IEEE Open Journal of the Communications Society

Received XX Month, XXXX; revised XX Month, XXXX; accepted XX Month, XXXX; Date of publication XX Month, XXXX; date of current version 11 January, 2024.

Digital Object Identifier 10.1109/OJCOMS.2024.011100

Multi-cluster Frequency-angle Diffuse Scattering Model and its Estimation

Zihang Cheng[1] ***(Student Member, IEEE)*** **and Andreas F. Molisch**[1] ***(Fellow, IEEE)***

[1]Ming-Hsieh Department of Electrical and Computer Engineering, Wireless Devices and Systems (WiDeS) group, University of Southern California, Los Angeles, CA 90089, USA

CORRESPONDING AUTHOR: Zihang Cheng (e-mail: zihangch@usc.edu).

ABSTRACT For the development and deployment of upcoming 6G systems, propagation channel measurements in new scenarios, new frequency bands, and for new antenna arrangements will be required. The diffuse multipath component (DMC) contains a non-negligible portion of the channel impulse response, and thus must be considered in channel evaluations and modeling. Previous work observed that multiple DMC clusters exist, which are associated with specular multipath component (MPC) clusters, yet the widely used Kronecker assumption for the delay/angle structure of these DMC clusters rarely holds in reality. In this paper, we propose an estimation algorithm for a parametric multi-cluster DMC model that more accurately models the connection between delay and angular parameters for each cluster, assuming that the Kronecker model is only valid *within* each cluster instead of for the whole channel. This avoids creating ghost modes that would reduce the DMC covariance matrix estimation accuracy. Our proposed algorithm is incorporated into the framework of the RiMax algorithm, an iterative maximum-likelihood estimation scheme. The effectiveness and correctness of the proposed algorithm are verified on synthetic channels, since they have a known ground truth. Simulations demonstrate improvements not only in the accuracy of the DMC, but also the associated specular MPC compared to the state-of-the-art uni-modal DMC model. Furthermore, the application of this algorithm to real-world sub-THz channel measurements is demonstrated.

INDEX TERMS Channel estimation, covariance matrix estimation, diffuse scattering, high-resolution parameter extraction algorithm

I. INTRODUCTION

WHILE fifth-generation (5G) systems are being deployed all over the world, researchers have started to move their focus to Beyond 5G (B5G, often called 6G), systems. Demand for new applications, such as transmission of high-fidelity holograms, increased efficiency for connectivity of things, and support for extremely latency-sensitive applications motivate new research from the physical layer to the network layer [1]. The B5G networks will operate in new frequency bands (e.g., upper midband and sub-THz), new environments (e.g., drone-mounted "flying base stations"), and sport new antenna arrangements (e.g., ultra-massive multiple-input multiple-output (MIMO)). Since both the fundamental performance limits and the practical performance of any wireless system are determined by the characteristics of the channels, we need new channel models that are valid in these new settings.

Any channel model needs to be based on, or verified by, accurate measurements (*i.e.*, *channel sounding*). However, each channel sounding system, regardless of the hardware details or the sounding signal design, is normally limited in its Fourier resolution - the inverse of the bandwidth is the delay resolution and the antenna beamwidth is the angular resolution. High-resolution parameter estimation (HRPE) algorithms are capable to overcome this limitation and achieve higher resolution, by postulating a parametric model for the propagation channel and estimating the parameters from the received signals. The most common assumption is that the signal at transmitter (Tx) and receiver (Rx) can be described as a sum of planar waves. With the given signal model, it is possible to evaluate the channel parameters, often in an iterative manner, such as in the widely used CLEAN [2], SAGE [3], and RiMax algorithm [4]. It is obvious that the accuracy of HRPE depends on the validity of the postulated parametric model [5], [6].

While the "sum of plane waves" model is a popular description of multipath channels, it is not the most general one. Plane waves arise from a far-field representation of

waves undergoing specular reflections. However, other signal contributions might arise from diffuse scattering on rough surfaces, and are known as DMC. Furthermore, other weak signals, including contributions from small objects such as leaves, and possibly even wavefront curvature, might be subsumed into the DMC as well [7]. Considering the large amount of the aforementioned weak signals, the DMC is modeled as a statistical process. Based on the Central Limit Theorem, the impulse response is assumed to follow a complex circularly symmetric Gaussian distribution with zero mean and specific covariance matrix. Various measurements have shown that diffuse scattering not only exists, but contributes a significant portion of the communication channel, e.g., [7]–[10], ranging from 5% to 50% of the total channel power, with that percentage higher for indoor and industrial rich scattering environments, e.g., [11], [12].

Since the DMCs compose an important part of the channel, a model for them should be incorporated into parameter estimation. Among the widely used HRPE algorithms, only RiMax actually does so. Specifically, it enables the estimation of the DMC parameters iteratively: given estimated (specular) MPC parameters, the residue (measured signal minus signal created by the estimated MPCs) is modeled by a parametric DMC model, and its parameters are evaluated. Similarly, the MPC parameter estimation can be improved based on the fitted DMC. Both the MPC and DMC descriptions are optimized in an iterative, alternating manner. Thus, the performance of MPC and DMC are interdependent—bad estimation of the DMC not only leads to inaccurate DMC model parameters, but also reduces the accuracy of the specular MPC parameters, and *vice versa*. Therefore, it is crucial to design methods for accurate DMC covariance matrix estimation and reconstruction [5], [6], [13]–[16]. This in turn requires both finding an accurate generic model for the DMC [16], and designing parameter estimation methods for it.

A. STATE OF THE ART

The first model for DMC, proposed in [4], was a single-sided exponential decay in the delay domain and uniform distribution in the angular domain (both at Tx and Rx). The base delay, which corresponds to a sharp onset of the single-sided exponential, is equal to or larger than the line-of-sight distance between the Tx and Rx. The decay time constant is related to the coherence bandwidth of the channel.

An improved model for the angular distribution, namely a *von Mises distribution*, was used in [9], [17], [18]. The von Mises distribution is equivalent to a normal distribution in the angular range $[-\pi, \pi]$. It is described with two parameters—(i) the average direction, and (ii) the variance or the degree of spread. Depending on the choice of the variance, its shape can range from highly concentrated around the mean direction, to uniformly distributed over all angles.

In [19], [20], the angular distributions are further modeled as multi-modal distributions, motivated by observations from several measurement campaigns [21]–[24]. Similar to the specular components, the DMC might occur in clusters. Thus, there might exist multiple local maxima in the angular spectrum, which deviates from the von Mises distribution that has only one single maximum. These papers therefore establish a multi-modal von Mises distribution (*i.e.*, a weighted sum of a set of von Mises distributions with different means and possibly different variances). Moreover, the weights are normalized to sum to unity. The joint delay-angle distribution has a Kronecker structure: in other words, the delay behavior is a single-exponential decay common to all clusters.

However, the Kronecker model usually does not hold for channels with well-separated clusters [4]. In practice, the relationship between propagation delay and angular information is of significant importance. When there are multiple separate clusters, the Kronecker model creates clusters for *any* combination of measured cluster delay and cluster angle; some of those combinations have no correspondence in the actual (measured) channels and are thus "ghost" clusters. The accuracy of the MPC parameter estimation is heavily influenced by the knowledge of the DMC covariance matrix. These ghost clusters introduce inaccuracies in the inversion of the covariance matrix, resulting in the sub-optimal estimation of the MPC parameters, exemplifying the above-mentioned interdependence between the estimation of DMC and specular MPCs.

To remedy these issues, multiple-cluster DMC models have been proposed. In [25], multiple DMC clusters are observed after the removal of specular components estimated by SAGE algorithm; a strong correlation between specular and diffuse clusters is also observed. However, no iterative refinement between the estimation of DMC and specular MPCs is done. Independent multi-modal DMC models are also mentioned in [26], [27], but no corresponding estimator that can be incorporated into an iterative solver like RiMax was developed.

B. OUR CONTRIBUTIONS

This paper deals with a multi-cluster, non-Kronecker-based DMC model. The main contribution is the development of the gradient-descent estimation algorithm for such a model that efficiently extracts parameters of this model within RiMax's framework. Based on the observations in the literature, we propose a multi-cluster frequency-angle DMC model. Each cluster center has a distinct combination of delay, direction of departure (DoD), and direction of arrival (DoA). The Kronecker model is only assumed to be valid *within* each cluster to formulate the joint distribution from each domain, while the total covariance matrix does *not* have a Kronecker structure. Note that the Kronecker model within a cluster, while mathematically convenient, may also not be strictly fulfilled, as, e.g., single-scattering processes still lead to a unique coupling between DoA, DoD, and delay. However, the error introduced by this assumption is significantly lower than the error caused by the Kronecker assumption for the total channel. Furthermore,

the base delay and the average direction of arrival/departure generally coincide with the corresponding MPCs/clusters, yet the model can also account for DMC-only clusters.

Exploiting the fact that the starting delay and mean direction of the DMC cluster coincide with one of the MPCs, the raw channel observation can be used to initialize DMC parameters with a larger dynamic range and better robustness against noise; alternatively, we propose a modified second-order derivative method for determining cluster start times. The parameters for each separate DMC cluster can then be estimated with the gradient descent method of [4], [20]. The total covariance matrix for the full channel is inverted through a whitening method that avoids the numerically expensive inversion of a correlation matrix for which the Kronecker assumption does not hold, but rather performs inversion separately for each dimension (delay, DoA, DoD). The DMC estimation method can be seamlessly integrated into the framework of RiMax [4].

The correctness and effectiveness of the proposed DMC estimation are verified in synthetic channels, which allows assessment against a "ground truth", and also comparison of the algorithm versus the Cramér-Rao lower bound (CRLB). A real-world example is also provided, demonstrating the application of the proposed DMC estimation algorithm to outdoor measurements in the 145 GHz frequency band.

C. PAPER ORGANIZATION

This paper is organized as follows: The description of the signal model and the frequency-angle DMC model are presented in Section II. In Section III, we derive our new gradient-descent-based parameter estimation algorithm for this DMC model. Section IV conducts the numerical verifications of DMC estimation and joint estimation of specular and diffuse components in the synthetic and measured scenarios, followed by concluding remarks in Section V.

Notation: scalars are written as normal-face letters, vectors as non-cap bold letters, and matrices as capital bold letters. Capital letters with calligraphic font, e.g., $\mathcal{H}$, represent matrices with higher dimensionalities. Roman subscripts are descriptors (e.g., "n" for noise), while italic subscripts are indices (e.g., n indexing discrete MPCs). Superscript $(\cdot)^T$ is transpose and $(\cdot)^\dagger$ stands for its Hermitian (complex conjugate) transpose. The transformation matrix of discrete Fourier transform (DFT) with proper dimension is expressed as $\mathbf{F}$. The operator with subscript, $\mathbb{E}_{(\cdot)}$, denotes the expectation applied over a variable or an index.

II. SIGNAL MODEL

The raw measured channel transfer functions are the ratio of the received signal divided by the transmitted (sounding) signal. Assuming excitation from a suitably designed multi-carrier signal [4], [28], the sounding signal magnitude is constant over the bandwidth of interest, so that the division by it does not change the noise statistics. Under our data model, the measured channel, $\mathbf{H}_{\mathrm{mea}}(f_k)$, at a certain subcarrier frequency f_k can be regarded as a sum of the superposition of N_{path} planar waves, the DMC $\mathbf{N}_{\mathrm{DMC}}(f_k)$, and additive white Gaussian noise (AWGN) $\mathbf{N}(f_k)$. Thus, the MIMO channel between the N_{T} and N_{R} antenna elements at the Tx and Rx end, respectively, can be expressed as

$$\begin{aligned}\mathbf{H}_{\mathrm{mea}}(f_k) &= \mathbf{H}(f_k) + \mathbf{N}_{\mathrm{DMC}}(f_k) + \mathbf{N}(f_k) \in \mathbb{C}^{N_{\mathrm{R}}\times N_{\mathrm{T}}} \\ &= \sum_{n=1}^{N_{\mathrm{path}}} \gamma_n \tilde{\mathbf{g}}_{\mathrm{R}}(\mathbf{\Omega}_{\mathrm{R},n}; f_k)\tilde{\mathbf{g}}_{\mathrm{T}}^T(\mathbf{\Omega}_{\mathrm{T},n}; f_k)e^{-j2\pi f_k \tau_n} \\ &\quad + \mathbf{N}_{\mathrm{DMC}}(f_k) + \mathbf{N}(f_k), \end{aligned} \tag{1}$$

where the second equality relates the MIMO channel matrix to the parameters of the double-directional channel model [29] that we wish to extract: γ is the complex amplitude, τ refers to the propagation delay, and $\mathbf{\Omega}_{\mathrm{R}}$ and $\mathbf{\Omega}_{\mathrm{T}}$ stand for the DoA and DoD, respectively, of the n-th MPC. The complex Tx and Rx antenna element pattern, $\tilde{\mathbf{g}}_{\mathrm{T}} \in \mathbb{C}^{N_{\mathrm{T}}\times 1}$ and $\tilde{\mathbf{g}}_{\mathrm{R}} \in \mathbb{C}^{N_{\mathrm{R}}\times 1}$, are measured as function of direction and frequency point, e.g., in an anechoic chamber.

Note that our model ignores the fact that (in the farfield) the MPCs as well as the DMC consist of two orthogonal polarizations. We adopt this approach for two main reasons: (i) it simplifies notation, and (ii) it is in line with many measurements especially at high frequencies that use single-polarized (usually vertical-to-vertical impulse responses only). It is important to acknowledge, though, that neglecting of the polarization may reduce the accuracy of the estimates even of the chosen polarization [6]. Inclusion of dual polarization into the DMC model in our algorithm does not pose fundamental problems because the DMC cluster structure is generally assumed (based on physical motivations) to be the same for all polarization combinations, with just the power being different.

The full description of the measured MIMO channel can be expressed as a three-dimensional (3D) matrix, $\mathcal{H}_{\mathrm{mea}} \in \mathbb{C}^{N_{\mathrm{f}}\times N_{\mathrm{R}}\times N_{\mathrm{T}}}$, in which the first dimension is the subcarrier frequencies; however, such 3D matrices are complicated to handle. For the sake of simplification in the following discussion and analysis, we introduce the vector form of the measured channel, $\mathbf{h}_{\mathrm{mea}}$, as

$$\begin{aligned}\mathbf{h}_{\mathrm{mea}} \triangleq [&[\mathcal{H}_{\mathrm{mea}}]_{1,1,1}, \cdots, [\mathcal{H}_{\mathrm{mea}}]_{N_{\mathrm{f}},1,1}, \cdots\cdots, \\ &[\mathcal{H}_{\mathrm{mea}}]_{N_{\mathrm{f}},N_{\mathrm{R}},1}, [\mathcal{H}_{\mathrm{mea}}]_{1,1,2}, \cdots\cdots]^T \\ &= \mathbf{h} + \mathbf{n}_{\mathrm{DMC}} + \mathbf{n} \in \mathbb{C}^{N_{\mathrm{f}}N_{\mathrm{R}}N_{\mathrm{T}}\times 1}, \end{aligned} \tag{2}$$

where $[\mathcal{H}_{\mathrm{mea}}]_{i,j,k}$ denotes the (i,j,k)-th element in the 3D channel observation matrix, $\mathcal{H}_{\mathrm{mea}}$. Similarly, the vector form of the contribution from specular MPC, DMC, and additive noise are thus defined as $\mathbf{h}$, $\mathbf{n}_{\mathrm{DMC}}$ and $\mathbf{n}$, respectively.

The analysis in this paper is based on the following assumptions:

a) the additive noise is assumed to be an i.i.d. zero-mean circularly symmetric complex Gaussian process;
b) the DMC contribution can be modeled as zero-mean complex Gaussian process with specific covariance matrix;

c) each DMC cluster is independent of other DMC clusters and noise.

Assumption a) corresponds to the standard model of thermal noise. The noise is assumed to be a random realization at each Tx-Rx antenna combination.[1] Due to the identical design for each radio frequency chain, the distributions are identical with zero mean and the same variance, also known as the noise level; it is independent of all signal components.

Assumption b) comes from the superposition of a large number of small contributions with independent complex amplitude. Based on the Central Limit Theorem, the DMC process can be modeled as a zero-mean circularly symmetric process due to the superposition of many complex vectors with uniformly distributed random phases. The structural covariance matrix will be discussed in the next section. Both assumptions a) and b) are standard in the DMC literature [7].

Assumption c) corresponds to the independence of diffuse contributions stemming from interactions with different clusters of objects, which leads to independence in distributions.

A. FREQUENCY/DELAY DOMAIN CHARACTERIZATION OF THE DMC

The traditional model for the DMC in the literature is a single-sided exponential decay in the delay domain. Specifically, the parametric Power Delay Profile (PDP) of the DMC process, $\zeta_{\mathrm{DMC}}(\tau)$, can be modeled as

$$\zeta_{\mathrm{DMC}}(\tau) = \begin{cases} 0 & , \quad \tau < \tau_{\mathrm{d}} \\ \frac{1}{2}\alpha & , \quad \tau = \tau_{\mathrm{d}} \\ \alpha e^{-\beta_{\mathrm{d}}(\tau-\tau_{\mathrm{d}})} & , \quad \tau > \tau_{\mathrm{d}} \end{cases}, \tag{3}$$

where α is the maximal power of the DMC in the delay domain, and τ_d stands for the base delay. Equation (3) incorporates a *Heaviside step function* since the diffuse scatterings always add extra delay to the base delay τ_{d}. The coherence bandwidth is β_{d}. The Fourier transform of (3) can be expressed as

$$\psi(f_k) = \frac{\alpha}{\beta_{\mathrm{d}} + j2\pi f_k} e^{-j2\pi f_k \tau_{\mathrm{d}}}. \tag{4}$$

Based on the aforementioned assumption, the statistical description of the DMC in the delay domain can be modeled as a random vector that follows the multivariate complex Gaussian distribution. The covariance matrix of an arbitrary transfer function vector at each Tx-Rx antenna combination from the observation, $\mathbf{u}(f) \in \mathbb{C}^{N_{\mathrm{f}} \times 1}$, can be expressed as

$$\begin{aligned} \mathbf{R}_{\mathrm{f}} &\triangleq \mathbb{E}\left[\mathbf{u}(f)\mathbf{u}^{\dagger}(f)\right] \\ &= \begin{bmatrix} \psi(0) & \cdots & \psi(-(N_{\mathrm{f}}-1)\Delta f) \\ \psi(\Delta f) & \cdots & \psi(-(N_{\mathrm{f}}-2)\Delta f) \\ \vdots & \ddots & \vdots \\ \psi((N_{\mathrm{f}}-1)\Delta f) & \cdots & \psi(0) \end{bmatrix}, \end{aligned} \tag{5}$$

where $\mathbb{E}\left[u(f_1)u^*(f_2)\right] = \psi(f_1 - f_2)$ holds due to the stationarity of the process. The equal frequency spacing, $\Delta f = \mathrm{BW}/N_{\mathrm{f}}$, is the frequency separation between two adjacent frequency points within the bandwidth, BW.

In contrast, in our model we consider the PDP to be the sum of different DMC clusters, *each* of which has a single-sided exponential decay with different base delays and possibly different decay constants. The base delay for a cluster can be modeled as the same delay as the corresponding MPC that gives rise to the DMC cluster. From the assumption c) on independence between any DMC clusters, the DMC covariance matrix in the frequency domain can be reconstructed by the summation from each single DMC delay cluster. With the consideration of the additive Gaussian noise in the channel, the total covariance matrix of the random processes (noise and DMC) can be written as

$$\mathbf{R}_{\mathrm{f}} = \sum_i \mathbf{R}_{\mathrm{f},i} + \sigma_{\mathrm{n}}^2 \mathbf{I} = \mathrm{toep}(\mathbf{v}(\boldsymbol{\theta}), \mathbf{v}^{\dagger}(\boldsymbol{\theta})), \tag{6}$$

where σ_{n}^2 is the noise power level. The matrix, $\mathbf{R}_{\mathrm{f},i}$, denotes the covariance matrix for each DMC cluster appearing in the delay domain. The Toeplitz operator is defined as

$$\begin{aligned} \mathrm{toep}(\mathbf{x}, \mathbf{y}) &= \mathrm{toep}\left(\begin{bmatrix} x_1 \\ \vdots \\ x_M \end{bmatrix}, [y_1, \cdots, y_M] \right) \\ &= \begin{bmatrix} x_1 & y_2 & \cdots & y_M \\ x_2 & x_1 & \ddots & y_{M-1} \\ \vdots & \ddots & \ddots & \vdots \\ x_M & x_{M-1} & \cdots & x_1 \end{bmatrix}, \end{aligned} \tag{7}$$

where the first elements are real-valued and equal, $x_1 = y_1$. The covariance matrix with the special structure is a *Toeplitz* matrix [13], where its generating vector can be expressed as

$$\mathbf{v}(\boldsymbol{\theta}) \triangleq \sigma_{\mathrm{n}}^2 \mathbf{e}_0 + \sum_i \alpha_i \left[\frac{1}{\beta_{\mathrm{d},i}}, \cdots, \frac{e^{-j2\pi \frac{N_{\mathrm{f}}-1}{N_{\mathrm{f}}} \tau_{\mathrm{d},i}}}{\beta_{\mathrm{d},i} + j2\pi \frac{N_{\mathrm{f}}-1}{N_{\mathrm{f}}}} \right]^T, \tag{8}$$

where the descriptive parameter vector is defined as $\boldsymbol{\theta} = [\boldsymbol{\alpha}^T, \boldsymbol{\tau}_{\mathrm{d}}^T, \boldsymbol{\beta}_{\mathrm{d}}^T, \sigma_{\mathrm{n}}^2]^T$. The vector $\mathbf{e}_0 \triangleq [1, 0, \cdots, 0]$ is used to generate the identity matrix with Toeplitz structure.

B. ANGULAR DOMAIN CHARACTERIZATION OF THE DMC

The antenna correlation can be used to characterize the angular domain distribution. According to several channel measurements and investigations, a multi-modal von Mises distribution [20], [24] is a suitable description for signals arriving at the Rx from different directions forming different clusters, which is expressed as

$$f_{\mathrm{m-VMD}}(\varphi) = \frac{1}{2\pi} \sum_i w_i \frac{e^{\kappa_i \cos(\varphi - \mu_i)}}{I_0(\kappa_i)} \tag{9}$$

with the constraint $w_i \geq 0, \forall i$ and $\sum_i w_i = 1$ on the weights. The average direction is denoted by μ_i for a certain cluster. κ_i is the parametric description of how dispersed the cluster is—a larger value corresponds to a higher concentration in

[1] This is because in channel sounding different Tx antennas are excited on orthogonal resources.

distribution. A set of the first kind of modified Bessel function with order zero, $I_0(\kappa_i)$, is introduced for the purpose of normalization. Thus, the antenna correlation matrix at the Rx can be written as

$$\mathbf{R}_{\mathrm{Rx}} = \sum_{\phi} \tilde{\mathbf{g}}_{\mathrm{R}}(\phi)\tilde{\mathbf{g}}_{\mathrm{R}}^{\dagger}(\phi) f_{\mathrm{m-VMD}}(\phi) = \mathbf{G}_{\mathrm{R}}\mathbf{K}\mathbf{G}_{\mathrm{R}}^{\dagger}, \quad (10)$$

where $\mathbf{K}$ is a diagonal matrix whose elements correspond to PDF value $f_{\mathrm{m-VMD}}(\phi)$. The matrix $\mathbf{G} \in \mathbb{C}^{N_{\mathrm{R}} \times N_{\phi}}$ is the aggregation of array pattern $\tilde{\mathbf{g}}_{\mathrm{R}}(\phi)$ from N_{ϕ} different directions into columns. In our examples, we adopt $N_{\phi} = 36$, corresponding to the angular sampling spacing of 10 degrees for a rotating horn antenna (this sampling spacing will be used in the real-world measurement in Sec V.F).[2] Note that for the MIMO case, a similar definition can be used for the Tx.

The Rx correlation matrix is an implicit function of the angular distribution $f_{\mathrm{m-VMD}}$. Besides, the correlation matrix suffers from the limitation that the resolution cannot be better than the antenna array beamwidth. Therefore, the angular description of the DMC clusters in the channel can be acquired from the angular power spectrum (APS), which can be expressed as

$$\begin{aligned} \mathrm{APS}(\phi_{\mathrm{R}}) &= \frac{1}{N_{\mathrm{f}}} \sum_{k=1}^{N_{\mathrm{f}}} |[\mathbf{H}]_{k,n_{\mathrm{R}}}|^2 \\ &= \int_{\phi} |\tilde{\mathbf{g}}_{\mathrm{R},\mathrm{n_R}}(\varphi - \phi_{\mathrm{R}})|^2 f_{\mathrm{m-VMD}}(\varphi) d\varphi \quad (11) \\ &\approx |\tilde{\mathbf{g}}_{\mathrm{R},\mathrm{n_R}}(\phi_{\mathrm{R}})|^2 f_{\mathrm{m-VMD}}(\phi_{\mathrm{R}}), \end{aligned}$$

where $[\mathbf{H}]_{k,n_{\mathrm{R}}}$ denotes the (k, n_{R})-th element in the 2D observation matrix $\mathbf{H} \in \mathbb{C}^{N_{\mathrm{f}} \times N_{\mathrm{R}}}$. It is assumed that the channel observation is conducted using a rotating horn channel sounder. The horn antenna rotates and steers towards the direction ϕ_{R}, acting as the n_{R}-th antenna in the array during the measurement. The approximation error is small when the DMC cluster angular spread is larger than the antenna beamwidth—there is less variation in the angular spectrum within the beamwidth so that the sample at ϕ_{R} can characterize its distribution. The APS can be approximately regarded as the diagonal entries of correlation matrix $\mathbf{R}_{\mathrm{Rx}}$ with antenna angular sampling $\{\boldsymbol{\phi}_{\mathrm{R}}\}$. The local maxima in the angular power spectrum indicate the horn steering direction that is closest to the directions of the impinging signals at the Rx end.

The above description was for the Rx side; for ease of exposition, much of the following section will deal with the single-input multiple-output (SIMO) case. For the multiple-input single-output (MISO) and MIMO case, we need equations for the Tx, which are completely analogous to the Rx case. Furthermore, for the MIMO case, we will use the assumption that *within each cluster*, the angular distribution of the DMC at the Tx and Rx ends are independent of each other—in other words, the Kronecker model only holds *within each cluster*. The angular covariance matrix for each DMC cluster $\mathbf{R}_{\mathrm{ang}} \in \mathbb{C}^{N_{\mathrm{T}}N_{\mathrm{R}} \times N_{\mathrm{T}}N_{\mathrm{R}}}$ can be thus decomposed into the Kronecker product

$$\mathbf{R}_{\mathrm{ang}} = \mathbf{R}_{\mathrm{Tx}} \otimes \mathbf{R}_{\mathrm{Rx}}, \quad (12)$$

where $\mathbf{R}_{\mathrm{Tx}} \in \mathbb{C}^{N_{\mathrm{T}} \times N_{\mathrm{T}}}$ and $\mathbf{R}_{\mathrm{Rx}} \in \mathbb{C}^{N_{\mathrm{R}} \times N_{\mathrm{R}}}$ stand for the Tx and Rx correlation matrix, respectively. Note however that the DoAs and DoDs of the overall channel do not fulfill this Kronecker condition; this will be further discussed in Sec III.F.

C. FREQUENCY-ANGLE DMC MODEL

Multiple MIMO channel measurements [21], [25], [30]–[33] have shown that each DMC cluster can be typically associated with a specular cluster. The physical interpretation of this phenomenon could be attributed to the fact that DMC stems from rough surfaces that give rise to both specular and diffuse components, or from interactions with smaller objects surrounding larger specular clusters. For example, if one of the MPCs corresponds to the reflection from a building exterior, there might exist more weak reflections coming from the same building, e.g., from the rough surfaces, corners, glasses, metal door/window frames, *etc*. The delays and directions of the weak components are similar to the specular components since they come from the same building.

Based on observation from those measurements, *each* DMC cluster in the delay domain shows an exponential-decay behavior (as opposed to the conventional model assuming *only one* exponential decay of the *total* DMC PDP). Especially if the clusters are well-separated, the resulting *total* channel cannot be described by a Kronecker model: there is a clear relationship between the base delay of a cluster and its angle (related to the object creating the scattering), while the Kronecker model is based on the assumption of independence among different parametric domains—delay, DoA, and DoD.

For this reason, in the multiple DMC cluster model, only the *individual* DMC clusters follow the Kronecker model.[3] In a SIMO scenario as an example:

$$\mathbf{R} = \sum_{i} \mathbf{R}_{\mathrm{Rx},i} \otimes \mathbf{R}_{\mathrm{f},i} + \sigma_{\mathrm{n}}^2 \mathbf{I}, \quad (13)$$

where $\mathbf{R}_{\mathrm{Rx},i}$ and $\mathbf{R}_{\mathrm{f},i}$ stand for the DMC covariance matrix in DoA and delay domain, respectively. The independence between DMC clusters and the noise results in the summation of covariance matrices, while the covariance matrix for each DMC cluster is expressed by the Kronecker product.

III. DMC ESTIMATION

The covariance matrix of each DMC process in (13) is a square matrix with $N_{\mathrm{f}}N_{\mathrm{R}}$ dimensionality. It is almost infeasible to solve the problem of fitting a parametric description

[2]While it is not strictly necessary to choose N_{ϕ} equal to the sampling spacing of measurements, it does mitigate the impact of potential calibration errors in the horn antenna pattern.

[3]This behavior is similar to the model for the MPCs in the COST 259 model [34].

to it with limited computational resources. We thus perform fitting sequentially in the two dimensions, frequency/delay and angle. Since delay filtering can be implemented more easily than spatial filtering, the estimation first focuses on the delay domain. We first estimate the cluster start times, and then use delay gating to isolate the contributions of this cluster, whose corresponding DMC angular descriptions can be evaluated consequently. Thus, the full description of DMC clusters with appropriate combination of the delay and angular parameters is built up. Note that even though we are treating the two domains sequentially, we *do not* make a Kronecker assumption; the angular characteristics can be different for each estimated delay cluster. This strategy is demonstrated in Sec III.A in delay cluster identification and initialization. The delay parameters of the DMC are then estimated via the gradient descent method, which is reproduced from [4] for the sake of convenience in the later discussion, in Sec III.B, and the angular estimation is developed in Sec III.D.

The number of DMC processes will be determined in the initialization by the detection of the number of peaks, sequentially in PDP and angular spectrum. Since the estimated DMC (*i.e.*, residue after subtraction of the specular MPCs from the channel) has a lower dynamic range, we use the raw *total* average Power Delay Profile (APDP) as the basis for the estimation. This is justified by the fact that the locations of specular and DMC clusters usually coincide, while at the same time, it retains the capability to identify strong DMC clusters that are not associated with a specular component (e.g., originating from vegetation).

Algorithm 1: DMC delay cluster parameter estimation

function `DMC_delay_est(`$\mathbf{H}$, $\hat{\mathbf{H}}_{\mathrm{sp}}$`)`

```
// DMC delay cluster initialization
// Focus on raw channel observations H
```

Compute APDP $\boldsymbol{\zeta}(\tau)$, (16);

Detect all local maxima outside clear-out regions, $\{\tau_{\mathrm{c}}\}$, (14);

Initialize $\{\breve{\alpha}\}$, $\{\breve{\tau}_{\mathrm{d}}\}$, $\{\breve{\boldsymbol{\beta}}_{\mathrm{d}}\}$, $\breve{\sigma}_{\mathrm{n}}^2$, (17).

```
// estimate DMC delay parameters
// Focus on extracted observations
```

$\mathbf{H}_r = \mathbf{H} - \hat{\mathbf{H}}_{\mathrm{sp}}$

while `non-convergence` **do**

- Compute residual APDP, $\boldsymbol{\zeta}_{\mathrm{r}}(\tau)$, (19);
- Compute Jacobian matrix, (24);
- Calculate the error function, (26);
- Gradient-descend parameter update, (27);

return $\{\hat{\alpha}\}$, $\{\hat{\tau}_{\mathrm{d}}\}$, $\{\hat{\boldsymbol{\beta}}_{\mathrm{d}}\}$, $\hat{\sigma}_{\mathrm{n}}^2$

A. DELAY CLUSTER IDENTIFICATION AND INITIALIZATION

Since later computation steps (gradient descent method) do not provide the ability to add or remove DMC clusters, it is essential to detect all possible DMC clusters in the initialization. The basic detection mechanism is based on the observation that at the base delay of each DMC cluster, there is a sudden change in the PDP. In order to provide robustness, and exclude the impact of the time-decay constant of the cluster, we base the identification on the peaks of the second-order derivative of the original PDP [35]. The set of the selected delay indices, S, is thus

$$S \triangleq \left\{ \tau_{\mathrm{c}} \;\middle|\; -\frac{d^2}{d\tau^2}\boldsymbol{\zeta}(\tau)\Big|_{\tau=\tau_{\mathrm{c}}} \geq p_{\mathrm{thresh}} \right\}, \tag{14}$$

where τ_{c} refers to the base delay candidates, and p_{thresh} is the threshold value. The second-order derivative of the discrete PDP can be approximated by

$$-\frac{d^2}{d\tau^2}\boldsymbol{\zeta}(\tau)\Big|_{\tau=\tau_{\mathrm{c}}} \approx \frac{2\boldsymbol{\zeta}(\tau_{\mathrm{c}}) - \boldsymbol{\zeta}(\tau_{\mathrm{c}}-\Delta\tau) - \boldsymbol{\zeta}(\tau_{\mathrm{c}}+\Delta\tau)}{\Delta\tau^2}, \tag{15}$$

where $\Delta\tau$ is the spacing between two adjacent delay samples. The average PDP of the *raw* channel observations is defined as

$$\boldsymbol{\zeta}(\tau) \triangleq \mathbb{E}_{\ell}\left[|\mathbf{F}^{\dagger}\mathbf{h}_{\ell}|^2\right], \tag{16}$$

where $\mathbf{h}_{\ell} \in \mathbb{C}^{N_{\mathrm{f}}\times 1}$ stands for the frequency domain sampling at ℓ-th antenna port.

It is possible that some of the detected peaks do not correspond to starts of new clusters, but rather arise from ripples in the PDP that are either caused by the fact that real-world PDPs might not decay in a strictly monotonic way, or even if they do, their *estimates* suffer from random fluctuations. Thus, to avoid the over-estimation of the number of clusters, we create a *clear-out region*: once the earlier cluster is detected, the next detection is forced to occur at least several delay bins later. While this could prevent separating two closely spaced DMC clusters, it is often preferable anyway to regard two such close clusters as a "merged" DMC cluster and only determine the first base delay. The size of the clear-out region should be on the order of the typical cluster decay constant. Note that if two clusters have similar delays but different angles, they will be separated in the subsequent steps.

Based on this cluster identification, the initialization of the parameters is conducted. The detected local maxima correspond to the base delay for each DMC cluster, and their height to the maximum power of the exponential decay. The decay parameter, β_{d}, can be initialized by fitting the PDP between the base delay of the current cluster and its following one. Assuming $\{\tau_{\mathrm{c}}\}$ are sorted in the ascending order and $\tau_{\mathrm{c},i}$ stands for the i-th component from the detection, the initialization can be expressed as

$$\begin{aligned}
\breve{\alpha}_i &= \boldsymbol{\zeta}(\tau_{\mathrm{c},i}) \\
\breve{\tau}_{\mathrm{d},i} &= \tau_{\mathrm{c},i} \\
\breve{\beta}_{\mathrm{d},i} &= \frac{\ln\boldsymbol{\zeta}(\tau_{\mathrm{c},i}) - \ln\boldsymbol{\zeta}(\tau_{\mathrm{c},i+1})}{\tau_{\mathrm{c},i+1} - \tau_{\mathrm{c},i}} \\
\breve{\sigma}_{\mathrm{n}}^2 &= \bar{\boldsymbol{\zeta}}_0,
\end{aligned} \tag{17}$$

where $\breve{(\cdot)}$ refers to the initialization of parameters, and the average noise level can be estimated, e.g., from measurements where no signal power is present, denoted by $\bar{\zeta}_0$. The summary of the initialization can be found in Alg. 1.

B. DELAY PARAMETER ESTIMATION

Based on the assumption mentioned in Sec II, the DMC process and noise are both modeled as multivariate complex Gaussian distribution. A maximum-likelihood estimator is employed to estimate the descriptive parameters in the channel under the assumptions. The log-likelihood function of the DMC clusters in the delay domain can be expressed as

$$\begin{aligned}\mathcal{L}(\boldsymbol{\theta}|\mathbf{H}) = &- N_{\mathrm{R}} N_{\mathrm{f}} \ln \pi - N_{\mathrm{R}} \ln \det(\mathbf{R}_{\mathrm{f}}(\boldsymbol{\theta})) \\ &- \sum_{\ell=1}^{N_{\mathrm{R}}} (\mathbf{h}_\ell - \hat{\mathbf{h}}_{\mathrm{sp},\ell})^\dagger \mathbf{R}_{\mathrm{f}}^{-1}(\boldsymbol{\theta})(\mathbf{h}_\ell - \hat{\mathbf{h}}_{\mathrm{sp},\ell}). \quad (18)\end{aligned}$$

The reconstructed contributions of the specular MPCs at the same antenna combination, $\hat{\mathbf{h}}_{\mathrm{sp},\ell}$, is built up based on either initialization or the estimation from the previous RiMax iteration. Thus, the APDP of the DMC calculated from the residue after the MPC removal, $\boldsymbol{\zeta}_{\mathrm{r}}(\tau)$, can be written as

$$\boldsymbol{\zeta}_{\mathrm{r}}(\tau) = \mathbb{E}_\ell \left[|\mathbf{F}^\dagger (\mathbf{h}_\ell - \hat{\mathbf{h}}_{\mathrm{sp},\ell})|^2 \right] = \mathrm{diag}\left[\mathbf{F}^\dagger \mathbf{R}_{\mathrm{f}} \mathbf{F} \right], \quad (19)$$

where the covariance matrix in the delay domain is the residual averaged over all "realizations" based on the definition, $\mathbf{R}_{\mathrm{f}} \triangleq \mathbb{E}_\ell \left[(\mathbf{h}_\ell - \hat{\mathbf{h}}_{\mathrm{sp},\ell})(\mathbf{h}_\ell - \hat{\mathbf{h}}_{\mathrm{sp},\ell})^\dagger \right]$. The APDP can be obtained from the diagonal elements of the Fourier transform of the covariance matrix of the total channel.

Conversely, the log-likelihood function can be rewritten by approximating the covariance matrix $\mathbf{R}_{\mathrm{f}}$ using the APDP vector, $\boldsymbol{\zeta}(\boldsymbol{\theta})$. Note that the parametric APDP is denoted by $\boldsymbol{\zeta}(\boldsymbol{\theta})$, while the channel residual APDP is denoted by $\boldsymbol{\zeta}_{\mathrm{r}}(\tau)$. Consequently, the log-likelihood function can be expressed as (20) on the next page, where $\zeta_k(\boldsymbol{\theta})$ and $\zeta_{\mathrm{r},k}(\tau)$ denote the reconstructed APDP and residual APDP for the k-th delay-domain sample, respectively, and (19) is used. A computationally efficient method to reduce matrix multiplications and simplify the expression is introduced as

$$\boldsymbol{\zeta}(\boldsymbol{\theta}) = \mathbf{F}^\dagger (\mathbf{W}_1 \mathbf{v}(\boldsymbol{\theta}) + \mathbf{W}_2 \mathbf{v}^\dagger(\boldsymbol{\theta})), \quad (21)$$

using the weight matrices $\mathbf{W}_1$, $\mathbf{W}_2$ defined in [4, p. 135] and $\mathbf{v}(\boldsymbol{\theta})$ is the generating vector to the Toeplitz matrix (8).

The APDP is now expressed as a function of DMC parameters in a vector, $\mathbf{v}(\boldsymbol{\theta})$. The gradient descent method requires the partial derivative with respect to each parameter to optimize the estimation. Since $\mathbf{W}_1$ and $\mathbf{W}_2$ are not a function of any DMC parameters, the first-order derivative of (21) with respect to DMC parameters can be expanded as

$$\frac{\partial}{\partial \theta_q} \boldsymbol{\zeta}(\boldsymbol{\theta}) = \mathbf{F}^\dagger (\mathbf{W}_1 \frac{\partial \mathbf{v}(\boldsymbol{\theta})}{\partial \theta_q} + \mathbf{W}_2 \frac{\partial \mathbf{v}^\dagger(\boldsymbol{\theta})}{\partial \theta_q}), \quad (22)$$

where θ_q refers to an arbitrary DMC parameter. The first-order derivative of the Toeplitz generating vector with respect to parameters, $\frac{\partial}{\partial \theta_q} \mathbf{v}(\boldsymbol{\theta})$, can thus be expressed as (23). The vector derivatives can be obtained using the chain rule by substituting (23) into (22).

The modified Jacobian matrix, $\mathbf{D}(\boldsymbol{\theta})$, which is the aggregation of $3N_{\mathrm{DMC}}+1$ vectors (N_{DMC} is the number of DMC clusters) of the first-order derivatives with respect to DMC parameters in the delay domain, is defined as

$$\begin{aligned}\mathbf{D}(\boldsymbol{\theta}) = &\mathrm{diag}[\boldsymbol{\zeta}(\boldsymbol{\theta})]^{-1} \cdot \\ &\left[\frac{\partial}{\partial \boldsymbol{\alpha}} \boldsymbol{\zeta}(\boldsymbol{\theta}), \frac{\partial}{\partial \boldsymbol{\tau}_{\mathrm{d}}} \boldsymbol{\zeta}(\boldsymbol{\theta}), \frac{\partial}{\partial \boldsymbol{\beta}_{\mathrm{d}}} \boldsymbol{\zeta}(\boldsymbol{\theta}), \frac{\partial}{\partial \sigma_{\mathrm{n}}^2} \boldsymbol{\zeta}(\boldsymbol{\theta}) \right], \quad (24)\end{aligned}$$

where the matrix inversion of the diagonal matrix whose elements are the reconstructed parametric APDP comes from the derivative of the log-likelihood function (20) as

$$\begin{aligned}\frac{\partial}{\partial \theta_q} \mathcal{L}(\boldsymbol{\theta}|\mathbf{H}) &= -N_{\mathrm{R}} \sum_{k=1}^{N_{\mathrm{f}}} \left[\frac{1}{\zeta_k(\boldsymbol{\theta})} - \frac{\zeta_{\mathrm{r},k}(\tau)}{\zeta_k^2(\boldsymbol{\theta})} \right] \frac{\partial \zeta_k(\boldsymbol{\theta})}{\partial \theta_q} \\ &= N_{\mathrm{R}} \mathbf{D}_q^T(\boldsymbol{\theta}) \left[\mathrm{diag}[\boldsymbol{\zeta}(\boldsymbol{\theta})]^{-1} \boldsymbol{\zeta}_{\mathrm{r}}(\tau) - \mathbb{1}_{N_{\mathrm{f}}} \right], \quad (25)\end{aligned}$$

where $\mathbf{D}_q(\boldsymbol{\theta})$ stands for the q-th column in the matrix and $\mathbb{1}_{N_{\mathrm{f}}}$ refers to the all-one vector with dimensionality N_{f}. A detailed proof that the estimator is unbiased and approximation of Hessian matrix with the modified Jacobian matrix can be found in [4].

The term within the square brackets in (25) can be regarded as the difference between the reconstructed DMC and the residual PDP (after removing specular MPCs). Therefore, the error function is defined as

$$\boldsymbol{\epsilon} \triangleq \mathrm{diag}[\boldsymbol{\zeta}(\boldsymbol{\theta})]^{-1} \boldsymbol{\zeta}_{\mathrm{r}}(\tau) - \mathbb{1}_{N_{\mathrm{f}}}. \quad (26)$$

The estimated DMC parameters are the solution to $\nabla_{\boldsymbol{\theta}} \mathcal{L}(\boldsymbol{\theta}|\mathbf{H}) = \mathbf{0}$, which are also the minimizer to the error function.

In this paper, we adopt a Gauss-Newton method for all gradient descent steps due to its simplicity. Thus, the estimated parameters update, $\Delta\boldsymbol{\theta}$, along the gradient descent direction is computed as [4]

$$\Delta \hat{\boldsymbol{\theta}}^{\{p+1\}} = \arg\min_{\Delta\boldsymbol{\theta}} \| \mathbf{D}(\hat{\boldsymbol{\theta}}^{\{p\}}) \Delta\boldsymbol{\theta} - \boldsymbol{\epsilon}^{\{p\}} \|_{\mathrm{F}}^2, \quad (27)$$

where the superscript $(\cdot)^{\{p\}}$ refers to the estimated parameters in the p-th iterative loop. The update for the next iteration, $\Delta\hat{\boldsymbol{\theta}}^{\{p+1\}}$, is based on the Jacobian matrix $\mathbf{D}(\hat{\boldsymbol{\theta}}^{\{p\}})$ and the corresponding error function $\boldsymbol{\epsilon}^{\{p\}}$ with the evaluated parameters from the previous iteration. The update of the parameters is added to the previous estimation, $\hat{\boldsymbol{\theta}}^{\{p+1\}} = \hat{\boldsymbol{\theta}}^{\{p\}} + \eta \Delta \hat{\boldsymbol{\theta}}^{\{p+1\}}$, with a step size variable η. The step size is chosen to satisfy $\mathcal{L}(\hat{\boldsymbol{\theta}}^{\{p+1\}}|\mathbf{H}) > \mathcal{L}(\hat{\boldsymbol{\theta}}^{\{p\}}|\mathbf{H})$ so that the convergence and the convergence speed can be guaranteed. The parameter estimation with proper update will be iterated until convergence, which could be judged to be reached either due to relative update magnitude falling below a threshold, or a maximum number of iterations performed. The pseudo-code of the iterative gradient-based method can be found in Alg. 1.

$$\mathcal{L}(\boldsymbol{\theta}|\mathbf{H}) \propto -N_{\mathrm{R}} \ln\det(\mathbf{F}\mathrm{diag}[\boldsymbol{\zeta}(\boldsymbol{\theta})]\mathbf{F}^{\dagger}) - \sum_{\ell=1}^{N_{\mathrm{R}}} (\mathbf{h}_{\ell} - \hat{\mathbf{h}}_{\mathrm{sp},\ell})^{\dagger}\mathbf{F}\mathrm{diag}[\boldsymbol{\zeta}(\boldsymbol{\theta})]^{-1}\mathbf{F}^{\dagger}(\mathbf{h}_{\ell} - \hat{\mathbf{h}}_{\mathrm{sp},\ell})$$
$$= -N_{\mathrm{R}} \sum_{k=1}^{N_{\mathrm{f}}} \left[\ln \zeta_k(\boldsymbol{\theta}) + \frac{\zeta_{\mathrm{r},k}(\tau)}{\zeta_k(\boldsymbol{\theta})} \right], \tag{20}$$

$$\frac{\partial}{\partial \alpha}\mathbf{v}(\boldsymbol{\theta}) = \left[\frac{1}{\beta_{\mathrm{d}}} \quad \frac{e^{-j2\pi\frac{1}{N_{\mathrm{f}}}\tau_{\mathrm{d}}}}{\beta_{\mathrm{d}} + j2\pi\frac{1}{N_{\mathrm{f}}}} \quad \cdots \quad \frac{e^{-j2\pi\frac{N_{\mathrm{f}}-1}{N_{\mathrm{f}}}\tau_{\mathrm{d}}}}{\beta_{\mathrm{d}} + j2\pi\frac{N_{\mathrm{f}}-1}{N_{\mathrm{f}}}} \right]^T$$
$$\frac{\partial}{\partial \tau_{\mathrm{d}}}\mathbf{v}(\boldsymbol{\theta}) = \frac{\alpha}{N_{\mathrm{f}}} \left[0 \quad \frac{-j2\pi e^{-j2\pi\frac{1}{N_{\mathrm{f}}}\tau_{\mathrm{d}}}}{\beta_{\mathrm{d}} + j2\pi\frac{1}{N_{\mathrm{f}}}} \quad \cdots \quad \frac{-j2\pi(N_{\mathrm{f}}-1)e^{-j2\pi\frac{N_{\mathrm{f}}-1}{N_{\mathrm{f}}}\tau_{\mathrm{d}}}}{\beta_{\mathrm{d}} + j2\pi\frac{N_{\mathrm{f}}-1}{N_{\mathrm{f}}}} \right]^T \tag{23}$$
$$\frac{\partial}{\partial \beta_{\mathrm{d}}}\mathbf{v}(\boldsymbol{\theta}) = -\alpha \left[\frac{1}{\beta_{\mathrm{d}}^2} \quad \frac{e^{-j2\pi\frac{1}{N_{\mathrm{f}}}\tau_{\mathrm{d}}}}{(\beta_{\mathrm{d}} + j2\pi\frac{1}{N_{\mathrm{f}}})^2} \quad \cdots \quad \frac{e^{-j2\pi\frac{N_{\mathrm{f}}-1}{N_{\mathrm{f}}}\tau_{\mathrm{d}}}}{(\beta_{\mathrm{d}} + j2\pi\frac{N_{\mathrm{f}}-1}{N_{\mathrm{f}}})^2} \right]^T$$
$$\frac{\partial}{\partial \sigma_{\mathrm{n}}^2}\mathbf{v}(\boldsymbol{\theta}) = \mathbf{e}_0,$$

C. DELAY-ANGLE CLUSTER IDENTIFICATION

To estimate the angular distribution corresponding to a particular delay cluster, it is crucial to reduce the "interference" from other clusters. Each estimated delay cluster is captured by delay gating, based on which the corresponding angular spectrum will be analyzed. In other words, the angular descriptions are evaluated for each estimated DMC cluster in the delay domain. The delay gating can thus be expressed as

$$\mathbf{H}_{\mathrm{gate}} = \mathbf{F}\left[\mathbf{win}_i \odot \mathbf{F}^{-1}\mathbf{H}_{\mathrm{r}}\right], \tag{28}$$

where $\odot$ denotes the Hadamard product and the concentration window, $\mathbf{win}_i \in \mathbb{R}^{N_{\mathrm{f}}\times 1}$, is defined as

$$\mathbf{win}_i = \begin{cases} 1, & \hat{\tau}_{\mathrm{d},i} \le \tau < \hat{\tau}_{\mathrm{d},i+1} \\ 0, & \text{otherwise} \end{cases}, \tag{29}$$

where $\hat{\tau}_{\mathrm{d},i}$ stands for the base delay for i-th DMC cluster which has already been estimated. The rectangular window function is employed to isolate the DMC cluster from the others based on the estimated delay parameters. Note that the rectangular window is applied for the purpose of isolating clusters only.

The angular information of the delay-gated channel observation described by the APS, $\boldsymbol{\nu}_{\tau}(\phi)$, is obtained by substituting $\mathbf{H}$ with $\mathbf{H}_{\mathrm{gate}}$ in (11) and sampling at the horn antenna steering direction, $\{\boldsymbol{\phi}\} = \{\boldsymbol{\phi}_{\mathrm{R}}\}$. The contribution from previous estimated DMC clusters whose "tails" fall into the current delay gate is excluded as

$$\boldsymbol{\nu}_{\tau,i+1}(\phi) = \boldsymbol{\nu}_{\tau}(\phi) - \sum_{j=1}^{i} p_j \boldsymbol{\nu}_{\tau,j}(\phi), \tag{30}$$

where $p_j = \|\mathbf{win}_{i+1} \odot \mathbf{F}^{-1}\hat{\mathbf{H}}_{\mathrm{DMC},j}\|$ is the reconstructed DMC "leakage" power and $\hat{\mathbf{H}}_{\mathrm{DMC},j}$ is the reconstructed DMC cluster using its estimated parameters.

The angular parameters are initialized from the delay-gated observations. The local maxima of the angular spectrum, $\{\boldsymbol{\phi}_{\mathrm{c}}\}$, can be found by

$$\{\boldsymbol{\phi}_{\mathrm{c}}\} \triangleq \Big\{ \phi_s \Big| \boldsymbol{\nu}_{\tau}(\phi_s) > \boldsymbol{\nu}_{\tau}(\phi_{s-1}) \cap \boldsymbol{\nu}_{\tau}(\phi_s) > \boldsymbol{\nu}_{\tau}(\phi_{s+1}) \cap \boldsymbol{\nu}_{\tau}(\phi_s) > p_{\mathrm{thresh}} \Big\}. \tag{31}$$

From the assumption of the DMC angular model, *i.e.*, the multi-modal von Mises distribution, the APS is proportional to the probability density function. The detected local maxima correspond to the mean directions of the DMC angular cluster when clusters are well-separated; the peaks are shifted from the direction of the cluster center towards each other if the clusters overlap. The initialization of κ does not have an explicit solution, so we consider using an empirical value [36]. The mixture weights, w_i, are proportional to the total power ratio. Consequently, the initialization can be expressed as

$$\begin{aligned} \breve{\mu}_i &= \phi_{\mathrm{c},i} \\ \breve{\kappa}_i &= 5 \\ \breve{w}_i &= \frac{|\boldsymbol{\nu}_{\tau}(\phi_{\mathrm{c},i})|}{\sum_j |\boldsymbol{\nu}_{\tau}(\phi_{\mathrm{c},j})|}. \end{aligned} \tag{32}$$

However, there exists the possibility that two or more clusters have similar yet non-identical delays and different angular information.[4] With the delay initialization and estimation mechanism mentioned in Section III.A and B, the clusters are erroneously modeled to have a single delay that is then estimated. This reconstruction error dominates the total estimation performance especially when those clusters are relatively strong compared to other clusters.

Although the clusters are merged in estimation, it is still possible to determine the distinct clusters in angular-delay power spectrum (ADPS), which is computed by

$$\mathrm{ADPS}(\boldsymbol{\tau}, \boldsymbol{\phi}) \triangleq |\mathbf{F}^{-1}\mathbf{H}_{\mathrm{r}}|^2 \in \mathbb{R}^{N_{\mathrm{f}}\times N_{\phi}}. \tag{33}$$

[4]Identical delays and different angular information can be seen as a special case.

Algorithm 2: DMC angular cluster parameter estimation

```
function DMC_angle_est(H, Ĥ_sp, {τ̂_d})
    s = 1, r = 1, n_cluster = |{τ̂_d}|;
    for s ≤ n_iter do
        for r ≤ n_cluster do
            // DMC delay-angle cluster initialization
            // Focus on extracted channel observations H_r = H − Ĥ_sp
            Compute residual ADPS, (36), smoothed residual ADPS with safety margin, (34,37), and local maxima, (35)
            if one local maximum then
                Delay-gate observation, H_gate, (28);
                Compute angular spectrum, ν_τ(φ);
                Detect all local maxima ,{φ_c}, (31);
                Initialize {μ̆}, {κ̆}, {w̆}, (32).
            else
                Determine antenna/beam subset, A;
                Re-evaluate delay parameters;
                Delay-gate observation, H_gate, (28);
                Compute angular spectrum, ν_τ(φ);
                Detect all local maxima ,{φ_c}, (31);
                Initialize {μ̆}, {κ̆}, {w̆}, (32);
                n_cluster ← n_cluster + 1.
            // estimate DMC angular parameters
            // Focus on extracted channel observations H_r = H − Ĥ_sp
            while non-convergence do
                Compute Jacobian matrix, (39);
                Calculate the error function (40);
                Gradient-descend parameter update;
            r ← r + 1;
        s ← s + 1;
    return {μ̂}, {κ̂}, {ŵ}
```

In order to remedy the above-mentioned defect in the estimation mechanism, it is necessary to determine the locations of the clusters in the joint delay-angle spectrum. Moreover, each delay-angle bin may have different fading status, which causes difficulties in the detection of start delay and mean direction of the DMC cluster in ADPS. Therefore, averaging in a local region is used to smooth the fading dips. The smoothed ADPS is computed by

$$\widetilde{\text{ADPS}} = \frac{1}{IJ}\text{ADPS} \circledast_2 \mathbb{1}_{I\times J}, \tag{34}$$

where the operator $\circledast_2$ denotes the two-dimensional (2D) circular convolution. $\mathbb{1}_{I\times J}$ is the all-one matrix with dimensionality $I \times J$, where I and J are the window length of moving average in delay and angular domain, repectively.

The local maxima detected from the smooth ADPS, $\widetilde{\text{ADPS}}$, is expressed as

$$\begin{aligned}\text{LM} \triangleq\{(\tau_{\text{LM}}, \phi_{\text{LM}})| \\ \widetilde{\text{ADPS}}(\tau,\phi) > \widetilde{\text{ADPS}}(\tau \pm \Delta\tau, \phi \pm \Delta\phi)\},\end{aligned} \tag{35}$$

where $\Delta\tau$ and $\Delta\phi$ is the delay and angular bin spacing, respectively.

If several local maxima within a certain delay gate are observed, it indicates the existence of several clusters that can be separated. For each candidate cluster, it is feasible to select a subset of the steering antennas/beams, $\mathcal{A}$, that isolates the cluster from others by limiting the receiving angular range. This angular isolation enables the DMC delay parameter estimation for each cluster without interference from other clusters in this gating interval. Therefore, the delay estimation algorithm can be employed to re-evaluate its parameters. The initialization is $\breve{\tau} = \tau_{\text{LM}} - I/2$, in which τ_{LM} is the detected delay of local maximum from the smoothed ADPS; the start delay deviates by a half of the 2D convolution window length to the local maximum, which is compensated in the initialization.

To further eliminate the "inter-cluster" interference, each cluster contribution is removed from the ADPS after its estimation, giving the residual ADPS

$$\text{ADPS}_{\text{res},i} = \text{ADPS} - \sum_{j\neq i}\widehat{\text{ADPS}}_j, \tag{36}$$

where $\widehat{\text{ADPS}}_j$ denotes the ADPS reconstructed from its estimation. However, the removed ADPS is the estimate of the statistical average, while the actually observed signal includes the fading status. This might lead to imperfect subtraction and introduce "ghost" clusters. Thus, a safety margin is added in the removal, so that only a cluster above the safety margin is determined as a distinct cluster. This can be expressed as

$$\text{ADPS}^{\text{sm}}_{\text{res},i} = \max(\text{ADPS} - \sum_{j\neq i}\widehat{\text{ADPS}}_j \cdot \xi, 0), \tag{37}$$

where ξ is the safety margin. The safety margin is defined similarly to the fading margin and is introduced to ensure that the detected peaks correspond to actual clusters rather than spurious fading peaks. The determination of the safety margin relies on two key factors: the statistical distribution and the admissible false alarm probability. Since each ADPS bin is assumed to follow Rayleigh fading, the average over multiple delay-angle bins follows the Nakagami-m distribution, which effectively captures the amplitude statistical characteristics. Since the antenna sampling interval (10 degrees) is close to the 3-dB beamwidth (13 degrees), $J = 1$ is selected. If the angular sampling in the channel observation is denser, a different value of J can be chosen. The false alarm probability is set to 1%.

It may require several iterations n_{iter} to let the estimation converge to stable results. It is worthwhile to mention that at the first iteration, $s = 1$, it is impossible to remove the

contribution from clusters in the $j > i$ gating interval, which are therefore ignored on the computation of $\mathrm{ADPS}_{\mathrm{res},i}$. The summary of the initialization can be found in Alg. 2.

D. ANGULAR PARAMETER ESTIMATION

Based on (10), the first-order derivative of the antenna correlation matrix with respect to angular parameters can be expressed as

$$\begin{aligned}\frac{\partial}{\partial \mu_i} f_{\mathrm{m-VMD}}(\varphi) &= w_i \kappa_i \sin(\varphi - \mu_i) \frac{e^{\kappa_i \cos(\varphi-\mu_i)}}{2\pi I_0(\kappa_i)} \\ \frac{\partial}{\partial \kappa_i} f_{\mathrm{m-VMD}}(\varphi) &\approx w_i \cos(\varphi - \mu_i) \frac{e^{\kappa_i \cos(\varphi-\mu_i)}}{2\pi I_0(\kappa_i)} \\ \frac{\partial}{\partial w_i} f_{\mathrm{m-VMD}}(\varphi) &= \frac{e^{\kappa_i \cos(\varphi-\mu_i)}}{2\pi I_0(\kappa_i)}\end{aligned} \quad , \quad (38)$$

where the normalization $I_0(\kappa_i)$ is regarded as a constant when taking the derivative, resulting in a simplified (though approximate) expression.

The modified angular Jacobian matrix, $\mathbf{D}(\boldsymbol{\theta})$, which is the aggregation of the first-order derivative vectors with respect to angular cluster parameters, is defined in (39). The linearity property of the diagonalization operation—the derivative of the diagonal matrix is equivalent to the derivative of the diagonal elements, $\frac{\partial}{\partial \theta_q}\mathbf{K} = \mathrm{diag}(\frac{\partial}{\partial \theta_q}\mathbf{f}_{\mathrm{m-VMD}})$—is used. The variable, $\boldsymbol{\theta}$, is extended to include arbitrary angular parameters describing DMC clusters.

The estimation method is analogous to the estimation of DMC parameters in the delay domain. The error function of the angular distribution is defined as

$$\boldsymbol{\epsilon}_\phi \triangleq \mathrm{diag}(\boldsymbol{\nu}_\tau(\boldsymbol{\theta}))^{-1}\boldsymbol{\nu}_\tau(\phi) - \mathbb{1}_{N_\mathrm{R}}, \quad (40)$$

where $\boldsymbol{\nu}_\tau(\boldsymbol{\theta})$ denotes the reconstructed APS with the estimated angular parameters and $\boldsymbol{\nu}_\tau(\phi)$ corresponds to the actual APS to be fitted. With the gradient descent method, the error function is minimized to provide the parameter update on estimation.

E. COVARIANCE MATRIX RECONSTRUCTION AND INVERSION

The computational complexity of matrix inversion or reconstruction goes approximately with the third power of the matrix dimensions. In the existing RiMax DMC models, the DMC covariance matrix inversion and reconstruction take advantage of the Kronecker product property—the inverse of the covariance matrix can be decomposed into the inverse of matrices with smaller scales. Specifically, instead of a direct inversion of a square matrix with the dimensionality of $N_\mathrm{f}N_\mathrm{R}$, inversion of square matrices with dimension N_f and N_R is done instead. However, since the Kronecker model does not hold for the more accurate DMC model used here, this complexity saving cannot be achieved.

Consequently, we use a subspace (the term "subspace" refers to the parametric domains) whitening method with proper normalization to deal with the structural covariance matrix [4], where whitening is performed sequentially in the different parametric domains. The DMC clusters in the delay domain can be whitened, creating a flat process in the frequency domain. Moreover, it is also possible to whiten the APS in both Rx and Tx spectra. Therefore, all the diffuse scattering processes can be whitened, which renders the covariance matrix for each DMC cluster an identity matrix.

Based on the independence assumption between different DMC clusters, the approximation of the total covariance matrix in the delay domain is the summation of each component as in (6). Then the angular covariance matrix is written as

$$\hat{\mathbf{R}}_\mathrm{Rx} \approx \sum_i P_i \hat{\mathbf{R}}_{\mathrm{Rx},i}, \quad (41)$$

where the power for each reconstructed DMC cluster in the delay domain is given by

$$P_i = \mathrm{trace}\left(\mathrm{diag}(\hat{\mathbf{R}}_{\mathrm{f},i})\right), \quad (42)$$

in which $\hat{\mathbf{R}}_{\mathrm{f},i}$ is the reconstructed covariance matrix for DMC cluster i. A lower triangular matrix, $\mathbf{L}$, which is obtained from Cholesky decomposition of the covariance matrix in a certain parametric domain, e.g., $\hat{\mathbf{R}}_\mathrm{f} = \mathbf{L}_\mathrm{f}\mathbf{L}_\mathrm{f}^\dagger$, with proper dimension can be applied to the channel residue to whiten the process. Once the delay/frequency domain has been whitened, the angular covariance matrix of the total channel is altered, as described by the following:

$$\tilde{\mathbf{R}}_\mathrm{Rx} \approx \sum_i \tilde{P}_i \hat{\mathbf{R}}_{\mathrm{Rx},i} \propto \sum_i \hat{\mathbf{R}}_{\mathrm{Rx},i}, \quad (43)$$

since the power for each DMC cluster in the delay domain is whitened and normalized as

$$\tilde{P}_i = \mathrm{trace}\left(\mathrm{diag}(\mathbf{L}_\mathrm{f}^{-1}\hat{\mathbf{R}}_{\mathrm{f},i}\mathbf{L}_\mathrm{f}^{-\dagger})\right) \approx \mathrm{const.} \quad (44)$$

With our proposed DMC estimation algorithm, it is possible to avoid the duplicate consideration and account for the change of total covariance matrix due to the subspace whitening in other parametric domains. The estimation, not relying on the Kronecker model, can provide $\hat{\mathbf{R}}_{\mathrm{Rx},i}$ to incorporate potential covariance matrix changes from $\hat{\mathbf{R}}_\mathrm{Rx}$ to $\tilde{\mathbf{R}}_\mathrm{Rx}$. If the approximation works ideally, the whitening will show a flat spectrum in both delay and DoA domain. Thus, the flatness of the whitening results can be regarded as a measure for the accuracy of the matrix inversion approximation.

F. GENERALIZATION TO THE MIMO CASE

We now discuss the extension of DMC estimation from SIMO to MIMO scenarios, assuming that multiple clusters start at the same delay, though with different DoDs and DoAs. This assumption of an identical base delay for multiple clusters simplifies the subsequent discussion, as it avoids the need for angular isolation and the re-evaluation of the base delays, but of course incorporation of different base delays is easily possible according to the above-outlined procedure.

On one hand, a 2D gradient-descent method aiming at the estimation of the *joint* angular spectrum has high com-

$$\mathbf{D}(\boldsymbol{\theta}) = \mathrm{diag}(\mathbf{f}_{\mathrm{m-VMD}})^{-1} \cdot \left[\mathrm{diag}(\mathbf{G}_{\varphi}\mathrm{diag}(\frac{\partial}{\partial \boldsymbol{\mu}}\mathbf{f}_{\mathrm{m-VMD}})\mathbf{G}_{\varphi}^{\dagger}) \quad \mathrm{diag}(\mathbf{G}_{\varphi}\mathrm{diag}(\frac{\partial}{\partial \boldsymbol{\kappa}}\mathbf{f}_{\mathrm{m-VMD}})\mathbf{G}_{\varphi}^{\dagger}) \quad \mathrm{diag}(\mathbf{G}_{\varphi}\mathrm{diag}(\frac{\partial}{\partial \mathbf{w}}\mathbf{f}_{\mathrm{m-VMD}})\mathbf{G}_{\varphi}^{\dagger})\right]. \quad (39)$$

plexity. On the other hand, the traditional decomposition of this joint spectrum into a Kronecker product of DoD and DoA spectra will create ghost clusters in the reconstruction. Consistent with the key idea of the current paper to avoid the "ghost modes" from the Kronecker product in channels with multiple clusters, we propose the following method to generalize to MIMO channels.

First, the (marginal) DoD and DoA spectra are generated separately from the observation. With the same procedure and techniques from initialization to estimation, the parameters are estimated accordingly for both Tx and Rx side—average direction $\{\hat{\mu}_{\mathrm{D}}\}$, spread $\{\hat{\kappa}_{\mathrm{D}}\}$, and weights $\{\hat{w}_{\mathrm{D}}\}$ for DoD spectrum, and average direction $\{\hat{\mu}_{\mathrm{A}}\}$, spread $\{\hat{\kappa}_{\mathrm{A}}\}$, and weights $\{\hat{w}_{\mathrm{A}}\}$ for DoA spectrum, respectively. A Cartesian product creates all element combinations between two sets, $\{\hat{\mu}_{\mathrm{D}}\}$ and $\{\hat{\mu}_{\mathrm{A}}\}$, though this now includes ghost clusters.

Using the smoothed 2D joint APS, computed from 2D spectrum $\boldsymbol{\nu}(\phi_{\mathrm{D}}, \phi_{\mathrm{A}})$, using 2D moving average, it is possible to eliminate the ghost clusters. The true angular cluster's location coincides with the local maximum of the 2D joint spectrum above a certain threshold. The joint APS at the location of ghost clusters either does not exhibit local maxima or exhibits such low power that it falls under the threshold. Therefore, the angular cluster locations $\{(\hat{\mu}_{\mathrm{D},m}, \hat{\mu}_{\mathrm{A},n})\}$ are determined by filtering according to the aforementioned conditions.

Another problem is to allocate proper weights to each cluster from two estimated sets of weights, $\{\hat{w}_{\mathrm{D}}\}$ and $\{\hat{w}_{\mathrm{A}}\}$. With the determined cluster centers, $\{(\hat{\mu}_{\mathrm{D},m}, \hat{\mu}_{\mathrm{A},n})\}$, the APS contribution from each cluster, $\hat{\boldsymbol{\nu}}_{i,j}$, can be constructed with the corresponding spread, $\hat{\kappa}_{\mathrm{D},i}$ and $\hat{\kappa}_{\mathrm{A},j}$. The actual APS is the weighted sum of all angular clusters,

$$\boldsymbol{\nu}(\phi_{\mathrm{D}}, \phi_{\mathrm{A}}) = \sum_{\{(i,j)\}} w_{i,j}\hat{\boldsymbol{\nu}}_{i,j}(\phi_{\mathrm{D}}, \phi_{\mathrm{A}}), \quad (45)$$

which can be solved by a linear estimator as $\hat{\mathbf{w}} = \hat{\mathbf{A}}^{+}\mathbf{a}$, where $\mathbf{a} = \mathrm{vec}(\boldsymbol{\nu}(\phi_{\mathrm{D}}, \phi_{\mathrm{A}}))$ is the vectorized 2D joint APS. The pseudo-inverse of the matrix, $\hat{\mathbf{A}}$, whose columns are the vectorized reconstructed APSs, $\mathrm{vec}(\hat{\boldsymbol{\nu}}_{i,j}(\phi_{\mathrm{D}}, \phi_{\mathrm{A}}))$, is applied to solve the linear problem.

When elevation is incorporated into the SIMO case, the algorithm can be generalized by using a similar step as for the generalization from (azimuth-only) SIMO to MIMO. Specifically, the DoA-DoD azimuthal joint spectrum will then become a DoA azimuth-elevation joint spectrum, while the core estimation process and weight assignment remain unchanged. This method can also be applied to the more general scenario where both azimuth and elevation angles are incorporated in MIMO channel measurements.

IV. NUMERICAL RESULTS

To verify our proposed algorithm, this section presents numerical results based on a synthetic channel model, which ensures that the ground truth is known. Unless stated otherwise, we assume that the sounding system measures the transfer function by scanning over 101 frequency points within a 1GHz bandwidth. The sounding system is equipped with rotating horn antennas at Tx and Rx ends with beamwidth of approximately 13 degrees; measurements are taken with the horns rotating in 10-degree steps. For ease of explanation, we first show the results of a SIMO system, with parameters chosen to emulate a rotation-horn setup at the Rx end; an omnidirectional antenna is deployed at the Tx end. MIMO angular estimation results are shown in Sec IV.C. Note that Figs. 1–9 show results for a particular realization of the DMC, which serve to illustrate the specifics of the evaluation. The results of statistical evaluations in Figs. 10–12 are based on 200 synthetic pure DMC channels. Fig. 13, which presents the estimation accuracy, then takes the ensemble average over 50 synthetic channels with 20 DMC and noise realizations each. For the sake of convenience, the PDP's horizontal axis is expressed in terms of propagation distance, which is equivalent to the time delay multiplied by the speed of light.

The propagation channel is assumed to have five MPCs (widely separated, so that each specular MPC "cluster" consists of just a single component) and five corresponding DMC clusters, as shown in Fig. 1. The channel parameters are randomly generated and shown in Table 1. The base delay and mean direction for each DMC coincide with the corresponding MPC's delay and DoA. The two closest (in the delay domain) MPCs and their corresponding DMC clusters have a runlength difference of 0.5 m, which is higher than the Fourier delay resolution of 0.3 m. The CLEAN algorithm is used to initialize the estimation of the specular components, which are then used as input to the RiMax algorithm. Also shown in this figure are the MPCs estimated with CLEAN when 25 discrete MPCs are assumed; the estimated MPCs contribute to the main peak in each cluster. The CLEAN algorithm tends to account for the residue from imperfect subtraction, as well as the DMC contribution that is not in the signal model, by placing more specular components in the solution. The number 25 is empirically selected such that the detection of all MPC clusters is guaranteed.

The CLEAN algorithm is commonly employed for the initialization of MPC parameters in RiMax, due to its simplicity in implementation. By applying the principle of successive interference cancellation, the estimated contribution of each MPC is subtracted after being determined. The residue, remaining after MPC extraction from the channel observation, is considered as unexplained by specular

components and is thus used for the initialization of the DMC. Both the initialized MPC and DMC cluster parameters are fed into RiMax for parameter estimation enhancement under its algorithmic framework, with our proposed DMC estimation algorithm integrated into the process. In RiMax, the optimization of MPC and DMC parameters proceeds alternatingly—DMC estimation is performed using the pre-estimated MPCs, followed by updating the MPC parameters based on the reconstructed covariance matrix of DMC clusters and the noise and adjusting the residue accordingly. If the ratio between the MPC power error bound and its power falls below a predefined threshold [4], the MPCs are excluded from the solution.

If consecutive snapshots (with potentially small movement of Tx and/or Rx) are available, the MPC and DMC parameters exhibit minor variations across snapshots. As a result, RiMax can be executed for the first snapshot, with the initialization provided by the CLEAN algorithm. For the subsequent snapshots, the previously estimated MPC and DMC parameters can serve as the initialization, being considered no longer requisite for the CLEAN algorithm in the later snapshots.

Figure 1 illustrates the estimated MPCs using CLEAN under the assumption of 25 discrete MPCs, where the estimated MPCs correspond to the main peaks within each cluster. The CLEAN algorithm tends to account for the residuals from imperfect subtraction, as well as for the contribution of the DMC, by introducing additional specular components, also known as *ghost MPCs*, into the solution. The selection of 25 MPCs is made empirically to ensure the detection of all clusters. The residue after the subtraction, which cannot be ascribed to specular MPCs, is used to initialize the DMC contribution.

TABLE 1: Synthetic channel parameters

		1	**2**	**3**	**4**	**5**
MPC	τ[m]	2.5	7.7	14.3	14.8	22.1
	ϕ[°]	175.2	4.8	177.6	-20.9	-55.9
	γ[dB]	10.1	3.3	0.1	-2.1	-6.9
DMC	α	-8.8	-16.6	-15.5	-20.9	-22.4
	β_d	0.3	0.4	0.4	0.3	0.2
	κ	5.9	2.8	4.4	4.5	5.5

A. DMC ESTIMATION IN DELAY DOMAIN

The first step is cluster initialization. As discussed in Sec III.A, its second-order derivative may suffer from dynamic range reduction, as illustrated by the blue solid line in the lower subplot in Fig. 2, so that noise fluctuation may lead to false alarms and missed detections of cluster starts.

As shown in Fig. 2, the modified second-order derivatives are much stronger, with peak magnitudes increased by at least 15dB, and thus much less sensitive to noise. The location of the local maxima in the derivatives could be used in the determination of the DMC base delay, which shows a good match to the rising slopes in the synthetic channel.

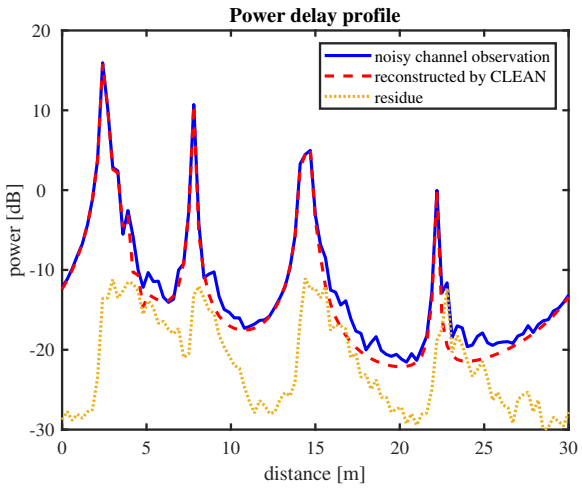

FIGURE 1: The APDP comparison between synthetic channel observation (blue solid line) and the estimation from CLEAN algorithm with 25 MPCs: estimated MPCs (red dashed line) and residue (yellow dotted line).

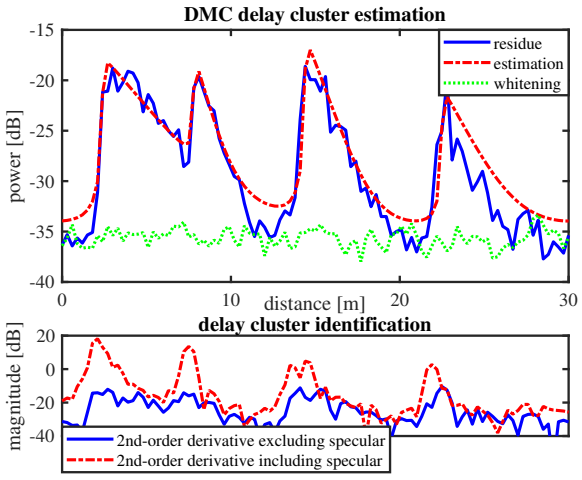

FIGURE 2: The initialization and estimation of DMC delay clusters. Upper subplot: comparison of APDP of residual channel observations (blue solid line) and its estimate (red dot-dash line) is shown. Lower subplot: second-order derivatives with (red dot-dash) /without (blue solid) MPC contribution.

Estimation of base delay, peak power, and decay constant is done as described in Sec III.B.

In the upper subplot in Fig. 2, four DMC clusters in the delay domain are estimated, and their parameters are refined by the gradient-descent method; results are plotted with the red dashed line to compare with the residue from extracted channel observation to be fitted as DMC clusters in Fig. 2. Note that two clusters in the delay domain are merged—while their associated MPCs are separated more than a Fourier width, the DMC clusters overlap, and thus the cluster identification in the delay domain considers them as a single cluster. Furthermore, the total covariance matrix is reconstructed with the estimated parameters. The whitening

of the channel residue, using the reconstructed covariance matrix, shows a flat process (green dotted line in Fig. 2). The flatness of the whitening result illustrates the accuracy of the reconstruction.

B. DMC ESTIMATION IN ANGULAR DOMAIN

In the delay-gated fourth DMC cluster starting at 14.8 meters, the angular spectrum, normalized to its maximum, has two peaks at -20 and 180 degrees, which corresponds to its ground truth and the interference from the third cluster. To accurately identify clusters, it is crucial to incorporate interference cancellation of the other clusters in the cluster identification. Otherwise, ghost peaks may occur, where a late contribution from a cluster that was already identified in a previous delay gate is counted as another cluster. In Fig. 3, only one local maximum, closer to -20 degrees, *should* be detected from the angular spectrum. However, there is another peak close to 180 degrees, which can be attributed to the tail from the third cluster falling into the current delay gate, manifesting as interference. The contribution from this cluster can be removed properly based on the estimated delay and angular parameters, shown as black solid line. The estimation, denoted by the red dot-dash line in Fig. 3, shows a good match to the delay-gated angular spectrum, indicating faithful covariance matrix reconstruction.

In the delay-gated third DMC cluster starting at 14 meters, two clusters are well-separated with DoA at -20 and 180 degrees. With our proposed angular cluster isolation method in the cluster identification and initialization, it is possible to separate the clusters that are merged in the PDP and re-evaluate parameters for each cluster. As shown in Fig. 4, the third and the fourth cluster are separated and show different start delays, as opposed to Fig. 2.

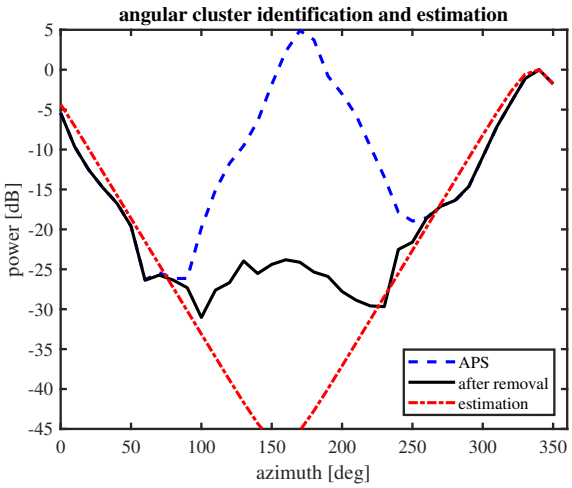

FIGURE 3: APS of the fourth DMC cluster. The angular spectrum after the removal (black solid line) is compared with the estimation (red dot-dash line).

In Fig. 5, the total DoA APS of the residual channel observation is represented by the blue solid line. Using the estimated angular and corresponding delay parameters,

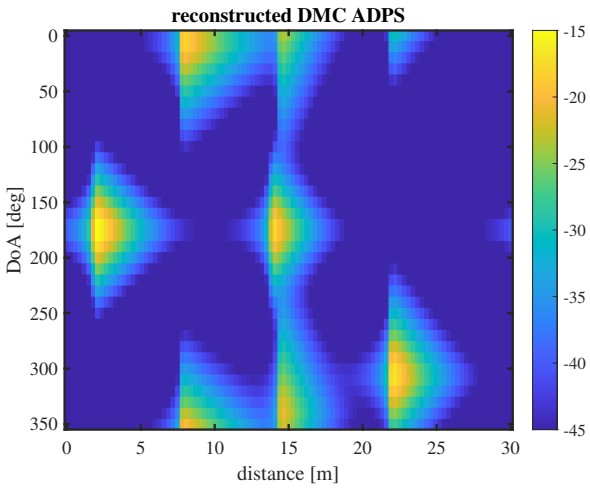

FIGURE 4: Reconstructed DMC ADPS with estimated parameters.

converted to power weights based on equation (43), the total Rx covariance matrix can be reconstructed. The red dot-dash line shows the power-weighted APS reconstruction using all estimated angular parameters. The angular estimation closely matches the angular power spectrum, which can serve as a validation of the angular parameter estimation. The flatness of the whitening, indicated by no more than a 2 dB variation in the green dotted line, is used as a metric for the quality of the estimation.

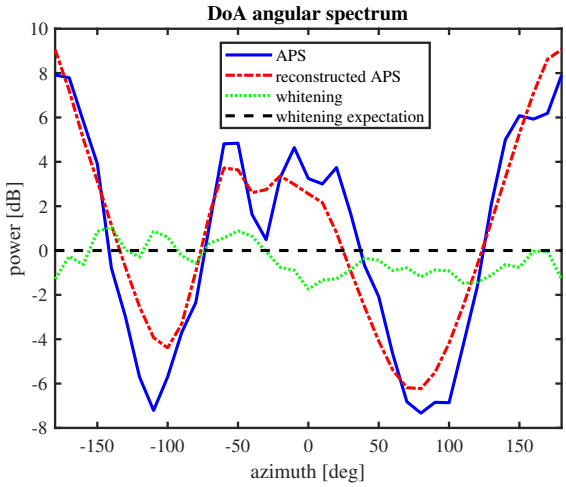

FIGURE 5: The comparison of the angular spectrum between the residual observation and the reconstruction from the estimated DMC angular parameters. The "whitening expectation" refers to the ideal performance of whitening, *i.e.*, flat angular spectrum

C. MIMO GENERALIZATION

In order to test the correctness and effectiveness of the proposed generalization to the MIMO case, a joint angular spectrum is generated. We use a different synthetic channel

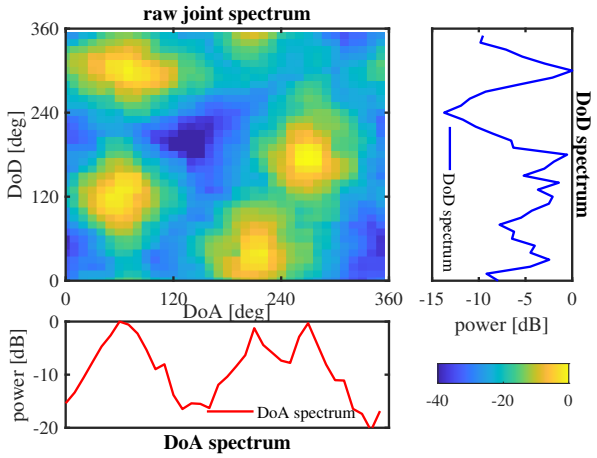

FIGURE 6: The joint angular spectrum and its DoD and DoA spectrum. The spectrum and the joint spectrum are normalized to their maximum.

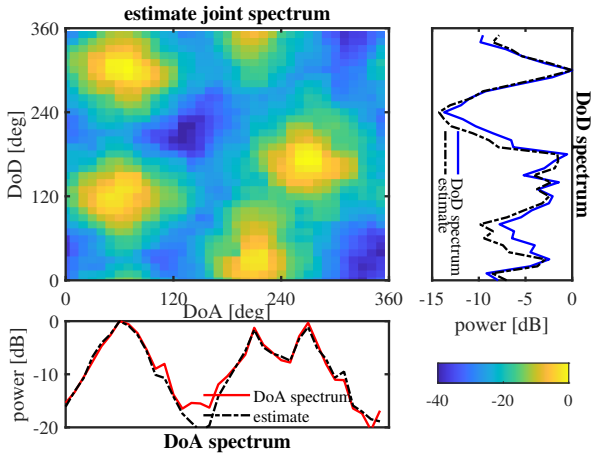

FIGURE 7: The reconstruction of the estimated joint angular spectrum and its DoD and DoA spectrum. The spectrum and the joint spectrum are normalized to their maximum. The reconstructed spectrum estimated from gradient-descent method (black dot-dash line) in each marginal spectrum for comparison.

compared to that used in the previous examples, with angular parameters listed in Table 2; all clusters are assumed to be observed within the same delay gate. As shown in Fig. 6, there are four angular clusters in the joint spectrum, with two of them having similar DoAs around 60 deg and having equal weights. Thus, there are three peaks in the DoA spectrum, while four peaks can be observed in the DoD spectrum. Furthermore, we confirm that our estimation method eliminates the ghost components of the Kronecker assumption. The estimation of DoD and DoA spectrum can be seen in Fig. 7. Instead of the 12 combinations between DoAs and DoDs that one gets under the Kronecker model assumption, there are only four estimated clusters. The estimated angular parameters for all four clusters are shown

TABLE 2: MIMO synthetic channel parameters

	1	**2**	**3**	**4**
μ_{A}[deg]	60	65	210	270
κ_{A}	3.73	5.19	4.56	2.46
μ_{D}[deg]	120	300	40	180
κ_{D}	2.70	3.89	2.35	4.21
w	0.25	0.25	0.25	0.25

TABLE 3: Estimated MIMO parameters

	1	**2**	**3**	**4**
$\hat{\mu}_{\mathrm{A}}$[deg]	60	60	210	270
$\hat{\kappa}_{\mathrm{A}}$	4.68	4.68	4.04	5.31
$\hat{\mu}_{\mathrm{D}}$[deg]	122	300	40	182
$\hat{\kappa}_{\mathrm{D}}$	3.22	4.92	2.86	4.95
$\hat{w}$	0.256	0.245	0.246	0.253

in Table 3. The final estimated weights are almost identical among the four clusters, which shows a good match to the ground truth.

D. INCORPORATION IN RIMAX FRAMEWORK

In order to test the impact of the improved DMC estimation on the specular MPCs, we integrate the algorithm into our existing RiMax program and use it to evaluate the synthetic channel. Results are shown in Fig. 8. The red square markers stand for the parameters of the specular MPCs in the delay-power description, where there are five MPCs in the synthetic channel. The green circle and black cross markers denote the estimation from the initialization process and RiMax, respectively. Since the parameters are randomly generated and not located on the search grids for the initialization (which is done by CLEAN), the imperfect estimation leads to residue after the "actual" MPCs have been successfully removed from the channel observation. This results in ghost paths, which are the reason for the over-estimation of the number of MPCs, as well as errors in the delay and angular parameters and the complex amplitude.

RiMax takes DMC into account in both the signal model and the estimation. With a proper DMC covariance matrix reconstruction, it is possible to eliminate (or at least greatly mitigate) the ghost paths [4, (4.100)], according to the ratio between MPC magnitude error bound and its power magnitude, $\mathrm{var}(|\gamma|)/|\gamma|^2 < 1$ as the criterion; otherwise, the "over-estimation" remains in the solution. With the proposed covariance matrix estimation, the redundant 15 MPCs are eliminated in the iterations, as shown in Fig. 8, showing that the DMC estimation is successful and the accuracy is satisfactory. As a comparison in Fig. 9, an implementation of DMC estimation [20] with the state-of-the-art DMC model (one single exponential decay in delay and multi-modal von Mises distribution) is used to evaluate the same synthetic channel. However, due to the model mismatch, the ghost paths are not eliminated, and a larger error occurs in the MPC magnitude estimation as well.

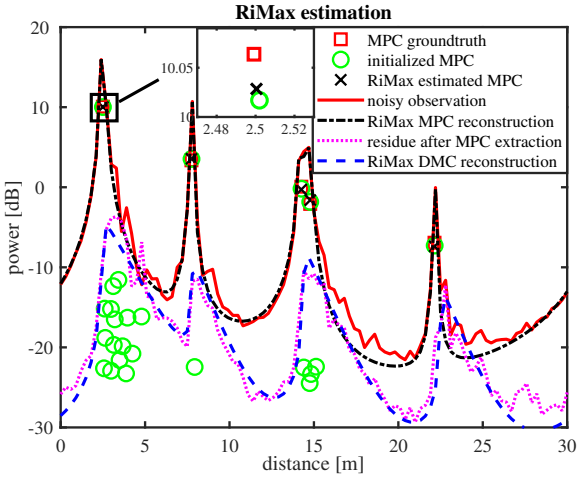

FIGURE 8: APDP and MPC parameters: noisy channel observation with diffuse scattering (red solid line), reconstruction from estimated specular MPCs (black dot-dash line); residue (blue dashed line), and DMC estimation (magenta dotted line). Specular components are denoted by discrete markers for ground truth (red squares), initialization (green circles) and RiMax estimation (black crosses).

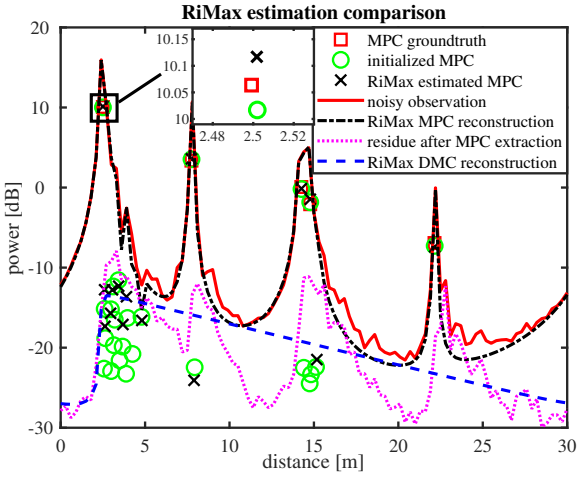

FIGURE 9: RiMax estimation with uni-modal DMC model [20]. The lines and markers are defined as in Fig. 8.

E. STATISTICAL EVALUATION—DMC

To measure the algorithm's performance in terms of DMC parameter estimation, Monte Carlo simulations are conducted with *pure* DMC synthetic SIMO channels, whose parameters are generated based on a Saleh-Valenzuela model [37]. The DMC cluster arrival rate is modeled as a Poisson arrival process with a fixed rate Λ for a particular synthetic channel. The ray arrival rate is assumed to be so high that each resolvable delay bin is Rayleigh fading; the power of the bins as a function of the delay from the cluster start is exponential with the DMC cluster decay β_{d}. The mean of the DMC cluster magnitude shows an exponential decay with a constant Γ. An independent, uniformly distributed (on a dB scale) random attenuation is multiplied with the mean magnitude of each cluster. The noise level is fixed at -25 dB. The detailed model parameters can be found in Table 4.

TABLE 4: Parameter setup based on Saleh-Valenzuela model.

parameter	distribution	range
$1/\Lambda$	uniform	[20,40]
Γ	uniform	[0.04,0.07]
β_{d}	uniform	[0.2,0.5]
σ_n^2 [dB]	-	-25
μ [rad]	uniform	[0,2π]
κ	uniform	[2,4]

To evaluate the goodness-of-fit between the DMCs, shown in magenta dotted line, and its estimation, shown in blue dashed line in Fig. 8, we consider to use the correlation coefficient as the metric. To measure the fitness of DMC estimation in both delay and angular domain, the ADPS is evaluated. The correlation coefficient between two ADPSs is expressed as

$$\text{corr coef} \triangleq \frac{|\mathrm{vec}(\mathrm{ADPS}_u)^T \mathrm{vec}(\mathrm{ADPS}_v)|}{\|\mathrm{vec}(\mathrm{ADPS}_u)\| \|\mathrm{vec}(\mathrm{ADPS}_v)\|}, \tag{46}$$

where $\mathrm{ADPS}_{u/v} \in \mathbb{R}^{N_{\mathrm{f}} \times N_\phi}$ with N_ϕ different directions is vectorized, with the defined operator $\mathrm{vec}(\cdot)$, to convert into a longer vector. A larger correlation value means the estimation is closer to the goal.

We compute the correlation coefficient between the estimation of the DMC and the experimentally observed DMC contribution, to measure the algorithm estimation accuracy. There are two types of error sources in the DMC estimation. The first type of error is that the expectation of the DMC ADPS reconstructed from the channel estimation deviates from the "ground truth" expectation due to the noise and errors in the estimation procedure. Secondly, errors come from the fact that the experimental observations provide us with individual *realizations* of the DMC. The power of such a realization in each delay or angle bin might deviate considerably from the *expectation* of the DMC power in this bin. If we only have a single observation of the fading state, this deviation will be especially pronounced and impact the DMC parameter estimation. Consequently, there are two types of ADPSs regarded as the baseline—one is the ADPS generated directly from the generating or estimated covariance matrices of DMC clusters describing the expectation of DMC, corresponding to "ideal DMC" and "est DMC", respectively; the other is the particular realization of the DMC, namely "DMC rea", also known as the instantaneous channel state of the DMC contribution. The ADPSs generated directly from DMC covariance matrices can be approximated by averaging over adequate number of snapshots. Thus, the correlation between estimation based on 10 snapshots and the ideal DMC case in 200 different synthetic channels shows a consistent level of goodness of fit as shown in Fig. 10. Smaller correlation is observed

in estimation based on 1 DMC realization, indicating a larger difference between the actual (fading) DMC and the estimation of the ADPS, which tries to reconstruct the average ADPS. With multiple snapshots, the gap between blue line and red line becomes smaller, which verifies our statement that an adequate number of ADPSs can be used to approximate the generating covariance matrix.

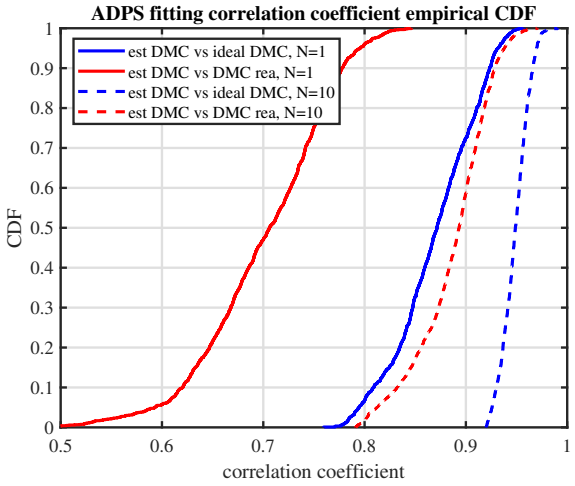

FIGURE 10: DMC ADPS goodness-of-fit. The correlation coefficient empirical CDF between ADPS of multiple DMC clusters from either 1 realization (solid lines) or average over 10 snapshots (dashed lines) with its generating covariance matrix ADPS (blue lines) and DMC realization ADPS (red lines) is shown.

As the correlation coefficient may lose some detail information about the estimation error, we then measure the spectrum reconstruction error, comparing between uni-modal and multi-modal DMC model. The maximal deviation in ADPS, PDP, and APS on a *logarithmic* scale are defined, respectively, as

$$\begin{aligned}\Delta\mathrm{ADPS}(\tau, n_{\mathrm{R}}) &\triangleq |\widehat{\mathrm{ADPS}}^{(\mathrm{dB})}(\tau, n_{\mathrm{R}}) - \mathrm{ADPS}^{(\mathrm{dB})}(\tau, n_{\mathrm{R}})| \\ d_{\mathrm{ADPS}} &= \max_{\tau, n_{\mathrm{R}}} \Delta\mathrm{ADPS}(\tau, n_{\mathrm{R}}) \\ d_{\mathrm{PDP}} &= \max_{\tau} \frac{1}{N_{\mathrm{R}}} \sum_{n_{\mathrm{R}}} \Delta\mathrm{ADPS}(\tau, n_{\mathrm{R}}) \\ d_{\mathrm{APS}} &= \max_{n_{\mathrm{R}}} \frac{1}{N_{\mathrm{f}}} \sum_{\tau} \Delta\mathrm{ADPS}(\tau, n_{\mathrm{R}}),\end{aligned}$$

where $\widehat{\mathrm{ADPS}}^{(\mathrm{dB})}$ and $\mathrm{ADPS}^{(\mathrm{dB})}$ denotes the reconstructed and "ground truth" ADPS in dB with delay and antenna index of τ and n_{R}, respectively. The maximal deviation in ADPS, d_{ADPS}, captures the largest discrepancy in the delay-angle plane, while d_{PDP} and d_{APS} represent the difference in the separate parametric domains. These maximal deviations serve as alternative indicators for measuring the goodness-of-fit. The cumulative distribution function of the maximal deviation across 200 synthetic *pure* DMC channels, each with 10 different snapshots to mitigate the "fading" effect, in ADPS (blue lines), PDP (red lines), and APS (magenta lines) is shown in Fig. 11. Compared to the uni-modal (dashed lines), the deviations in the reconstruction in the multi-modal (solid lines) DMC model are significantly reduced, as the model mismatch is alleviated. The statistical metrics—the mean and the standard deviation—of the maximum deviation in the spectra are summarized in Table 5.

However, the absolute values of the reconstruction error, *i.e.*, up to 15dB in ADPS, in our proposed multi-modal DMC model appear very high and thus seem to be in contradiction to the good agreement shown by the correlation coefficient in Fig. 10. A closer comparison of the ADPS reveals that the larger deviations typically occur in weak delay-angle bins, and the use of a logarithmic scale may blur the actual power level information.[5] To better understand the deviations' distribution, we computed the mean and standard deviation of the reconstruction error at varying relative power levels, shown in Fig.12. The relative power level is the delay-angle bin power, $\mathrm{ADPS}_{\mathrm{rel}}(\tau, n_{\mathrm{R}})$, normalized by the maximum power in the ADPS. The relative power bin with step size of 5dB collects the bins whose relative power level falls into the range centered at the power level, *i.e.*, -10dB relative power bin collects the bins within $[-12.5, -7.5)$dB. In Fig. 12, the deviation in uni-modal case, whose mean and standard deviation is denoted by blue dashed lines and their error bars, respectively, shows almost uniform mean value among different power levels, indicating that the deviation may occur in each delay-angle ADPS bin due to the DMC model mismatch. In contrast, the power difference reconstructed from multi-modal DMC model in red solid line, shows mean and standard deviation of the reconstruction error decreasing with increasing relative power, *i.e.*, stronger delay-angle bins tend to have less reconstruction error. Since the stronger delay-angle bins make the larger contributions to the correlation coefficient, this also explains the good match of that coefficient.

TABLE 5: Maximum deviation statistics in spectrum.

[dB]	uni-modal	multi-modal
$\bar{d}_{\mathrm{ADPS}}$	24.92	10.44
std(d_{ADPS})	4.06	2.46
$\bar{d}_{\mathrm{PDP}}$	11.49	5.01
std(d_{PDP})	3.18	1.60
$\bar{d}_{\mathrm{APS}}$	9.83	4.10
std(d_{APS})	2.04	1.09

F. STATISTICAL EVALUATION—MPC

To further verify the algorithm's correctness and performance, Monte Carlo simulations are conducted on 50 different randomly generated synthetic channels with 20 DMC and noise realizations each. The Cramér-Rao lower bound

[5]However, use of a linear scale would over-emphasize the reconstruction of the strongest bins only, which is why we use the dB scale here.

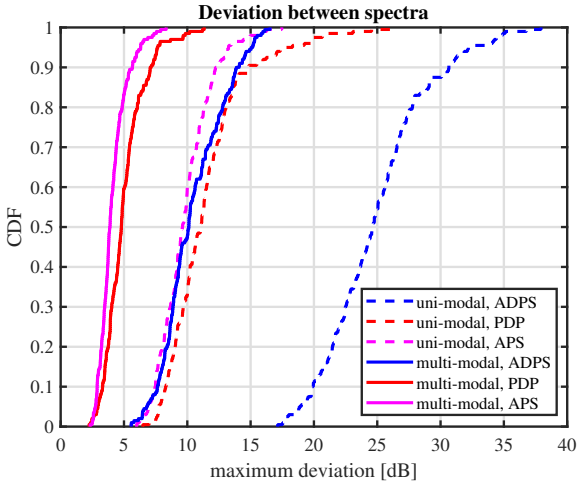

FIGURE 11: Maximal deviation in ADPS, PDP and APS compared between fitting with uni-modal and multi-modal DMC model.

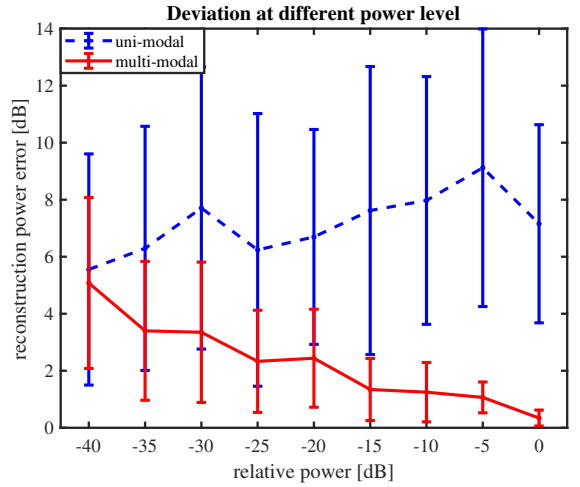

FIGURE 12: Power deviation in ADPS at different relative power levels.

(CRLB) of the parameter estimation root mean square error (RMSE) [4] for the average over *all* MPCs are computed and regarded as the benchmark. The heights of DMC clusters are controlled by scaling them with the same constant such that the DMC power percentage from the channel varies while the other DMC parameters and the noise level stay unchanged. The DMC power percentage is defined as

$$\text{DMC}\% \triangleq \frac{P_{\text{DMC}}}{P_{\text{H}}} = \frac{\|\mathbf{N}_{\text{DMC}}\|_{\text{F}}^2}{\|\mathbf{H}_{\text{mea}}\|_{\text{F}}^2} \times 100\%, \qquad (47)$$

where P_{DMC} is the total power from the diffuse components, and P_{H} stands for the total channel power. Note that for the purpose of estimating the specular MPCs, the DMC acts like colored noise [4, (4.16)], even though it does carry signal information. The derivation starts with the log-likelihood function as shown in equation (18).

We found in our examples that the estimation error is about 2–3dB above the CRLB, see Fig. 13, for DMC power percentage below 20%. Furthermore, the results show that our method has lower parameter estimation error compared to the state-of-the-art DMC model [20] ("uni-modal" is named after the single exponential decay in the delay domain). This is consistent with the expectation from theory that a reduced model mismatch leads to reduction of the MPC parameter estimation error.

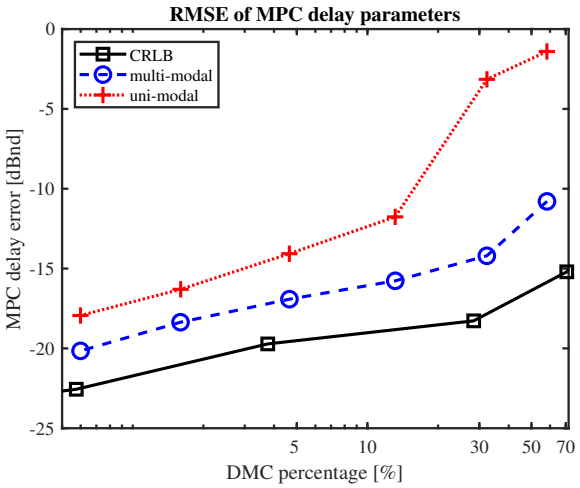

FIGURE 13: Average RMSE of MPC delay parameters (*not* the base delay error for estimated DMC) from all synthetic channels in unit of normalized delay (in short 'nd'), defined as $\tau_{\text{nd}} = \tau \cdot \text{BW}$. CRLB (solid line), estimation error with multi-modal (dashed line) and uni-modal (dotted line) for all MPCs.

G. EVALUATION OF OUTDOOR MEASUREMENT

To further validate the proposed DMC estimation and its compatibility with the RiMax algorithm scheme, we performed one example evaluation of a MIMO channel measurement in the sub-THz regime. The measurement is performed in an outdoor device-to-device setup with line-of-sight (LOS) connection between transmitter and receiver, which are separated by 2.4 meters. The sounding system measures the transfer function over a 1GHz bandwidth centered on 145.5 GHz by scanning over 1001 frequency points with a Vector Network Analyzer. A horn antenna is rotated in 10-degree steps, similar to the setup previously discussed for the simulations. The detailed measurement scenarios and the channel sounder descriptions can be found in [38].

Fig. 14 shows, as the blue dotted line, the raw measured PDP. After estimation and subtraction of the contribution of the specular MPCs (green cross markers), the residue (red solid line) indicates the presence of multiple clusters. Using the proposed DMC estimation algorithm, seven DMC clusters are detected and analyzed. It is worthwhile to notice that five out of seven clusters, located within 30 meters, are associated with MPCs. The remaining two clusters, one starting at 35 meters and the one at 70 meters, are fitted for the unexplained components, modeled as two stand-alone DMC clusters.

The identified DMC clusters, for example, the first stand-alone DMC cluster between 35 meters and 70 meters, may exhibit a shape that deviates significantly from the expected exponential model. Since we have only a single observation of the fading behavior of these clusters, it is difficult to determine whether this deviation is due to fading or to a model mismatch, where the actual physical clusters have an average PDP that differs from the assumed exponential form. With the proposed multi-cluster DMC estimation method, the RiMax algorithm successfully eliminates most of the ghost paths; this is indicated by the fact that while 250 MPCs are detected in the initialization procedure (which assumes an absence of DMC), the final result identifies only 22 specular MPCs in addition to the 7 DMC clusters.

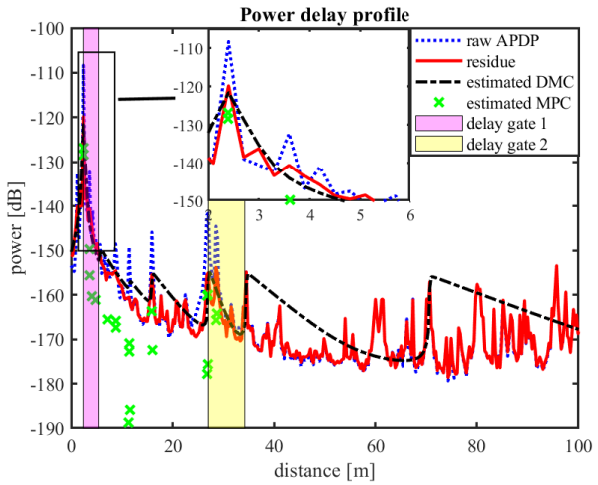

FIGURE 14: APDP and MPC parameters evaluated from LOS measurement: noisy channel observation (blue dotted line), residue after MPC extraction (red solid line), and DMC estimation (black dot-dash line). Estimated MPCs are denoted by green discrete markers. The two shaded area stand for two delay gatings.

In Fig. 15, the DoA-DoD joint angular spectrum of the residue, after the extraction of MPCs from the noisy channel observations, is presented. The LOS cluster is positioned at 0 degree for both DoA and DoD domain, for simplicity written as the tuple (0,0). In the joint spectrum, five strong clusters are observable, each corresponding to a distinct MPC cluster. The residuals from the channel observations can be used to validate the non-applicability of the Kronecker model. The magenta and yellow color-shaded regions in Fig. 14 represent two different delay gates, whose respective DoA-DoD joint angular spectra are shown in Fig. 16 and 17. In the first delay gate, it contains the LOS cluster, and part of another cluster centered at (340,120). The second delay gate includes two clusters with nearly identical DoAs, centered at (90, 90) and (90, 270), respectively. Residuals from the other three clusters—LOS cluster, cluster centered at (0,180) and (340,120)—are also observable, as their weak tails, decaying gradually and thus still partly falling into the delay-gate region, contribute to their "visibility".

The reconstruction of the DMC ADPS is demonstrated in Fig. 18. The five dominant DMC clusters are evaluated and reconstructed, though several small and weak clusters are not fully detected. Notably, the ADPS reconstruction avoids the creation of ghost clusters. The correlation coefficient between the reconstruction and the original ADPS is 0.8079, which is not in contradiction with the blue solid curve in Fig. 10, (though it must be remembered that Fig. 10 is based on a different, idealized channel), as the ADPS of the experimental observations provides us with individual *realizations* of the DMC, namely "DMC rea" by definition.

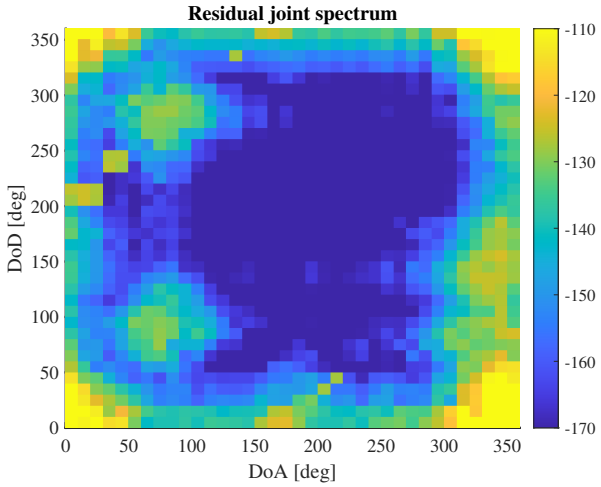

FIGURE 15: The DoA-DoD joint angular spectrum for residue after the MPC extraction.

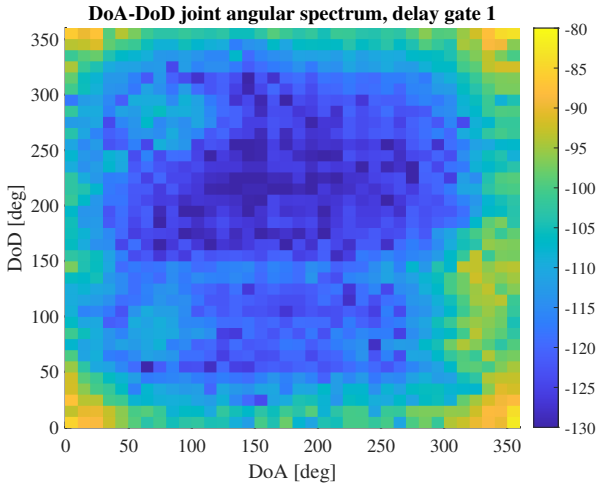

FIGURE 16: The DoA-DoD joint angular spectrum for residue after the MPC extraction in delay gate 1.

For comparison, Fig. 19 shows the results of implementing the uni-modal DMC model from [20], as mentioned for Fig. 9, on the same real-world outdoor measurement campaign. Due to mismatch in the DMC model, ghost paths are not effectively eliminated, leaving over 200 MPCs after the convergence of the RiMax algorithm. Consequently, the residue after MPC extraction is higher than the residual APDP shown

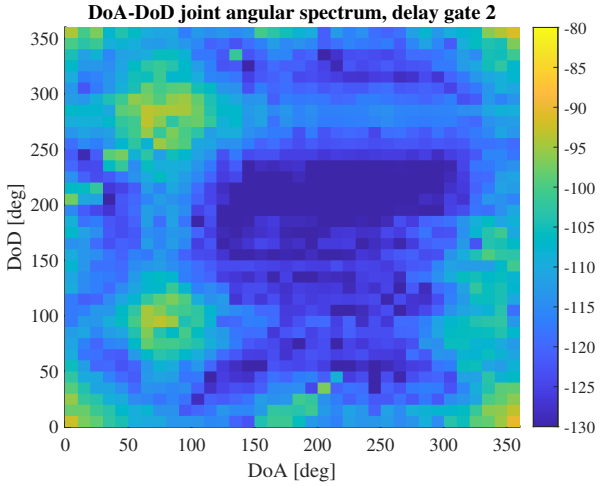

FIGURE 17: The DoA-DoD joint angular spectrum for residue after the MPC extraction in delay gate 2.

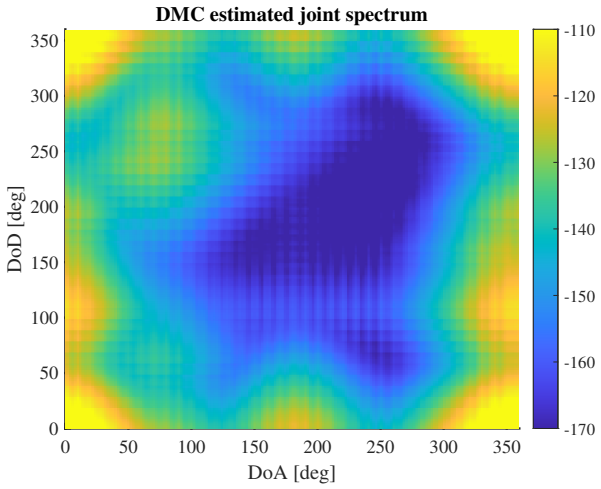

FIGURE 18: The DoA-DoD joint angular spectrum reconstructed from DMC estimation.

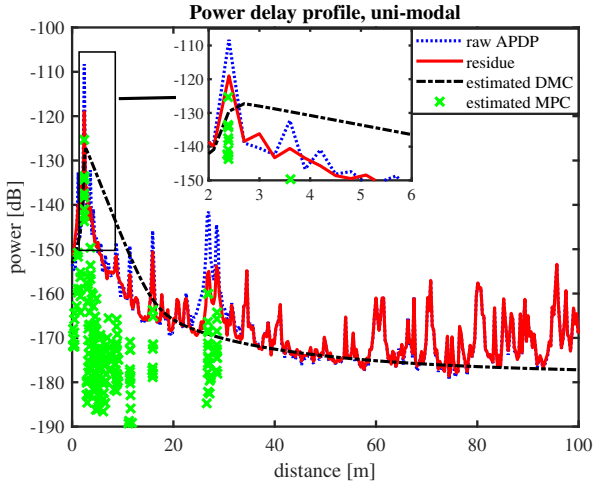

FIGURE 19: LOS measurement RiMax estimation with uni-modal DMC model [20]. The lines and markers are defined as in Fig. 14.

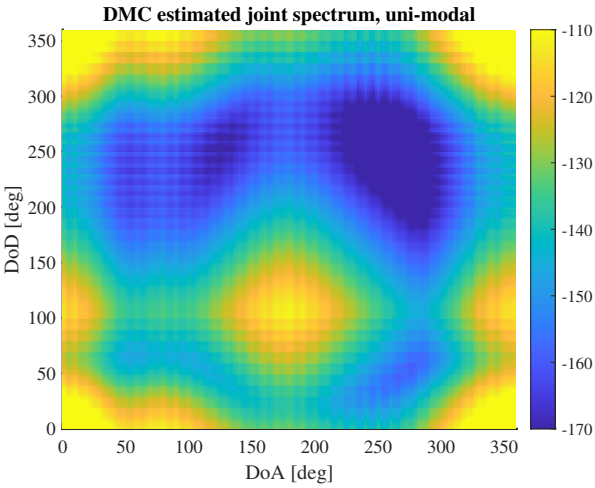

FIGURE 20: The DoA-DoD joint angular spectrum reconstructed from uni-modal DMC estimation.

in Fig. 14. Additionally, Fig. 20 depicts the joint DoA-DoD spectrum reconstruction under the Kronecker model assumption. Due to model mismatch and inaccuracies in the delay domain, the reconstructed spectrum significantly deviates from the objective, *i.e.*, the residual ADPS shown in Fig. 15. This leads to the creation of ghost clusters and a pronounced deviation in the ADPS, resulting in a correlation coefficient of less than 0.3.

V. CONCLUSIONS

The phenomenon of multiple diffuse scattering clusters' existence in the wireless channel has been observed from measurements, but has up to now not been included in parameter estimation algorithms. Rather, the Kronecker model is widely adopted, which however creates ghost clusters that do not correspond to physical reality. This paper considers high-resolution parameter estimation with a multi-cluster frequency-angle non-Kronecker DMC model that better describes the diffuse scattering clusters. Key points of the algorithm include (i) initialization for DMC parameters from the total channel response due to the association between MPC and DMC, (ii) sequential estimation of the different DMC clusters, where the gradient-descent method is deployed to improve the parameter estimation, and (iii) use of dimension-by-dimension computation of the DMC parameters, using spectral whitening, to improve computational efficiency. Our proposed algorithm improves not only the estimation of the DMC itself, but also the specular MPCs. The RMSE of the estimated parameters are within a few dB of the CRLB in the considered examples as long as the percentage of the DMC energy is within 20% of the total signal energy.

The accurate extraction of the DMC forms the basis of better MIMO channel modeling. Since experiments have shown DMC energy to be a significant percentage of the total signal energy, its proper incorporation is important. This is all the more the case as DMC can significantly impact system performance: generally, DMC impacts the skew of the eigenvalue distribution of the channel Grammian, and

thus the MIMO capacity. The amount of this change depends significantly on the modeling of the angular structure. On the other hand, the DMC can impact the interference to other users, and thus the multi-user capacity in particular when analog beamforming is used. The Kronecker assumption in the angular domain leads to an over-estimation of both of these effects. Furthermore, the delay-angle Kronecker assumption might distort the assessment of analog beamforming (which can only create beams that are independent of delay/frequency) compared to digital beamforming.

Investigation of these system implications will be part of our future work. Furthermore, we will apply the proposed multi-cluster DMC model to the evaluation of real-world channel measurements with the RiMax algorithm.

REFERENCES

[1] H. Tataria, M. Shafi, A. F. Molisch, M. Dohler, H. Sjöland, and F. Tufvesson, "6G wireless systems: vision, requirements, challenges, insights, and opportunities," *Proc. IEEE*, vol. 109, no. 7, pp. 1166–1199, 2021.

[2] V. Kristem *et al.*, "3D MIMO outdoor-to-indoor propagation channel measurement," *IEEE Trans. Wireless Commun.*, vol. 16, no. 7, pp. 4600–4613, 2017.

[3] B. H. Fleury, M. Tschudin, R. Heddergott, D. Dahlhaus, and K. I. Pedersen, "Channel parameter estimation in mobile radio environments using the SAGE algorithm," *IEEE J. Sel. Areas Commun.*, vol. 17, no. 3, pp. 434–450, 1999.

[4] A. Richter, *Estimation of radio channel parameters: Models and algorithms*. Ph.D dissertation, Electr. Eng. Inform. Techn., Ilmenau University of Technology, 2005.

[5] D. P. Gaillot *et al.*, "Accuracy of specular path estimates with ESPRIT and RiMAX in the presence of measurement-based diffuse multipath components," in *2011 5th European Conf. Antennas Propag. (EuCAP 2011)*, pp. 3619–3622, IEEE, 2011.

[6] M. Landmann, M. Kaske, and R. S. Thoma, "Impact of incomplete and inaccurate data models on high resolution parameter estimation in multidimensional channel sounding," *IEEE Trans. Antennas Propag.*, vol. 60, no. 2, pp. 557–573, 2011.

[7] S. Jiang, W. Wang, Y. Miao, W. Fan, and A. F. Molisch, "A survey of dense multipath and its impact on wireless systems," *IEEE Open J. Antennas Propag.*, vol. 3, pp. 435–460, 2022.

[8] A. Richter, J. Salmi, and V. Koivunen, "Distributed scattering in radio channels and its contribution to MIMO channel capacity," in *2006 First European Conf. Antennas Propag. (EuCAP 2006)*, pp. 1–7, IEEE, 2006.

[9] S. Sangodoyin *et al.*, "Cluster characterization of 3-D MIMO propagation channel in an urban macrocellular environment," *IEEE Trans. Wireless Commun.*, vol. 17, no. 8, pp. 5076–5091, 2018.

[10] F. Challita *et al.*, "On the contribution of dense multipath components in an intrawagon environment for 5G mmW massive MIMO channels," *IEEE Antennas Wireless Propag. Lett.*, vol. 18, no. 12, pp. 2483–2487, 2019.

[11] E. Tanghe, D. P. Gaillot, M. Liénard, L. Martens, and W. Joseph, "Experimental analysis of dense multipath components in an industrial environment," *IEEE Trans. Antennas Propag.*, vol. 62, no. 7, pp. 3797–3805, 2014.

[12] B. Hanssens *et al.*, "Measurement-based analysis of dense multipath components in a large industrial warehouse," pp. 1–5, 2018.

[13] A. Richter, J. Salmi, and V. Koivunen, "ML estimation of covariance matrix for tensor valued signals in noise," in *2008 IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP 2008)*, pp. 2349–2352, IEEE, 2008.

[14] F. Mani, F. Quitin, and C. Oestges, "Directional spreads of dense multipath components in indoor environments: Experimental validation of a ray-tracing approach," *IEEE Trans. Antennas Propag.*, vol. 60, no. 7, pp. 3389–3396, 2012.

[15] J. Li, B. Ai, R. He, M. Yang, and Z. Zhong, "On modeling of dense multipath component for indoor massive MIMO channels," *IEEE Antennas Wireless Propag. Lett.*, vol. 18, no. 3, pp. 526–530, 2019.

[16] E. Tanghe, D. P. Gaillot, W. Joseph, M. Liénard, P. Degauque, and L. Martens, "Robustness of high-resolution channel parameter estimators in presence of dense multipath components," *Electron. Lett.*, vol. 48, no. 2, pp. 130–132, 2012.

[17] M. Kaske, M. Landmann, and R. Thoma, "Modelling and synthesis of dense multipath propagation components in the angular domain," in *2009 3rd European Conf. Antennas Propag. (EuCAP 2009)*, pp. 2641–2645, IEEE, 2009.

[18] M. Käske and R. Thomä, "Maximum-likelihood based estimation of angular parameters of dense-multipath-components," in *2015 9th European Conf. Antennas Propag. (EuCAP 2015)*, pp. 1–6, IEEE, 2015.

[19] C. B. Ribeiro, E. Ollila, and V. Koivunen, "Propagation parameter estimation in MIMO systems using mixture of angular distributions model," in *2005 IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP 2005) Proc.*, vol. 4, pp. iv–885, IEEE, 2005.

[20] C. B. Ribeiro, A. Richter, and V. Koivunen, "Joint angular- and delay-domain MIMO propagation parameter estimation using approximate ML method," *IEEE Trans. Signal Process.*, vol. 55, no. 10, pp. 4775–4790, 2007.

[21] J. Poutanen, F. Tufvesson, K. Haneda, L. Liu, C. Oestges, and P. Vainikainen, "Adding dense multipath components to geometry-based MIMO channel models," *IEICE Proc. Series*, vol. 52, no. 2WC1-4, 2010.

[22] J. Poutanen, J. Salmi, K. Haneda, V.-M. Kolmonen, F. Tufvesson, and P. Vainikainen, "Propagation characteristics of dense multipath components," *IEEE Antennas Wireless Propag. Lett.*, vol. 9, pp. 791–794, 2010.

[23] J. Poutanen, J. Salmi, K. Haneda, V.-M. Kolmonen, and P. Vainikainen, "Angular and shadowing characteristics of dense multipath components in indoor radio channels," *IEEE Trans. Antennas Propag.*, vol. 59, no. 1, pp. 245–253, 2010.

[24] B. Hanssens, K. Saito, E. Tanghe, L. Martens, W. Joseph, and J.-I. Takada, "Modeling the power angular profile of dense multipath components using multiple clusters," *IEEE Access*, vol. 6, pp. 56084–56098, 2018.

[25] F. Quitin, C. Oestges, F. Bellens, S. Van Roy, F. Horlin, and P. De Doncker, "Extracting specular-diffuse clusters from MIMO channel measurements," in *2011 IEEE 22nd Int. Symp. Pers. Indoor Mobile Radio Commun. (PIMRC)*, pp. 940–944, IEEE, 2011.

[26] J. Salmi *et al.*, "Incorporating diffuse scattering in geometry-based stochastic MIMO channel models," in *2010 4th European Conf. Antennas Propag. (EuCAP 2010)*, pp. 1–5, IEEE, 2010.

[27] S. Schieler, M. Döbereiner, S. Semper, and M. Landmann, "Estimating multi-modal dense multipath components using auto-encoders," in *2022 30th European Signal Process. Conf. (EUSIPCO)*, pp. 1716–1720, IEEE, 2022.

[28] R. Wang, O. Renaudin, C. U. Bas, S. Sangodoyin, and A. F. Molisch, "On channel sounding with switched arrays in fast time-varying channels," *IEEE Trans. Wireless Commun.*, vol. 18, no. 8, pp. 3843–3855, 2019.

[29] M. Steinbauer, A. F. Molisch, and E. Bonek, "The double-directional radio channel," *IEEE Antennas Propag. Mag.*, vol. 43, pp. 51–63, 2001.

[30] F. Quitin, C. Oestges, F. Horlin, and P. De Doncker, "Diffuse multipath component characterization for indoor MIMO channels," in *2010 4th European Conf. Antennas Propag. (EuCAP 2010)*, pp. 1–5, IEEE, 2010.

[31] K. Saito, J.-I. Takada, and M. Kim, "Characteristics evaluation of dense multipath component in 11GHz-band indoor environment," in *2016 10th European Conf. Antennas Propag. (EuCAP 2016)*, pp. 1–3, IEEE, 2016.

[32] K. Saito, J.-i. Takada, and M. Kim, "Dense multipath component parameter estimation in 11GHz-band indoor environment," in *2016 IEEE 27th Int. Symp. Pers. Indoor Mobile Radio Commun. (PIMRC)*, pp. 1–6, IEEE, 2016.

[33] M. Kim, T. Iwata, K. Umeki, K. Wangchuk, J.-i. Takada, and S. Sasaki, "Mm-wave outdoor-to-indoor channel measurement in an open square smallcell scenario," in *2016 Int. Symp. Antennas Propag. (ISAP)*, pp. 614–615, IEEE, 2016.

[34] A. F. Molisch, H. Asplund, R. Heddergott, M. Steinbauer, and T. Zwick, "The COST259 directional channel model-part I: overview and methodology," *IEEE Trans. Wireless Commun.*, vol. 5, no. 12, pp. 3421–3433, 2006.

[35] R. He *et al.*, "On the clustering of radio channel impulse responses using sparsity-based methods," *IEEE Trans. Antennas Propag.*, vol. 64, no. 6, pp. 2465–2474, 2016.
[36] M. Käske and R. S. Thomä, "Analysis of angular parameters of dense multipath components in an urban macro-cell scenario," in *2011 5th European Conf. Antennas Propag. (EuCAP 2011)*, pp. 3429–3433, IEEE, 2011.
[37] A. A. Saleh and R. Valenzuela, "A statistical model for indoor multipath propagation," *IEEE J. Sel. Areas Commun.*, vol. 5, no. 2, pp. 128–137, 1987.
[38] N. A. Abbasi *et al.*, "THz band channel measurements and statistical modeling for urban D2D environments," *IEEE Trans. Wireless Commun.*, vol. 22, no. 3, pp. 1466–1479, 2022.