimated as part of the necessary steps for communication purposes. Orthogonal frequency-division multiplexing (OFDM) radar [4] can be used to exploit this information by means of the reflections from objects illuminated by the transmitted signal to estimate their ranges and angles of arrivals (AoAs). Among the possible estimation techniques, we consider the MUltiple SIgnal Classification (MUSIC) algorithm [5], also because it allows coping with generic antenna array shapes.

Ideally, range and AoA are estimated jointly. This offers the major advantage that targets can be resolved in two dimensions and consequently improves the ability to detect closely spaced targets compared to 1D estimation. In general, joint estimation of multiple parameters is a well-researched field and has been proposed e. g., in [6] and [7]. For the particular case of jointly estimating range and AoA with MUSIC and OFDM signals, *SpotFi* [8] has made use of the spatial smoothing technique [9]. Differently from [8], which deals with WiFi signals, we consider 5G New Radio (NR) numerologies [10], requiring a large number of subcarriers to achieve the necessary bandwidth for a high range resolution. The increased CSI matrix dimension would lead to a computational complexity that is too high to be handled in real-time applications.

To that end, we enhance the spatial smoothing technique by properly decimating the estimated CSI, adopting the approach used in [11] for joint range-Doppler estimation in automotive radars. Our proposal enables high resolution in 5G sensing applications while drastically reducing the computational complexity. We also formulate a necessary condition for detecting multiple targets with multi-dimensional spatial smoothing. Similar to the motivation of earlier works ([12], [13]), we want to avoid an inefficient grid search in the resulting 2D MUSIC spectrum. We therefore present a peak search routine using Powell's algorithm [14] that can detect multiple targets at once. Moreover, we can reduce the missed detection probability by coherently removing the contribution of previously detected targets utilizing subspace tracking methods [15]. All our algorithms can straightforwardly be extended to higher dimensions for estimating additional parameters and can directly be implemented on top of 5G NR architectures.

The paper is structured as follows: we first define the underlying signal model in Section II before presenting our algorithms in Section III. In Section IV, we introduce the simulation scenario and assumptions, and discuss the benefits of our method based on the results. The paper is wrapped up with a conclusion (Section V).

II. Signal Model

We consider a single TX antenna illuminating the environment by transmitting phase-modulated complex symbols $\mathbf{s} = \begin{bmatrix} s_1, & s_2, & \dots, & s_N \end{bmatrix}$ modulated onto N subcarriers of an OFDM signal with subcarrier spacing Δf. The carrier wavelength of the signal is denoted by λ. While we only investigate static targets in this paper, the extension to Doppler estimation could be accomplished by considering multiple OFDM symbols.

The RX uniform linear array (ULA) is co-located with the transmitter and comprises K antennas with element spacing d. Let the received symbols at array element *k* be written as a row vector $\mathbf{y}_k = \begin{bmatrix} y_{k,1}, & y_{k,2}, & \dots, & y_{k,N} \end{bmatrix}$. Assuming knowledge

of the TX symbols, the frequency domain CSI at each RX antenna is obtained by carrying out an element-wise division

$$c_{k,n} = \frac{y_{k,n}}{s_n}. \tag{1}$$

Stacking the CSI of all antenna elements vertically yields the $K \times N$ CSI matrix

$$\mathbf{C} = \begin{bmatrix} \mathbf{c}_1, \ \mathbf{c}_2, \ \ldots, \ \mathbf{c}_K \end{bmatrix}^{\mathrm{T}}. \tag{2}$$

Assuming Q impulsive scatterers as targets, each of them generates a reflection according to azimuth AoA θ_q and range $r_q = \frac{\tau_q c}{2}$, where τ_q is the delay of the reflection caused by the q-th target and c the speed of light. The CSI matrix can be expressed as the superposition of the channel contributions of all Q targets

$$\mathbf{C} = \sum_{q=1}^{Q} h_q \mathbf{b}(\theta_q)\mathbf{a}(r_q)^{\mathrm{T}} + \mathbf{Z}, \tag{3}$$

where h_q is the complex coefficient of the q-th target and $\mathbf{Z}$ the $K \times N$ random complex additive white Gaussian noise (AWGN) matrix. Under the assumption that $N\Delta f \ll f_c$ (carrier frequency), the vectors $\mathbf{b}(\theta_q)$ and $\mathbf{a}(r_q)$ are given as

$$\mathbf{b}(\theta_q) = \left[1, \ e^{j2\pi \frac{d}{\lambda}\sin(\theta_q)}, \ \ldots, \ e^{j2\pi(K-1)\frac{d}{\lambda}\sin(\theta_q)}\right]^{\mathrm{T}} \tag{4}$$

$$\mathbf{a}(r_q) = \left[1, \ e^{-j2\pi\Delta f \cdot 2\frac{r_q}{c}}, \ \ldots, \ e^{-j2\pi(N-1)\Delta f \cdot 2\frac{r_q}{c}}\right]^{\mathrm{T}} \tag{5}$$

and represent the linear phase shifts in the antenna and subcarrier dimension induced by θ_q and r_q of the q-th target.

III. Proposed Algorithm

A. Spatial Smoothing with CSI Decimation

Applying MUSIC in two dimensions couples range and AoA estimates to the same target and therefore improves the ability of discriminating different targets. However, previous work [9] states that for detecting multiple targets, a necessary condition is that the number of independent measurements for computing the sample covariance matrix $\hat{\mathbf{R}}_y$ must be greater or equal than Q. To achieve this with a single snapshot, [8] makes use of the spatial smoothing technique by generating L sub-arrays $\begin{bmatrix}\mathbf{C}_{s,1}, \ \mathbf{C}_{s,2}, \ \ldots, \ \mathbf{C}_{s,L}\end{bmatrix}$ from $\mathbf{C}$. The $M \times L$ smoothed CSI matrix $\tilde{\mathbf{C}}$ is then constructed as

$$\begin{aligned}\tilde{\mathbf{C}} &= \begin{bmatrix}\mathrm{vec}(\mathbf{C}_{s,1}), \ \mathrm{vec}(\mathbf{C}_{s,2}), \ \ldots, \ \mathrm{vec}(\mathbf{C}_{s,L})\end{bmatrix} \\ &= \begin{bmatrix}\mathbf{c}_{s,1}, \ \mathbf{c}_{s,2}, \ \ldots, \ \mathbf{c}_{s,L}\end{bmatrix},\end{aligned} \tag{6}$$

with M being the number of samples per sub-array and vec$(\cdot)$ the vectorization operator applied in row-major order. As an example, *all* elements contained in the orange and blue boxes of Fig. 1 make up two independent sub-arrays. The sample covariance matrix can be computed as

$$\hat{\mathbf{R}}_y = \frac{1}{M}\tilde{\mathbf{C}}\tilde{\mathbf{C}}^{\mathrm{H}}. \tag{7}$$

After performing the eigenvector decomposition (EVD) of $\hat{\mathbf{R}}_y$, the eigenvectors are partitioned into the $M \times Q$ signal subspace $\mathbf{U}_S$ corresponding to the Q strongest eigenvalues and the complementary $M \times (M-Q)$ noise subspace $\mathbf{U}_N$. For estimating Q, the minimum description length method [16] is used. The 2D MUSIC spectrum is obtained by computing

$$P_{\mathrm{MU}}(r,\theta) = \frac{1}{(\mathbf{b}(\theta)\otimes\mathbf{a}(r))^{\mathrm{H}}\mathbf{U}_N\mathbf{U}_N^{\mathrm{H}}(\mathbf{b}(\theta)\otimes\mathbf{a}(r))}, \tag{8}$$

where $\mathbf{b}(\theta)$ and $\mathbf{a}(r)$ are the steering vectors for the trial range-azimuth pair (r,θ) and $\otimes$ is the Kronecker product. Preserving the domain apertures of $\mathbf{C}$ leads to the dimensions of the sub-arrays growing rapidly, rendering the approach computationally expensive. Obviously, one could simply consider small sub-arrays and limit the complexity in this way. However, in modern 5G systems, even with Δf = 60 kHz, the resulting number of subcarriers N could already be above one thousand. To preserve the range resolution, it is prerequisite to have sub-arrays with a large frequency domain aperture A_f, as the range resolution Δr of an OFDM radar is

$$\Delta r = \frac{c}{2A_f\Delta f}. \tag{9}$$

Therefore, the range resolution is inversely proportional to the frequency aperture. With fixed Δf, one can only increase A_f, thus M, to improve ranging performance, quickly leading to computationally infeasible situations.

As introduced in [11] for range-Doppler estimation, we decimate the sub-arrays, enabling us to keep a large aperture in both frequency (i. e., resulting effective bandwidth) and spatial (i. e., aperture of the antenna array) dimension. Besides apertures A_f and A_a, we parametrize the generation of sub-arrays by defining decimation and stride between them, which hereinafter shall be denoted as D and S, respectively, with subscripts f and a representing frequency and spatial dimension. The antenna and subcarrier indices of the ℓ-th sub-array of the smoothed CSI matrix $\tilde{\mathbf{C}}$ are denoted as vectors of length M

$$\mathbf{a}_\ell = \mathbf{1}_{\tilde{N}_f} \otimes \begin{bmatrix}\tilde{a}_\ell, \ \tilde{a}_\ell + D_a, \ \ldots, \ \tilde{a}_\ell + (\tilde{N}_a - 1)D_a\end{bmatrix}^{\mathrm{T}} \tag{10}$$

$$\mathbf{f}_\ell = \begin{bmatrix}\tilde{f}_\ell, \ \tilde{f}_\ell + D_f, \ \ldots, \ \tilde{f}_\ell + (\tilde{N}_f - 1)D_f\end{bmatrix}^{\mathrm{T}} \otimes \mathbf{1}_{\tilde{N}_a} \tag{11}$$

to be used to sample $\mathbf{c}_{s,\ell}$ from $\mathbf{C}$ as

$$\mathbf{c}_{s,\ell} = \begin{bmatrix}\mathbf{C}_{a_{\ell,1},f_{\ell,1}}, \ \mathbf{C}_{a_{\ell,2},f_{\ell,2}}, \ \ldots, \ \mathbf{C}_{a_{\ell,M},f_{\ell,M}}\end{bmatrix}^{\mathrm{T}}, \tag{12}$$

where $\tilde{f}_\ell$ and $\tilde{a}_\ell$ are the respective initial indices, $\tilde{N}_f = \lceil A_f/D_f \rceil$ and $\tilde{N}_a = \lceil A_a/D_a \rceil$ the numbers of subcarriers and antennas per sub-array, and $\mathbf{1}_{\tilde{N}_f}$ and $\mathbf{1}_{\tilde{N}_a}$ are all-ones vectors of length $\tilde{N}_f$ and $\tilde{N}_a$. In total, $L = \tilde{S}_f\tilde{S}_a$ sub-arrays can be obtained, where $\tilde{S}_f = \lfloor (N - A_f)/S_f + 1 \rfloor$ and $\tilde{S}_a = \lfloor (K - A_a)/S_a + 1 \rfloor$ denote the number of different subcarrier and antenna sets as a result of striding. Referring to Fig. 1, one could now subsample the initial full sub-arrays by selecting only the *colored* elements. After generating $\tilde{\mathbf{C}}$ in this way, the 2D MUSIC spectrum is estimated using (6) -

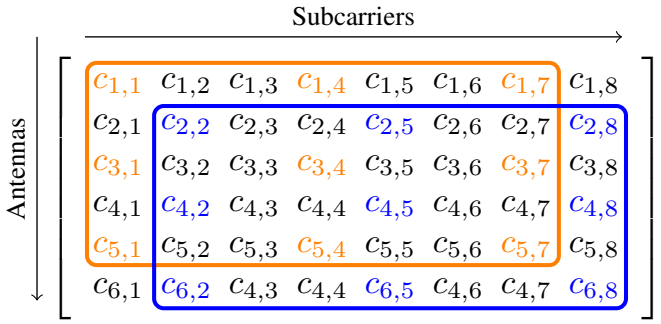


Fig. 1: Illustration of two different sub-arrays obtained by proposed sub-array generation with apertures $A_f = 7$ and $A_a = 5$, decimations $D_f = 3$ and $D_a = 2$, and strides $S_f = S_a = 1$.

(8), with the steering vectors being dependent on the sub-array definition parameters

$$\tilde{\mathbf{b}}_{\phi_a}(\theta) = \left[1,\ e^{j\phi_a \sin(\theta)},\ \ldots,\ e^{j\phi_a \tilde{N}_a \sin(\theta)}\right]^{\mathrm{T}} \tag{13}$$

$$\tilde{\mathbf{a}}_{\phi_f}(r) = \left[1,\ e^{j\phi_f \cdot 2\frac{r}{c}},\ \ldots,\ e^{j\phi_f \tilde{N}_f \cdot 2\frac{r}{c}}\right]^{\mathrm{T}}, \tag{14}$$

where $\phi_a = 2\pi D_a \frac{d}{\lambda}$ and $\phi_f = -2\pi D_f \Delta f$ to ease notation. Due to decimation, the number of elements per sub-array M is reduced from $A_f A_a$ to $\tilde{N}_f \tilde{N}_a$. Therefore, an equal frequency domain aperture, and thus resolution, can be obtained with sub-arrays with ca. $D_f D_a$ times less elements, drastically reducing the required computational complexity. However, decimating in the subcarrier domain reduces the unambiguous range $r_{\max} = \frac{c}{2D_f \Delta f}$ by D_f. We account for this by choosing the ranges in (14) within $r_{\max}$ to avoid aliasing.

Given that we want to discriminate Q targets in a multi-dimensional space, at least Q independent sub-arrays are necessary to compute $\hat{\mathbf{R}}_y$. If the targets are not resolvable in one domain, e. g., they have the same range, we notice that striding only in the corresponding dimension, in that case subcarriers, does not generate independent measurements. We can then formalize a necessary condition for the separation of Q targets in a multi-dimensional space.

Theorem 1. *Consider signals generated by periodically sampling across V different dimensions of interest, e. g., subcarriers and antennas in Eq. (3). To separate Q targets, that can be solved in at least one dimension of interest, with multi-dimensional spatial smoothing, at least Q independent sub-arrays must be used with respect to* ***each dimension of interest****.*

Proof. We provide a proof without considering additive noise, since the extension of the proof for noisy scenarios is straightforward. We need to separate Q targets in a V-dimensional MUSIC spectrum. This means that the sample covariance matrix $\hat{\mathbf{R}}_y$ must be of rank Q, i. e., it has Q non-zero eigenvalues. We consider the worst-case scenario, that is that the Q targets have exactly equal coordinates in $V-1$ dimensions. However, they can be discriminated in dimension v, e. g., angle. Assume we have $Q' < Q$ independent sub-arrays, obtained by striding over dimension v, plus an arbitrary number of sub-arrays obtained by striding over the other $V-1$ dimensions. However, if the Q targets have exactly the same coordinates over the $V-1$ dimensions, all the sub-arrays generated by striding over the $V-1$ dimensions will be equal to $\mathbf{q}$, that is one of the Q' sub-arrays with only constant phase shifts, according to (3). Therefore, when we use these sub-arrays to estimate the sample covariance matrix in (7), their contribution will be the same of $\mathbf{q}$. This leads to a covariance matrix of rank Q'. Recalling that we need to have covariance matrices with at least Q non-zero eigenvalues, we must have that $Q' \geq Q$.

Let us consider a 2D example with Q targets at equal range r. According to (3)

$$\mathbf{C} = \sum_{q=1}^{Q} h_q \mathbf{b}(\theta_q)\mathbf{a}(r)^{\mathrm{T}} = \left(\sum_{q=1}^{Q} h_q \mathbf{b}(\theta_q)\right)\mathbf{a}(r)^{\mathrm{T}}\ .$$

If we consider two sub-arrays $\mathbf{C}_{s,1}$ and $\mathbf{C}_{s,2}$ obtained by selecting the fist sub-array and by then striding of S_f indices across subcarriers, respectively, we would have

$$\mathbf{C}_{s,1} = \left(\sum_{q=1}^{Q} h_q \tilde{\mathbf{b}}(\theta_q)\right)\tilde{\mathbf{a}}(r)^{\mathrm{T}}\ ,$$

$$\mathbf{C}_{s,2} = \left(\sum_{q=1}^{Q} h_q \tilde{\mathbf{b}}(\theta_q)\right)\tilde{\mathbf{a}}(r)^{\mathrm{T}} e^{-j2\pi S_f \Delta f \cdot 2\frac{r}{c}} = \mathbf{C}_{s,1} e^{jW}\ ,$$

where the dependency on ϕ_a and ϕ_f has been dropped from the sub-array steering vectors defined in (13)-(14) to improve readability. The real phase constant is defined as $W = -2\pi S_f \Delta f \cdot 2\frac{r}{c}$. Then, if these two sub-arrays are used to estimate $\hat{\mathbf{R}}_y$, one would have

$$\hat{\mathbf{R}}_y = \frac{1}{2}\left(\mathbf{C}_{s,1}\mathbf{C}_{s,1}^{\mathrm{H}} + \mathbf{C}_{s,2}\mathbf{C}_{s,2}^{\mathrm{H}}\right) =$$
$$= \frac{1}{2}\mathbf{C}_{s,1}\mathbf{C}_{s,1}^{\mathrm{H}}\left(1 + e^{jW}e^{-jW}\right) = \mathbf{C}_{s,1}\mathbf{C}_{s,1}^{\mathrm{H}}\ .$$

It is clear that the contribution of the two sub-arrays is completely dependent, thus they will contribute to generate a single eigenvalue. It follows that one would need to stride across the antenna dimension at least Q times to generate Q independent contributions, thus the necessary Q eigenvalues to separate them in the angular domain. □

B. Peak Search Routine

In order to find targets in the 2D space, the maxima of the MUSIC spectrum defined in (8) need to be found. Avoiding an exhaustive grid search and allowing to detect multiple targets at once are the motives behind a proper peak search routine. To find peaks in the spectrum, Powell's algorithm [14] implemented in Python's SciPy library is utilized. For the algorithm to converge to the 2D MUSIC spectrum maxima, the proper selection of the search starting points is critical. For this, we first evaluate (8) in a coarse grid, where the sampling period in each dimension is equal to one half of the corresponding domain resolution. The locations of the N_{start} highest MUSIC spectrum values are selected as starting points. Each of them is

used by the Powell search algorithm to determine a peak that is compared against a constant false alarm rate (CFAR) noise threshold γ defined by a predetermined probability of false alarm p_{FA} [17]. If the peak value is below γ, it is discarded.

C. Coherent Target Cancelation

To check the spectrum iteratively for remaining targets, their contribution to the MUSIC spectrum should be removed. We first construct the channel contribution of the detected target

$$\mathbf{c}_q = \tilde{\mathbf{b}}_{\phi_a}(\hat{\theta}_q) \otimes \mathbf{a}_{\phi_f}(\hat{r}_q), \tag{15}$$

where $\hat{\theta}_q$ and $\hat{r}_q$ are the AoA and range estimates of q. Then, $\mathbf{c}_q$ is orthogonalized w. r. t. to all eigenvectors spanning $\mathbf{U}_N$

$$\tilde{\mathbf{c}}_q = \mathbf{c}_q - \mathbf{U}_N \mathbf{U}_N^{\mathrm{H}} \mathbf{c}_q. \tag{16}$$

After normalizing the L2-norm

$$\tilde{\mathbf{c}}_{q,\mathrm{norm}} = \frac{\tilde{\mathbf{c}}_q}{||\tilde{\mathbf{c}}_q||}, \tag{17}$$

$\tilde{\mathbf{c}}_{q,\mathrm{norm}}$ is added to $\mathbf{U}_N$ to obtain the updated noise subspace $\tilde{\mathbf{U}}_N = \left[\mathbf{u}_1,\ \mathbf{u}_2,\ \ldots,\ \tilde{\mathbf{c}}_{q,\mathrm{norm}}\right]$ to be used in (8) to re-estimate the spectrum. Note that this subspace tracking approach does not require to re-calculate the EVD of $\hat{\mathbf{R}}_y$ [15]. Fig. 2a shows an example, where the detected target of the initial spectrum on the left (red cross) is canceled out coherently and thus its contribution does not show up in the updated spectrum on the right. However, in Fig. 2b it can be observed that if two targets are neither resolvable in range nor in angle, removing the contribution of one target leads to a slight displacement of the remaining target w. r. t. to its actual location. While this peak can still be detected, the displacement naturally degrades the estimation accuracy. Using multiple starting points helps circumventing this issue, as it enables detecting close peaks. Nonetheless, cancelation enhances the detection capabilities, as it facilitates detecting weak targets after removing the contribution of much stronger ones.

D. Investigated Peak Selection Routines

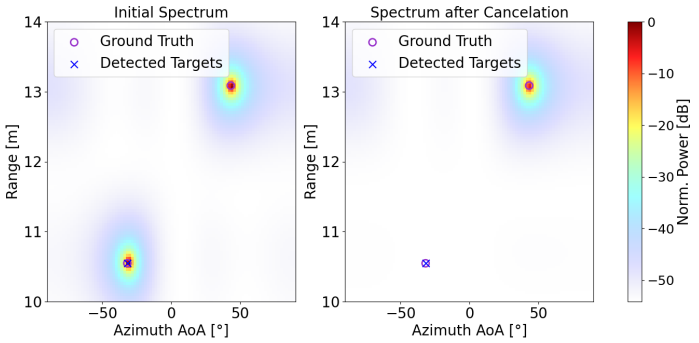
Combining the optionalities to either detect one or multiple targets at once and to iteratively remove their contributions, we define three different peak selection routines to be investigated.

• *Single.* Only the strongest point (i. e., $N_{\mathrm{start}} = 1$) of the pre-computed grid is used as a starting point for the peak search. If this leads to a detection, the contribution of the target is canceled out before re-computing the spectrum for the next iteration.

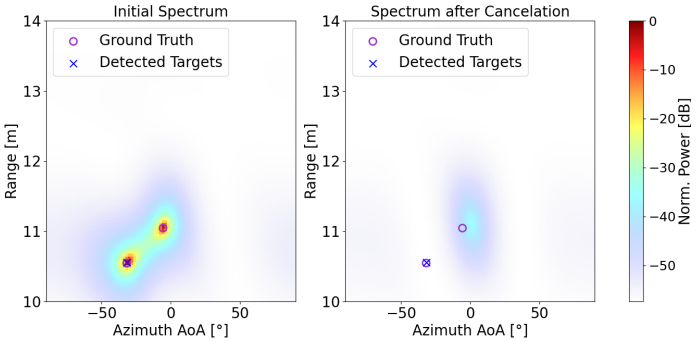
• *Multiple.* To be able to detect multiple targets at each iteration, the N_{start} strongest points of the coarse grid are used as starting points. All detected targets are then removed successively before re-estimating the spectrum.

• *Off.* This mode also considers $N_{\mathrm{start}} > 1$, but no coherent target cancelation is performed. Therefore, computation time is saved at the expense of detection capability.

Routines *single* and *multiple* iterate until no more targets are found, whereas *off* only performs a single peak detection iteration.



(a) Resolvable targets.



(b) Non-resolvable targets.

Fig. 2: Comparison of 2D MUSIC spectra before (left) and after (right) cancelation of a target. In (a), targets are resolvable, whereas in (b) the residual spectrum is displaced from the true target due to the targets being non-resolvable.

IV. Results and Discussion

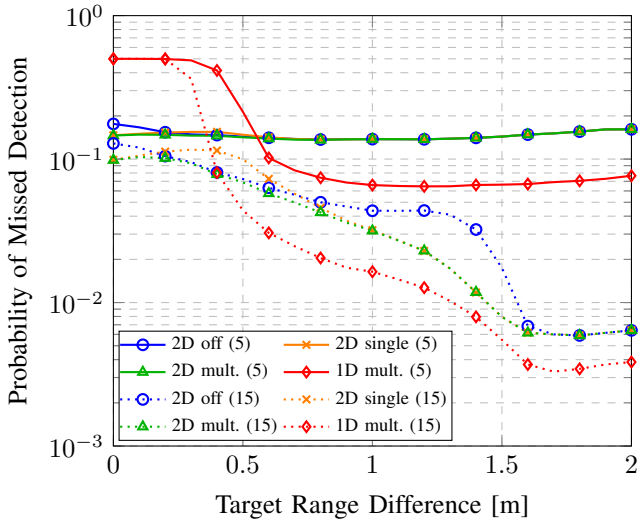
A. 1D vs. 2D Estimation

In our simulation study, we consider two targets initially placed at J = 10000 random positions resulting in the same range between target and RX ULA (within $r_{\max}$). From there, one target is incrementally moved away from the RX such that the target range difference is increased by 0.1 m. For each range difference, the same random AoAs uniformly drawn between -60° and 60° are kept. The signal-to-noise ratio (SNR) at RX anntena k is defined as SNR $= \frac{\mathbb{E}(\mathbf{y}_k^2)}{\sigma_N^2}$, where σ_N^2 is the noise variance. Moreover, we consider free-space path loss attenuation in our simulations. Table I lists the 5G compliant OFDM signal specification, the chosen parameters for our proposed algorithm and the resulting radar characteristics. For our simulations we keep to 4 antennas in the ULA with $\lambda/2$ element spacing and only decimate in the frequency domain (i. e., D_a = 1) to avoid spatial aliasing. According to Theorem 1, we choose $A_a = 3$ to allow striding in the spatial domain.

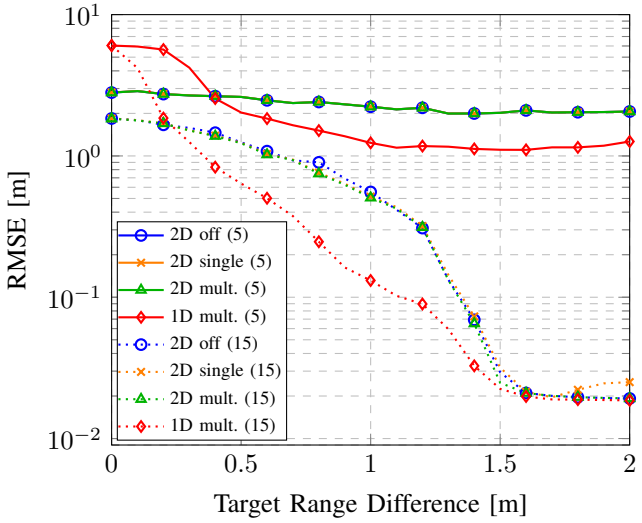
We define the root-mean-square error (RMSE) for parameter Θ (either range r or azimuth AoA θ) as

$$\mathrm{RMSE} = \sqrt{\frac{1}{2J}\sum_{j=1}^{J}\sum_{q=1}^{2}|\hat{\Theta}_q - \Theta_q|^2}, \tag{18}$$

where $\hat{\Theta}_q$ and Θ_q are estimate and true value of the respective parameter for target q. For computing the RMSE, we put the



(a) Probability of Missed Detection.



(b) Range RMSE.

Fig. 3: Comparison of missed detection probability and range RMSE performance between 1D and 2D estimation for increasing range differences between two targets at SNRs of 5 dB and 15 dB (in parentheses of legend).

TABLE I: Simulation Parameters.

Parameter	Value
Number of subcarriers N	1500
Carrier frequency f_c	3.5 GHz
Subcarrier spacing Δf	60 kHz
Number of RX antennas K	4
Antenna spacing d	$\lambda/2 \approx 0.05$ m
Frequency aperture A_f	1401
Frequency decimation D_f	100
Antenna aperture A_a	3
Antenna decimation D_a	1
Frequency stride S_f	1
Antenna stride S_a	1
Starting points for peak search N_{start}	10
Range resolution Δr	1.78 m
Unambiguous range $r_{\max}$	25 m

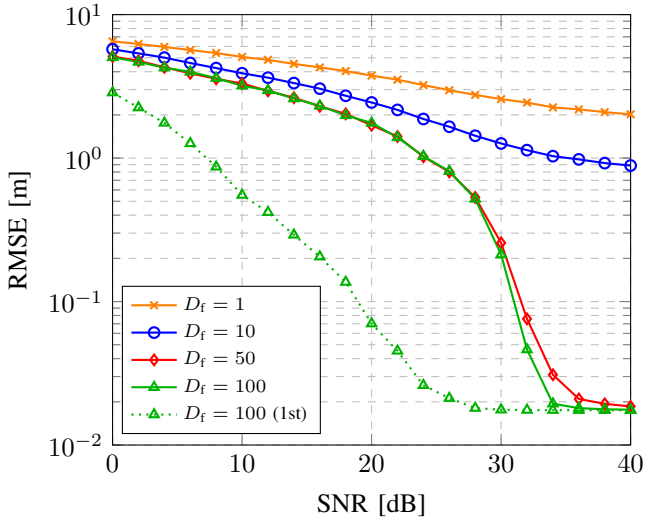
main focus on the target closer to the ULA (first target), i. e., in case of a single detection we compute the error to the first target. If both targets are detected, the estimate $\hat{r}_q$ that is closer to the ULA is assigned to the first target and the remaining estimate to the second target. Note that we assign detected targets to true targets only based on the range estimate. To be able to compute errors in case of missed target detections, we either use the maximum of the pre-computed coarse grid as the estimate for the first target (in case of zero detections) or remove the contribution of the first detected target and use the maximum of the resulting coarse grid as the estimate for the second target. Furthermore, we discard the 1% highest and lowest errors for the RMSE calculation, as outliers can be removed with proper tracking techniques [2].

First, the general advantages of joint estimation are shown by outlining the performance differences of 1D range-only estimation with MUSIC to our proposed algorithm. For the 1D estimation we used the same parameters of Table I, apart from choosing $A_a = 1$.

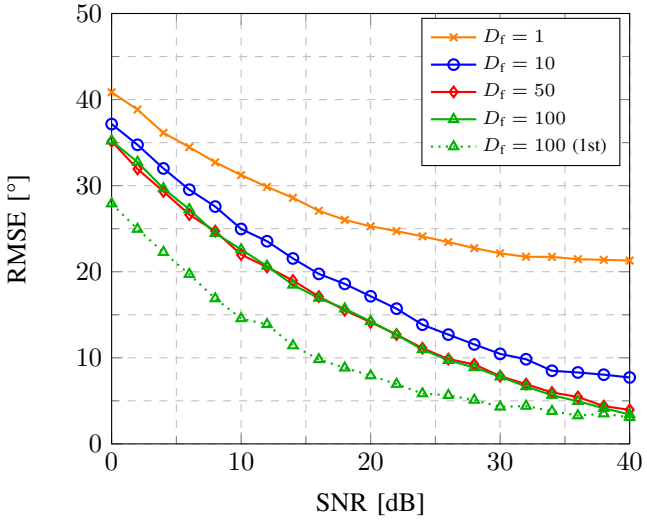
Fig. 3a shows the probability of missed detection for 1D range-only and 2D estimation at SNRs of 5 dB and 15 dB. To limit the number of curves we only plot peak selection routine *multiple* for 1D estimation. It can be seen that for both SNRs, the 2D estimation methods already achieve low probabilities of missed detection for targets with the same range, whereas the 1D range-only estimation can only detect a single target in such cases. Nonetheless, 1D estimation is more robust for higher target range differences, especially at an SNR of 5 dB. Observing the performance at 15 dB clarifies the benefit of recomputing the spectrum after removing the contributions of detected targets, as *2D off* (no target cancelation) shows the worst performance. Once the targets are fully resolvable in range, the 2D estimation routines attain missed detection probabilities of roughly 0.006, while *1D multiple* performs slightly better and achieves circa 0.004. Overall, *2D multiple* exhibits the best performance of the investigated 2D routines. The corresponding range RMSE curves in Fig. 3b show similar progressions. All 2D peak selection routines display a comparable RMSE performance. At 15 dB, both 1D and 2D techniques converge to an error floor of roughly 0.02 m.

B. Benefits of CSI Decimation

To demonstrate the practical benefits of decimating in the subcarrier domain, we compare the parametrization in Table I against setups with i) $A_f = 15$, $D_f = 1$, ii) $A_f = 141$, $D_f = 10$, and iii) $A_f = 701$, $D_f = 50$. The number of samples per sub-array is identical ($M = 45$), but the achievable range resolution is reduced by a factor equal to the reduction in A_f. To keep the computational effort equal, $L = 200$ sub-arrays were used for all setups. For this experiment we place the two targets at random positions within $r_{\max}$ without a fixed range difference and only investigate peak selection routine *multiple*.



(a) Range RMSE.



(b) Azimuth RMSE.

Fig. 4: Range and azimuth RMSE curves for decimations $D_f = 1, D_f = 10, D_f = 50$, and $D_f = 100$ and routine *multiple*.

Figs. 4 and 5 show that a higher range resolution improves both the RMSE and the missed detection probability performance significantly. Nonetheless, one can observe that a higher SNR is necessary to achieve similiar missed detection probability and RMSE capabilities as in Fig. 3. This can be explained by trials where the targets are either placed so close that they can not be resolved in either domain, or so far apart that the second target can not be detected due to the path-loss attenuation making them fall below the detection threshold. We therefore as a reference included the curves for the first target only (*1st*), which for $D_f = 100$ (corresponding to *2D multiple* in Fig. 3) converge to about 0.01 m and 3°, respectively.

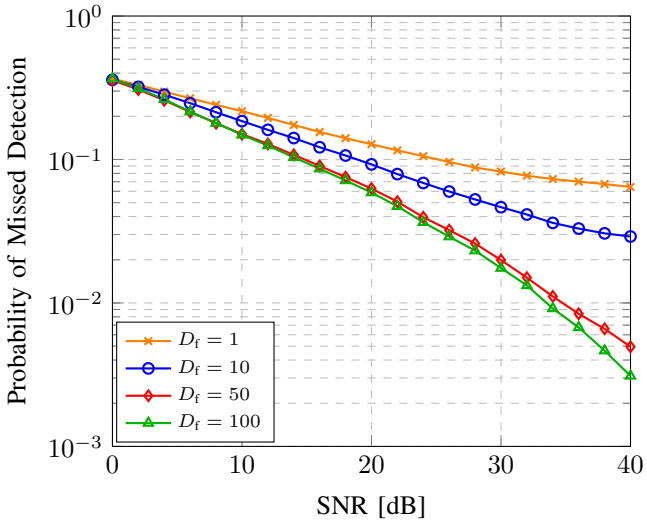


Fig. 5: Probability of missed detection for decimations $D_f = 1, D_f = 10, D_f = 50$, and $D_f = 100$ and routine *multiple*.

C. Computational Complexity Comparison

Finally, we demonstrate the savings in computational complexity due to decimation by comparing the signal parametrization in Table I (i. e., $D_f = 100$) with the smoothing technique without decimation (i. e., $D_f = 1$) [8]. The two setups result in $M = 45$ and $M = 4203$ elements per sub-array, respectively. The number of floating-point operations (FLOPs), i. e., complex multiplications and summations, for a single evaluation of (8) can be approximated as $2M^2(M-Q)$, thus $O(M^3)$, given that Q is small enough. This leads to ≈ 148 GFLOPs for $D_f = 1$ and ≈ 174 kFLOPs for $D_f = 100$. The numbers show that we can reduce the required FLOPs by a factor of $\approx 8.5 \cdot 10^5$ in a typical 5G system operating at 3.5 GHz with 90 MHz bandwidth. The above-mentioned gain comes at the price of a reduced unambiguous range, which is why the system must be parametrized according to the a-priori confidence of the targets' rough locations. Moreover, a slight processing gain enhancement due to the higher number of sub-array elements could be achieved by not decimating.

V. Conclusion

In this paper we presented an algorithm for the joint estimation of range and azimuth AoA using MUSIC and OFDM signals. CSI decimation significantly reduces the required number of computations for the considered 5G scenario and therefore makes joint multi-dimensional estimation feasible in practical systems. We demonstrated that our joint approach improves the probability of detection for closely spaced targes compared to 1D range-only estimation while achieving similar range errors. Moreover, we have proposed efficient single and multi-target detection algorithms, leading to an enhanced detection capability.

Acknowledgments

The authors gratefully acknowledge helpful discussions with Thorsten Wild, Stephan Saur and Traian Emanuel Abrudan.

REFERENCES

[1] T. Wild, V. Braun, and H. Viswanathan, "Joint design of communication and sensing for beyond 5G and 6G systems," *IEEE Access*, vol. 9.

[2] S. Saur, M. Mizmizi, J. Otterbach, T. Schlitter, R. Fuchs, and S. Mandelli, "5GCAR demonstration: Vulnerable road user protection through positioning with synchronized antenna signal processing," in *WSA 2020; 24th International ITG Workshop on Smart Antennas*.

[3] C. Geng, T. E. Abrudan, V.-M. Kolmonen, and H. Huang, "Experimental study on probabilistic ToA and AoA joint localization in real indoor environments," *arXiv preprint arXiv:2102.11233*, 2021.

[4] C. Sturm, T. Zwick, and W. Wiesbeck, "An OFDM system concept for joint radar and communications operations," in *VTC Spring 2009-IEEE 69th Vehicular Technology Conference*.

[5] R. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE transactions on antennas and propagation*, vol. 34, no. 3.

[6] M. C. Vanderveen, C. B. Papadias, and A. Paulraj, "Joint angle and delay estimation (JADE) for multipath signals arriving at an antenna array," *IEEE Communications letters*, vol. 1, no. 1.

[7] L. Taponecco, A. A. D'Amico, and U. Mengali, "Joint TOA and AOA estimation for UWB localization applications," *IEEE Transactions on Wireless Communications*, vol. 10, no. 7.

[8] M. Kotaru, K. Joshi, D. Bharadia, and S. Katti, "SpotFi: Decimeter level localization using WiFi," in *Proceedings of the 2015 ACM Conference on Special Interest Group on Data Communication*.

[9] T.-J. Shan and T. Kailath, "Adaptive beamforming for coherent signals and interference," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 33, no. 3.

[10] 3GPP, "NR; Physical channels and modulation," Technical Specification (TS) 38.211, 07, version 16.2.0.

[11] J. B. Sanson, P. M. Tomé, D. Castanheira, A. Gameiro, and P. P. Monteiro, "High-resolution delay-doppler estimation using received communication signals for OFDM radar-communication system," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 11.

[12] K. V. Rangarao and S. Venkatanarasimhan, "Gold-MUSIC: A variation on MUSIC to accurately determine peaks of the spectrum," *IEEE Transactions on Antennas and Propagation*, vol. 61, no. 4.

[13] J.-f. Chen and H. Ma, "An accurate real-time algorithm for spectrum peaks search in 2D MUSIC," in *2011 International Conference on Multimedia Technology*.

[14] M. J. Powell, "An efficient method for finding the minimum of a function of several variables without calculating derivatives," *The computer journal*, vol. 7, no. 2.

[15] B. Yang, "Projection approximation subspace tracking," *IEEE Transactions on Signal processing*, vol. 43, no. 1.

[16] J. Rissanen, "Modeling by shortest data description," *Automatica*, vol. 14, no. 5, pp. 465–471, 1978.

[17] K. M. Braun, "OFDM radar algorithms in mobile communication networks," Ph.D. dissertation, KIT-Bibliothek, 2014.