Multilinear Singular Value Decomposition for Millimeter Wave Channel Parameter Estimation

MACEY RUBLE, AND İSMAIL GÜVENÇ, (Senior Member, IEEE)
Department of Electrical and Computer Engineering, North Carolina State University, Raleigh, NC 27695, USA
Corresponding author: Macey Ruble (mcruble@ncsu.edu)

This work was supported in part by the National Aeronautics and Space Administration through the Federal Award under Grant NNX17AJ94A.

ABSTRACT Fifth generation (5G) cellular standards are set to utilize millimeter wave (mmWave) frequencies, which enable data speeds greater than 10 Gbps and sub-centimeter localization accuracy. These capabilities rely on accurate estimates of the channel parameters, which we define as the angle of arrival, angle of departure, and path distance for each path between the transmitter and receiver. Estimating the channel parameters in a computationally efficient manner poses a challenge because it requires estimation of parameters from a high-dimensional measurement—particularly for multi-carrier systems since each subcarrier must be estimated separately. Additionally, channel parameter estimation must be able to handle hybrid beamforming, which uses a combination of digital and analog beamforming to reduce the number of required analog to digital converters. This paper introduces a channel parameter estimation technique based on the multilinear singular value decomposition (MSVD), a Tucker form tensor analog of the singular value decomposition, for massive multiple input multiple output (MIMO) multi-carrier systems with hybrid beamforming. The MSVD tensor estimation approach is more computationally efficient than methods such as the canonical polyadic decomposition (CPD) and the Tucker form of the MSVD enables paths to be extracted based on signal energy. The algorithms performance is compared to the CPD method and shown to closely match the Cramer-Rao bound (CRB) of channel parameter estimates through simulations. Additionally, limitations of channel parameter estimation and communication waveform effects are studied.

INDEX TERMS AOA, AOD, channel estimation, MSVD, massive MIMO, mmWave, OFDM, SVD, tensor estimation.

I. INTRODUCTION Fifth generation (5G) cellular networks are set to deploy millimeter wave (mmWave) technology with carrier frequencies ranging from 30 GHz to 300 GHz [1]. The small wavelengths associated with mmWave frequencies allow massive multiple-input-multiple-output (MIMO) arrays to be deployed in small spaces [2]. This along with the large available bandwidth enables mmWave technology to reach data rates over 10 Gbps [1] while also offering sub-centimeter receiver positioning capability [3].

High performance communication and localization for mmWave technology are both dependent on accurate estimates of the channel parameters, which we define as the angle of arrival (AOA), angle of departure (AOD), and total transmitter to receiver path distance for each significant path. This dependence is a result of highly directional beams and large path reflection attenuation, which leads to few significant received paths and a sparse channel that is completely characterized by the channel parameters [4]. Channel estimation for mmWave is accomplished by transmitting a known training sequence so that the environmental impacts on the channel can be determined and accounted for when unknown data sequences are transmitted [3], [5]–[8]. The channel parameters from multiple paths can also be simultaneously utilized to estimate a receiver’s position and orientation while also mapping the scatterers in the environment [3], [9].

5G networks will use orthogonal frequency division multiplexing (OFDM) or OFDM-like waveforms, which enable broadband communication by distributing the transmitted data over many subcarrier frequencies [3], [5], [10], [11]. Channel parameter estimation is particularly challenging for MIMO OFDM systems since the channel is different at each subcarrier. An antenna with an analog to digital
converter (ADC) at every single antenna element becomes too costly for large arrays. Thus, a channel parameter estimation method must function with hybrid digital/analog beamforming, which uses digital beamforming as well as phased array analog beamforming to reduce the overall number of required ADCs [5], [6], [12], [13]. Adding to the challenge is that the number of significant received paths is unknown and needs to first be estimated. Furthermore, a collection of AOA, AOD, and path distance estimates is not a complete solution: a complete solution must also link each estimated channel parameter to particular paths.

A. LITERATURE REVIEW
Massive MIMO channel parameter estimation has received much recent attention in the literature [3], [4], [6], [14]–[22], but it still remains a challenge to utilize all of the subcarriers simultaneously for OFDM systems and process the high-dimensional data that streams from the massive MIMO arrays. The work focuses on channel parameter estimation where the transmitter and receiver are all on the same plane. However, these methods may not scale to higher dimensional scenarios when both the azimuthal and elevation angles for AOD and AOA must be considered.

High accuracy mmWave localization enables a new era of user tracking and augmented reality [4], [23]–[26]. The work in [3] estimates channel parameters and shows that the channel parameters for a few paths are sufficient to estimate a receiver’s position and orientation. Additionally, mmWave non-line of sight (NLOS) paths are not treated as interference, but rather as additional paths that carry useful information [4], [27]. This enables the reflection locations to be estimated from the channel parameters; making simultaneous localization and mapping (SLAM) possible, where the receiver is localized while the environment is mapped in parallel [9], [26].

B. RELATED WORK
It has been shown that tensor decomposition provides an efficient method for high dimensional parameter estimation in many applications [28]–[31]. Of particular importance is the multilinear singular value decomposition (MSVD), which is a tensor analogue of the singular value decomposition (SVD) commonly seen in linear algebra. The MSVD allows computationally efficient parameter estimation if the tensor is low rank and is often used in machine learning [30], [31]. Another common decomposition for low rank tensors is the canonical polyadic decomposition (CPD) [32]. The CPD form essentially guarantees uniqueness provided that Kruskal’s criteria are met [33]. On the other hand, the MSVD finds a Tucker tensor form, which is not unique. An introduction and comparison of the MSVD and CPD decompositions can be found in [30], [34] along with applications to signal processing and machine learning.

The work in [35] shows that OFDM MIMO receiver measurements fit a low rank tensor form and channel parameter estimation is achieved by estimating a best-fit CPD tensor form. However, the CPD tensor form is constructed for a particular rank, or number of paths, which is unknown. Additionally, the CPD form does not allow multiple paths to share the same channel parameters. Thus, two paths with the same AOA can not be represented by the CPD form.

C. CONTRIBUTIONS
In this work, we:

- Propose an alternative method to [35] for MIMO OFDM channel parameter estimation, where the measurements are not represented in CPD form, but instead a Tucker form achieved by truncating a MSVD of the measurements.
- Introduce a computationally efficient technique that determines the number of paths and estimates the channel parameters with super-resolution by separating each of them into an independent low dimensional subproblem.
- Provide a method for linking the estimated channel parameters to ensure that the correct sets of parameters are linked to each path.
- Derive the Cramer-Rao bound (CRB) performance bound for each of the path parameters.
- Compare the proposed channel parameter estimation method to the CRB bound and the CPD method.
- Study channel parameter estimation accuracy relative to waveform specifications including antenna numbers, OFDM symbol interval, and number of subcarriers.

While a truncated MSVD form is not guaranteed to obtain the optimal lower rank tensor estimate, it is often considered to be good enough in practice with the advantages of reduced computational complexity and speed of convergence [36]. Thus, it can provide similar results to the CPD method, but with significantly less computation. Our results show that the MSVD technique closely matches the CRB and CPD method. Another advantage of the Tucker form and MSVD estimation method over the CPD method is that it offers a simple method to estimate the channel rank and number of paths. Additionally, it allows different paths to share channel parameters.

D. PAPER ORGANIZATION
This paper is organized as follows. Section II lays out the MIMO OFDM channel model in Tucker tensor form. Section III-A formulates the channel parameter estimation problem from Tucker tensor form receiver measurements. Then, Section III introduces the MSVD channel parameter estimation technique. Following this, Section IV discusses how the channel parameters can be used for channel estimation and localization followed by how waveform parameters affect channel parameter estimation accuracy. Subsequently, Section V provides simulation results of channel parameter estimation and compares the results to the CRB bound. Finally, Section VI provides concluding remarks.
II. BROADBAND MIMO OFDM MODEL

This section covers the model of an OFDM MIMO system. First, the Tucker tensor form is introduced. Then, a channel model is given in a general format representative of hybrid digital/analog beamforming architectures. Finally, it is shown that the MIMO OFDM channel is naturally represented by a low rank Tucker tensor model.

A. TUCKER TENSOR FORM

To be self-contained, we first cover prerequisite tensor knowledge. Then, it is shown that the received signal in a MIMO OFDM system naturally groups into a Tucker tensor form. Our notation is consistent with [30], which can also be referred to for further information. We first define the tensor product or outer product. The third order tensor product results in a tensor, which is defined such that tensor elements resulting from the product between the vectors $a$ and $b$, and $c$ are $(a \otimes b \otimes c)(i_1, i_2, i_3) = a(i_1)b(i_2)c(i_3)$, where $i_1$, $i_2$, and $i_3$ are element indices with respect to the tensor or vector form. Tensor products of higher dimension follow similarly.

Any tensor can be represented in Tucker form. For a third order $M_1 \times M_2 \times M_3$ tensor $T$, the Tucker form is [30]:

$$T = \sum_{i_1}^{M_1} \sum_{i_2}^{M_2} \sum_{i_3}^{M_3} G(i_1, i_2, i_3) V_1(:, i_1) \otimes V_2(:, i_2) \otimes V_3(:, i_3),$$

where $G$ is a core tensor, and $V_1$, $V_2$, $V_3$ are matrices composed of column vectors such that $V(:, i)$ is used to represent the $i$th column of matrix $V$. The core element $G(i_1, i_2, i_3)$ corresponds with the strength of the interaction between column $i_1$ in $V_1$, column $i_2$ in $V_2$, and column $i_3$ in $V_3$. Fig. 3 visualizes $T$ as a series of vector tensor products. The Tucker tensor form in (1) can also be written in a shorthand notation as follows:

$$T = G \otimes_1 V_1 \otimes_2 V_2 \otimes_3 V_3,$$

where $\otimes_i$ represents a tensor product along the $i$th dimension with the core tensor $G$. Fig. 4 visualizes this form. Now that the Tucker form has been introduced, the following subsections show that the MIMO OFDM received signal naturally groups into a Tucker tensor form.

B. MIMO OFDM CHANNEL MODEL

A MIMO OFDM system is considered with $N_{tx}$ transmit antennas and $N_{rx}$ receive antennas. Array element numbers are assumed odd for notational simplicity, but results are equivalent for even numbers. The data is distributed over a bandwidth $B_{OFDM}$ between $N_s$ subcarriers at frequencies $f_k = f_c + k/T_{OFDM}$, for $k = 0, \ldots, N_s - 1$, where $f_c$ is the carrier frequency and $T_{OFDM} = N_s T_s$ is the OFDM symbol duration, where $T_s = 1/B_{OFDM}$ is the sampling interval.

Fig. 1 shows the transmitter, channel, and receiver model, where the data $X[k] \in \mathbb{C}^{N_{tx} \times N_s}$ is divided into $L_{tx}$ data streams of length $N_T$ and precoded with a digital precoder $F_D \in \mathbb{C}^{N_{tx} \times N_{tx}}$. A $N_T$-point inverse fast Fourier transform (IFFT) is applied to convert the data to the time domain where the output of an IFFT block is

$$s(t) = \sum_{k=0}^{N_s-1} x(k) e^{2\pi j (f_c + k/T_{OFDM}) t}, \quad 0 \leq t \leq T_{OFDM}. \quad (3)$$

Following this, a cyclic prefix is added to suppress inter-symbol interference (ISI), but is not shown in Fig. 1. Then, the transmitter employs an analog precoder $F_{RF} \in \mathbb{C}^{N_{tx} \times N_{tx}}$ and the signal is transmitted over $N_{tx}$ antennas. The signal is received by $N_{rx}$ receiving antennas followed by an analog combiner $W_{RF} \in \mathbb{C}^{N_{rx} \times N_{tx}}$. A fast Fourier transform (FFT) is then employed as the inverse of the IFFT block before a digital combiner $W_D \in \mathbb{C}^{N_{rx} \times N_{rx}}$ converts the signal to $L_{rx} \leq N_{rx}$ data streams to obtain the received signal $Y$.

The model in Fig. 1 can be represented at each subcarrier by [3], [5]:

$$Y[k] = W[k]^H H[k] F[k] X[k] + n[k]. \quad (4)$$

for $k = 1, \ldots, N_c$, where $k$ is the subcarrier and $Y[k] \in \mathbb{C}^{L_{rx} \times N_s}$ is the received signal at each subcarrier. The matrix $F[k] = F_{RF} F_D \in \mathbb{C}^{N_{tx} \times L_{tx}}$ is the precoding matrix, $H[k] \in \mathbb{C}^{N_s \times N_{tx}}$ is the channel matrix, $W[k] = W_{RF} W_D[k] \in \mathbb{C}^{L_{tx} \times N_{rx}}$ is the combiner matrix, and $n[k] \in \mathbb{C}^{N_{rx} \times N_{rx}}$ is noise. The precoder and combiner matrices $F[k]$ and $W[k]$ can be designed to improve channel estimation [6]. The received signal to noise ratio (SNR) is defined as:

$$\text{SNR} = \frac{\|W[k]^H H[k] F[k] X[k]\|_F^2}{\|n[k]\|_F^2}, \quad (5)$$

where $\| \cdot \|_F$ is the Frobenius norm [11]. It is noted that the received SNR depends on beamforming as well as the channel and can change even with a fixed transmitted SNR.

The highly directional beams in large scale MIMO systems result in few received significant paths [37]. Furthermore,
NLOS paths can be assumed single bounce because multiple bounce NLOS paths will have large attenuation and much lower signal strength [15]. This results in a sparse channel, which can be expanded in terms of the individual received paths [4]. We assume $N_p$ significant paths are received, where the $n^{th}$ path geometry is seen in Fig. 2. The path begins at the transmitter array located at $q$ with array orientation angle $\theta_\text{tx}$, reflects at location $r$, and ends at a receiver array with location $p$ and array orientation angle $\theta_\text{rx}$. The path geometry created by the transmitter, reflector, and receiver locations dictate the parameters $\theta_\text{tx,n}$: the angle of departure at the transmitter, $\theta_\text{rx,n}$: the angle of arrival at the receiver, and $d_n$: the total distance traveled by the path from transmitter to receiver. It is noted that the path in Fig. 2 characterizes both LOS and NLOS paths since a LOS path can be characterized by a reflector anywhere on the line segment between $q$ and $p$. The channel from (4) is expressed in terms of the $N_p$ significant paths [3], [18]:

$$H[k] = \sum_{n=0}^{N_p-1} h_n(k) \exp\left\{ -j2\pi k \tau_n \right\} a_\text{tx}(\theta_\text{tx,n})^H a_\text{rx}(\theta_\text{rx,n}),$$  \hspace{1cm} (6)$$

where $h_n(k)$ is the path loss and $\tau_n = d_n/c$ is the time to travel from the center of the transmitter array to the center of the receiver array, with $c$ as the speed of light. It is assumed that both transmitter and receiver have antenna spacing $a = \lambda_c/2$, where $\lambda_c$ is the wavelength associated with the center frequency. Then, the beamforming vectors $a_\text{tx}$ and $a_\text{rx}$ are:

$$a_\text{tx}(\theta_\text{tx,n}) = \left[ e^{-j2\pi a_N (\sin(\theta_\text{tx,n}) + 1)} \ldots e^{j2\pi a_N (\sin(\theta_\text{tx,n}) + 1)} \ldots e^{j2\pi a_N (\sin(\theta_\text{tx,n}) + 1)} \right]^T,$$

$$a_\text{rx}(\theta_\text{rx,n}) = \left[ e^{-j2\pi a_N (\sin(\theta_\text{rx,n}) + 1)} \ldots e^{j2\pi a_N (\sin(\theta_\text{rx,n}) + 1)} \ldots e^{j2\pi a_N (\sin(\theta_\text{rx,n}) + 1)} \right]^T.$$

For sufficiently long symbols ($T_{\text{OFDM}}$) the path loss coefficient is equivalent at each subcarrier [3], [18]. Thus, it is assumed that $h_n(k) \approx h_n$ for all $k$. This assumption is discussed in further detail in Section IV-A.

By substituting $\tau_n = d_n/c$ in (6), it is seen that besides the path loss $h_n$, the channel is completely characterized by the channel parameters $\theta_\text{tx,n}, \theta_\text{rx,n}$, and $d_n$:

$$H[k] = \sum_{n=0}^{N_p-1} h_n a_\text{tx}(\theta_\text{tx,n}) a_\text{rx}(\theta_\text{rx,n}) H \phi(d_n)[k],$$  \hspace{1cm} (7)$$

where,

$$\phi(d_n)[k] = \exp\left\{ -j2\pi k d_n \right\}. \hspace{1cm} (8)$$

### C. THE CHANNEL IN TUCKER TENSOR FORM

Channel estimation and localization are typically performed during a training sequence interval, where the data $X$ is known [3], [38]. With knowledge of the data, the only unknowns are the channel parameters. Without loss of generality, we let $X$ be all zero, besides ones on the diagonal. Substituting $X$ and the channel representation from (7) into (4), the received signal for each subcarrier is

$$Y[k] = \sum_{n=0}^{N_p-1} h_n w_{a,n}(\theta_\text{tx,n}) f_{a,n}(\theta_\text{rx,n}) \phi(d_n)[k] + n[k],$$  \hspace{1cm} (9)$$

for $k = 1, \ldots, N_t$ where

$$w_{a,n}(\theta_\text{tx,n}) = W^H a_\text{tx}(\theta_\text{tx,n}),$$  \hspace{1cm} (10)$$

and

$$f_{a,n}(\theta_\text{rx,n}) = a_\text{rx}(\theta_\text{rx,n})^H F.$$

A vector is created from the terms $\phi(d_n)[k]$ that contains $\{\phi(d_n)[k]\}_{k=1}^{N_t}$ across all subcarrier frequencies:

$$\phi_n(d_n) = \left[ 1 \frac{e^{-j2\pi d_n}}{T_{\text{OFDM}}} \ldots \frac{e^{-j2\pi c N_t d_n}}{T_{\text{OFDM}}} \right]^T.$$  \hspace{1cm} (12)$$

Then, the measurement from (9) can be expanded to represent all subcarrier frequencies simultaneously as a third-order tensor with dimensions $L_T \times N_T \times N_t$ as follows:

$$Y = \sum_{n=0}^{N_p-1} h_n w_{a,n} \circ f_{a,n}^T \circ \phi_n + n,$$  \hspace{1cm} (13)$$

which is now in Tucker form with rank $N_p$ as seen in (1), where $n \in \mathbb{C}^{L_T \times N_T \times N_t}$ is the noise over all subcarriers. It is noted that the dependencies of the vectors on the parameters $\theta_\text{tx,n}, \theta_\text{rx,n}$, and $d_n$ have been dropped for simplicity in notation. The simplified Tucker form from (2) is obtained by grouping the vectors $w_{a,n}$ into the matrix $W_a = [w_{a,1} w_{a,2} \ldots ]$, the vectors $f_{a,n}$ into the matrix $F_a = [f_{a,1} f_{a,2} \ldots ]$, and the vectors $\phi_n$ into the matrix $\Phi = [\phi_1 \phi_2 \ldots ]$. Then, the measurement tensor is:

$$Y = \Psi \odot_1 W_a \otimes_2 F_a \otimes_3 \Phi + n,$$  \hspace{1cm} (14)$$

where $\Psi$ is a core tensor with only $N_p$ nonzero elements, corresponding with the path gain $h_n$ for each path. The Tucker form in (14) is different than the CPD form in [35] since it allows separate tensor components for path energy and each channel parameter.

This work focuses on third order tensors because we restrict paths to a plane and do not consider elevation angles.
for simplicity. The result is a third order measurement tensor where the column space corresponds with path AOA, the row space corresponds with path AOD, and the fiber space corresponds with path distance. A model that allows a three-dimensional path and considers both elevation angles and azimuthal path angles will lead to a five-dimensional measurement tensor. A significant advantage of the Tucker form is that it extends to higher dimensions. All of the derivations in this work are for three-dimensional tensors, but each step can easily be extended and applied to higher order tensors for path models that also consider elevation angle.

Accurate estimates of the channel parameters \( \theta_{l,n}, \theta_{l,n}, \) and \( d_n \) for \( n = 0, \ldots, N_p - 1 \) enable a number of capabilities for 5G networks. The most obvious capability is channel estimation, since \( H[k] \) can be reconstructed from the channel parameters for each subcarrier [6]. Additionally, the channel parameters can be exploited to estimate the receiver position and NLOS path reflection locations [9]. Thus, receiver localization and environmental mapping can be achieved with the knowledge of these parameters, where environmental mapping is obtained when the NLOS path reflection locations are estimated over several measurements. The estimate of \( d_n \) from \( \phi_n \) can also be used to estimate the time delay \( \tau_n \) between the transmitter and receiver, which can be used to assist synchronization between the transmitter and receiver [39].

III. MULTILINEAR SVD FOR mmWave CHANNEL PARAMETER ESTIMATION

This section first formulates the mmWave channel parameter estimation problem. Then, the MSVD is introduced and applied to a third order measurement tensor to create a reduced rank Tucker form and estimate the channel parameters.

A. PROBLEM FORMULATION FOR mmWave CHANNEL PARAMETER ESTIMATION

We are interested in estimating the channel parameters \( \theta_{l,n}, \theta_{l,n}, \) and \( d_n \) for \( n = 0, \ldots, N_p - 1 \) as well as the path gain \( h_n \) from the tensor form of the receiver measurement \( Y \) of a training sequence as in (14). This is accomplished by estimating \( \Psi, W_a, F_a, \) and \( \Phi \), which is the main goal of the rest of this paper. A number of challenges make this a difficult task. For example, the number of paths \( N_p \) is unknown and must also be estimated. Additionally, this is a high dimensional problem and the channel parameters must be estimated jointly.

Mathematically, since the measurement tensor is low rank, the channel parameter estimation problem can be posed in the following form:

\[
\hat{\Psi}, \hat{W}_a, \hat{F}_a, \hat{\Phi} = \arg \min_{\Psi, W_a, F_a, \Phi} f(\Psi, W_a, F_a, \Phi),
\]

\[
f(\Psi, W_a, F_a, \Phi) = ||Y - \Psi \otimes_1 W_a \otimes_2 F_a \otimes_3 \Phi||_F^2 + \lambda ||\text{vec}(\Psi)||_0,
\]

where \( || \cdot ||_0 \) is the \( \ell_0 \)-norm (number of non-zero terms), \( \lambda \) is a scaling parameter, and \( \Psi, W_a, F_a, \Phi \) are functions of \( h_n, \theta_{l,n}, \theta_{l,n}, \) and \( d_n \) respectively. The first term in (16) minimizes the Frobenius norm of the error and the second term enforces low rank in the estimated tensor reconstruction. The rank of the reconstructed tensor is the estimate for the number of paths and the second term ensures that a minimal number of paths are used for channel parameter estimation. Section III provides a method that uses the MSVD to estimate the number of significant paths and recover each of the channel parameters as well as the path gain in (13).

B. THE MULTILINEAR SINGULAR VALUE DECOMPOSITION

The MSVD is an extension of the singular value decomposition to tensors and reconstructs a third order tensor into a set of column, row, and fiber basis vectors. It reconstructs tensors into a Tucker form, so that the interaction energy between columns, row, and fibers can also be used to estimate the channel parameters as well as the path gain in (13).
It is noted that in the standard SVD, the column space and row space have the same rank. However, in the multilinear SVD, the column, row, and fiber spaces can have different ranks. For (17), the column rank is the number of columns in \(U^{(1)}\), the row rank is the number of columns in \(U^{(2)}\), and the fiber rank is the number of columns in \(U^{(3)}\).

C. RANK REDUCTION

It is known that few significant paths exist in the model from (13), which besides small noise interactions, makes \(Y\) a low rank tensor. On the other hand, the MSVD from (17) has a column rank of \(L_{rx}\) (number of data streams), a row rank of \(N_T\) (number of training symbols), and a fiber rank of \(N_s\) (number of subcarriers). A majority of the interactions from the MSVD must be eliminated to obtain a low rank representation of \(Y\). The strongest interactions mainly consist of energy from significant received paths and the weakest interactions are mainly composed of energy from non-significant paths and noise.

The rank of the MSVD (17) can be reduced by removing interactions that correspond with singular values below a threshold in each dimension. This is achieved by removing planes from \(\Sigma\), corresponding with the weak singular values along with the interacting row, column, and fiber vectors. This acts as a denoising process [28] and the remaining reduced rank Tucker form is

\[
Y = \Sigma_{s} \otimes_1 U_s^{(1)} \otimes_2 U_s^{(2)} \otimes_3 U_s^{(3)},
\]  

(18)

where if \(r_1\), \(r_2\), and \(r_3\) are the reduced ranks of the column, row, and fiber subspaces; then \(\Sigma_s\) is a \(r_1 \times r_2 \times r_3\) core tensor, \(U_s^{(1)}\) is a \(L_{rx} \times r_1\) matrix, \(U_s^{(2)}\) is a \(N_T \times r_2\) matrix, and \(U_s^{(3)}\) is a \(N_s \times r_3\) matrix.

While the MSVD does not guarantee a unique solution [30], it is often just as accurate, but much faster to compute than the CPD. Fig. 5 shows the time to compute the MSVD and CPD decompositions of randomly generated \(M \times M \times M\) tensors. Any desired rank can be obtained from the MSVD, while the CPD can only be computed for a single specified rank. Thus, Fig. 5 shows the computation time for a variety of CPD ranks, each of which must be computed from the start to obtain a desired rank CPD.

A variety of methods can be used to select singular value thresholds. The aim is to select thresholds such that the energy related to weak paths and noise are removed while the remaining terms in (18) consist of energy from significant paths. The work in [40] shows that under similar conditions, the optimal threshold for each dimension is \(s_{\text{thresh}} = 2.858 s_{\text{med}}\), where \(s_{\text{med}}\) is the median singular value along that dimension. We use this threshold value, but [40] discusses alternative thresholds. Unlike CPD estimation in [35], the MSVD approach allows multiple paths to share parameters. In this type of scenario, a single singular value is used to represent multiple paths. This is seen in Fig. 6, which shows an example of singular value thresholding where the AOA is equivalent for two paths. One singular value is accurately kept for AOA, while two are kept for AOD and path length. Fig. 7 shows the reduced rank Tucker tensor form after thresholding the singular values and truncating for an arbitrary number of thresholded singular values.

D. SEPARATING CHANNEL PARAMETER ESTIMATION INTO SEPARATE SUBSPACE PROBLEMS

The reduced Tucker form from the MSVD in (18) eliminates noise and estimates the number of paths \(N_p\), which is the rank of the measurement tensor. The denoised and rank reduced MSVD is then approximately equivalent to the noiseless...
signal in the receiver measurement (14):

\[ \Psi \otimes_1 W_a \otimes_2 F_a \otimes_3 \Phi \approx \Sigma \otimes_1 U_s^{(1)} \otimes_2 U_s^{(2)} \otimes_3 U_s^{(3)}. \]  

(19)

Channel parameter estimation is accomplished by estimating the unknown \( \Psi, W_a, F_a, \) and \( \Phi \) matrices. Assuming the reduced Tucker form from the MSVD sufficiently separates the noise from the signal, the basis vectors in (19) \( U_s^{(1)}, U_s^{(2)}, \) and \( U_s^{(3)} \) share the same subspaces as \( W_a, F_a, \) and \( \Phi, \) respectively. Furthermore, each of the subspaces are independent. Therefore, representing the channel in Tucker form enables channel parameter estimation to be separated into three independent sub-problems in the column, row, and fiber subspaces, which separately solve for \( \theta_{tx,n}, \theta_{tx,n}, \) and \( d_n, \) respectively. This significantly reduces the required computation, since otherwise every permutation of \( \theta_{tx,n}, \theta_{tx,n}, \) and \( d_n \) has to be considered jointly.

For the column subspace, the objective is to utilize the basis vectors from \( U_s^{(1)} \) to estimate \( W_a \) such that the columns of \( W_a \) correspond with physically realizable paths in the channel. This equates to finding a value of \( \theta_{tx,n} \) for each column of \( W_a. \) However, this is challenging since the basis vectors generated by the MSVD are not unique. Additionally, the basis vectors in \( U_s^{(1)} \) are orthonormal while the true path column vectors in \( W_a \) are not necessarily orthogonal.

The desired vectors in the matrix \( W_a \) are unknown, but a dictionary of possible vectors formulated from physical paths exists as follows:

\[ W_a^{(D)} = \left[ \hat{w}_{a,n}(\theta_{tx,i}) \cdots \hat{w}_{a,n}(\theta_{tx,D_1}) \right], \]

(20)

where \( \{\theta_{tx,n}\}_{m=1}^{D_1} \) is a set of \( D_1 \) dictionary values and

\[ \hat{w}_{a,n}(\theta_{tx,i}) = \frac{w_{a,n}(\theta_{tx,i})}{w_{a,n}(\theta_{tx,i})}, \quad \text{for} \quad i = 1, \ldots, D_1 \]

are normalized column vectors from (10), formulated from an over-complete dictionary of \( D_1 \) possible \( \theta_{tx,n} \) values. Fig. 8 shows an example with two basis vectors and two dictionary vectors. In this example, the orthonormal basis vectors from the MSVD \( U_s^{(1)} \) do not align with possible basis vectors in \( W_a^{(D)}. \) Note that Fig. 8 is a projection into two dimensions while the actual basis vectors are in \( N_{rx} \) space.

Estimating the columns of \( W_a \) essentially becomes a basis transformation where we seek the basis transformation matrix \( Q_1, \) such that:

\[ U_s^{(1)} = W_a^{(D)} Q_1. \]  

(21)

This is an overdetermined problem since the dictionary is overcomplete and there are many more dictionary terms than paths \( (D_1 \geq N_p). \) Thus, the solution is not guaranteed to be unique. However, the measurement tensor is low rank and the solution for \( Q_1, \) must have \( r_1 \) nonzero terms to match the rank of the column subspace. This leads to a sparse optimization problem, where the selection of dictionary basis vectors is accomplished by enforcing sparsity on \( Q_1. \) The row and column subspaces follow similarly.

**FIGURE 8.** A basis transformation converts from an orthogonal MSVD basis to a non-orthogonal dictionary basis.

**FIGURE 9.** Overdetermined problem to compute the column subspace basis transformation \( Q_1. \)

### E. SUBSPACE ESTIMATION

Similar sparse estimation problems to (21) are posed in [28], [41], [42] where sparsity is an outcome of vector selection from an overcomplete dictionary. In the work sparsity is enforced in \( Q_1 \) by solving the optimization problem:

\[ \arg \min_{Q_1} ||U_s^{(1)} - W_a^{(D)} Q_1||_F^2 + \lambda ||Q_1||_1, \]

(22)

where \( \lambda \) is a tuning parameter and \( Q_1^{(1)} \) is a vector containing the \( l_2 \) norm for each row of \( Q_1. \) Posing the sub-problem in this form results in the multiple measurement vector sparse estimation problem. Multiple solution techniques are discussed in [41], [43], but we choose to use the multiple measurement vector orthogonal matching pursuit (M-OMP) algorithm. M-OMP is chosen since it is a greedy algorithm and requires less computation. Details on M-OMP can be found in [43].

The sparsity of the column subspace \( Q_1 \) is set to the rank of the column subspace found from the MSVD \( (r_1), \) which is the number of columns in \( U_s^{(1)} \). The output of M-OMP is \( Q_1 \) with \( r_1 \) non-zero rows. This effectively eliminates all but \( r_1 \) dictionary vectors and allows the dictionary to be reduced to the following:

\[ W_a^{(D)} = W_a^{(D)} A, \]

(23)

where the reduced dictionary matrix \( \tilde{W}_a^{(D)} \) is \( N_{rx} \times r_1 \) and \( A \) is a sparse matrix identical to \( Q_1 \), but with any non-zero elements replaced with one. Then the reduced dictionary is used in replacement of (21) to obtain

\[ U_s^{(1)} = \tilde{W}_a^{(D)} Q_1. \]

(24)
where $\tilde{Q}_i$ is a $r_1 \times r_1$ matrix. Note that selecting $r_1$ dictionary terms at this step significantly reduces future computational effort. Similarly, the row subspace uses a reduced $r_2 \times r_2$ matrix $\tilde{Q}_2$ and a $N_T \times r_2$ matrix $\tilde{F}_a$ such that

$$U_s^{(2)} = \tilde{F}_a^{(D)} \tilde{Q}_2. \quad (25)$$

The fiber subspace uses a reduced $r_3 \times r_3$ matrix $\tilde{Q}_3$ and a $N_s \times r_3$ matrix $\tilde{\Phi}^{(D)}$ such that

$$U_s^{(3)} = \tilde{\Phi}^{(D)} \tilde{Q}_3. \quad (26)$$

At this point dictionary terms have been selected in each subspace. The transformation matrices $\tilde{Q}_i$ for $i = 1, 2, 3$ contain information about how each of the dictionary vectors align with the MSVD basis vectors, but dictionary vectors have not yet been explicitly chosen as estimates for the basis vectors.

**F. SUPER-RESOLUTION CHANNEL PARAMETER ESTIMATION**

The dictionaries used to obtain solutions to (24)-(26) may be coarse since the dictionaries are limited in size. This is especially true since the computational effort of M-OMP increases with the number of dictionary terms and smaller coarse dictionaries reduce computations. Higher resolution can be obtained by iteratively updating the dictionary as strong dictionary terms are identified. One approach for this is the K-SVD method, which is a dictionary learning algorithm that can be employed with sparse problems [44], [45]. We use K-SVD for our simulations as it significantly reduced computation time when compared to a single large dictionary. The implementation used in this context derives from [45] and is covered in Algorithm 1, which solves $Y = DX$. Here, $Y$ is the measurement, $D$ is the super-resolution dictionary we are trying to estimate, and $X$ is the sparse representation matrix. This approach is applied to each channel parameter estimation problem as seen in Section III-E. An iteration begins by solving the sparse estimation problem with the last dictionary set. Then a new dictionary is created that focuses around the dictionary terms chosen during the sparse step. We use the method from [45] for the dictionary update step.

**G. MSVD BASIS TRANSFORMATION**

The best dictionary terms have been selected from (24)-(26) along with their transformation matrices. These are now used to transform the MSVD in terms of the dictionary basis set in every subspace. Substituting $U_s^{(1)}$, $U_s^{(2)}$, and $U_s^{(3)}$ from (24), (25), and (26) into (18):

$$Y = \Sigma_s \otimes_1 (\bar{W}_a^{(D)} \tilde{Q}_1) \otimes_2 (\bar{F}_a^{(D)} \tilde{Q}_2) \otimes_3 (\tilde{\Phi}^{(D)} \tilde{Q}_3). \quad (27)$$

Or, equivalently:

$$Y = \Sigma'_s \otimes_1 \bar{W}_a^{(D)} \tilde{Q}_1 \otimes_2 \bar{F}_a^{(D)} \tilde{Q}_2 \otimes_3 \tilde{\Phi}^{(D)}, \quad (28)$$

where

$$\Sigma'_s = \Sigma_s \otimes_1 \tilde{Q}_1 \otimes_2 \tilde{Q}_2 \otimes_3 \tilde{Q}_3. \quad (29)$$

*Algorithm 1 Super-Resolution Channel Parameter Estimation With K-SVD (L Iterations): $[D, X] = \text{estSuperRes}(Y)$*

1. Create initial coarse dictionary with $K$ elements: $D = D_0$
   for $i = 1 : L$
   * $X = \text{M-OMP}(Y, D)$
   for $j = 1 : K$
   - $D(:, :, 0) = 0$
   - $l = \{\text{indices of signals in } Y \text{ that use } D(:, :, j)\}$
   - $E = Y - DX$
   - Dictionary update: $d, g = \arg \min_{d, g} \|E - dg^T\|_F^2$
   - $X_{i, j} = g^T$
   end for
   end for

The tensor in (28) expresses $Y$ in terms of the basis created by the dictionaries in each subspace. The core tensor $\Sigma'_s$ contains the interaction energy between the path channel parameters $\theta_{x,n}, \theta_{x,n}$, and $d_n$ from the dictionary. However, it may not be obvious which path channel parameters have the strongest interactions and should be linked together.

**H. LINKING CHANNEL PARAMETERS TO PATHS**

It is known that the tensor $Y$ has rank $N_p$. This low rank structure provides a means to select and link the dictionary terms to paths. One may expect $\Sigma'_s$ in (28) to have $N_p$ non-zero terms. However, this is not guaranteed after the basis transformation as the dictionary vectors in each subspace are generally not orthogonal. Therefore, we must eliminate interactions from $Y$ so that it again has rank $N_p$. To do this, we select the dictionary terms with the strongest interactions in $\Sigma'_s$. This is accomplished by first taking the MSVD of $\Sigma'_s$:

$$\Sigma'_s = S \otimes_1 D^{(1)} \otimes_2 D^{(2)} \otimes_3 D^{(3)}. \quad (30)$$

The MSVD of $\Sigma'_s$ results in the core tensor $S$, which organizes the interaction energies in decreasing order.

At this point the strongest $N_p$ elements in $S$ correspond with the $N_p$ paths. It is noted that distinct channel parameters for each path lead to $N_p = r_1 = r_2 = r_3$ and the strongest interactions will be the diagonal elements in $S$. However, there are scenarios when this is not the case. For example, if $N_p = 2$, but both paths have the same $\theta_{x,n}$ with different $d_n$ then $r_1 = 1$ while $r_2 = 2$ and $r_3 = 2$. In this case, one of the strongest interactions will be off-diagonal in $S$. The method in [35] does not offer a solution for this type of scenario since the CPD tensor form does not allow multiple paths to share parameter vectors.

A significant advantage to the MSVD in (30) is that it provides a simple means to determine the values of $\theta_{x,n}, \theta_{x,n}$, and $d_n$ for each path. In the form of (30) each column of $D^{(i)}$ for $i = 1, 2, 3$ corresponds with a path and has elements that correspond with the amount that each dictionary term is aligned for that subspace and path. Therefore, the channel
parameters for each path are desired by choosing the maximum value in each column of $D(i)$ for $i = 1, 2, 3$. Suppose that the strongest $N_p$ paths in $S$ have been chosen and the channel parameters are desired for path $n$ that correspond with the element $S(i_1, i_2, i_3)$. Then, the estimated channel parameters are:

$$\theta_{tx,n} = \tilde{\theta}_{tx,m_1}, \quad \theta_{tx,n} = \tilde{\theta}_{tx,m_2}, \quad d_n = \tilde{d}_{m_3}, \quad (31)$$

where,

$$m_1 = \arg \max_m D(m, i_1),$$

$$m_2 = \arg \max_m D(m, i_2),$$

$$m_3 = \arg \max_m D(m, i_3). \quad (32)$$

and $\{\tilde{\theta}_{tx}^{(D)}\}, \{\tilde{\theta}_{tx}^{(D)}\}, \{\tilde{d}^{(D)}\}$ are the set of dictionary terms from (24)-(26) with sizes $r_1$, $r_2$, and $r_3$, respectively. An example of the path linking process is seen in Fig. 10. In this example, the strongest two paths are the diagonal elements of $S$. The strongest elements in each column are selected and used to link and estimate the channel parameters.

**I. ESTIMATING PATH GAIN**

If desired, the path gain can be estimated for each path by vectorizing the measurement signal tensor as shown in [30]:

$$\text{vec}(Y) = \text{vec}(\Psi \odot W_a \odot F_a \odot \Phi), \quad (33)$$

$$= (\Phi \odot F_a \odot W_a) h, \quad (34)$$

where $\odot$ is the Khatri-Rao product and $h$ is a vector of the path gains. Let $A = \Phi \odot F_a \odot W_a$; then, the path gains are estimated by solving for $h$ such that

$$\hat{h} = \arg \min_h \|Y - Ah\|^2. \quad (35)$$

This is solved using the pseudoinverse or SVD methods [28].

**J. COMPUTATIONAL COMPLEXITY**

The computational complexity of the proposed channel estimation approach is dependent on the measurement tensor size and the number of significant paths. The computational complexity of the MSVD is discussed in [36] and rapidly increases with measurement tensor size. The rest of the estimation algorithm scales with the number of significant paths, or the rank of the measurement tensor, since each significant

path requires independent analysis. Figure 11 shows that the non-MSVD computations are independent of tensor size and scale with the number of significant paths. It is seen that the computational complexity of the MSVD significantly increases for larger tensor sizes. Therefore, the computations required to perform the MSVD based on the measurement tensor size will predominantly dictate computational complexity for larger measurement tensors.

**K. APPLICATIONS OF CHANNEL PARAMETER ESTIMATION**

Algorithm 2 summarizes the proposed channel parameter estimation algorithm. Once obtained, the estimates for the channel parameters can be used to estimate the channel matrix $H(k)$ for $k = 1, \ldots, N_s$ by substituting $\theta_{tx,n} \odot \theta_{tx,n}$, $\tau_n = d_n/c$ and $h_n$ for the $N_p$ paths into (7). The channel parameters from one LOS path or three NLOS paths are also

![FIGURE 10. Linking channel parameters for each path.](image)

![FIGURE 11. Computation time for the MSVD and total time of the MSVD estimation algorithm for a scenario with 100 subcarriers, 101 transmitting antennas, and 65 receiving antennas. The measurement tensor size is varied by increasing the OFDM symbol length.](image)
sufficient for estimating receiver location and orientation as well as mapping the environment [3], [9].

L. CRAMER-RAO BOUND
The CRB bound provides a RMSE performance bound for any estimator and is calculated using the inverse of the Fisher information matrix \( \mathbf{I}(\theta) \) [30], [46], which is derived in Appendix to give:

\[
C(\theta) = \sqrt{\mathbf{I}(\theta)}^{-1},
\]

where \( C(\theta) \) is the CRB bound matrix. The vector \( \theta \) is the collection of channel parameters and path gains for all paths:

\[
\theta = \left[ \theta_{tx,n}^{N_p}, \theta_{tx,n}^{N_p}, \{d_n\}^{N_p}_1, \{h\}^{N_p}_1 \right].
\]

The diagonal elements of \( C(\theta) \) are the RMSE bounds for each of the corresponding elements in \( \theta \).

Simulations in Section V are compared to the CRB bound through the root-mean square error (RMSE) for each channel parameter, which for path \( n \) is calculated as:

\[
\text{RMSE}(\theta_{tx,n}) = \sqrt{E[\theta_{tx,n} - \theta_{0,tx,n}^{0,n}]^2].
\]

\[
\text{RMSE}(\theta_{tx,n}) = \sqrt{E[\theta_{tx,n} - \theta_{0,tx,n}^{0,n}]^2].
\]

\[
\text{RMSE}(d_n) = \sqrt{E[d_n - d_n^0]^2},
\]

where \( \theta_{0,tx,n}^{0,n}, \theta_{0,tx,n}^{0,n}, \) and \( d_n^0 \) are the true channel parameters. The expectation value is simulated with \( N_{sim} \) Monte-Carlo simulations such that for the random variable \( x \),

\[
E[x^2] = \frac{\sum_{i=1}^{N_{sim}} x_i^2}{N_{sim}}.
\]

where each \( x_i \) is a Monte-Carlo simulated observation of \( x \).

IV. WAVEFORM CONSIDERATIONS FOR MIMO OFDM CHANNEL PARAMETER ESTIMATION
Channel parameter estimation accuracy depends on the communication waveform. Certain waveform parameters improve channel parameter estimation accuracy while other parameters can significantly degrade channel parameter estimation accuracy. This section provides details on how waveform parameters effect channel parameter estimation as well as the validity of the assumptions used to reach the measurement model in (14).

A. FREQUENCY SELECTIVITY
It is assumed in (7) that the path gain is constant over all subcarriers, or \( h_{0}(k) \approx h_{0} \) for all \( k \). This subsection discusses waveform requirements for this frequency non-selective assumption to be valid, which depends on the delay spread of the channel and channel dispersion effects.

\( a: \) DELAY SPREAD
Frequency selective fading occurs when the multipath delay extends over more than one symbol [11]. A cyclic prefix in each OFDM symbol aims to capture multipath from previous symbols. However, the frequency non-selective channel assumption requires the OFDM symbol length \( T_{OFDM} \) to be much larger than the delay spread \( \sigma_\tau \), or \( T_{OFDM} \gg \sigma_\tau \).

\( b: \) CHANNEL DISPERSION
For large OFDM bandwidths and large arrays, small phase differences observed at the antenna elements for each subcarrier can cause channel dispersion in space and time [3], [18]. Channel dispersion increases with the number of antenna elements and subcarriers. The frequency non-selective assumption holds if \( N_{rx} N_{s} / 2 T_{OFDM} \ll f_c \), where channel dispersion effects become more noticeable as the term to the left of the inequality increases. Larger center frequencies lead to less channel dispersion.

Whether dealing with delay spread or channel dispersion effects, a frequency non-selective assumption requires the symbol duration \( T_{OFDM} \) to be sufficiently large. On the other hand, larger OFDM bandwidths and additional antenna elements cause channel dispersion. Systems with channel dispersion can still achieve frequency non-selectivity by only utilizing a subset of subcarriers and dividing channel parameter estimation into blocks where the subcarriers in each block do not observe channel dispersion. Alternatively, the channel parameter estimation algorithm can be adjusted to account for the phase differences in the subcarriers in the model.

B. EFFECTIVE SNR
Channel parameter estimation improves as the number of subcarriers \( N_s \) increases and the symbol length of the training sequence \( N_T \) increases. Each of these adds dimensionality to the measurement in (14) and provides additional observations of the same channel parameters. The additional diversity of measurements increases the effective SNR and leads to improved parameter estimation. Also, increasing the numbers of antenna elements at the transmitter and receiver increases the measurement tensor dimensions and effective SNR.

C. DISTANCE REDUNDANCY
A limitation of channel parameter estimation is that the solution for the path distance is not unique. The path distance is estimated in the tensor fiber subspace. From (12), the fiber vector for each path contains the terms:

\[
\phi_n(d_n) = \left\{ e^{\frac{2\pi j k}{T_{OFDM}}} \right\}_{k=0}^{N_s-1} \text{ for } n = 0, \ldots, N_p - 1.
\]

For path \( n \), multiple values of \( d_n \) give the same \( \phi_n(d_n) \). To see this, first let \( d_{OFDM} = c T_{OFDM} \). Then, we can write:

\[
\phi_n(d_n + d_{OFDM}) = \phi_n(d_n).
\]

Therefore, the solution to the distance for each path is periodic with period \( d_{OFDM} \). A unique solution is only achieved if the distance search space is limited to a range of \( d_{OFDM} \). This guarantees that multiple distances are not found for each path. This effect can be mitigated by increasing \( T_{OFDM} \), which
V. SIMULATION RESULTS

This section compares simulations of the proposed MSVD channel parameter estimation to the CRB and the CPD approach from [35] under a variety of waveform parameters, numbers of antenna elements, and SNR. The range of waveform parameters considered are listed in Table 1. We compute the MSVD and CPD in our simulations using the Tensorlab software package for Matlab [31].

A simulation is conducted by first choosing locations and orientations for the transmitter and receiver. A LOS path is generated by connecting the transmitter and receiver. A NLOS path is generated by choosing a reflection point and connecting it to the transmitter and receiver. The channel parameters $\theta_{rx,n}$, $\theta_{tx,n}$, and $d_n$ are then dictated by the generated path geometries. Similar to [17], the precoding and combining matrices $F[k]$ and $W[k]$ from (4) are generated by uniformly sampling from $\{1, -1, i, -i\}$ for each element. Then, the measurement tensor $Y$ from (14) is constructed from the channel parameters, path gain, precoding and mixing matrices, and additive random noise to achieve a specified SNR. It should be noted that we do not explicitly simulate interfering paths and weaker reflecting paths. The random noise is used to represent interfering paths or many non-significant paths. We also fix the array orientations to $\phi_{rx} = 0$ and $\phi_{tx} = 0$ for simplicity.

A. SNR AND NUMBERS OF ANTENNA ELEMENTS

Channel parameter estimation error is first studied as a function of SNR and transmitter/receiver array sizes. Two receiver/transmitter array element numbers are used: $\{N_{rx} = 65, N_{tx} = 101\}$ and $\{N_{rx} = 11, N_{tx} = 21\}$. The number of
TABLE 1. Range of waveform specifications considered in simulations.

| Waveform Specifications |  |
|-------------------------|--|
| Center Frequency $f_c$  | 60 GHz |
| Number of Subcarriers $N_s$ | 10-100 |
| OFDM Bandwidth $B_{OFDM}$ | 100 MHz |
| OFDM Symbol Length $T_{OFDM}$ | 1 µs |
| Sampling Time $T_s$ | 10 ns |
| Subcarrier Spacing $Δf_s$ | 1 MHz |
| Training Sequence Length $N_T$ | 10-100 symbols |

FIGURE 13. Path geometry for channel parameter simulations with transmitter array at (0,0), receiver array at (15,-5), and path bounce at (35,15).

subcarriers is fixed to $N_s = 100$ and the number of training symbols is fixed at $N_T = 10$. A transmitting antenna is fixed at $(0,0)$ (m) and the receiver is randomly selected from points in the grid $-100 < x, y < 100$ (m). The SNR is varied and the average RMSE over the randomly selected points is calculated using both the MSVD and CPD estimators. Fig. 12 shows the average RMSE for both estimators and the CRB bound for the channel parameters and path gain.

The legend uses $\{N_{rx}, N_{tx}\}$ to convey the transmitter/receiver array sizes. Each of these plots show that the MSVD channel parameter estimation technique closely matches the CRB bound. Furthermore, it is seen that the MSVD and CPD approaches have similar accuracy. It is also seen that the larger set of transmitter/receiver antenna elements provides orders of magnitude of improvement in parameter estimation. Additionally, comparison between (a) and (b) in Fig. 12 shows more accurate AOD estimates than AOA estimates, since the transmitter has more antenna elements than the receiver.

B. TRAINING SYMBOL LENGTH

These simulations use a specified path geometry of one LOS path and one NLOS path between a transmitter and receiver as seen in Fig. 13 with 65 transmitting antennas and 101 receiving antennas. The SNR is fixed to 5 dB. The path gain for the first path is set to $h_1 = 1$ and the path gain for the second path is set to $h_2 = 0.5$ with the number of subcarriers set to $N_s = 50$. Fig. 14 shows the average RMSE of the channel parameters for both paths as the number of training symbols ($N_T$) is varied from 10 to 100 symbols. It is seen that longer training symbols improve parameter estimation with diminishing improvement as the symbol length increases. It is also observed that the CRB bound is not smooth as a function of the number of training symbols. This is because each additional training symbol adds a random column to the measurement tensor in the row subspace. While the improvement with symbol length is monotonically decreasing, the amount each symbol improves the estimate depends on how independent the extra columns are from the previous columns. It is also seen that high resolution estimates are obtained with short training symbols (10 symbols). The slope of the CRB bound is very steep below 10 symbols and channel estimation performance quickly degrades with decreasing training symbol length. It is also observed that the use of M-OMP does not always obtain the optimal estimate and meet the CRB. This is expected, particularly at low SNR, since M-OMP is a greedy algorithm. Other algorithms can replace M-OMP in the sparse estimation step for improved performance [41].

C. NUMBER OF SUBCARRIERS

This simulation studies the effect of subcarriers using the same path geometry from Fig. 13 with 65 transmitting antennas, 101 receiving antennas, and SNR of 5 dB. Fig. 15 shows the average RMSE of the channel parameters for both paths as the number of subcarriers ($N_s$) is varied from 30 to 100 with the number of training symbols fixed at $N_T = 50$ symbols. A subcarrier spacing of 1 MHz is used so the maximum OFDM bandwidth simulated is 100 MHz. Channel parameter estimation performance improves with the number of subcarriers as each subcarrier adds a column in the fiber subspace and increases the dimensionality of the measurement tensor. This also increases the effective SNR. Smaller tensor sizes than what is simulated in this paper may lead to low effective SNR values where M-OMP performance degrades and
non-greedy sparse estimation methods are needed to achieve the CRB [41].

VI. CONCLUSION
A channel parameter estimation technique has been proposed based on the MSVD, which is ideally suited for channel parameter estimation since the measurement tensor is naturally represented in Tucker tensor form. Simulations show that the RMSE obtained using the proposed technique closely matches the CRB. Comparison to a CPD tensor estimation approach shows that the MSVD method can achieve similar accuracy with reduced computational complexity. This paper considers scenarios where the transmitter/receiver are limited to a plane. However, the technique is easily extendable to provide a computationally efficient method for channel parameter estimation in three-dimensional coordinate systems where the elevation angle must also be considered. Channel parameter estimation has been studied as a function of waveform parameter specifications, where improvements are seen with increased numbers of subcarriers, longer training sequence length, and larger antenna array sizes. Future work will focus on utilizing channel parameter estimation and localization simultaneously for reduced overhead and improved estimation. Extending matrix techniques such as the generalized Least Squares Matrix Decomposition (GMD) to tensor-based MIMO channel estimation is also of interest because it enables prior information about structure or two-way dependencies of the channel to be incorporated.

APPENDIX
CRB DERIVATION FOR PARAMETER ESTIMATION
This derivation utilizes the Kronecker product (⊗) and Khatri-Rao product (⊙). The CRB bound is first derived by vectorizing the receiver measurement tensor as shown in [30]:

\[ \text{vec}(Y) = \Psi \otimes_1 W_a \otimes_2 F_a \otimes_3 \Phi + N, \]

where

\[ = (\Phi \otimes F_a \otimes W_a)I + \text{vec}(N), \]

\[ = X(\theta) + \text{vec}(N), \]

where \( X(\theta) = (\Phi \otimes F_a \otimes W_a)I \) is a function of the path gains and channel parameters

\[ \theta = [\{\theta_{tx,n}\}^N_p, \{\theta_{rx,n}\}^N_p, \{d_{tx}\}^N_p, \{h\}^N_p]. \]

The vectorized receiver measurement in (43) is in a linear Gaussian form where each of the elements in vec(\( N \)) are sampled from a \( \mathcal{N}(0, \sigma^2) \) and \( \sigma^2 \) is determined by the SNR. Therefore, the Fisher information matrix can be calculated as [30], [46]:

\[ I(\theta) = \frac{1}{\sigma^2} \left[ \frac{\partial X}{\partial \theta} \right]^T \left[ \frac{\partial X}{\partial \theta} \right]. \]

where \( \frac{\partial X}{\partial \theta} \) is the Jacobian of \( X \).

The Jacobian of \( X \) can be divided into four blocks such that

\[
\frac{\partial X}{\partial \theta} = \begin{bmatrix}
\frac{\partial X_1}{\partial \theta} & \frac{\partial X_2}{\partial \theta}
\end{bmatrix},
\]

(46)

The columns of each block are calculated by rearranging \( X \) as shown in [30] to the equivalent forms:

\[ X = [(\Phi' \otimes F_a) \otimes I_{N_s}] \text{vec}(W_a), \]

(47a)

\[ = K_{N_t, N_s} [(W_a \otimes \Phi') \otimes I_T] \text{vec}(F_a), \]

(47b)

\[ = K_{N_r, N_s} [(F_a \otimes W_a) \otimes I_{N_s}] \text{vec}(\Phi'), \]

(47c)

where \( K_{m,n} \) converts the vectorization of the \( m \times n \) matrix \( S \) to the vectorization of its transpose, or \( K_{m,n} \text{vec}(S) = \text{vec}(S^T) \).

The columns of the first block are calculated using (47a) to obtain:

\[ \frac{\partial X}{\partial \theta_{tx,n}} = \frac{\partial}{\partial \theta_{tx,n}} [(\Phi' \otimes F_a) \otimes I_{N_s}] \text{vec}(W_a), \]

(48)

The columns of the second block are calculated similarly using (47b) to obtain:

\[ \frac{\partial X}{\partial \theta_{tx,n}} = K_{N_t, N_s} [(W_a \otimes \Phi') \otimes I_T] \frac{\partial \text{vec}(F_a)}{\partial \theta_{tx,n}}, \]

(49)

The columns of the third block are obtained using (47c):

\[ \frac{\partial X}{\partial d_{tx}} = K_{N_r, N_s} [(F_a \otimes W_a) \otimes I_{N_s}] \frac{\partial \text{vec}(\Phi')}{\partial d_{tx}}, \]

(50)

The columns of the fourth block are calculated using (47a) to obtain:

\[ \frac{\partial X}{\partial h_{tx}} = K_{N_r, N_s} [(F_a \otimes W_a) \otimes I_{N_s}] \frac{\partial \text{vec}(\Phi')}{\partial h_{tx}}, \]

(51)

where

\[ \text{vec}(\Phi') = \text{vec}(\Phi \text{ Diag}(h) I_{N_p}), \]

(52)

\[ = (I_{N_p} \otimes \Phi)h, \]

(53)

where the second step follows from [30]. Then, the fourth block is simplified to:

\[ \frac{\partial X}{\partial h_{tx}} = K_{N_r, N_s} [(F_a \otimes W_a) \otimes I_{N_s} (I_{N_p} \otimes \Phi)] \frac{\partial h}{\partial h_{tx}}. \]

(54)
REFERENCES

[1] T. S. Rappaport, S. Sun, R. Mayzus, H. Zhao, Y. Azar, K. Wang, G. N. Wong, J. K. Schulz, M. Samimi, and F. Gutierrez, “Millimeter wave mobile communications for 5G cellular: It will work!” IEEE Access, vol. 1, pp. 335–349, May 2013.

[2] Y. Zhu, Z. Zhang, Z. Marzi, C. Nelson, U. Madhow, B. Y. Zhao, and H. Zheng, “Demystifying 60GHz outdoor picocells,” in Proc. 20th Annu. Int. Conf. Mobile Comput. Netw. (MobiCom), 2014, pp. 5–16.

[3] A. Shahmansoori, G. E. Garcia, G. Destino, G. Seco-Granados, and H. Wyfearsch, “Position and orientation estimation through millimeter-wave MIMO,” in Proc. IEEE Globecom Workshops (GC Wkshps), Dec. 2015, pp. 1–6.

[4] N. Garcia, H. Wyfearsch, and D. Slock, “Optimal precoders for tracking the AoD and AoA of a mm-wave path,” 2017, arXiv:1703.10978. [Online]. Available: http://arxiv.org/abs/1703.10978

[5] N. Garcia, H. Wyfearsch, E. G. Larsson, A. M. Haimovich, and M. Coulon, “Direct localization for massive MIMO,” IEEE Trans. Signal Process., vol. 65, no. 10, pp. 2475–2487, Jul. 2017.

[6] K. Witsal, P. Meissner, E. Leitinger, Y. Shen, C. Gustafson, F. Tufvesson, K. Haneda, D. Dardari, A. F. Molisch, A. Conti, and M. Z. Win, “High-accuracy localization for assisted living: 5G systems will turn multipath channels from foe to friend,” IEEE Signal Process. Mag., vol. 33, no. 2, pp. 59–70, Mar. 2016.

[7] R. Mendzlik, H. Wyfearsch, G. Bauch, and Z. Abu-Shaban, “Harnessing NLOS components for position and orientation estimation in 5G mmWave MIMO,” 2017, arXiv:1712.01445. [Online]. Available: http://arxiv.org/abs/1712.01445

[8] C. E. Hayes, J. H. McClellan, and W. R. Scott, “Low-rank physical model recovery from low-rank signal approximation,” in Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP), Mar. 2017, pp. 331–335.

[9] D. Pert, and N. D. Sidiropoulos, “Tensor algebra and multidimensional harmonic retrieval in signal processing for MIMO radar,” IEEE Trans. Signal Process., vol. 58, no. 11, pp. 5693–5705, Nov. 2010.

[10] N. D. Sidiropoulos, L. De Lathauwer, X. Fu, K. Huang, E. E. Papalexakis, and C. Faloutsos, “Tensor decomposition for signal processing and machine learning,” IEEE Trans. Signal Process., vol. 65, no. 13, pp. 3551–3582, Jul. 2017.

[11] N. Vervliet, O. Debals, L. Sorber, M. Van Barel, and L. De Lathauwer, (Mar. 2016). Tensorlab 3.0. [Online]. Available: http://www.tensorlab.net

[12] J. H. de M. Goulart, M. Boizard, R. Boyer, G. Favier, and P. Comon, “Tensor CP decomposition with structured factor matrices: Algorithms and performance,” IEEE J. Sel. Topics Signal Process., vol. 10, no. 4, pp. 757–769, Jun. 2016.

[13] A. Stégeman and N. D. Sidiropoulos, “On Kraskal’s uniqueness condition for the Candecomp/Parafac decomposition,” Linear Algebra Appl., vol. 420, nos. 2–3, pp. 540–552, Jan. 2007.

[14] S. Rabanser, O. Shchur, and S. Günnemann, “Introduction to tensor decompositions and their applications in machine learning,” 2017, arXiv:1711.10781. [Online]. Available: http://arxiv.org/abs/1711.10781

[15] Z. Zhou, J. Fang, L. Yang, H. Li, Z. Chen, and R. S. Blum, “Low-rank tensor decomposition-aided channel estimation for millimeter wave MIMO-OFDM systems,” IEEE J. Sel. Areas Commun., vol. 35, no. 7, pp. 1524–1538, Jul. 2017.

[16] P. Comon, X. Luciani, and A. L. F. de Almeida, “Tensor decompositions, alternating least squares and other tales,” J. Chemometrics, vol. 23, nos. 7–8, pp. 393–405, Jul. 2009.

[17] T. S. Rappaport, Y. Qiao, J. I. Tamir, J. N. Murdock, and E. Ben-Dor, “Cellular broadband millimeter wave propagation and angle of arrival for adaptive beam steering systems (invited paper),” in Proc. IEEE Radio Wireless Symp., Jan. 2012, pp. 151–154.

[18] J. W. Choi, B. Shim, Y. Ding, B. Rao, and D. I. Kim, “Compressed sensing for wireless communications: Useful tips and tricks,” IEEE Commun. Surveys Tuts., vol. 19, no. 3, pp. 1527–1550, 3rd Quart., 2017.

[19] M. Koivisto, M. Costa, A. Hakkarainen, K. Leppanen, and M. Valkama, “Joint 3D positioning and network synchronization in 5G ultra-dense networks using UKF and EKF,” in Proc. IEEE Globecom Workshops (GC Wkshps), Dec. 2016, pp. 1–7.

[20] M. Ganesh and D. L. Donoho, “The Optimal Hard Threshold for Singular Values is 4/√L,” IEEE Trans. Inf. Theory, vol. 60, no. 8, pp. 5040–5053, Aug. 2014.

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

[22] D. Malouifot, M. Cetin, and A. S. Willsky, “A sparse signal reconstruction perspective for source localization with sensor arrays,” IEEE Trans. Signal Process., vol. 53, no. 8, pp. 3010–3022, Aug. 2005.

[23] J. Chen and X. Huan, “Sparse representations for multiple measurement vectors (MMV) in an over-complete dictionary,” in Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP), Mar. 2005, p. 287.

M. Ruble, İ. Güvenc: MSVD for mmWave Channel Parameter Estimation

VOLUME 8, 2020 75605
[44] M. Aharon, M. Elad, and A. Bruckstein, “K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation,” *IEEE Trans. Signal Process.*, vol. 54, no. 11, pp. 4311–4322, Nov. 2006.

[45] R. Rubinstein, M. Zibulevsky, and M. Elad, “Efficient implementation of the K-SVD algorithm using batch orthogonal matching pursuit,” Comput. Sci. Dept., Technion, Haifa, Israel, Tech. Rep. CS-2008-08, 2008.

[46] S. M. Kay, *Statistical signal processing: Estimation Theory*, vol. 1. Upper Saddle River, NJ, USA: Prentice-Hall, 1993.

**MACEY RUBLE** received the bachelor’s degree in physics and mathematics from the University of North Carolina at Charlotte, the master’s degree in applied physics from Cornell University, and the Ph.D. degree in electrical engineering from North Carolina State University, in 2018. His research interests include artificial intelligence, estimation theory, 5G wireless systems, and localization. He was a recipient of the NSF Fellowship, in 2014.

**İSMAIL GÜVENÇ** (Senior Member, IEEE) received the Ph.D. degree in electrical engineering from the University of South Florida, in 2006. He was with Mitsubishi Electric Research Labs, in 2005, DOCOMO Innovations, from 2006 to 2012, and Florida International University, from 2012 to 2016. Since 2016, he has been an Associate Professor with the Department of Electrical and Computer Engineering, North Carolina State University. His recent research interests include 5G wireless systems, communications and networking with drones, and heterogeneous wireless networks. He has published more than 130 conference/journal articles and book chapters, and several standardization contributions. He has coauthored/co-edited three books for Cambridge University Press. He is also an inventor/coinventor some 30 U.S. patents. He was a recipient of the USF Outstanding Dissertation Award, in 2006, the Ralph E. Powe Junior Faculty Enhancement Award, in 2014, the NSF CAREER Award, in 2015, and the FIU College of Engineering Faculty Research Award, in 2016. He has served as an Editor for the IEEE *COMMUNICATIONS LETTERS*, from 2010 to 2015. He has been an Editor of the IEEE *WIRELESS COMMUNICATIONS LETTERS*, since 2011. He has served as a Guest Editor for several other journals.

* * *

---

**M. Ruble, İ. Güvenç: MSVD for mmWave Channel Parameter Estimation**

---

75606 VOLUME 8, 2020