Title	Cross-Domain Joint Channel and Data Estimation for mmWave MIMO-OFDM Systems with Low-Resolution ADCs
Author(s)	Ito, Kenta; Takahashi, Takumi; Ishibashi, Koji et al.
Citation	IEEE Journal on Selected Topics in Signal Processing. 2025
Version Type	VoR
URL	https://hdl.handle.net/11094/103253
rights	This article is licensed under a Creative Commons Attribution 4.0 International License.
Note	

This article has been accepted for publication in IEEE Journal of Selected Topics in Signal Processing. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JSTSP.2025.3589107

Cross-Domain Joint Channel and Data Estimation for mmWave MIMO-OFDM Systems With Low-Resolution ADCs

Kenta Ito, *Member, IEEE*, Takumi Takahashi, *Member, IEEE*,
Koji Ishibashi, *Senior Member, IEEE*, Shinsuke Ibi, *Senior Member, IEEE*, and Hideki Ochiai, *Fellow, IEEE*

***Abstract*—This paper proposes a novel cross-domain joint channel and data estimation (JCDE) algorithm based on expectation propagation (EP) for millimeter-wave (mmWave) multiple-input multiple-output orthogonal frequency-division multiplexing (MIMO-OFDM) systems equipped with low-resolution analog-to-digital converters (ADCs). In the design of an EP-based message-passing algorithm (MPA), messages are formulated by projecting the true posterior distribution of the unknown signal onto a Gaussian distribution through *moment matching (MM)*, which can be modeled as the additive white Gaussian noise (AWGN) channel output of the unknown signal. Consequently, executing domain conversion within this *message space* enables signal processing in a suitable domain unaffected by quantization distortion. Specifically, the proposed algorithm consists of an antenna-delay (AD)-domain de-quantization (DQ) process that reconstructs the original signal prior to ADCs from quantized observations, a beam-delay (BD)-domain channel estimation (CE) that leverages the stochastic sparse structure of wireless channels, and an antenna-frequency (AF)-domain data detection (DD) that aligns with the OFDM format. This cross-domain strategy, tailored to the EP framework, facilitates low-overhead and high-precision JCDE. Computer simulations show that the proposed method not only significantly outperforms the state-of-the-art (SotA) but also approaches the performance of an idealized Genie-aided scheme under certain system conditions.**

***Index Terms*—MmWave MIMO-OFDM, low-resolution ADCs, expectation propagation, frequency-selective fading.**

I. INTRODUCTION

Millimeter-wave (mmWave) multiple-input multiple-output orthogonal frequency-division multiplexing (MIMO-OFDM) technology is a promising solution for delivering high data rates and low latency in wireless communication services by utilizing the extensive spectrum accessible in the 30–300 GHz range [1]. Although compact antenna elements can be densely arranged in mmWave MIMO base stations (BSs), the primary challenges to their practical application are increased power consumption and implementation costs; this constraint inhibits the deployment of a complete radio-frequency (RF) chain for each antenna [2]. In particular, analog-to-digital converters (ADCs) and data interface cards are power-hungry devices. The utilization of high-resolution ADCs for each antenna necessitates the processing of numerous parallel Giga samples per second data streams, which substantially elevates the need for digital signal processing and leads to excessive power consumption [3]. One approach to addressing this issue is to reduce the resolution of the ADCs to a minimal level, *e.g.*, as low as one bit; this results in a trade-off between augmenting the number of RF chains and decreasing the number of quantization bits for ADCs [4]–[6]. The power consumption of interface cards is contingent upon the resolution; hence, decreasing the number of bits in the ADC decreases the power consumption of the baseband circuit [7].

Nevertheless, the low-resolution ADC results in significant information loss, complicating subsequent signal processing. In particular, to acquire high-precision channel state information (CSI), it is necessary to transmit excessively lengthy pilot sequences to mitigate information loss, resulting in inevitable system performance decrease owing to communication overhead. Moreover, the high directivity of mmWave signals renders their propagation exceedingly vulnerable to environmental variations [8], making the practical implementation of extended pilot-based channel estimation (CE) unrealistic [9]. Consequently, it is essential to acquire high-precision CSI using brief pilot sequences, and there are two primary directions to accomplish this objective.

One direction is to design an efficient CE algorithm. By exploiting the statistical properties of the mmWave wireless channel, we may reduce the effective unknown variable space and achieve high-precision CE with shorter pilot sequences [10]–[13]. In mmWave MIMO-OFDM systems, signals are received in the antenna-delay (AD) domain, but the beam-delay (BD) domain is the most efficient for CE. This is predicated on the observation that the mmWave wireless channel consists of a limited number of paths [14]. Projecting signals into the angular domain by full-digital beamforming (FDBF) results in a sparse channel impulse response (CIR) in the BD domain, which can be used to substantially decrease pilot overhead. Nevertheless, utilizing the low-resolution ADC complicates such *domain conversion*; hence, CE algorithms have been investigated to leverage this sparse structure in the presence of nonlinear quantization errors.

The predominant approach is linearization derived from Bussgang's theorem [15]–[17]. The Bussgang minimum mean square error (BMMSE) filtering, developed from the linear regression model following Bussgang decomposition [18], is one of the most common CE methods for 1-bit MIMO receivers. However, for CIR estimation, it necessitates the

This work was supported in part by JSPS KAKENHI under Grant JP23K13335, JP23K22754, and JP25H01111; in part by JST, CRONOS, Japan under Grant JPMJCS24N1; and in part by MIC/FORWARD under Grant JPMI240710001. *(Corresponding author: Takumi Takahashi.)*

K. Ito, T. Takahashi, and H. Ochiai are with Graduate School of Engineering, Osaka University, 2-1 Yamada-oka, Suita, 565-0871, Japan (e-mail: k-ito@wcs.comm.eng.osaka-u.ac.jp, {takahashi, ochiai}@comm.eng.osaka-u.ac.jp). K. Ishibashi is with Advanced Wireless & Communication Research Center (AWCC), The University of Electro-Communications, 1-5-1 Chofugaoka, Chofu, Tokyo 182-8585, Japan (e-mail: koji@ieee.org). S. Ibi is with Faculty of Science and Engineering, Doshisha University, 1-3 Tataramiyakodani, Kyotanabe, 610-0394, Japan (e-mail: sibi@mail.doshisha.ac.jp).

This article has been accepted for publication in IEEE Journal of Selected Topics in Signal Processing. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JSTSP.2025.3589107

prior acquisition of an accurate delay profile. Furthermore, when the number of quantization bits is minimal (≤ 3), the performance deteriorates markedly due to interference from quantization distortion, which hinders the full exploitation of channel sparsity [19]–[21]. A more elaborate alternative is the Bayesian CE algorithm based on message-passing de-quantization (MPDQ) [20]–[26]. This method iteratively alternates between de-quantization (DQ), which reconstructs the original pre-quantized signal from the quantized observations, and the estimation of unknown variables, thereby enabling sparsity-aware beam-domain CE even in the presence of quantization noise [27]–[31]. To implement FDBF on quantized observations and formulate the CE problem for channel sparsity exploitation, MPDQ must be applied to an augmented received signal model whose dimensionality is equal to the product of the number of receive antennas and the number of time slots, thereby considerably increasing the computational and processing complexity [27], [28]. Consequently, low-overhead and high-precision CE algorithms in the BD domain for mmWave MIMO-OFDM systems with low-resolution ADCs are currently under development[1].

The other direction involves the extension to a joint channel and data estimation (JCDE) algorithm based on Bayesian bilinear inference (BBI) [35]–[40]. This approach exploits the estimated data symbols as a soft pilot sequence, facilitating a notable reduction in pilot length without compromising estimation performance. The performance bound of the CE algorithms is fundamentally constrained by the pilot length, irrespective of the inference framework employed. To mitigate this limitation, it is essential to utilize not only the reference pilot signal but also the data signal. There are several studies of JCDE algorithms under quantization observations [20]–[24], but to our knowledge, there is a significant lack of research focused on the design of JCDE algorithms that incorporates channel sparsity within the BD domain. Reference [28] is one of the limited examples; however, it necessitates iterative calculations of the inverse matrix on the extended signal model described above, resulting in prohibitively high computational cost. In OFDM transmission, performing DD in the frequency domain is also desirable. However, this domain conversion is non-trivial due to inter-subcarrier interference caused by quantization noise [23], [41]. Therefore, the design of the Bayesian JCDE algorithm for mmWave MIMO-OFDM systems under low-resolution ADCs remains an unresolved challenge.

In light of the above, this paper proposes a novel cross-domain JCDE algorithm based on expectation propagation (EP) [42]–[46], designed to address and eliminate the limitations associated with existing methods. The primary characteristic of the proposed method is its execution of all domain conversion processing within the *message space*, focusing on the fact that in the EP framework derived on the generalized linear model (GLM), all messages exchanged on a factor graph (FG) are projected onto Gaussian distributions through moment matching (MM). This enables the execution of DQ processing, CE, and DD within the designated signal space without the need to expand the signal model as in [27], [28]. Our contribution is fundamentally algorithmic, such that our results are presented in terms of the normalized mean square errors (NMSEs) of estimated channel coefficients and bit error rates (BERs) computed from actual constellation points:

- A novel cross-domain Bayesian JCDE algorithm is proposed, employing a message-passing algorithm (MPA) designed within the EP framework for GLM. In this framework, the AD-domain DQ process, the BD-domain CE, and the antenna-frequency (AF)-domain DD are interconnected through domain converters within the message space. The key theoretical contribution is that all processing steps, including domain conversion, are consistently derived within the EP framework, without introducing any additional assumptions or approximations beyond those inherent in the original EP derivation. This unified formulation is made possible by extending the GLM-based EP framework to handle BBI. To the best of our knowledge, this is the first paper to present a systematic cross-domain JCDE mechanism based on message-space domain conversion that remains fully consistent with MM.
- To validate the proposed cross-domain JCDE algorithm in mmWave MIMO-OFDM systems with low-resolution ADCs, we evaluate its performance under various system parameters. Simulation results demonstrate that the proposed method consistently outperforms state-of-the-art (SotA) alternatives across all tested scenarios and closely approaches the performance of a Genie-aided scheme with perfect CSI. Notably, in our simulation setup, 3-bit quantization is sufficient to achieve performance close to that of the unquantized reference performance.
- New insights into the trade-off between spatial degrees of freedom (DoF) and ADC resolution are presented. When the total number of bits transferred to the data interface card is fixed, increasing the quantization resolution to 2 bits—at the expense of reduced spatial DoF—yields significant performance improvements. However, as the quantization resolution increases further, the marginal gains diminish rapidly, and increasing spatial DoF becomes more critical for performance enhancement.

Notation: A set of complex numbers is denoted by $\mathbb{C}$. Vectors and matrices are denoted in lower- and upper-case boldface fonts, respectively. The conjugate, transpose, and conjugate transpose operators are denoted by $(\cdot)^*$, $(\cdot)^{\mathrm{T}}$, and $(\cdot)^{\mathrm{H}}$, respectively. The real and imaginary parts of a complex variable a are denoted by $a^{\Re}$ $(= \Re\{a\})$ and $a^{\Im}$ $(= \Im\{a\})$, respectively, and defined as $\mathcal{C} \triangleq \{\Re, \Im\}$. Random variables and their outcomes are denoted in sans serif and italic fonts, respectively, as in a and a, such that the conditional probability density function (PDF), the conditional expectation, and conditional variance of the outcome a of a, given the occurrence b of b and a set $\mathcal{A}$ are respectively denoted by $p_{\mathsf{a}|\mathsf{b}}(a|b;\mathcal{A})$, $\mathbb{E}_{\mathsf{a}}[a|b;\mathcal{A}]$, and $\mathbb{V}_{\mathsf{a}}[a|b;\mathcal{A}]$. The real and complex Gaussian distributions with mean a and variance b that a

---

[1] While details are omitted here, various approaches have been studied for data detection (DD) with low-resolution ADCs. In addition to BMMSE filtering and MPDQ [26], which were also introduced in the context of CE, a range of techniques—particularly for 1-bit ADCs—have been proposed, including autoencoding-based methods [32], sphere decoding exploiting symbol discreteness [33], and iterative detection leveraging error correction [34].

variable x follows are respectively denoted by $\mathcal{N}(x;a,b)$ and $\mathcal{CN}(x;a,b)$, and in equations the probability density at $x=\dot{x}$ is denoted by $\mathcal{N}(\dot{x};a,b)$ and $\mathcal{CN}(\dot{x};a,b)$. The $a \times a$ square identity matrix is denoted by $\boldsymbol{I}_a$. The (i,j)-th element of a matrix $\boldsymbol{A}$ is denoted by $[\boldsymbol{A}]_{i,j}$. The diagonal matrix constructed by placing the elements of a vector $\boldsymbol{a}$ on its main diagonal is denoted by $\mathrm{diag}[\boldsymbol{a}]$. The vector generated by cyclically shifting a vector $\boldsymbol{a}$ to the down by p is denoted by $(\boldsymbol{a})_p$. The Frobenius norm is denoted by $|\cdot|_{\mathrm{F}}$. The Dirac delta function is denoted by $\delta(\cdot)$. The PDF and the cumulative distribution function (CDF) of the standard normal distribution are denoted by $\phi(a) \triangleq \frac{1}{\sqrt{2\pi}}e^{-\frac{a^2}{2}}$ and $\mathit{\Phi}(a) \triangleq \int_{-\infty}^{a} \phi(a')\mathrm{d}a'$, respectively, and the complementary function of $\mathit{\Phi}(a)$ is denoted by $\mathit{\Omega}(a) \triangleq 1-\mathit{\Phi}(a) = \int_a^{\infty} \phi(a')\mathrm{d}a'$. Finally, we use the notation $\int_a f(a) \triangleq \int_{-\infty}^{\infty} f(a)\mathrm{d}a$ and $\sum_{i\neq j}^{I} a_i \triangleq \sum_{i=1}^{I} a_i - a_j$.

II. SYSTEM MODEL

A. Signal Model

Consider an uplink massive multi-user MIMO-OFDM system composed of a BS having N receive (RX) antennas and serving M $(\leq N)$ single-antenna user equipment (UE) devices. The frequency-domain transmit (TX) symbol matrix from the m-th UE is expressed as

$$\boldsymbol{X}_m \triangleq [\boldsymbol{x}_{m,1},\ldots,\boldsymbol{x}_{m,t},\ldots,\boldsymbol{x}_{m,T}] \in \mathbb{C}^{K_{\mathrm{c}}\times T}, \tag{1a}$$

$$\boldsymbol{x}_{m,t} \triangleq [x_{m,1t},\ldots,x_{m,k_{\mathrm{c}}t},\ldots,x_{m,K_{\mathrm{c}}t}]^{\mathrm{T}} \in \mathbb{C}^{K_{\mathrm{c}}\times 1}, \tag{1b}$$

where K_{c} and T are the numbers of subcarriers and OFDM symbols, respectively. Each symbol represents a (known) pilot symbol or data symbol. The data symbol represents one among J constellation points $\mathcal{X} \triangleq \{\chi_1,\cdots,\chi_j,\cdots,\chi_J\}$, where the average power density of $\mathcal{X}$ is denoted by E_{s}. Denoting a K_{c}-point fast Fourier transform (FFT) matrix by $\boldsymbol{F}$, the time-domain OFDM symbol can be expressed as

$$\boldsymbol{s}_{m,t} = \boldsymbol{F}^{\mathrm{H}}\boldsymbol{x}_{m,t} \in \mathbb{C}^{K_{\mathrm{c}}\times 1}. \tag{2}$$

A cyclic prefix (CP) is then inserted into the time-domain OFDM symbols, and the resultant signal is transmitted.

At the receiver, the RX signal observed at each RX antenna is converted to a baseband signal, and then the CP is removed. Assuming frequency-selective fading channels, the time-domain RX matrix for the t-th OFDM symbol in the AD domain $\boldsymbol{Y}_t \in \mathbb{C}^{N\times K_{\mathrm{c}}}$ can be expressed as

$$\boldsymbol{Y}_t = \sum_{m=1}^{M}\sum_{p=1}^{P} \boldsymbol{a}_{p,m}\left[\boldsymbol{s}_{p,m,t}\right]^{\mathrm{T}} + \boldsymbol{W}_t, \tag{3}$$

where $p \in \{1,2,\cdots,P\}$ is a path delay index, $\boldsymbol{a}_{p,m} \triangleq [a_{p,1m},\ldots,a_{p,nm},\ldots,a_{p,Nm}]^{\mathrm{T}} \in \mathbb{C}^{N\times 1}$ is a CIR vector for the m-th UE at the p-th delay tap, and $\boldsymbol{s}_{p,m,t} \triangleq (\boldsymbol{s}_{m,t})_{p-1}$ is a cyclic TX vector. With slow TX power control, the total path gain, *i.e.*, $\xi^a = \sum_{p=1}^{P}\mathbb{E}_{\mathsf{a}_{p,nm}}\left[|a_{p,nm}|^2\right]$, is assumed to be known at the receiver side. $\boldsymbol{W}_t \triangleq [\boldsymbol{w}_{t,1},\ldots,\boldsymbol{w}_{t,k_{\mathrm{c}}}\ldots,\boldsymbol{w}_{t,K_{\mathrm{c}}}] \in \mathbb{C}^{N\times K_{\mathrm{c}}}$ is an additive white Gaussian noise (AWGN) matrix, where $\boldsymbol{w}_{t,k_{\mathrm{c}}}$ follows $\mathcal{CN}(\boldsymbol{w}_{t,k_{\mathrm{c}}};\boldsymbol{0},N_0\boldsymbol{I}_N)$, with the noise spectral density N_0, *i.e.*, $\mathbb{E}_{\mathsf{w}_{t,k_{\mathrm{c}}}}[\boldsymbol{w}_{t,k_{\mathrm{c}}}(\boldsymbol{w}_{t,k_{\mathrm{c}}})^{\mathrm{H}}] = N_0\boldsymbol{I}_N$.

Introducing an index $u \triangleq ((m-1)\cdot P + p) \in \{1,\cdots,U\}$, $(U \triangleq MP)$ consisting of the path index p and UE index m, (3) can be rewritten as

$$\boldsymbol{Y}_t = \sum_{m=1}^{M}\sum_{u\in\mathcal{U}_m} \boldsymbol{a}_u\left[\boldsymbol{s}_{u,t}\right]^{\mathrm{T}} + \boldsymbol{W}_t = \sum_{u=1}^{U}\boldsymbol{a}_u\left[\boldsymbol{s}_{u,t}\right]^{\mathrm{T}} + \boldsymbol{W}_t, \tag{4}$$

where $\boldsymbol{a}_u = \boldsymbol{a}_{p,m}$, $\boldsymbol{s}_{u,t} = \boldsymbol{s}_{p,m,t}$, and $\mathcal{U}_m \triangleq \{u \mid (m-1)\cdot P+1,\cdots,m\cdot P\}$. From (4), the spatio-temporal AD-domain RX matrix $\boldsymbol{Y} \in \mathbb{C}^{N\times K}$ can be expressed as

$$\boldsymbol{Y} \triangleq [\boldsymbol{Y}_1,\ldots,\boldsymbol{Y}_t,\ldots,\boldsymbol{Y}_T] = \sum_{u=1}^{U}\boldsymbol{a}_u\left[\boldsymbol{s}_u\right]^{\mathrm{T}} + \boldsymbol{W}, \tag{5}$$

with

$$\boldsymbol{s}_u \triangleq [\boldsymbol{s}_{u,1}^{\mathrm{T}},\ldots,\boldsymbol{s}_{u,t}^{\mathrm{T}},\ldots,\boldsymbol{s}_{u,T}^{\mathrm{T}}]^{\mathrm{T}} \in \mathbb{C}^{K\times 1}, \tag{6a}$$

$$\boldsymbol{W} \triangleq [\boldsymbol{W}_1,\ldots,\boldsymbol{W}_t,\ldots,\boldsymbol{W}_T] \in \mathbb{C}^{N\times K}, \tag{6b}$$

where $K \triangleq K_{\mathrm{c}}T$. Accordingly, the (n,k) element of $\boldsymbol{Y}$ can be expressed as

$$y_{nk} = \sum_{u=1}^{U} a_{nu}s_{uk} + w_{nk} = z_{nk} + w_{nk}, \tag{7}$$

where $a_{nu} \triangleq [\boldsymbol{a}_u]_{n,1}$, $s_{uk} \triangleq [\boldsymbol{s}_u]_{k,1}$, $z_{nk} \triangleq \sum_{u=1}^{U} a_{nu}s_{uk}$, and $w_{nk} \triangleq [\boldsymbol{W}]_{n,k}$, with the index $k \triangleq ((t-1)\cdot K_{\mathrm{c}} + k_{\mathrm{c}}) \in \mathcal{K} \triangleq \{1,\cdots,K\}$.

For later convenience, let $\mathcal{T}_{\mathrm{p}}$ denote the set of time indices to which pilot symbols are assigned, and $\mathcal{T}_{\mathrm{d}}$ denote the set of time indices to which data symbols are assigned, and further define the corresponding symbol index sets by $\mathcal{K}_{\mathrm{p}}$ and $\mathcal{K}_{\mathrm{d}}$, respectively. Denoting the set of symbol indices at the t-th time index by $\mathcal{K}_t \triangleq \{k \in \mathcal{K} \mid (t-1)\cdot K_{\mathrm{c}}+1,\cdots,t\cdot K_{\mathrm{c}}\}$, then $\mathcal{K}_{\mathrm{p}} \triangleq \bigcup_{t\in\mathcal{T}_{\mathrm{p}}}\mathcal{K}_t$ and $\mathcal{K}_{\mathrm{d}} \triangleq \bigcup_{t\in\mathcal{T}_{\mathrm{d}}}\mathcal{K}_t$, where we define $K_{\mathrm{p}} \triangleq |\mathcal{K}_{\mathrm{p}}|$ and $K_{\mathrm{d}} \triangleq |\mathcal{K}_{\mathrm{d}}|$, *i.e.*, $K = K_{\mathrm{p}} + K_{\mathrm{d}}$.

B. Quantization With Low-Resolution ADCs

The in-phase and quadrature components of the signal received at each antenna are individually quantized using an ADC with b-bit resolution. The b-bit ADC can be characterized by a set of 2^b+1 quantization thresholds $\mathcal{R}_b \triangleq \{\lambda_0,\lambda_1,\cdots,\lambda_{2^b}\}$, such that $-\infty < \lambda_0 < \lambda_1 < \cdots < \lambda_{2^b} < \infty$, and a set of 2^b representative values $\mathcal{L}_b \triangleq \{l_0,l_1,\ldots,l_{2^b-1}\}$, where $l_i \in (\lambda_i,\lambda_{i+1}]$. The joint operation of the N b-bit ADCs at the receiver can be represented by the quantization function $Q_b^{N\times K}(\cdot): \mathbb{R}^{N\times K} \to \mathcal{L}_b^{N\times K}$ that maps the AD-domain RX matrix $\boldsymbol{Y}$ to the corresponding quantized output matrix $\boldsymbol{R}$, defined by

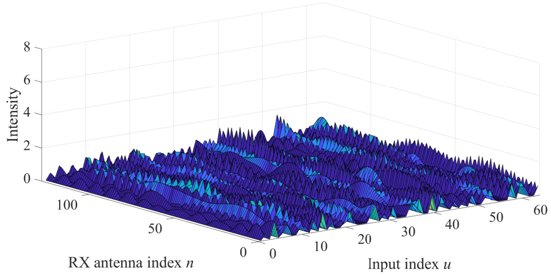
$$\boldsymbol{R} = Q_b^{N\times K}\left(\boldsymbol{Y}^{\Re}\right) + \mathrm{j}\cdot Q_b^{N\times K}\left(\boldsymbol{Y}^{\Im}\right), \tag{8}$$

where the (n,k) element of $\boldsymbol{R}$ is obtained from the corresponding quantization label as [21]

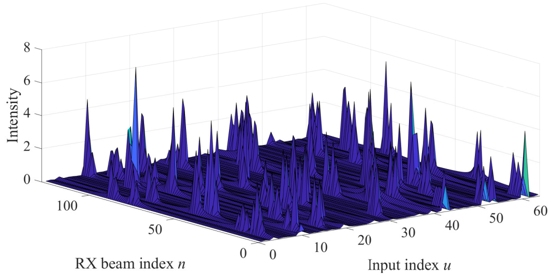
$$r_{nk}^{\Re} = Q_b^{1\times 1}\left(y_{nk}^{\Re}\right) = l_i,\ \text{if } y_{nk}^{\Re} \in (\lambda_i,\lambda_{i+1}], \tag{9a}$$

$$r_{nk}^{\Im} = Q_b^{1\times 1}\left(y_{nk}^{\Im}\right) = l_j,\ \text{if } y_{nk}^{\Im} \in (\lambda_j,\lambda_{j+1}]. \tag{9b}$$

For later convenience, let $\mathcal{L}_{nk}^{\Re} = \{\lambda_i,\lambda_{i+1}\}$ and $\mathcal{L}_{nk}^{\Re} = \{\lambda_j,\lambda_{j+1}\}$ denote the set of quantization thresholds corre-



(a) AD-domain channel matrix $\boldsymbol{A}$.



(b) BD-domain channel matrix $\boldsymbol{H}$.

Fig. 1. Intensity of the elements in the channel matrices, where the MIMO configuration was set to $(N, M, P) = (128, 16, 4)$, and the other parameters are the same as in the numerical studies in Section IV.

sponding to the in-phase and quadrature components of y_{nk}, respectively, where we define $\mathcal{L}_{nk} \triangleq \bigcup_{i \in \mathcal{C}} \mathcal{L}_{nk}^{i}$.

C. Appropriate Signal Domain for CE and DD

Before moving on to the proposed method, this subsection describes an appropriate signal domain for CE and DD based on the signal model without quantization in (5).

1) Domain conversion for CE: In mmWave wireless communications, only a small number of paths reach the receiver, so the MIMO channel observed by a BS with large arrays has sparsity in the angular domain [2]. Hence, using the discrete Fourier transform (DFT) basis for FDBF at the receiver makes it possible to express the beam-domain channel sparsely [47], [48]. Denoting the N-point DFT matrix by $\boldsymbol{D}$, the BD-domain RX matrix is expressed as

$$\boldsymbol{Y}^{\mathcal{BD}} = \boldsymbol{D}^{\mathrm{H}} \boldsymbol{Y} = \sum_{u=1}^{U} \boldsymbol{h}_u \left[\boldsymbol{s}_u \right]^{\mathrm{T}} + \boldsymbol{W}^{\mathcal{BD}}, \tag{10}$$

where $\boldsymbol{h}_u \triangleq \boldsymbol{D}^{\mathrm{H}} \boldsymbol{a}_u$ and $\boldsymbol{W}^{\mathcal{BD}} \triangleq \boldsymbol{D}^{\mathrm{H}} \boldsymbol{W}$ denote the equivalent CIR vector and noise matrix in the BD domain, respectively.

Fig. 1 compares the intensity of the elements in the AD-domain and BD-domain channel matrices, where $\boldsymbol{A} \triangleq [\boldsymbol{a}_1, \ldots, \boldsymbol{a}_u, \ldots, \boldsymbol{a}_U] \in \mathbb{C}^{N \times U}$ and $\boldsymbol{H} \triangleq [\boldsymbol{h}_1, \ldots, \boldsymbol{h}_u, \ldots, \boldsymbol{h}_U] \in \mathbb{C}^{N \times U}$. As shown in Fig. 1(a), the AD-domain channel has a dense structure, and there is no significant difference in the strength among the elements. In contrast, the BD-domain channel shown in Fig. 1(b) has a *sparse-like* structure in which only a small number of elements corresponding to the arrival paths have large values. Hence, accurately estimating only the dominant channel coefficients is sufficient. By formulating CE as a sparse reconstruction problem, the variable space can be efficiently reduced, enabling high-precision CE with shorter pilots [49]–[51].

2) Domain conversion for DD: In OFDM transmission, the orthogonality between subcarriers is guaranteed by inserting a proper CP. The received signal is converted to the frequency domain, and then DD is performed for each subcarrier. From (3), the AF-domain RX matrix is expressed as

$$\begin{aligned} \boldsymbol{Y}_t^{\mathcal{AF}} &= \boldsymbol{Y}_t \boldsymbol{F} = \sum_{u=1}^{U} \boldsymbol{a}_u \left[\boldsymbol{s}_{u,t} \right]^{\mathrm{T}} \boldsymbol{F} + \boldsymbol{W}_t \boldsymbol{F} \\ &\triangleq \left[\boldsymbol{y}_{1,t}^{\mathcal{AF}}, \ldots, \boldsymbol{y}_{k_{\mathrm{c}},t}^{\mathcal{AF}}, \ldots, \boldsymbol{y}_{K_{\mathrm{c}},t}^{\mathcal{AF}} \right] \in \mathbb{C}^{N \times K_{\mathrm{c}}}, \end{aligned} \tag{11}$$

with

$$\boldsymbol{y}_{k_{\mathrm{c}},t}^{\mathcal{AF}} = \sum_{m=1}^{M} \boldsymbol{c}_{k_{\mathrm{c}},m} x_{m,k_{\mathrm{c}}t} + \boldsymbol{w}_{k_{\mathrm{c}},t}^{\mathcal{AF}}, \tag{12}$$

where $\boldsymbol{c}_{k_{\mathrm{c}},m} \triangleq \left[c_{k_{\mathrm{c}},1m}, \ldots, c_{k_{\mathrm{c}},nm}, \ldots, c_{k_{\mathrm{c}},Nm} \right]^{\mathrm{T}} \in \mathbb{C}^{N \times 1}$ is the AF-domain channel vector at the k_{c}-th subcarrier, the n-th element of which can be expressed as

$$c_{k_{\mathrm{c}},nm} = \sqrt{K_{\mathrm{c}}} \sum_{u \in \mathcal{U}_m} f_{k_{\mathrm{c}},\upsilon(u)} a_{nu}, \tag{13}$$

with $\upsilon(i) \triangleq (i - 1 \mod P) + 1$, $f_{i,j} \triangleq [\boldsymbol{F}]_{i,j}$, and $|f_{i,j}|^2 = 1/K_{\mathrm{c}}$. $\boldsymbol{w}_{k_{\mathrm{c}},t}^{\mathcal{AF}}$ denotes the AF-domain AWGN vector. From (12), the received signal $\boldsymbol{Y}_t$ can be separated into K_{c} AF-domain vectors $\boldsymbol{y}_{k_{\mathrm{c}},t}^{\mathcal{AF}}, \forall k_{\mathrm{c}}$, via FFT, so subsequent DD can be performed for each subcarrier.

3) Problem statements and motivation: As described above, it is appropriate to perform CE and DD within different signal domains, and furthermore, the domain conversion process in (10) and (11) cannot be directly applied to the signal after quantization by the low-resolution ADC in (8). In the following section, we will provide a framework that enables domain conversion even in the presence of quantization errors, and propose a cross-domain JCDE algorithm that can perform each signal processing within the designated signal space.

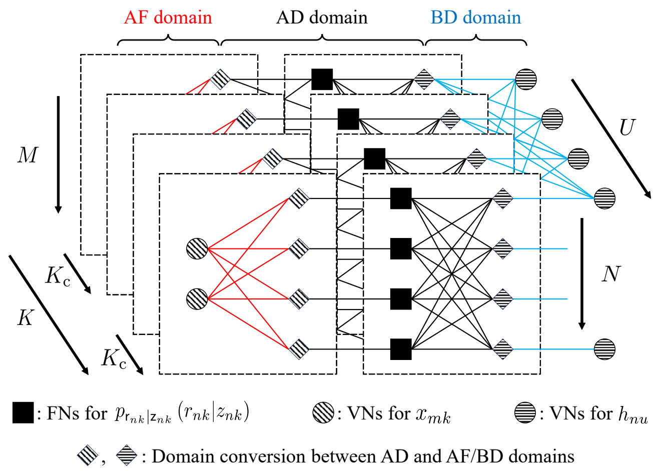
III. CROSS-DOMAIN JCDE ALGORITHM BASED ON EXPECTATION PROPAGATION

In this section, we describe a novel cross-domain JCDE algorithm based on the EP framework for GLM, which detects the intended data symbols $x_{mk}, \forall(n, k \in \mathcal{K}_{\mathrm{d}})$, and estimates the channel coefficients $a_{nu}, \forall(n, u)$, out of the quantized outputs $r_{nk}, \forall(n, k)$, quantization thresholds $\mathcal{L}_{nk}, \forall(n, k)$, and the pilot symbols $x_{mk}, \forall(n, k \in \mathcal{K}_{\mathrm{p}})$, based on (8).

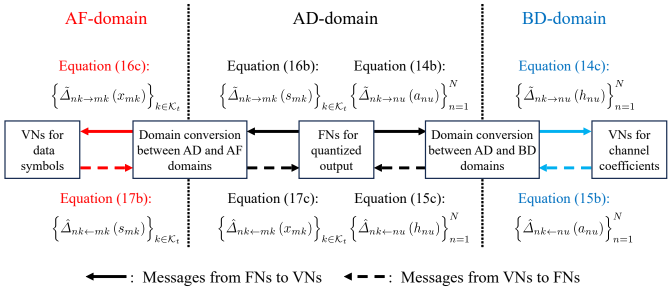
A. Overview of Cross-Domain JCDE Algorithm

The proposed algorithm is derived by approximating the classical sum-product algorithm (SPA) according to the EP framework [43]–[46] and then introducing a domain conversion mechanism. In SPA, unknown variables are estimated by iteratively exchanging messages (probability information reflecting the estimation accuracy of unknown variables) between the factor nodes (FNs) corresponding to the observation signal (*i.e.*, quantized outputs) and the variable nodes (VNs) corresponding to the unknown variables (*i.e.*, channel coefficients and data symbols).

Fig. 2(a) shows the tripartite FG of the proposed cross-domain JCDE algorithm, where the square marks indicate FNs,



(a) Illustration of the tripartite FG designed for the proposed cross-domain JCDE algorithm, with $N = 4$, $M = 2$, $U = 4$, $K_c = 2$, and $T = 2$.



(b) Illustration of the message-passing flow between FNs and VNs.

Fig. 2. Schematic illustration of the proposed cross-domain JCDE algorithm.

the circle marks indicate VNs, and the diamond marks indicate domain conversion within the message space, respectively. The SPA based on this FG can be expressed as follows, including the projection operation of the message distribution onto the Gaussian distribution, which is performed under the assumption of subsequent application of the GLM-based EP framework.

Messages from FNs to VNs for CE:

$$\overline{\Delta}_{nk}(a_{nu}) = \int_{\{a_{ni}\}_{i\neq u}^{U}} \int_{\{s_{uk}\}_{u=1}^{U}} p_{\mathsf{r}_{nk}|\mathsf{z}_{nk}}(r_{nk}|z_{nk}) \prod_{i\neq u}^{U} \hat{\Delta}_{nk\leftarrow ni}(a_{ni}) \prod_{u=1}^{U} \hat{\Delta}_{nk\leftarrow uk}(s_{uk}), \tag{14a}$$

$$\tilde{\Delta}_{nk\to nu}(a_{nu}) = \frac{\mathrm{Proj}_{\mathrm{G}}\left[\overline{\Delta}_{nk}(a_{nu})\,\hat{\Delta}_{nk\leftarrow nu}(a_{nu})\right]}{\hat{\Delta}_{nk\leftarrow nu}(a_{nu})}, \tag{14b}$$

$$\tilde{\Delta}_{nk\to nu}(h_{nu}) = \mathrm{Proj}_{\mathrm{A}\to\mathrm{B}}\left[\tilde{\Delta}_{nk\to nu}(a_{nu})\right], \tag{14c}$$

Messages from VNs to FNs for CE:

$$\Delta_{nu}(h_{nu}) = p_{\mathsf{h}_{nu}}(h_{nu}) \prod_{k=1}^{K} \tilde{\Delta}_{nk\to nu}(h_{nu}), \tag{15a}$$

$$\hat{\Delta}_{nk\leftarrow nu}(h_{nu}) = \frac{\mathrm{Proj}_{\mathrm{G}}\left[\Delta_{nu}(h_{nu})\right]}{\tilde{\Delta}_{nk\to nu}(h_{nu})}, \tag{15b}$$

$$\hat{\Delta}_{nk\leftarrow nu}(a_{nu}) = \mathrm{Proj}_{\mathrm{A}\leftarrow\mathrm{B}}\left[\hat{\Delta}_{nk\leftarrow nu}(h_{nu})\right]. \tag{15c}$$

Messages from FNs to VNs for DD:

$$\overline{\Delta}_{nk}(s_{mk}) = \int_{\{a_{ni}\}_{u=1}^{U}} \int_{\{\{s_{uk}\}_{u\in\mathcal{U}_i}\}_{i\neq m}^{M}} p_{\mathsf{r}_{nk}|\mathsf{z}_{nk}}(r_{nk}|z_{nk}) \prod_{u=1}^{U} \hat{\Delta}_{nk\leftarrow nu}(a_{ni}) \prod_{i\neq m}^{M} \prod_{u\in\mathcal{U}_i} \hat{\Delta}_{nk\leftarrow uk}(s_{uk}), \tag{16a}$$

$$\tilde{\Delta}_{nk\to mk}(s_{mk}) = \frac{\mathrm{Proj}_{\mathrm{G}}\left[\overline{\Delta}_{nk}(s_{mk}) \prod_{u\in\mathcal{U}_m} \hat{\Delta}_{nk\leftarrow uk}(s_{uk})\right]}{\prod_{u\in\mathcal{U}_m} \hat{\Delta}_{nk\leftarrow uk}(s_{uk})}, \tag{16b}$$

$$\tilde{\Delta}_{nk\to mk}(x_{mk}) = \mathrm{Proj}_{\mathrm{D}\to\mathrm{F}}\left[\tilde{\Delta}_{nk\to mk}(s_{mk})\right], \tag{16c}$$

Messages from VNs to FNs for DD:

$$\Delta_{mk}(x_{mk}) = p_{\mathsf{x}_{mk}}(x_{mk}) \prod_{n=1}^{N} \tilde{\Delta}_{nk\to mk}(x_{mk}), \tag{17a}$$

$$\hat{\Delta}_{nk\leftarrow mk}(x_{mk}) = \frac{\mathrm{Proj}_{\mathrm{G}}\left[\Delta_{mk}(x_{mk})\right]}{\tilde{\Delta}_{nk\to mk}(x_{mk})}, \tag{17b}$$

$$\hat{\Delta}_{nk\leftarrow uk}(s_{uk}) = \mathrm{Proj}_{\mathrm{D}\leftarrow\mathrm{F}}\left[\hat{\Delta}_{nk\leftarrow mk}(x_{mk})\right], \forall (u,k) \in \mathcal{I}_{mk}^{s}. \tag{17c}$$

To clarify the relationship between the tripartite FG shown in Fig. 2(a) and the above SPA algorithm, Fig. 2(b) presents a schematic illustration of the message-passing flow, explicitly depicting the messages exchanged between the nodes in the FG. As you can see, since the MPAs for CE and DD are symmetrical, we will first describe the CE in (14) and (15) in detail and then describe the DD in (16) and (17) by focusing only on the points that differ from the CE case.

First, in (14a), we compute the likelihood information for the AD-domain channel coefficient a_{nu} by using the message propagated from VNs as prior information and marginalizing over all variables other than a_{nu} from $z_{nk} \triangleq \sum_{u=1}^{U} a_{nu} s_{uk}$. Next, in (14b), the message distribution obtained in (14a) is projected onto a Gaussian distribution based on the EP framework, and extrinsic information $\tilde{\Delta}_{nk\to nu}(a_{nu})$ that excludes the self-feedback component through MM is generated, where $q(a) = \mathrm{Proj}_{\mathrm{G}}[p(a)]$ is an operator that projects the input arbitrary distribution $p(a)$ onto Gaussian distribution $q(a)$ such that the Kullback-Leibler divergence (KLD) $D_{\mathrm{KL}}(p(a)\,\|\,q(a))$ between $p(a)$ and $q(a)$ is minimized, *i.e.*,

$$q(a) = \underset{q'(a)}{\mathrm{argmin}}\, D_{\mathrm{KL}}(p(a)\,\|\,q'(a)). \tag{18}$$

This projection of the posterior distribution onto a Gaussian distribution serves as the central approximation in the derivation of a practical algorithm within the EP framework [42]–[44]. It has been shown that the Gaussian distribution that minimizes the KLD from the true posterior, *i.e.*, the solution to (18), can be obtained through MM. The following (14c) is a domain conversion process within the message space, where $p(h) = \mathrm{Proj}_{\mathrm{A}\to\mathrm{B}}[p(a)]$ is an operator that projects the distribution of the AD-domain channel coefficients a onto the distribution of the corresponding BD-domain channel coefficients h via DFT. As we will describe in more detail later, since the message has already been projected onto a

Gaussian distribution in (14b), this operation is a simple DFT on the message. Now that the FN processing is complete, let us move on to the VN processing. First, in (15a), the posterior information of h_{nu} is computed by combining the messages propagated from FNs. Next, in (15b), the extrinsic information $\hat{\Delta}_{nk\leftarrow nu}(h_{nu})$ is computed in the same way as in (14b), and then in (15c), the domain conversion (*i.e.*, reverse operation of (14c)) is performed within the message space again. Finally, the message obtained in (15c) is fed back to the FNs.

Next, focusing on the DD in (16) and (17), they are basically the same as the CE, but the difference is that they include a joint operation over $u \in \mathcal{U}_m$ in (16a) and (16b). This is because the DD-domain data symbol s_{uk} contains P overlapping s_{mk} due to the cyclic shift operation, where we define the set of P two-dimensional (2D) indices (u,k) corresponding to s_{mk} as $\mathcal{I}^s_{mk}$ ($|\mathcal{I}^s_{mk}| = P$). There is no difference in the VN processing, where $p(x) = \mathrm{Proj}_{\mathrm{D}\to\mathrm{F}}[p(s)]$ in (16c) is an operator that projects the distribution of the delay-domain symbol s onto the distribution of the frequency-domain symbol x.

In the following subsections, we derive the EP-based MPA in a closed form that avoids the multiple integrals in (14a) and (16a) by sequentially approximating the messages from (14) to (17) using the scalar Gaussian approximation (SGA) [52] in conformity to the central limit theorem (CLT).

For later convenience, let us define the soft estimates of the AD-domain channel coefficients a_{nu} and TX symbols s_{uk}, as

$$\hat{a}_{k,nu} \triangleq \mathbb{E}_{\mathsf{a}_{nu}}\left[a_{nu} \mid \hat{\Delta}_{nk\leftarrow nu}(a_{nu})\right], \forall k, \tag{19a}$$

$$\hat{s}_{n,uk} \triangleq \mathbb{E}_{\mathsf{s}_{uk}}\left[s_{uk} \mid \hat{\Delta}_{nk\leftarrow uk}(s_{uk})\right], \forall n, \tag{19b}$$

respectively, such that their mean square errors (MSEs) can be respectively expressed as

$$\hat{\psi}^a_{k,nu} \triangleq \mathbb{V}_{\mathsf{a}_{nu}}\left[a_{nu} \mid \hat{\Delta}_{nk\leftarrow nu}(a_{nu})\right], \forall k, \tag{20a}$$

$$\hat{\psi}^s_{n,uk} \triangleq \mathbb{V}_{\mathsf{s}_{uk}}\left[s_{uk} \mid \hat{\Delta}_{nk\leftarrow uk}(s_{uk})\right], \forall n. \tag{20b}$$

B. Derivation of EP-based MPA for CE via SGA

1) Messages from FNs to VNs in (14)*:* The derivation consists of the following three steps: a) Approximate the message in (14a) with a Gaussian distribution based on SGA. b) Reconstruct the message in (14b) with a Gaussian distribution through MM. c) Replace the domain conversion operation from AD-domain to BD-domain in (14c) with DFT for the message.

Consider the estimation of an arbitrary a_{un}. For the large-system regime ($U \gg 1$) [53], the PDF of z_{nk} approaches a complex scalar Gaussian distribution in conformity to CLT. The conditional PDF of z_{nk}, given a_{nu}, can be expressed as

$$p_{\mathsf{z}_{nk}|\mathsf{a}_{nu}}(z_{nk}|a_{un}) \propto \exp\left[-\frac{|z_{nk}-\hat{p}_{u,nk}|^2}{\hat{\xi}^p_{u,nk}}\right], \tag{21}$$

where

$$\begin{aligned}&\hat{p}_{u,nk}\\&= \mathbb{E}_{\mathsf{z}_{nk}}\left[z_{nk} \,|\, a_{nu}, \{\hat{\Delta}_{nk\leftarrow ni}(a_{ni})\}^U_{i\neq u}, \{\hat{\Delta}_{nk\leftarrow uk}(s_{uk})\}^U_{u=1}\right]\\&= a_{nu}\hat{s}_{n,uk} + \tilde{p}_{u,nk} = g^a_{k,nu} + \tilde{p}_{u,nk},\end{aligned} \tag{22a}$$

$$\begin{aligned}&\hat{\xi}^p_{u,nk}\\&= \mathbb{V}_{\mathsf{z}_{nk}}\left[z_{nk} \,|\, a_{nu}, \{\hat{\Delta}_{nk\leftarrow ni}(a_{ni})\}^U_{i\neq u}, \{\hat{\Delta}_{nk\leftarrow uk}(s_{uk})\}^U_{u=1}\right]\\&= |a_{nu}|^2 \hat{\psi}^s_{n,uk} + \tilde{\xi}^p_{u,nk} \approx \hat{\xi}^a_u \hat{\psi}^s_{n,uk} + \tilde{\xi}^p_{u,nk},\end{aligned} \tag{22b}$$

with

$$g^a_{k,nu} \triangleq a_{nu}\hat{s}_{n,uk}, \tag{23a}$$

$$\tilde{p}_{u,nk} \triangleq \sum^U_{i\neq u} \hat{g}_{k,ni},\;\; \hat{g}_{k,nu} \triangleq \hat{a}_{k,nu}\hat{s}_{n,uk}, \tag{23b}$$

$$\tilde{\xi}^p_{u,nk} \triangleq \sum^U_{i\neq u} \left\{|\hat{a}_{k,ni}|^2 \hat{\psi}^s_{n,ik} + \hat{\psi}^a_{k,ni}\left(|\hat{s}_{n,ik}|^2 + \hat{\psi}^s_{n,ik}\right)\right\}. \tag{23c}$$

In (22b), the instantaneous channel gain $|a_{nu}|^2$ is not available; hence, by approximately replacing these with long-term statistics, *i.e.*, the variance $\mathbb{E}_{\mathsf{a}_{nu}}\left[|a_{nu}|^2\right] = \hat{\xi}^a_u$ [36]. Accordingly, in (14a), the likelihood information for a_{nu} is estimated by computing the marginalization based on the messages propagated from VNs on the basis of the predictive distribution rule: $p_{\mathsf{a}|\mathsf{b}}(a|b) = \int_c p_{\mathsf{a}|\mathsf{c}}(a|c)\cdot p_{\mathsf{c}|\mathsf{b}}(c|b)$. From (21), the message of (14a) can be approximated by [35], [54]

$$\begin{aligned}\overline{\Delta}_{nk}(a_{nu}) \approx \int_{z_{nk}} &p_{\mathsf{r}_{nk}|\mathsf{z}_{nk}}(r_{nk}|z_{nk})\\&\times \mathcal{CN}\left(z_{nk}; g^a_{k,nu} + \tilde{p}_{u,nk}, \hat{\xi}^p_{u,nk}\right).\end{aligned} \tag{24}$$

Next, $\hat{\Delta}_{nk\leftarrow nu}(a_{nu}) = \mathcal{CN}\left(g^a_{k,nu}; \hat{g}_{k,nu}, \hat{\psi}^g_{k,nu}\right)$ is computed through MM, where $\hat{\psi}^g_{k,nu} \triangleq |\hat{s}_{n,uk}|^2 \hat{\psi}^a_{k,nu}$ [43], [44]. From (24), the posterior probability information within the projection function in (14b) can be expressed as

$$\begin{aligned}\overline{\Delta}_{nk}(a_{nu})\hat{\Delta}_{nk\leftarrow nu}(a_{nu}) = \int_{z_{nk}} p_{\mathsf{r}_{nk}|\mathsf{z}_{nk}}(r_{nk}|z_{nk})\\ \times \underbrace{\mathcal{CN}\left(z_{nk}; g^a_{k,nu} + \tilde{p}_{u,nk}, \hat{\xi}^p_{u,nk}\right)\mathcal{CN}\left(g^a_{k,nu}; \hat{g}_{k,nu}, \hat{\psi}^g_{k,nu}\right)}_{= \mathcal{CN}\left(g^a_{k,nu}; z_{nk}-\tilde{p}_{u,nk}, \xi^p_{u,nk}\right)\mathcal{CN}\left(g^a_{k,nu}; \hat{g}_{k,nu}, \hat{\psi}^g_{k,nu}\right)},\end{aligned} \tag{25}$$

which can be rewritten using the Gaussian-PDF multiplication rule: $\mathcal{CN}(x; a, b)\,\mathcal{CN}(x; A, B) = \mathcal{CN}\left(x; \frac{aB+Ab}{b+B}, \frac{bB}{b+B}\right)\mathcal{CN}(0; a-A, b+B)$, as follows:

$$\begin{aligned}\overline{\Delta}_{nk}(a_{nu})\hat{\Delta}_{nk\leftarrow nu}(a_{nu}) = \int_{z_{nk}} p_{\mathsf{r}_{nk}|\mathsf{z}_{nk}}(r_{nk}|z_{nk})\\ \times \mathcal{CN}\left(z_{nk}; \overline{z}_{nk}, \overline{\xi}^z_{u,nk}\right)\mathcal{CN}\left(g^a_{k,nu}; \overline{g}_{k,nu}, \overline{\xi}^g_{k,nu}\right),\end{aligned} \tag{26}$$

with

$$\overline{z}_{nk} \triangleq \tilde{p}_{u,nk} + \hat{g}_{k,nu} = \sum^U_{u=1} \hat{g}_{k,nu}, \tag{27a}$$

$$\overline{\xi}^z_{u,nk} \triangleq \hat{\xi}^p_{u,nk} + \hat{\psi}^g_{k,nu}, \tag{27b}$$

$$\overline{g}_{k,nu} \triangleq \overline{\xi}^g_{k,nu}\left(\frac{z_{nk}-\tilde{p}_{u,nk}}{\hat{\xi}^p_{u,nk}} + \frac{\hat{g}_{k,nu}}{\hat{\psi}^g_{k,nu}}\right), \tag{27c}$$

$$\overline{\xi}^g_{k,nu} \triangleq \left(\frac{1}{\hat{\xi}^p_{u,nk}} + \frac{1}{\hat{\psi}^g_{k,nu}}\right)^{-1}. \tag{27d}$$

This article has been accepted for publication in IEEE Journal of Selected Topics in Signal Processing. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JSTSP.2025.3589107

From (26), using Pearl's normalization constant, the posterior probability of z_{nk} can be expressed as

$$p_{\mathsf{z}_{nk}}\left(z_{nk}\right)=\frac{p_{\mathsf{r}_{nk}|\mathsf{z}_{nk}}\left(r_{nk}|z_{nk}\right)\mathcal{CN}\left(z_{nk};\overline{z}_{nk},\overline{\xi}^{z}_{u,nk}\right)}{\int_{z'_{nk}}p_{\mathsf{r}_{nk}|\mathsf{z}_{nk}}\left(r_{nk}|z'_{nk}\right)\mathcal{CN}\left(z'_{nk};\overline{z}_{nk},\overline{\xi}^{z}_{u,nk}\right)}, \tag{28}$$

which is used to rewrite (26) as follows:

$$\begin{aligned}&\overline{\varDelta}_{nk}\left(a_{nu}\right)\hat{\varDelta}_{nk\leftarrow nu}\left(a_{nu}\right)\\&\quad\propto\underbrace{\int_{z_{nk}}p_{\mathsf{z}_{nk}}\left(z_{nk}\right)\mathcal{CN}\left(g^{a}_{k,nu};\overline{g}_{k,nu},\overline{\xi}^{g}_{k,nu}\right)}_{\triangleq\, p_{\mathsf{g}^{a}_{k,nu}}\left(g^{a}_{k,nu}\right)}.\end{aligned} \tag{29}$$

Following [43], [44], the distribution after projection is given by a Gaussian distribution whose parameters are the conditional expectation and variance calculated from the posterior probability information before projection, *i.e.*,

$$\mathcal{CN}\left(g^{a}_{k,nu};\check{g}^{a}_{k,nu},\check{\xi}^{a}_{k,nu}\right)=\mathrm{Proj}_{\mathrm{G}}\left[\overline{\varDelta}_{nk}\left(a_{nu}\right)\hat{\varDelta}_{nk\leftarrow nu}\left(a_{nu}\right)\right], \tag{30}$$

where

$$\begin{aligned}\check{g}^{a}_{k,nu}&=\int_{g^{a}_{k,nu}}g^{a}_{k,nu}p_{\mathsf{g}^{a}_{k,nu}}\left(g^{a}_{k,nu}\right)\\&=\frac{\hat{\psi}^{g}_{k,nu}\left(\check{z}_{u,nk}-\tilde{p}_{u,nk}\right)+\hat{\xi}^{p}_{u,nk}\hat{g}_{k,nu}}{\overline{\xi}^{z}_{u,nk}},\end{aligned} \tag{31a}$$

$$\begin{aligned}\check{\xi}^{a}_{k,nu}&=\int_{g^{a}_{k,nu}}\left|g^{a}_{k,nu}\right|^{2}p_{\mathsf{g}^{a}_{k,nu}}\left(g^{a}_{k,nu}\right)-\left|\check{g}^{a}_{k,nu}\right|^{2}\\&=\frac{\hat{\psi}^{g}_{k,nu}}{\overline{\xi}^{z}_{u,nk}}\frac{\hat{\psi}^{g}_{k,nu}\check{\xi}^{z}_{u,nk}+\hat{\xi}^{p}_{u,nk}\overline{\xi}^{z}_{u,nk}}{\overline{\xi}^{z}_{u,nk}},\end{aligned} \tag{31b}$$

with

$$\check{z}_{u,nk}=\int_{z_{nk}}z_{nk}p_{\mathsf{z}_{nk}}\left(z_{nk}\right), \tag{32a}$$

$$\check{\xi}^{z}_{u,nk}=\int_{z_{nk}}\left|z_{nk}\right|^{2}p_{\mathsf{z}_{nk}}\left(z_{nk}\right)-\left|\check{z}_{u,nk}\right|^{2}. \tag{32b}$$

Substituting $\hat{\varDelta}_{nk\leftarrow nu}\left(a_{nu}\right)=\mathcal{CN}\left(g^{a}_{k,nu};\hat{g}_{k,nu},\hat{\psi}^{g}_{k,nu}\right)$ and (30) into (14b), $\tilde{\varDelta}_{nk\rightarrow nu}\left(a_{nu}\right)$ can be expressed as

$$\begin{aligned}\tilde{\varDelta}_{nk\rightarrow nu}\left(a_{nu}\right)&=\frac{\mathcal{CN}\left(g^{a}_{k,nu};\check{g}^{a}_{k,nu},\check{\xi}^{a}_{k,nu}\right)}{\mathcal{CN}\left(g^{a}_{k,nu};\hat{g}_{k,nu},\hat{\psi}^{g}_{k,nu}\right)}\\&\propto\mathcal{CN}\left(g^{a}_{k,nu};\tilde{g}^{a}_{k,nu},\tilde{\xi}^{a}_{k,nu}\right),\end{aligned} \tag{33}$$

and we have

$$\begin{aligned}\tilde{\xi}^{a}_{k,nu}&=\left(\frac{1}{\check{\xi}^{g}_{u,nk}}-\frac{1}{\hat{\psi}^{g}_{k,nu}}\right)^{-1}\\&=\frac{\overline{\xi}^{z}_{u,nk}}{\overline{\xi}^{z}_{u,nk}-\check{\xi}^{z}_{u,nk}}\overline{\xi}^{z}_{u,nk}-\hat{\psi}^{g}_{k,nu},\end{aligned} \tag{34a}$$

$$\begin{aligned}\tilde{g}^{a}_{k,nu}&=\tilde{\xi}^{a}_{k,nu}\left(\frac{\check{g}^{a}_{k,nu}}{\check{\xi}^{g}_{u,nk}}-\frac{\hat{g}_{k,nu}}{\hat{\psi}^{g}_{k,nu}}\right)\\&=\frac{\overline{\xi}^{z}_{u,nk}}{\overline{\xi}^{z}_{u,nk}-\check{\xi}^{z}_{u,nk}}\left(\check{z}_{u,nk}-\overline{z}_{nk}\right)+\hat{g}_{k,nu}.\end{aligned} \tag{34b}$$

Based on the quantization operation in Section II-B, from (9) and (28), (32b) can be expressed as [29]

$$\check{\xi}^{z}_{u,nk}=\frac{N_0\overline{\xi}^{z}_{u,nk}}{\overline{\xi}^{z}_{u,nk}+N_0}+\left(\frac{\overline{\xi}^{z}_{u,nk}}{\overline{\xi}^{z}_{u,nk}+N_0}\right)^{2}\check{\xi}^{y}_{u,nk}, \tag{35}$$

with

$$\check{\xi}^{y}_{u,nk}=\mathbb{V}_{\mathsf{y}_{nk}}\left[y_{nk}\mid\overline{z}_{nk},\overline{\xi}^{z}_{u,nk}+N_0;\mathcal{L}_{nk}\right]. \tag{36}$$

From (35), the following relationship holds:

$$\left(\frac{\overline{\xi}^{z}_{u,nk}}{\overline{\xi}^{z}_{u,nk}-\check{\xi}^{z}_{u,nk}}\right)^{-1}=\frac{\overline{\xi}^{z}_{u,nk}}{\overline{\xi}^{z}_{u,nk}+N_0}\left(1-\frac{\check{\xi}^{y}_{u,nk}}{\overline{\xi}^{z}_{u,nk}+N_0}\right), \tag{37}$$

and therefore (34a) can be rewritten as

$$\tilde{\xi}^{a}_{k,nu}=\left(1-\frac{\check{\xi}^{y}_{u,nk}}{\overline{\xi}^{z}_{u,nk}+N_0}\right)^{-1}\left(\overline{\xi}^{z}_{u,nk}+N_0\right)-\hat{\psi}^{g}_{k,nu}. \tag{38}$$

Similar to (35), (32a) can be obtained from [29]

$$\check{z}_{u,nk}=\frac{N_0\overline{z}_{nk}+\overline{\xi}^{z}_{u,nk}\check{y}_{u,nk}}{\overline{\xi}^{z}_{u,nk}+N_0}, \tag{39}$$

with

$$\check{y}_{u,nk}=\mathbb{E}_{\mathsf{y}_{nk}}\left[y_{nk}\mid\overline{z}_{nk},\overline{\xi}^{z}_{u,nk}+N_0;\mathcal{L}_{nk}\right], \tag{40}$$

and from (37) and (39), (34b) can be rewritten as

$$\tilde{g}^{a}_{k,nu}=\left(1-\frac{\check{\xi}^{y}_{u,nk}}{\overline{\xi}^{z}_{u,nk}+N_0}\right)^{-1}\left(\check{y}_{u,nk}-\overline{z}_{nk}\right)+\hat{g}_{k,nu}. \tag{41}$$

Finally, under SGA of the message, the domain conversion in (14c) is replaced with a simple DFT. From (33), $\tilde{g}^{a}_{k,nu}$ can be modeled as the AWGN channel of $g^{a}_{k,nu}$ $(=a_{nu}\hat{s}_{n,uk})$, *i.e.*,

$$\tilde{g}^{a}_{k,nu}=a_{nu}\hat{s}_{n,uk}+w^{a}_{k,nu}, \tag{42}$$

where $w^{a}_{k,nu}$ follows $\mathcal{CN}\left(w^{a}_{k,nu};0,\tilde{\xi}^{a}_{k,nu}\right)$. By applying FDBF to $\tilde{g}^{a}_{k,nu},\forall n$, (42) can be approximated by a linear model for the BD-domain channel coefficient h_{nu}, as

$$\begin{aligned}\tilde{g}^{h}_{k,nu}&=\sum_{i=1}^{N}d^{*}_{n,i}\tilde{g}^{a}_{k,iu}\approx\check{s}_{uk}\sum_{i=1}^{N}d^{*}_{n,i}a_{ni}+\sum_{i=1}^{N}d^{*}_{n,i}w^{a}_{k,iu}\\&=h_{nu}\check{s}_{uk}+w^{h}_{k,nu},\end{aligned} \tag{43}$$

where $w^{h}_{k,nu}\triangleq\sum_{i=1}^{N}d^{*}_{n,i}w^{a}_{k,iu}$, $d_{i,j}\triangleq[\boldsymbol{D}]_{i,j}$, and $|d_{i,j}|^{2}=1/N$. $\check{s}_{uk}$ is the posterior estimate of s_{uk}, as described later. Assuming that the effective Gaussian noise in $\tilde{g}^{a}_{k,nu},\forall n$, is not correlated with each other, from the reproductive property of a Gaussian distribution, the conditional PDF of $\tilde{g}^{h}_{k,nu}$, given $h_{nu}\check{s}_{uk}$, can be expressed as

$$p_{\tilde{\mathsf{g}}^{h}_{k,nu}|\mathsf{h}_{nu}}\left(\tilde{g}^{h}_{k,nu}|h_{nu}\right)=\mathcal{CN}\left(h_{nu}\check{s}_{uk};\tilde{g}^{h}_{k,nu},\tilde{\xi}^{h}_{k,nu}\right), \tag{44}$$

with

$$\tilde{\xi}^{h}_{k,nu}=\mathbb{E}_{\{\mathsf{w}^{a}_{k,nu},\forall n\}}\left[\left|w^{h}_{k,nu}\right|^{2}\right]=\frac{1}{N}\sum_{i=1}^{N}\tilde{\xi}^{a}_{k,iu},\ \forall n. \tag{45}$$

Clearly, (44) is the message for the BD-domain channel coefficient h_{nu} after domain conversion, and is propagated to the VNs as $\tilde{\Delta}_{nk\to nu}\left(h_{nu}\right)$.

2) Messages from VNs to FNs in (15)*:* First, the joint likelihood information for h_{nu} is generated by combining the messages propagated from the FNs. From (15a) and (44), $\Delta_{nu}\left(h_{nu}\right)$ can be expressed as

$$\begin{aligned}\Delta_{nu}\left(h_{nu}\right) &= p_{\mathsf{h}_{nu}}\left(h_{nu}\right)\prod_{k=1}^{K}\tilde{\Delta}_{nk\to nu}\left(h_{nu}\right)\\ &\propto p_{\mathsf{h}_{nu}}\left(h_{nu}\right)\mathcal{CN}\left(h_{nu};\mu_{nu}^{h},\psi_{nu}^{h}\right),\end{aligned} \tag{46}$$

where

$$\mu_{nu}^{h}=\psi_{nu}^{h}\sum_{k=1}^{K}\frac{\tilde{g}_{u,nk}^{h}\check{s}_{uk}^{*}}{\tilde{\xi}_{u,nk}^{h}},\ \psi_{nu}^{h}=\left(\sum_{k=1}^{K}\frac{\left|\check{s}_{uk}\right|^{2}}{\tilde{\xi}_{u,nk}^{h}}\right)^{-1}. \tag{47}$$

Assuming that the effective Gaussian noise in $\mu_{nu}^{h},\forall(n,u)$, is not correlated with each other, using Bayes' rule, the posterior estimate and its MSE can be in general obtained from the symbol-wise conditional expectation as [36]

$$\check{h}_{nu}=\int_{h_{nu}}h_{nu}\frac{\Delta_{nu}\left(h_{nu}\right)}{\int_{h'_{nu}}\Delta_{nu}\left(h'_{nu}\right)}, \tag{48a}$$

$$\check{\psi}_{nu}^{h}=\int_{h_{nu}}\left|h_{nu}\right|^{2}\frac{\Delta_{nu}\left(h_{nu}\right)}{\int_{h'_{nu}}\Delta_{nu}\left(h'_{nu}\right)}-\left|\check{h}_{nu}\right|^{2}. \tag{48b}$$

Similarly to (30), the posterior distribution of h_{nu} is given by a Gaussian distribution based on (48) through MM. Substituting the projected distribution and (44) to (15b) yields

$$\begin{aligned}\hat{\Delta}_{nk\leftarrow nu}\left(h_{nu}\right) &= \frac{\mathcal{CN}\left(h_{nu};\check{h}_{nu},\check{\psi}_{nu}^{h}\right)}{\mathcal{CN}\left(h_{nu}\check{s}_{uk};\tilde{g}_{k,nu}^{h},\tilde{\xi}_{k,nu}^{h}\right)}\\ &\propto\mathcal{CN}\left(h_{nu};\hat{h}_{k,nu},\hat{\psi}_{k,nu}^{h}\right),\end{aligned} \tag{49}$$

and we have

$$\hat{\psi}_{k,nu}^{h}=\left(\frac{1}{\check{\psi}_{nu}^{h}}-\frac{\left|\check{s}_{uk}\right|^{2}}{\tilde{\xi}_{u,nk}^{h}}\right)^{-1}, \tag{50a}$$

$$\hat{h}_{k,nu}=\hat{\psi}_{k,nu}^{h}\left(\frac{\check{h}_{nu}}{\check{\psi}_{nu}^{h}}-\frac{\tilde{g}_{u,nk}^{h}\check{s}_{uk}^{*}}{\tilde{\xi}_{u,nk}^{h}}\right). \tag{50b}$$

Finally, similarly to (43), the domain conversion in (15c) is replaced with a simple inverse DFT. From (49) and (50), the message propagated to the FNs can be expressed as

$$\hat{\Delta}_{nk\leftarrow nu}\left(a_{nu}\right)=\mathcal{CN}\left(a_{nu};\hat{a}_{k,nu},\hat{\psi}_{k,nu}^{a}\right), \tag{51}$$

with

$$\hat{a}_{k,nu}=\mathbb{E}_{\mathsf{a}_{nu}}\left[a_{nu}\middle|\left\{\hat{h}_{k,nu},\hat{\psi}_{k,nu}^{h},\forall n\right\}\right]=\sum_{i=1}^{N}d_{i,n}\hat{h}_{k,iu}, \tag{52a}$$

$$\hat{\psi}_{k,nu}^{a}=\mathbb{V}_{\mathsf{a}_{nu}}\left[a_{nu}\middle|\left\{\hat{h}_{k,nu},\hat{\psi}_{k,nu}^{h},\forall n\right\}\right]=\frac{1}{N}\sum_{i=1}^{N}\hat{\psi}_{k,iu}^{h}. \tag{52b}$$

3) Design of Channel Sparsity-Aware Prior: In order to exploit the sparsity of the BD-domain mmWave MIMO channel shown in Fig. 1(b) of Section II-C, it is essential in (46) to choose an appropriate prior distribution $p_{\mathsf{h}_{nu}}\left(h_{nu}\right)$ that reflects the statistical property. In this paper, we employ a Laplace distribution as the prior distribution that can capture sparsity [55]. Since the Laplace distribution only supports real values, the prior distribution can be expressed as

$$p_{\mathsf{h}_{u}}\left(\boldsymbol{h}_{u};\beta_{u}^{(\tau)}\right)=\prod_{n=1}^{N}p_{\mathsf{h}_{nu}}\left(h_{nu};\beta_{u}^{(\tau)}\right), \tag{53a}$$

$$p_{\mathsf{h}_{nu}}\left(h_{nu};\beta_{u}^{(\tau)}\right)=p_{\mathsf{h}_{nu}^{\Re}}\left(h_{nu}^{\Re};\beta_{u}^{(\tau)}\right)p_{\mathsf{h}_{nu}^{\Im}}\left(h_{nu}^{\Im};\beta_{u}^{(\tau)}\right), \tag{53b}$$

where $p_{\mathsf{h}_{nu}^{i}}\left(h_{nu}^{i};\beta_{u}^{(\tau)}\right),\forall i\in\mathcal{C}$ denotes the Laplace distribution, which is expressed as

$$p_{\mathsf{h}_{nu}^{i}}\left(h_{nu}^{i};\beta_{u}^{(\tau)}\right)=\frac{1}{2\beta_{u}^{(\tau)}}\exp\left[-\frac{\left|h_{nu}^{i}\right|}{\beta_{u}^{(\tau)}}\right], \tag{54}$$

where $\beta_{u}^{(\tau)}$ denotes the scale parameter at the τ-th iteration step, which is updated at each iteration step using the expectation-maximization (EM) algorithm [55]–[57].

Next, let us rewrite (48) in a closed form based on (53). Assuming independence of the real and imaginary parts, $\Delta_{nu}\left(h_{nu}\right)$ in (46) can be rewritten as

$$\Delta_{nu}\left(h_{nu}\right)=\Delta_{nu}\left(h_{nu}^{\Re}\right)\Delta_{nu}\left(h_{nu}^{\Im}\right), \tag{55}$$

with

$$\Delta_{nu}\left(h_{nu}^{i}\right)=p_{\mathsf{h}_{nu}^{i}}\left(h_{nu}^{i};\beta_{u}^{(\tau)}\right)\mathcal{N}\left(h_{nu}^{i};\mu_{nu}^{i},\psi_{nu}^{i}\right),\forall i\in\mathcal{C}, \tag{56}$$

where $\mu_{nu}^{\Re}\triangleq\Re\{\mu_{nu}^{h}\}$, $\mu_{nu}^{\Im}\triangleq\Im\{\mu_{nu}^{h}\}$, and $\psi_{nu}^{\Re}=\psi_{nu}^{\Im}=\psi_{nu}^{h}/2$, respectively.

From (54) and (56), the posterior estimate and its MSE can be obtained in a closed form by (48), as [55]

$$\begin{aligned}\check{h}_{nu} &= \mathbb{E}_{\mathsf{h}_{nu}}\left[h_{nu}\mid\mu_{nu}^{h},\psi_{nu}^{h};\beta_{u}^{(\tau)}\right]\\ &= \zeta_{1}\left(\mu_{nu}^{\Re},\psi_{nu}^{\Re},\beta_{u}^{(\tau)}\right)+\mathrm{j}\cdot\zeta_{1}\left(\mu_{nu}^{\Im},\psi_{nu}^{\Im},\beta_{u}^{(\tau)}\right),\end{aligned} \tag{57a}$$

$$\begin{aligned}\check{\psi}_{nu}^{h} &= \mathbb{V}_{\mathsf{h}_{nu}}\left[h_{nu}\mid\mu_{nu}^{h},\psi_{nu}^{h};\beta_{u}^{(\tau)}\right]\\ &= \epsilon_{1}\left(\mu_{nu}^{\Re},\psi_{nu}^{\Re},\beta_{u}^{(\tau)}\right)+\epsilon_{1}\left(\mu_{nu}^{\Im},\psi_{nu}^{\Im},\beta_{u}^{(\tau)}\right)-\left|\check{h}_{nu}\right|^{2},\end{aligned} \tag{57b}$$

where

$$\zeta_{1}\left(\mu,\psi,\beta\right)\triangleq\frac{\eta^{+}\gamma^{+}+\eta^{-}\gamma^{-}}{\eta^{+}+\eta^{-}}, \tag{58a}$$

$$\begin{aligned}\epsilon_{1}\left(\mu,\psi,\beta\right)\triangleq{}&\frac{\eta^{+}\left[\psi+\left(\gamma^{+}\right)^{2}\right]+\eta^{-}\left[\psi+\left(\gamma^{-}\right)^{2}\right]}{\eta^{+}+\eta^{-}}\\ &-\frac{2\ \beta^{-1}\psi^{2}\cdot\mathcal{N}\left(0;\mu,\psi\right)}{\eta^{+}+\eta^{-}},\end{aligned} \tag{58b}$$

with

$$\alpha^{+}\triangleq\frac{\mu}{\beta}-\frac{\psi}{2\beta^{2}},\ \alpha^{-}\triangleq-\frac{\mu}{\beta}-\frac{\psi}{2\beta^{2}}, \tag{59a}$$

$$\gamma^{+}\triangleq\mu-\frac{\psi}{\beta},\ \gamma^{-}\triangleq\mu+\frac{\psi}{\beta}, \tag{59b}$$

$$\eta^{+}\triangleq\exp\left[-\alpha^{+}\right]\mathit{\Omega}\left(-\frac{\gamma^{+}}{\sqrt{\psi}}\right),\ \eta^{-}\triangleq\exp\left[-\alpha^{-}\right]\mathit{\Omega}\left(\frac{\gamma^{-}}{\sqrt{\psi}}\right). \tag{59c}$$

Finally, we derive the update rule for $\beta_{u}^{(\tau)}$ via EM algorithm. The EM algorithm allows updating $\beta_{u}^{(\tau)}$ for every iteration step so as to maximize likelihood function $p_{\mathsf{R}|\beta_{u}}\left(\boldsymbol{R}|\beta_{u}^{(\tau)}\right)$, which is achieved by alternating between an *E-step*, which

This article has been accepted for publication in IEEE Journal of Selected Topics in Signal Processing. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JSTSP.2025.3589107

Algorithm 1 - Proposed Cross-Domain JCDE Algorithm

```
Input: $\boldsymbol{R}$, $\{x_{mk}, \forall(m, k \in \mathcal{K}_{\mathrm{p}})\}$, $\beta_u^{(1)}$, $\{\tau_{1,\max}, \tau_{2,\max}\}$
Output: $\{\check{a}_{nu}, \forall(n, u)\}$, $\{\check{x}_{mk}, \forall(m, k \in \mathcal{K}_{\mathrm{d}})\}$
    /* ——— Initialization for Pilot Part ——— */
 1: $\forall(k \in \mathcal{K}_{\mathrm{p}}, n, u, \tau) : \hat{s}_{n,uk}^{(\tau)} = s_{uk},\ \psi_{n,uk}^{s,(\tau)} = 0$
 2: $\forall(k \in \mathcal{K}_{\mathrm{p}}, n, u) : \hat{a}_{k,nu}^{(1)} = 0,\ \psi_{k,nu}^{a,(1)} = \xi^a/P$
    /* ——— Phase I: CE Phase ——— */
 3: for $\tau = 1$ to $\tau_{1,\max}$ do
 4:     $\forall k \in \mathcal{K}_{\mathrm{p}}$ : Executing Algorithm 2
 5: end for
    /* ——— Initialization for Data Part ——— */
 6: $\forall(k \in \mathcal{K}_{\mathrm{d}}, m, u) : \hat{s}_{n,uk}^{(1)} = 0,\ \psi_{n,uk}^{s,(1)} = E_{\mathrm{s}}$
 7: $\forall(k \in \mathcal{K}_{\mathrm{d}}, n, u) : \hat{a}_{k,nu}^{(1)} = \check{a}_{nu},\ \psi_{k,nu}^{a,(1)} = \check{\psi}_{nu}^{a}$
    /* ——— Phase II: JCDE Phase ——— */
 8: for $\tau = 1$ to $\tau_{2,\max}$ do
 9:     $\forall k \in \mathcal{K}$ : Executing Algorithm 2
10:     $\forall k \in \mathcal{K}_{\mathrm{d}}$ : Executing Algorithm 3
11: end for
    // Termination and then output estimates
```

Algorithm 2 - Channel Estimation Mechanism: $\forall(n, u)$

```
Input: $\boldsymbol{R}$, $\{\hat{a}_{k,nu}^{(\tau)}, \hat{\psi}_{k,nu}^{a,(\tau)}\}$, $\{\hat{s}_{n,uk}^{(\tau)}, \hat{\psi}_{n,uk}^{s,(\tau)}\}$, $\{\check{s}_{uk}^{(\tau)}\}$, $\beta_u^{(\tau)}$
Output: $\{\hat{a}_{k,nu}^{(\tau+1)}, \hat{\psi}_{k,nu}^{a,(\tau+1)}\}$, $\{\check{a}_{nu}^{(\tau+1)}\}$, $\beta_u^{(\tau+1)}$
    /* ——— FN: AD-domain DQ ——— */
 1: $\bar{z}_{nk}^{(\tau)} = \sum_{u=1}^{U} \hat{a}_{k,nu}^{(\tau)} \hat{s}_{n,uk}^{(\tau)}$
 2: $\bar{\xi}_{nk}^{z,(\tau)} = \sum_{u=1}^{U} \Big\{ \big|\hat{a}_{k,nu}^{(\tau)}\big|^2 \hat{\psi}_{n,uk}^{s,(\tau)}$
        $+ \hat{\psi}_{k,nu}^{a,(\tau)} \Big( \big|\hat{s}_{n,uk}^{(\tau)}\big|^2 + \hat{\psi}_{n,uk}^{s,(\tau)} \Big) \Big\}$
 3: $\bar{\xi}_{nk}^{y,(\tau)} = \bar{\xi}_{nk}^{z,(\tau)} + N_0$
 4: $\check{y}_{nk}^{(\tau)} = \mathbb{E}_{\mathsf{y}_{nk}} \left[ y_{nk} \mid \bar{z}_{nk}^{(\tau)}, \bar{\xi}_{nk}^{y,(\tau)}; \mathcal{L}_{nk} \right]$
 5: $\check{\xi}_{nk}^{y,(\tau)} = \mathbb{V}_{\mathsf{y}_{nk}} \left[ y_{nk} \mid \bar{z}_{nk}^{(\tau)}, \bar{\xi}_{nk}^{y,(\tau)}; \mathcal{L}_{nk} \right]$
 6: $\bar{y}_{nk}^{(\tau)} = \check{y}_{nk}^{(\tau)} - \bar{z}_{nk}^{(\tau)}$
 7: $\tilde{g}_{u,nk}^{a,(\tau)} = \left( 1 - \frac{\check{\xi}_{nk}^{y,(\tau)}}{\bar{\xi}_{nk}^{y,(\tau)}} \right)^{-1} \bar{y}_{nk}^{(\tau)} + \hat{a}_{k,nu}^{(\tau)} \hat{s}_{n,uk}^{(\tau)}$
 8: $\tilde{\xi}_{u,nk}^{a,(\tau)} = \left( 1 - \frac{\check{\xi}_{nk}^{y,(\tau)}}{\bar{\xi}_{nk}^{y,(\tau)}} \right)^{-1} \bar{\xi}_{nk}^{y,(\tau)} - \hat{\psi}_{k,nu}^{a,(\tau)} \big|\hat{s}_{n,uk}^{(\tau)}\big|^2$
 9: $\tilde{g}_{u,nk}^{h,(\tau)} = \sum_{i=1}^{N} d_{n,i}^{*} \tilde{g}_{u,ik}^{a,(\tau)},\ \tilde{\xi}_{u,nk}^{h,(\tau)} = \frac{1}{N} \sum_{i=1}^{N} \tilde{\xi}_{u,ik}^{a,(\tau)}$
    /* ——— VN: BD-domain CE ——— */
10: $\nu_{k,nu}^{h,(\tau)} = \eta \cdot \frac{\big|\check{s}_{uk}^{(\tau)}\big|^2}{\tilde{\xi}_{u,nk}^{h,(\tau)}} + (1-\eta) \cdot \nu_{k,nu}^{h,(\tau-1)}$
11: $\gamma_{k,nu}^{h,(\tau)} = \eta \cdot \frac{\tilde{g}_{u,nk}^{h,(\tau)} \left(\check{s}_{uk}^{(\tau)}\right)^{*}}{\tilde{\xi}_{u,nk}^{h,(\tau)}} + (1-\eta) \cdot \gamma_{k,nu}^{h,(\tau-1)}$
12: $\psi_{nu}^{h,(\tau)} = \left( \sum_{k=1}^{K} \nu_{k,nu}^{h,(\tau)} \right)^{-1}$
13: $\mu_{nu}^{h,(\tau)} = \psi_{nu}^{h,(\tau)} \sum_{k=1}^{K} \gamma_{k,nu}^{h,(\tau)}$
14: $\check{h}_{nu}^{(\tau)} = \mathbb{E}_{\mathsf{h}_{nu}} \left[ h_{nu} \mid \mu_{nu}^{h,(\tau)}, \psi_{nu}^{h,(\tau)}; \beta_u^{(\tau)} \right]$
15: $\check{\psi}_{nu}^{h,(\tau)} = \mathbb{V}_{\mathsf{h}_{nu}} \left[ h_{nu} \mid \mu_{nu}^{h,(\tau)}, \psi_{nu}^{h,(\tau)}; \beta_u^{(\tau)} \right]$
16: $\hat{\psi}_{k,nu}^{h,(\tau)} = \left( \frac{1}{\check{\psi}_{nu}^{h,(\tau)}} - \nu_{k,nu}^{h,(\tau)} \right)^{-1}$
17: $\hat{h}_{k,nu}^{(\tau)} = \hat{\psi}_{k,nu}^{h,(\tau)} \left( \frac{\check{h}_{nu}^{(\tau)}}{\check{\psi}_{nu}^{h,(\tau)}} - \gamma_{k,nu}^{h,(\tau)} \right)$
18: $\check{a}_{nu}^{(\tau+1)} = \sum_{i=1}^{N} d_{i,n} \check{h}_{iu}^{(\tau)}$
19: $\hat{a}_{k,nu}^{(\tau+1)} = \sum_{i=1}^{N} d_{i,n} \hat{h}_{k,iu}^{(\tau)},\ \hat{\psi}_{k,nu}^{a,(\tau+1)} = \frac{1}{N} \sum_{i=1}^{N} \hat{\psi}_{k,iu}^{h,(\tau)}$
20: $\forall u$ : Update $\beta_u^{(\tau)}$ by using EM algorithm.
```

updates the posterior information for h_{nq} to minimize the KLD between the likelihood function and its lower bound, and an *M-step*, which estimates the distribution parameters that maximize the lower bound [56]. Since the posterior estimates can be computed in (56), only the *M-step* solving the following maximization problem needs to be considered: [57]

$$\beta_u^{(\tau+1)} = \underset{\beta_u}{\operatorname{argmax}} \underbrace{\sum_{n=1}^{N} \mathbb{E}_{\mathsf{h}_{nu}} \left[\ln p_{\mathsf{h}_{nu}} \left(h_{nu}; \beta_u \right) \middle| \mu_{nu}^{h}, \psi_{nu}^{h}; \beta_u^{(\tau)} \right]}_{\triangleq J(\beta_u)}, \tag{60}$$

where β_u is the true distribution parameter. The update rule can be obtained by solving for $\frac{\partial}{\partial \beta_u} J(\beta_u) = 0$, as

$$\beta_u^{(\tau+1)} = \frac{1}{2N} \sum_{n=1}^{N} \left[\nu_1 \left(\mu_{nu}^{\Re}, \psi_{nu}^{\Re}, \beta_u^{(\tau)} \right) + \nu_1 \left(\mu_{nu}^{\Im}, \psi_{nu}^{\Im}, \beta_u^{(\tau)} \right) \right], \tag{61}$$

with

$$\nu_1 \left(\mu, \psi, \beta \right) = \frac{\eta^{+} \gamma^{+} + \eta^{-} \gamma^{-} + 2\ \psi \cdot \mathcal{N}\left(0; \mu, \psi\right)}{\eta^{+} + \eta^{-}}. \tag{62}$$

C. Derivation of EP-based MPA for DD via SGA

The above discussion on CE can be directly applied to the detection of an arbitrary $x_{mk}, k \in \mathcal{K}_{\mathrm{d}}$. Based on (16) and (17), the derivation of MPA for DD can be performed in a manner similar to the CE described in Section III-B, such that further details are omitted and offered only in a summarized form in the pseudo-code (Algorithm 3) shown later.

D. Algorithm Description

For the sake of completeness, the pseudo-code of the proposed algorithm is given in Algorithms 1–3. Algorithm 1 shows the overall picture of the proposed cross-domain JCDE algorithm, where Algorithm 2 shows the CE mechanism presented in Section III-B and Algorithm 3 shows the DD mechanism presented in Section III-C.

As can be seen in Algorithm 1, the proposed algorithm has two phases. In **Phase I**, only the CE process in Algorithm 2 is iteratively executed using the knowledge of the quantized observations $r_{nk}, \forall(n, k \in \mathcal{K}_{\mathrm{p}})$ and the *pilot* symbols $s_{uk}, \forall(u, k \in \mathcal{K}_{\mathrm{p}})$. This is because when JCDE is run from the beginning, the initial CE accuracy is poor and hence the iterative convergence behavior becomes unstable. After the CE accuracy is improved in **Phase I**, JCDE is executed in **Phase II** by alternating between Algorithms 2 and 3. The accuracy of the CE and DD can be further improved by exploiting the data symbol estimates as the equivalent soft pilots.

A well-known belief damping [39] is introduced in lines 10 and 11 of Algorithm 2 and lines 11 and 12 of Algorithm 3, summarized as follows. Let a quantity z be calculated by a

Algorithm 3 - Data Detection Mechanism: $\forall(n,m)$

```
Input: $\boldsymbol{R}$, $\{\hat{s}^{(\tau)}_{n,uk}, \hat{\psi}^{s,(\tau)}_{n,uk}\}$, $\{\hat{a}^{(\tau+1)}_{k,nu}, \hat{\psi}^{a,(\tau+1)}_{k,nu}\}$, $\{\check{a}^{(\tau+1)}_{nu}\}$
Output: $\{\hat{s}^{(\tau+1)}_{n,uk}, \hat{\psi}^{s,(\tau)}_{n,uk}\}$, $\{\check{s}^{(\tau+1)}_{uk}\}$
    /* ———— FN: AD-domain DQ ———— */
 1: $\overline{z}'^{(\tau)}_{nk} = \sum_{u=1}^{U} \hat{a}^{(\tau+1)}_{k,nu} \hat{s}^{(\tau)}_{n,uk}$
 2: $\overline{\xi}'^{z,(\tau)}_{nk} = \sum_{u=1}^{U} \Big\{ \big|\hat{a}^{(\tau+1)}_{k,nu}\big|^2 \hat{\psi}^{s,(\tau)}_{n,uk} + \hat{\psi}^{a,(\tau+1)}_{k,nu} \Big( \big|\hat{s}^{(\tau)}_{n,uk}\big|^2 + \hat{\psi}^{s,(\tau)}_{n,uk} \Big) \Big\}$
 3: $\overline{\xi}'^{y,(\tau)}_{nk} = \overline{\xi}'^{z,(\tau)}_{nk} + N_0$
 4: $\check{y}'^{(\tau)}_{nk} = \mathbb{E}_{\mathsf{y}_{nk}} \Big[ y_{nk} \mid \overline{z}'^{(\tau)}_{nk}, \overline{\xi}'^{y,(\tau)}_{nk}; \mathcal{L}_{nk} \Big]$
 5: $\check{\xi}'^{y,(\tau)}_{nk} = \mathbb{V}_{\mathsf{y}_{nk}} \Big[ y_{nk} \mid \overline{z}'^{(\tau)}_{nk}, \overline{\xi}'^{y,(\tau)}_{nk}; \mathcal{L}_{nk} \Big]$
 6: $\overline{y}'^{(\tau)}_{nk} = \check{y}'^{(\tau)}_{nk} - \overline{z}'^{(\tau)}_{nk}$
 7: $\tilde{g}^{s,(\tau)}_{m,nk} = \Big( 1 - \frac{\check{\xi}'^{y,(\tau)}_{nk}}{\overline{\xi}'^{y,(\tau)}_{nk}} \Big)^{-1} \overline{y}'^{(\tau)}_{nk} + \sum_{u \in \mathcal{U}_m} \hat{a}^{(\tau+1)}_{k,nu} \hat{s}^{(\tau)}_{n,uk}$
 8: $\tilde{\xi}^{s,(\tau)}_{m,nk} = \Big( 1 - \frac{\check{\xi}'^{y,(\tau)}_{nk}}{\overline{\xi}'^{y,(\tau)}_{nk}} \Big)^{-1} \overline{\xi}'^{y,(\tau)}_{nk} - \sum_{u \in \mathcal{U}_m} \big|\hat{a}^{(\tau+1)}_{k,nu}\big|^2 \hat{\psi}^{s,(\tau)}_{n,uk}$
 9: $\tilde{g}^{x,(\tau)}_{m,nk} = \sum_{i \in \mathcal{K}_t} f_{\kappa(k),\kappa(i)} \tilde{g}^{s,(\tau)}_{m,ni}$, $\tilde{\xi}^{x,(\tau)}_{m,nk} = \frac{1}{N} \sum_{i \in \mathcal{K}_t} \tilde{\xi}^{s,(\tau)}_{m,ni}$
10: $\check{c}^{(\tau+1)}_{k,nm} = \sqrt{K_\mathrm{c}} \sum_{u \in \mathcal{U}_m} f_{\kappa(k),\upsilon(u)} \check{a}^{(\tau+1)}_{nu}$
    /* ———— VN: AF-domain DD ———— */
11: $\nu^{x,(\tau)}_{n,mk} = \eta \cdot \frac{\big|\check{c}^{(\tau+1)}_{k,nm}\big|^2}{\tilde{\xi}^{x,(\tau)}_{m,nk}} + (1-\eta) \cdot \nu^{x,(\tau-1)}_{n,mk}$
12: $\gamma^{x,(\tau)}_{n,mk} = \eta \cdot \frac{\big(\check{c}^{(\tau+1)}_{k,nm}\big)^* \tilde{g}^{x,(\tau)}_{m,nk}}{\tilde{\xi}^{x,(\tau)}_{m,nk}} + (1-\eta) \cdot \gamma^{x,(\tau-1)}_{n,mk}$
13: $\psi^{x,(\tau)}_{mk} = \Big( \sum_{n=1}^{N} \nu^{x,(\tau)}_{n,mk} \Big)^{-1}$
14: $\mu^{x,(\tau)}_{mk} = \psi^{x,(\tau)}_{mk} \sum_{n=1}^{N} \gamma^{x,(\tau)}_{n,mk}$
15: $\check{x}^{(\tau)}_{mk} = \mathbb{E}_{\mathsf{x}_{mk}} \Big[ x_{mk} \mid \mu^{x,(\tau)}_{mk}, \psi^{x,(\tau)}_{mk} \Big]$
16: $\check{\psi}^{x,(\tau)}_{mk} = \mathbb{V}_{\mathsf{x}_{mk}} \Big[ x_{mk} \mid \mu^{x,(\tau)}_{mk}, \psi^{x,(\tau)}_{mk} \Big]$
17: $\hat{\psi}^{x,(\tau)}_{n,mk} = \Big( \frac{1}{\check{\psi}^{x,(\tau)}_{mk}} - \nu^{x,(\tau)}_{n,mk} \Big)^{-1}$
18: $\hat{x}^{(\tau)}_{n,mk} = \hat{\psi}^{x,(\tau)}_{n,mk} \Big( \frac{\check{x}^{(\tau)}_{mk}}{\check{\psi}^{x,(\tau)}_{mk}} - \gamma^{x,(\tau)}_{n,mk} \Big)$
19: $\forall (n,(u,k) \in \mathcal{I}^s_{mk}) : \check{s}^{(\tau+1)}_{uk} = \sum_{i \in \mathcal{K}_t} f^*_{\kappa(i),\kappa(k)} \check{x}^{(\tau+1)}_{mi}$
20: $\forall (n,(u,k) \in \mathcal{I}^s_{mk}) : \hat{s}^{(\tau+1)}_{n,uk} = \sum_{i \in \mathcal{K}_t} f^*_{\kappa(i),\kappa(k)} \hat{x}^{(\tau+1)}_{n,mi}$
21: $\forall (n,(u,k) \in \mathcal{I}^s_{mk}) : \hat{\psi}^{s,(\tau+1)}_{n,uk} = \frac{1}{K_\mathrm{c}} \sum_{i \in \mathcal{K}_t} \hat{\psi}^{x,(\tau+1)}_{n,mi}$
```

function f_z. Then the τ-th damped value of z, here denoted $z^{(\tau)}$, is computed as the weighted average of $z^{(\tau-1)}$ and f_z, with weights set by the damping factor $\eta \in [0,1]$, *i.e.*,

$$z = f_z \quad \xrightarrow{\text{damping}} \quad z^{(\tau)} = \eta f_z + (1-\eta)\, z^{(\tau-1)}. \tag{63}$$

In lines 4 and 5 of Algorithms 2 and 3, the conditional expectation of y_{nk} and its MSE can be computed by using the truncated Gaussian distribution. Since the truncated Gaussian distribution only supports real values, in a similar manner to (57), they can be expressed as [20], [21]

$$\begin{aligned}
\check{y}_{nk} &= \mathbb{E}_{\mathsf{y}_{nk}} \Big[y_{nk} \mid \overline{z}_{nk}, \overline{\xi}^y_{nk}; \mathcal{L}_{nk} \Big] \\
&= \zeta_2 \left(\overline{z}^{\Re}_{nk}, \overline{\sigma}^{\Re}_{nk}, \lambda_i, \lambda_{i+1} \right) + \mathrm{j} \cdot \zeta_2 \left(\overline{z}^{\Im}_{nk}, \overline{\sigma}^{\Im}_{nk}, \lambda_j, \lambda_{j+1} \right), \quad \text{(64a)} \\
\check{\xi}^y_{nk} &= \mathbb{V}_{\mathsf{y}_{nk}} \Big[y_{nk} \mid \overline{z}_{nk}, \overline{\xi}^y_{nk}; \mathcal{L}_{nk} \Big] \\
&= \epsilon_2 \left(\overline{z}^{\Re}_{nk}, \overline{\sigma}^{\Re}_{nk}, \lambda_i, \lambda_{i+1} \right) + \epsilon_2 \left(\overline{z}^{\Im}_{nk}, \overline{\sigma}^{\Im}_{nk}, \lambda_j, \lambda_{j+1} \right), \quad \text{(64b)}
\end{aligned}$$

where $\overline{\sigma}^i_{nk} = \sqrt{\overline{\xi}^y_{nk}/2}, \forall i \in \mathcal{C}$, and

$$\zeta_2 \left(\mu, \sigma, a, b \right) = \mu + \sigma \frac{\phi(\alpha) - \phi(\beta)}{\Phi(\beta) - \Phi(\alpha)}, \tag{65a}$$

$$\begin{aligned}
&\epsilon_2 \left(\mu, \sigma, a, b \right) \\
&= \sigma^2 \left[1 + \frac{\alpha\phi(\alpha) - \beta\phi(\beta)}{\Phi(\beta) - \Phi(\alpha)} - \left(\frac{\phi(\alpha) - \phi(\beta)}{\Phi(\beta) - \Phi(\alpha)} \right)^2 \right], \quad \text{(65b)}
\end{aligned}$$

with

$$\alpha = \frac{a - \mu}{\sigma}, \ \beta = \frac{b - \mu}{\sigma}. \tag{66}$$

In lines 7 and 8 of Algorithms 2 and 3, a lower limit is introduced to reduce the numerical instability from calculating the expectation value of the truncated Gaussian distribution as follows: $\max\left(1 - \check{\xi}^y_{nk}/\overline{\xi}^y_{nk}, 10^{-2}\right)$.

In lines 15 and 16 of Algorithm 3, the conditional expectation of the data symbol x_{mk}, given μ^x_{mk} and ψ^x_{mk}, and its MSE are computed. When using quadrature amplitude modulation (QAM) signaling (*i.e.*, $p_{\mathsf{x}_{mk}}(\chi_j) = 1/J, \ \forall \chi_j \in \mathcal{X}$), the soft estimate $\check{x}_{mk}$ and its MSE $\check{\psi}^x_{mk}$ can be expressed as

$$\begin{aligned}
\check{x}_{mk} &= \mathbb{E}_{\mathsf{x}_{mk}} \left[x_{mk} \mid \mu^x_{mk}, \psi^x_{mk} \right] \\
&= \sum_{\chi_j \in \mathcal{X}} \chi_j \frac{\mathcal{CN}\left(\chi_j; \mu^x_{mk}, \psi^x_{mk}\right)}{\sum_{\chi'_j \in \mathcal{X}} \mathcal{CN}\left(\chi'_j; \mu^x_{mk}, \psi^x_{mk}\right)}, \quad \text{(67a)} \\
\check{\psi}^x_{mk} &= \mathbb{V}_{\mathsf{x}_{mk}} \left[x_{mk} \mid \mu^x_{mk}, \psi^x_{mk} \right] \\
&= \sum_{\chi_j \in \mathcal{X}} |\chi_j|^2 \frac{\mathcal{CN}\left(\chi_j; \mu^x_{mk}, \psi^x_{mk}\right)}{\sum_{\chi'_j \in \mathcal{X}} \mathcal{CN}\left(\chi'_j; \mu^x_{mk}, \psi^x_{mk}\right)} - |\check{x}_{mk}|^2. \quad \text{(67b)}
\end{aligned}$$

In lines 9, 10, 19, and 20 of Algorithm 3, we define the index $\kappa(i) \triangleq (i - 1 \mod K_\mathrm{c}) + 1$.

IV. PERFORMANCE ASSESSMENT

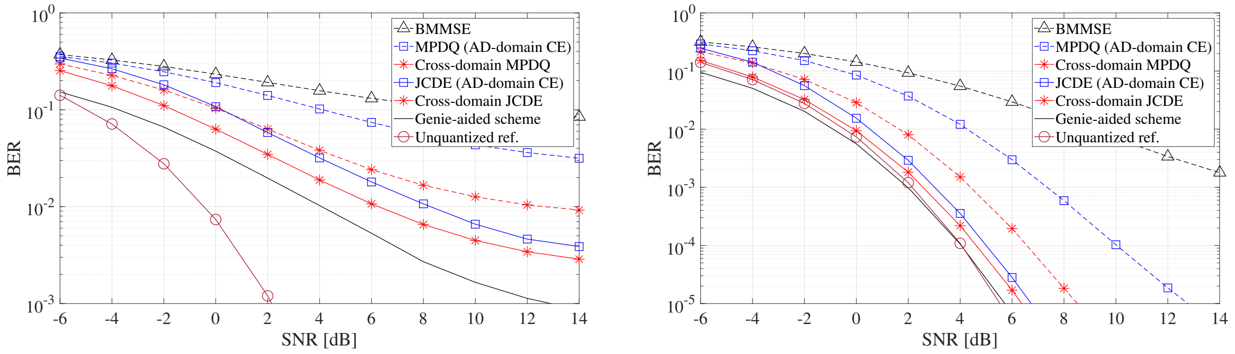
Numerical studies were conducted to demonstrate the performance of the proposed JCDE algorithm for uplink multi-user MIMO-OFDM systems with low-resolution ADCs. The MIMO configuration was set to $(N, M) = (128, 16)$. The OFDM format was set to $(K_\mathrm{c}, T) = (64, 6)$. The resource index sets assigned to the pilot symbol and data symbol were set to $\mathcal{T}_\mathrm{p} = \{1\}$ and $\mathcal{T}_\mathrm{d} = \{2, 3, \cdots, 6\}$, respectively. The Gray-coded 4- and 16QAM were employed for symbol mapping. The random pilots $x_{mk} = \sqrt{E_\mathrm{s}} \exp\left[\mathrm{j}\pi\upsilon_{mk}\right]$ were used, where υ_{mk} is chosen uniformly at random from the range $[-1, 1]$. In addition, the optimal quantization labels are determined via the Lloyd-Max algorithm [58].

The well-known clustered mmWave channel model with L clusters, each having C_l rays with $l \in \{1, 2, \cdots, L\}$ is considered [14], [31], [55]. The corresponding channel vector can then be expressed as

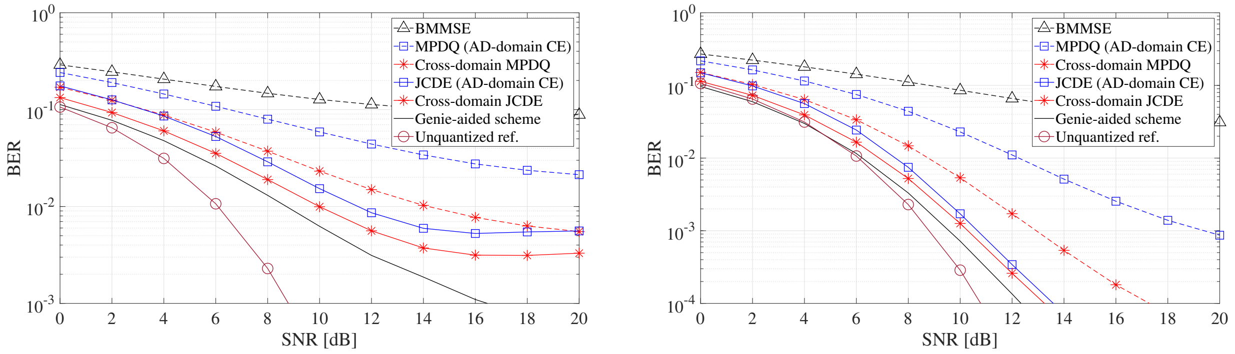
$$\boldsymbol{a}_{p,m} = \frac{1}{\sqrt{L}} \sum_{l=1}^{L} \frac{\delta\left(p - p_{l,m}\right)}{\sqrt{C_l}} \sum_{c=1}^{C_l} \zeta_{l,c,m} \cdot \boldsymbol{\alpha}\left(\Gamma_{l,c,m}\right), \tag{68}$$

where $p_{l,m}$ is the delay corresponding to the l-th cluster of the m-th UE, and $\zeta_{l,c,m} \sim \mathcal{CN}(\zeta_{l,c,m}; 0, \xi^a_{l,m})$ is the complex path gain along the (l,c)-th ray of the m-th UE. Assuming that the receiver is equipped with a uniform linear array (ULA) in

This article has been accepted for publication in IEEE Journal of Selected Topics in Signal Processing. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JSTSP.2025.3589107



(a) Gray-coded 4QAM, $b = 1$. (b) Gray-coded 4QAM, $b = 3$.



(c) Gray-coded 16QAM, $b = 2$. (d) Gray-coded 16QAM, $b = 3$.

Fig. 3. BER performance of uncoded MIMO-OFDM systems, where $(N, M, P, K_{\mathrm{p}}, K_{\mathrm{d}}) = (128, 16, 4, 64, 320)$.

which the antenna elements are arranged at half-wavelength intervals, the array response vector in (68) is given by

$$\boldsymbol{\alpha}\left(\mathit{\Gamma}_{l,c,m}\right) \triangleq \left[1, \exp\left[\mathrm{j}\pi \mathit{\Gamma}_{l,c,m}\right], \ldots, \exp\left[\mathrm{j}\pi\left(N-1\right)\mathit{\Gamma}_{l,c,m}\right]\right]^{\mathrm{T}}, \tag{69}$$

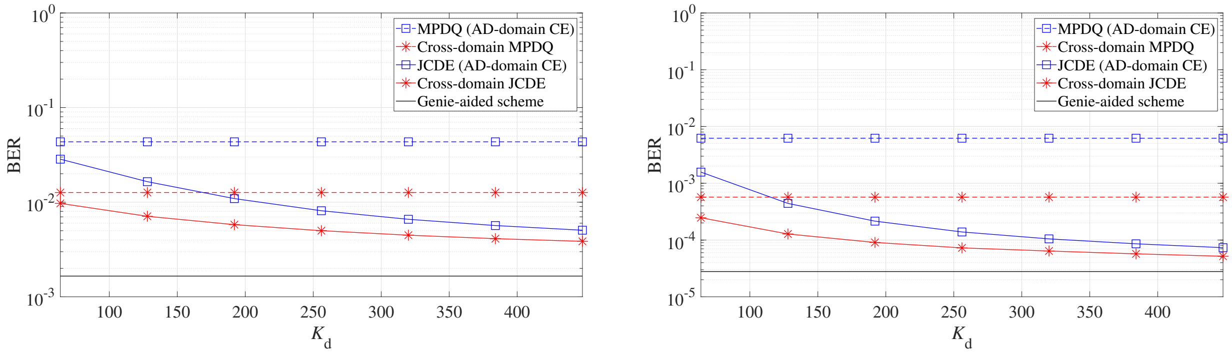
where $\mathit{\Gamma}_{l,c,m} \triangleq \sin\theta_{l,c,m}$ with $\theta_{l,c,m}$ denoting the angle of arrival (AoA) corresponding to the (l, c)-th ray of the m-th UE. The mmWave wireless communication channel consists of $L = 4$ clusters with $C_l = 15$ rays-per-cluster. A sector antenna of 120 degrees opening was considered, and the UE devices are randomly dropped in the above angular region around the receiver. The maximum number of delay taps was set to $P = 4$, and all delay paths shall fall within the CP length. The realization of $p_{l,m}$ was randomly selected from $\{1, 2, \cdots, P\}$ and $\xi^a_{l,m} = \xi^a/L$. The initial value $\beta_u^{(1)}$ was set to $\sqrt{\xi^a/(4P)}$ so that the variance of the Laplace distribution coincides with $\xi^a/(2P)$. The damping factor η was set to 1.0 at $\tau = 1$ and 0.5 otherwise, and the maximum number of iterations was set constant to $(\tau_{1,\max}, \tau_{2,\max}) = (16, 16)$.

A. BER Performance in Uncoded Case

Our first set of results is given in Fig. 3, where the BER performance as a function of the signal-to-noise ratio (SNR) of the MIMO-OFDM systems are compared:

- *BMMSE:* SotA linear receiver based on BMMSE [16], where CE and DD are performed separately in the AD and AF domains, respectively. This method cannot fully exploit the channel sparsity.
- *MPDQ (AD-domain CE):* MPDQ receiver based on a SotA Bayesian inference algorithm under nonlinear observations [44], where CE is performed in the AD domain. Used as a baseline to evaluate the performance gain achieved by exploiting channel sparsity in the beam domain through a cross-domain strategy.
- *Cross-domain MPDQ:* Proposed MPDQ receiver, where DD is performed once using Algorithm 3 instead of **Phase II**, after **Phase I**. This method serves as a baseline for assessing the performance gain achieved by the JCDE mechanism, which exploits detected data symbols as equivalent soft pilots.
- *JCDE (AD-domain CE):* JCDE receiver derived by extending "MPDQ (AD-domain CE)," where CE is performed in the AD domain using a Bernoulli-Gaussian (BG) prior [27], [39]. The SotA alternatives proposed in [20] and [21] are expected to achieve performance comparable to that of this method when applied to MIMO-OFDM systems. This method considers only the sparsity in the delay domain and serves as a baseline for evaluating the performance gain achieved by additionally exploiting channel sparsity in the beam domain.
- *Cross-domain JCDE:* Proposed JCDE receiver presented in Algorithms 1–3.
- *Genie-aided scheme:* Idealized scheme with perfect CSI at the receiver. It provides an absolute lower bound on the performance achievable by the proposed method.
- *Unquantized ref.:* Reference performance of the cross-domain JCDE algorithm under the unquantized signal model in (5), serving as a baseline for evaluating information loss due to low-resolution ADCs. The relative performance against the Genie-aided scheme is determined by whether CE error or quantization error dominates.

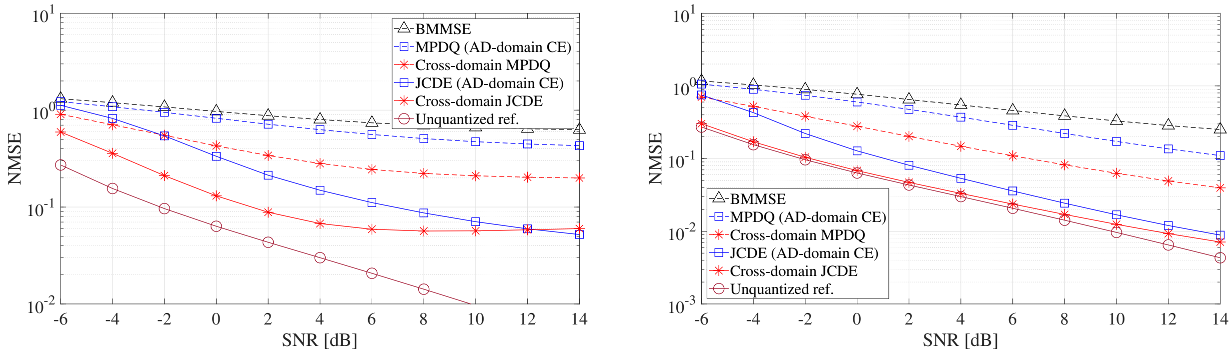
This article has been accepted for publication in IEEE Journal of Selected Topics in Signal Processing. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JSTSP.2025.3589107



(a) Gray-coded 4QAM, $b = 1$, SNR $= 10$ dB.

(b) Gray-coded 4QAM, $b = 3$, SNR $= 5$ dB.

Fig. 4. BER performances of uncoded MIMO-OFDM systems with Gray-coded 4QAM versus the data symbol length K_d.



(a) Gray-coded 4QAM, $b = 1$.

(b) Gray-coded 4QAM, $b = 3$.

Fig. 5. NMSE performance of MIMO-OFDM systems, where $(N, M, P, K_\mathrm{p}, K_\mathrm{d}) = (128, 16, 4, 64, 320)$.

Let us take a look at the methods based on CE that use only pilots. "BMMSE," which cannot exploit the channel sparsity, is unable to sufficiently suppress interference caused by quantization errors and suffers from a high-level error floor in all setups. Although "MPDQ (AD-domain CE)" improves performance by mitigating quantization distortion via the DQ process and iterative estimation, its accuracy remains limited. In contrast, "Cross-domain MPDQ," which corresponds to Phase I of the proposed method, can significantly improve estimation performance by exploiting the sparseness of the BD-domain channel through signal reconstruction via the DQ process and domain conversion. Next, focusing on the JCDE methods, we can confirm the effectiveness of using data symbols in CE due to the significant performance improvement. In addition, the performance difference between "Cross-domain JCDE" and "JCDE (AD-domain CE)" demonstrates the gain obtained by performing CE in the BD domain, where the sparsity of the mmWave wireless channel is maximized. This highlights the importance of domain conversion, which enables each process to be performed within the designated signal space. More specifically, as shown in Fig. 3(a), the proposed method achieves BER $= 10^{-2}$ even with $b = 1$ for 4QAM. Similarly, as shown in Fig. 3(c), it achieves BER $= 10^{-2}$ even with 16QAM for $b = 2$. Notably, in Figs. 3(b) and 3(d) for $b = 3$, the proposed method asymptotically approaches the Genie-aided reference, with performance degradation within 1 dB at BER $= 10^{-3}$. Furthermore, comparison with "Unquantized ref." confirms that, in the given MIMO configuration, the proposed method with $b = 3$ can almost compensate for the information loss caused by low-resolution ADCs.

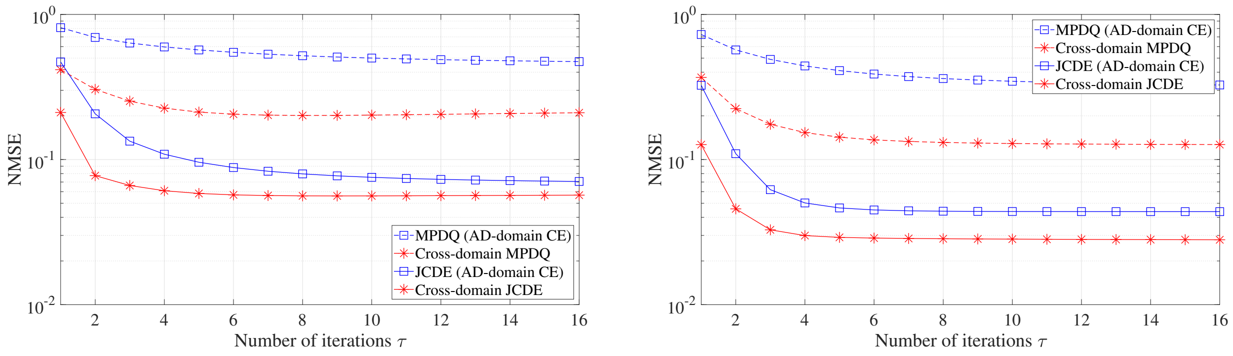
B. Effect of Data Symbol Length on BER Performance

The results in Fig. 3 clearly demonstrate the effectiveness of JCDE in compensating for the limited resolution of ADCs by exploiting detected data symbols as additional pilots. To further investigate the performance improvement enabled by the JCDE mechanism, Fig. 4 presents the BER as a function of the data length K_d used in JCDE. The SNR was set to 10 dB and 5 dB, and the number of quantization bits was set to $b = 1$ and $b = 3$, respectively. The Gray-coded 4QAM was used, and the other simulation parameters were the same as those used in Fig. 3.

The results clearly show that increasing K_d significantly reduces the BER and brings the performance closer to that of the Genie-aided scheme. Another key observation in JCDE is that the performance gap between the blue curve (*i.e.*, AD-domain CE) and the red curve (*i.e.*, BD-domain CE) narrows as K_d increases. This is because the JCDE mechanism increases the number of observations available for CE, thereby reducing reliance on channel sparsity-based dimensionality reduction. Nevertheless, even for large K_d, a noticeable performance gap remains between the two methods. This residual gap is likely due to the limited accuracy of the initial CE based solely on pilot symbols, which corresponds to the performance difference between "MPDQ (AD-domain CE)" and "Cross-domain MPDQ" shown by the dashed lines.

C. NMSE Performance

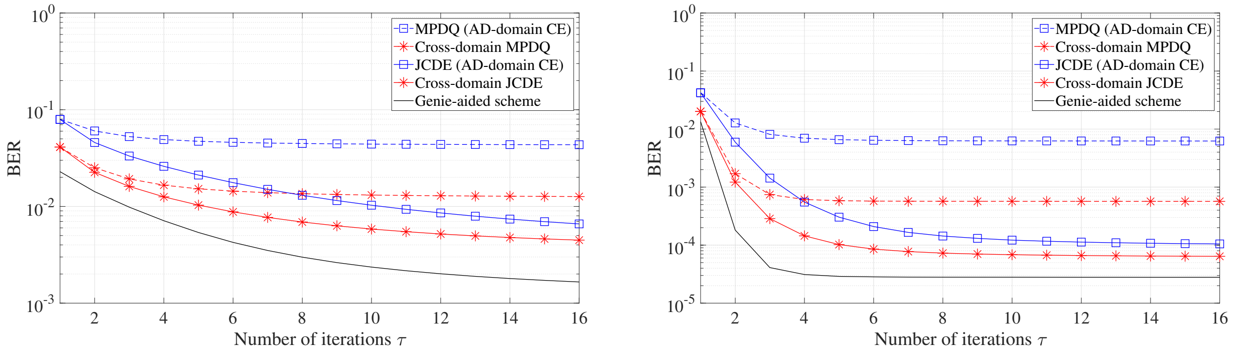
Fig. 5 shows the NMSE performance, where the system parameters are the same as in Figs. 3(a) and 3(c), where the



(a) Gray-coded 4QAM, $b = 1$, SNR $= 10$ dB.

(b) Gray-coded 4QAM, $b = 3$, SNR $= 5$ dB.

Fig. 6. NMSE performances of uncoded MIMO-OFDM systems with Gray-coded 4QAM versus the number of iterations τ.



(a) Gray-coded 4QAM, $b = 1$, SNR $= 10$ dB.

(b) Gray-coded 4QAM, $b = 3$, SNR $= 5$ dB.

Fig. 7. BER performance of uncoded MIMO-OFDM systems with Gray-coded 4QAM versus the number of iterations τ.

NMSE of the estimated AD-domain channel is defined as

$$\mathrm{NMSE} \triangleq \mathbb{E}_{\mathsf{A}}\left[\left|\boldsymbol{A} - c \cdot \check{\boldsymbol{A}}\right|_{\mathrm{F}}^2 / \left|\boldsymbol{A}\right|_{\mathrm{F}}^2\right], \qquad (70)$$

where $\left[\check{\boldsymbol{A}}\right]_{n,u} \triangleq \check{a}_{nu}$ and $c \triangleq \sqrt{\left|\boldsymbol{A}\right|_{\mathrm{F}}^2 / \left|\check{\boldsymbol{A}}\right|_{\mathrm{F}}^2}$.

As expected, the proposed method (red curves), which performs CE in the designated BD domain via the cross-domain strategy, significantly improves NMSE performance—especially in the low SNR region—compared to the conventional alternative (blue curves) that performs CE in the AD domain. These results strongly support the importance of conducting CE in the BD domain. In the high SNR region of Fig. 5(a), a slight reversal is observed between "Cross-domain JCDE" and "JCDE (AD-domain CE)," which appears inconsistent with the BER results in Fig. 3(a). This discrepancy arises because amplitude information of the estimated CSI is lost due to 1-bit quantization [31], [59]. Such differences between the NMSE and BER performance trends are also commonly reported in the literature (*e.g.*, [60]).

D. Convergence Analysis of Proposed Algorithms

We now turn to the analysis of the iterative behavior of the proposed algorithms. Fig. 6 shows the NMSE performance as a function of the number of iterations τ. In evaluating the JCDE algorithms, $\tau_{1,\max}$ was fixed at 16. All other simulation parameters are identical to those used in Fig. 4. As clearly observed from the results, BD-domain CE (red curves) not only achieves significantly better NMSE performance than AD-domain CE (blue curves) but also converges faster in terms of iteration count by effectively leveraging the sparsity of the mmWave MIMO-OFDM channel in the BD domain. This difference in convergence speed becomes particularly evident when the JCDE technique is applied. In both system configurations considered, the cross-domain JCDE scheme converges within approximately $\tau = 6$ iterations.

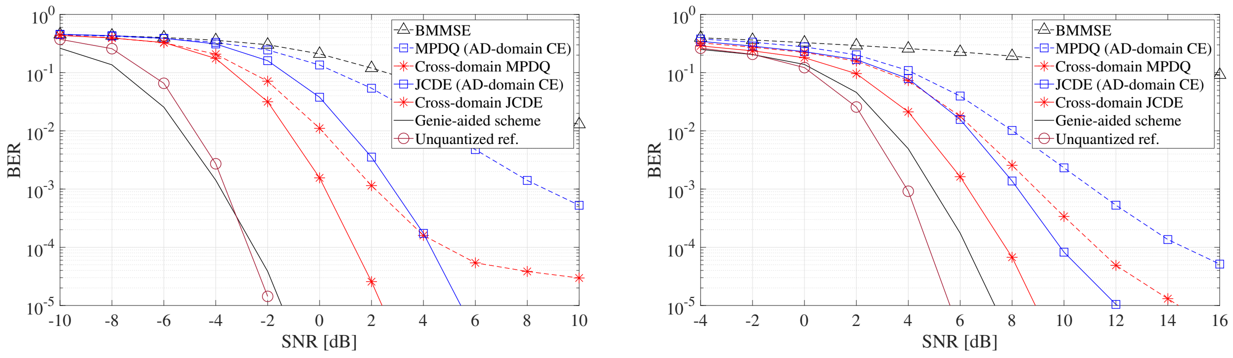
Next, Fig. 7 shows the BER performance obtained from the same computer simulation, exhibiting a trend similar to that of the NMSE results. The proposed methods (red curves) achieve reliable estimation accuracy with fewer iterations compared to the conventional alternatives (blue curves). Even in the challenging case of $b = 1$ shown in Fig. 7(a), where convergence tends to be slower, both "Cross-domain MPDQ" and "Cross-domain JCDE" are observed to converge within approximately $\tau = 16$ iterations.

E. BER Performance in Coded Case

Fig. 8 shows the BER performance in the channel-coded systems, where the other system parameters are the same as those in Figs. 3(a) and 3(c). Using low-density parity-check (LDPC) code that conforms to the fifth generation (5G) new radio (NR) specification, the coding rate was set to $1/3$ for 4QAM and $2/3$ for 16QAM, respectively. The SPA is employed as the soft decision decoder, and error correction by the channel decoder is performed only once after DD.

As expected from the results in Fig. 3, "BMMSE" fails to detect MIMO signals, with BER remaining above 10^{-2} even with the aid of error correction. On the other hand, "Cross-domain MPDQ" can significantly improve performance, but there is still an error floor. In contrast, the JCDE methods suppress the error floor and achieve BER $= 10^{-5}$ in both

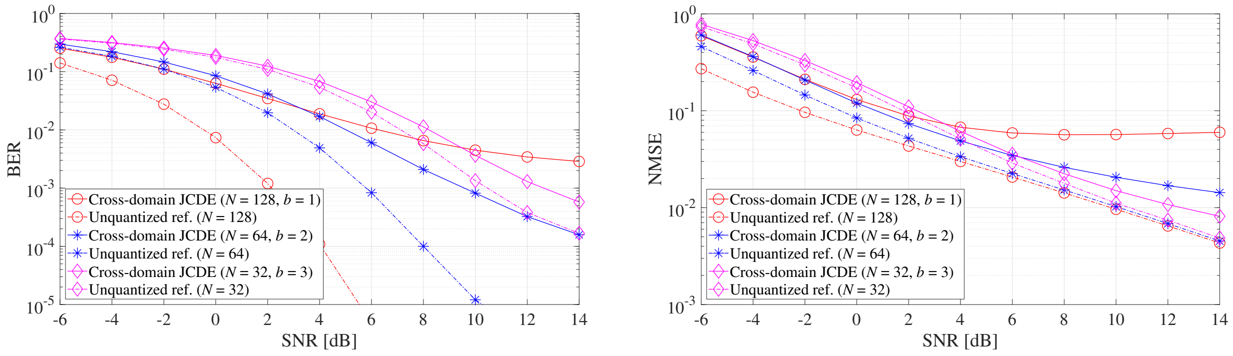
This article has been accepted for publication in IEEE Journal of Selected Topics in Signal Processing. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JSTSP.2025.3589107



(a) Gray-coded 4QAM, $b = 1$, coding rate: $1/3$.

(b) Gray-coded 16QAM, $b = 2$, coding rate: $2/3$.

Fig. 8. BER performance of coded MIMO-OFDM systems, where $(N, M, P, K_{\mathrm{p}}, K_{\mathrm{d}}) = (128, 16, 4, 64, 320)$.



(a) BER performance.

(b) NMSE performance.

Fig. 9. BER and NMSE performances of uncoded MIMO-OFDM systems with Gray-coded 4QAM, where $(M, P, K_{\mathrm{p}}, K_{\mathrm{d}}) = (16, 4, 64, 320)$.

configurations. The performance gap between the proposed method and the idealized reference is within 4 dB for $b = 1$ in Fig. 8(a) and within 2 dB for $b = 2$ in Fig. 8(b) at BER $= 10^{-4}$. Notably, the performance difference between "Cross-domain JCDE" and "JCDE (AD-domain CE)" is larger than in the uncoded case shown in Fig. 3. This is because, as observed in Fig. 5, performing CE in the BD domain, which exhibits higher sparsity, improves the CE accuracy—particularly in the low SNR region—and enhances the quality of the likelihood information fed into the channel decoder.

F. Insights on Optimal Configuration of Systems

To gain insight into the optimal system configuration, we analyze the BER performance of the proposed method in response to changes in the number of quantization bits when the total number of bits transferred to the data interface card is fixed. Fig. 9 compares the BER and NMSE performances for systems with $(N, b) = (128, 1)$, $(64, 2)$, and $(32, 3)$, where the product of the number of RX antennas and the number of quantization bits is fixed at $N \times 2^b = 256$, and the other system parameters are the same as in Fig. 3(a). The performances depicted by the dotted curve ($N = 32$, 64, and 128 in the legend) are unquantized references when the algorithm is run based on the pre-quantized signal model in (5).

First, we can see that even with 4QAM, it is quite difficult to achieve reliable communications with $b = 1$. Even if there is sufficient spatial DoF provided by a large number of RX antennas, as shown in Fig. 9(b), a significant decrease in CE accuracy is unavoidable with $b = 1$, where the phase information is completely lost. This is the cause of the high-level error floor in the BER performance of Fig. 9(a). What is interesting here is that the trend changes significantly when b increases to 2. Since the phase information is now retained and CE accuracy is greatly improved, the BER floor is also significantly suppressed. However, when b is increased further (*i.e.*, $b = 3$), the BER performance degrades due to the reduction in spatial DoF. This observation reveals a clear trade-off between spatial DoF and ADC resolution under a fixed total number of transferred bits. The above results indicate that, in this system configuration, the number of quantization bits b should be set to at least 2. Beyond this point, the performance improvement saturates rapidly, suggesting that increasing the spatial DoF becomes more critical[2].

G. Complexity Analysis

First, we summarize the computational complexity of each component of the proposed method in Tab. I. Since the DQ process, CE, and DD in the proposed method consist solely of scalar-by-scalar operations, their computational complexity is on the order of $\mathcal{O}(NUK)$ per iteration. Assuming that the PDF and CDF of the standard Gaussian distribution, *i.e.*, $\phi(a)$ and $\Phi(a)$, are pre-tabulated, the calculation of the conditional expectation for the Laplace prior in the DQ process requires $\mathcal{O}(NU)$ operations per iteration, which is

---

[2] Such trade-offs involving ADC resolution are not limited to spatial DoF; they also arise with respect to other available observation dimensions, such as the number of temporal samples obtained through oversampling [61], [62]. Therefore, system design must account for the specific trade-offs associated with the assumed system configuration.

This article has been accepted for publication in IEEE Journal of Selected Topics in Signal Processing. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JSTSP.2025.3589107

TABLE I
COMPUTATIONAL COMPLEXITY

CE (including DQ)	$\mathcal{O}\left(NUK\right)$ per iteration
Expectation calc. (Laplace prior)	$\mathcal{O}\left(NU\right)$ per iteration
Expectation calc. (BG prior)	$\mathcal{O}\left(NU\right)$ per iteration
Domain conversion in CE	$\mathcal{O}\left(\log_2\left(N\right)NUK\right)$ per iteration
DD (including DQ)	$\mathcal{O}\left(NUK\right)$ per iteration
Domain conversion in DD	$\mathcal{O}\left(\log_2\left(K_\mathrm{c}\right)NMK\right)$ per iteration

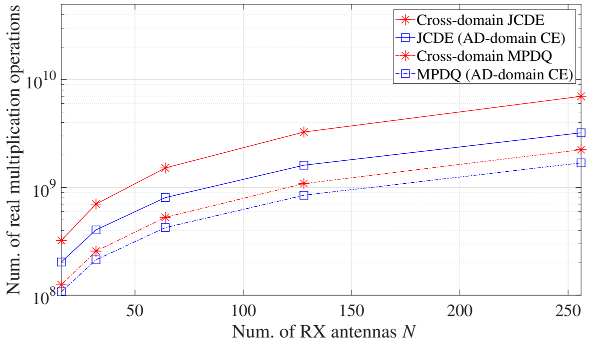


Fig. 10. The number of real-valued multiplication operations as a function of N (ranging from 32 to 256), with the parameters fixed as $M = 16$, $K_\mathrm{p} = 64$, $K_\mathrm{d} = 320$, $P = 4$, and $(\tau_{1,\max}, \tau_{2,\max}) = (16, 16)$.

equivalent to the complexity of the expectation calculation based on the BG prior. The dominant computational cost of the proposed method arises from the domain conversion required at each iteration. Specifically, the complexity is on the order of $\mathcal{O}\left(\log_2\left(N\right)NUK\right)$ per iteration for CE and $\mathcal{O}\left(\log_2\left(K_\mathrm{c}\right)NMK\right)$ per iteration for DD.

Next, to assess the impact of domain conversion on the overall computational cost, we compare the total number of real-valued multiplications required by each algorithm. To estimate the approximate number of real-valued multiplication operations, we adopt the basic assumptions presented in [47].

Fig. 10 shows the number of real-valued multiplication operations as a function of the number of RX antennas N (from 32 to 256). The other parameters are fixed and identical to those used in Fig. 3. As shown in the results, JCDE naturally incurs a higher computational cost than MPDQ, as it iteratively refines the CE by treating the estimated data symbols as additional pilots. However, an important observation is that, for both MPDQ and JCDE, the difference in computational cost between the blue curves (without domain conversion) and the red curves (with the proposed cross-domain strategy) remains relatively small. This is because the domain conversion in the message space does not require any modification to the received signal model, and the additional computational overhead is limited to FFT operations. Consequently, the proposed method can significantly enhance estimation accuracy without increasing the overall computational complexity order. Specifically, at $N = 128$, "Cross-domain MPDQ" operates at approximately 1.3 times the computational cost of "MPDQ (AD-domain CE)," while "Cross-domain JCDE" requires about twice the cost of "JCDE (AD-domain CE)." Considering the observed improvements in both BER and NMSE performance, it can be concluded that the proposed method achieves an excellent trade-off between computational complexity and estimation performance.

V. CONCLUSION

In this paper, we designed a novel cross-domain JCDE algorithm utilizing EP for mmWave MIMO-OFDM systems with low-resolution ADCs. The proposed method consists of the AD-domain DQ process, the BD-domain CE, and the AF-domain DD, and the domain converters within the message space that interconnect them. This domain conversion framework, which is based on the projection of messages through MM into Gaussian distributions, enables the execution of signal processing such as CE and DD within the relevant signal domain. The numerical results show that our proposed method significantly outperforms the SotA schemes in various system parameters and approaches the performance of the idealized scheme under certain conditions. In addition, a comparative study of system configurations with a fixed number of bits transferred to the data interface card provided insight into optimal system design using low-resolution ADCs.

REFERENCES

[1] T. Rappaport, R. Heath, R. Daniels, and J. Murdock, *Millimeter Wave Wireless Communications*, ser. Communications Engineering and Emerging Technology Series from Ted Rappaport Series. Prentice Hall, 2015.

[2] R. W. Heath, N. González-Prelcic, S. Rangan, W. Roh, and A. M. Sayeed, "An overview of signal processing techniques for millimeter wave MIMO systems," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 3, pp. 436–453, 2016.

[3] T. S. Rappaport, J. N. Murdock, and F. Gutierrez, "State of the art in 60-GHz integrated circuits and systems for wireless communications," *Proceedings of the IEEE*, vol. 99, no. 8, pp. 1390–1436, 2011.

[4] J. Mo, P. Schniter, N. G. Prelcic, and R. W. Heath, "Channel estimation in millimeter wave MIMO systems with one-bit quantization," in *Asilomar Conf. Signals Syst. Comput.*, 2014, pp. 957–961.

[5] J. Mo and R. W. Heath, "Capacity analysis of one-bit quantized MIMO systems with transmitter channel state information," *IEEE Trans. Signal Process.*, vol. 63, no. 20, pp. 5498–5512, 2015.

[6] C. Studer and G. Durisi, "Quantized massive MU-MIMO-OFDM uplink," *IEEE Trans. Commun.*, vol. 64, no. 6, pp. 2387–2399, 2016.

[7] C. Doan, S. Emami, D. Sobel, A. Niknejad, and R. Brodersen, "Design considerations for 60 GHz CMOS radios," *IEEE Commun. Mag.*, vol. 42, no. 12, pp. 132–140, 2004.

[8] K. Rikkinen, P. Kyosti, M. E. Leinonen, M. Berg, and A. Parssinen, "THz radio communication: Link budget analysis toward 6G," *IEEE Commun. Mag.*, vol. 58, no. 11, pp. 22–27, 2020.

[9] T. Takahashi, H. Iimori, K. Ishibashi, S. Ibi, and G. T. F. de Abreu, "Bayesian bilinear inference for joint channel tracking and data detection in millimeter-wave MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 23, no. 9, pp. 11 136–11 153, 2024.

[10] A. Adhikary, J. Nam, J. Ahn, and G. Caire, "Joint spatial division and multiplexing—The large-scale array regime," *IEEE Trans. Inf. Theory*, vol. 59, no. 10, pp. 6441–6463, Oct 2013.

[11] A. Alkhateeb, O. El Ayach, G. Leus, and R. W. Heath, "Channel estimation and hybrid precoding for millimeter wave cellular systems," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 831–846, 2014.

[12] J. Lee, G.-T. Gil, and Y. H. Lee, "Channel estimation via orthogonal matching pursuit for hybrid MIMO systems in millimeter wave communications," *IEEE Trans. Commun.*, vol. 64, no. 6, pp. 2370–2386, 2016.

[13] Z. Xiao, P. Xia, and X.-G. Xia, "Codebook design for millimeter-wave channel estimation with hybrid precoding structure," *IEEE Trans. Wireless Commun.*, vol. 16, no. 1, pp. 141–153, 2017.

[14] K. Venugopal, A. Alkhateeb, N. González Prelcic, and R. W. Heath, "Channel estimation for hybrid architecture-based wideband millimeter wave systems," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 9, pp. 1996–2009, 2017.

[15] S. Jacobsson, G. Durisi, M. Coldrey, U. Gustavsson, and C. Studer, "Throughput analysis of massive MIMO uplink with low-resolution ADCs," *IEEE Trans. Wireless Commun.*, vol. 16, no. 6, pp. 4038–4051, 2017.

This article has been accepted for publication in IEEE Journal of Selected Topics in Signal Processing. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JSTSP.2025.3589107

[16] Y. Li, C. Tao, G. Seco-Granados, A. Mezghani, A. L. Swindlehurst, and L. Liu, "Channel estimation and performance analysis of one-bit massive MIMO systems," *IEEE Trans. Signal Process.*, vol. 65, no. 15, pp. 4075–4089, 2017.

[17] I. Atzeni and A. Tölli, "Channel estimation and data detection analysis of massive MIMO with 1-bit ADCs," *IEEE Trans. Wireless Commun.*, vol. 21, no. 6, pp. 3850–3867, 2022.

[18] J. J. Bussgang, "Crosscorrelation functions of amplitude-distorted Gaussian signals," 1952.

[19] J. Zhang, X. Yuan, and Y. A. Zhang, "Blind signal detection in massive MIMO: Exploiting the channel sparsity," *IEEE Trans. Commun.*, vol. 66, no. 2, pp. 700–712, 2018.

[20] C. Wen, C. Wang, S. Jin, K. Wong, and P. Ting, "Bayes-optimal joint channel-and-data estimation for massive MIMO with low-precision ADCs," *IEEE Trans. Signal Process.*, vol. 64, no. 10, pp. 2541–2556, 2016.

[21] T. Takahashi, H. Iimori, K. Ando, K. Ishibashi, S. Ibi, and G. T. F. de Abreu, "Bayesian receiver design via bilinear inference for cell-free massive MIMO with low-resolution ADCs," *IEEE Trans. Wireless Commun.*, vol. 22, no. 7, pp. 4756–4772, 2023.

[22] S. S. Thoota and C. R. Murthy, "Variational Bayes' joint channel estimation and soft symbol decoding for uplink massive MIMO systems with low resolution ADCs," *IEEE Trans. Commun.*, vol. 69, no. 5, pp. 3467–3481, 2021.

[23] ——, "Massive MIMO-OFDM systems with low resolution ADCs: Cramér-Rao bound, sparse channel estimation, and soft symbol decoding," *IEEE Trans. Signal Process.*, vol. 70, pp. 4835–4850, 2022.

[24] Q. Zou, H. Zhang, and H. Yang, "Multi-layer bilinear generalized approximate message passing," *IEEE Trans. Signal Process.*, vol. 69, pp. 4529–4543, 2021.

[25] U. S. Kamilov, V. K. Goyal, and S. Rangan, "Message-passing de-quantization with applications to compressed sensing," *IEEE Trans. Signal Process.*, vol. 60, no. 12, pp. 6270–6281, 2012.

[26] I. Watanabe, T. Takahashi, S. Ibi, A. Tölli, and S. Sampei, "Gaussian belief propagation for mmWave large MIMO detection with low-resolution ADCs," in *2022 IEEE 23rd Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, 2022, pp. 1–5.

[27] J. Mo, P. Schniter, and R. W. Heath, "Channel estimation in broadband millimeter wave MIMO systems with few-bit ADCs," *IEEE Trans. Signal Process.*, vol. 66, no. 5, pp. 1141–1154, 2018.

[28] Y. Ding, S.-E. Chiu, and B. D. Rao, "Bayesian channel estimation algorithms for massive MIMO systems with hybrid analog-digital processing and low-resolution ADCs," *IEEE J. Sel. Topics Signal Process.*, vol. 12, no. 3, pp. 499–513, 2018.

[29] T.-C. Zhang, C.-K. Wen, S. Jin, and T. Jiang, "Mixed-ADC massive MIMO detectors: Performance analysis and design optimization," *IEEE Trans. Wireless Commun.*, vol. 15, no. 11, pp. 7738–7752, 2016.

[30] Y. Xiong, Z. Zhang, N. Wei, and B. Li, "A bilinear GAMP-based receiver for quantized mmWave massive MIMO using expectation maximization," *IEEE Commun. Lett.*, vol. 23, no. 1, pp. 84–87, 2019.

[31] L. Xu, C. Qian, F. Gao, W. Zhang, and S. Ma, "Angular domain channel estimation for mmWave massive MIMO with one-bit ADCs/DACs," *IEEE Trans. Wireless Commun.*, vol. 20, no. 2, pp. 969–982, 2021.

[32] S.-N. Hong, S. Kim, and N. Lee, "A weighted minimum distance decoding for uplink multiuser MIMO systems with low-resolution ADCs," *IEEE Trans. Commun.*, vol. 66, no. 5, pp. 1912–1924, 2018.

[33] Y.-S. Jeon, N. Lee, S.-N. Hong, and R. W. Heath, "One-bit sphere decoding for uplink massive MIMO systems with one-bit ADCs," *IEEE Trans. Wireless Commun.*, vol. 17, no. 7, pp. 4509–4521, 2018.

[34] Z. Shao, R. C. de Lamare, and L. T. N. Landau, "Iterative detection and decoding for large-scale multiple-antenna systems with 1-bit ADCs," *IEEE Wireless Commun. Letters*, vol. 7, no. 3, pp. 476–479, 2018.

[35] J. T. Parker, P. Schniter, and V. Cevher, "Bilinear generalized approximate message passing—part I: Derivation," *IEEE Trans. Signal Process.*, vol. 62, no. 22, pp. 5839–5853, Nov. 2014.

[36] K. Ito, T. Takahashi, S. Ibi, and S. Sampei, "Bilinear Gaussian belief propagation for massive MIMO detection with non-orthogonal pilots," *IEEE Trans. Commun.*, vol. 72, no. 2, pp. 1045–1061, 2024.

[37] H. Iimori, T. Takahashi, K. Ishibashi, G. T. F. de Abreu, and W. Yu, "Grant-free access via bilinear inference for cell-free MIMO with low-coherence pilots," *IEEE Trans. Wireless Commun.*, vol. 20, no. 11, pp. 7694–7710, 2021.

[38] H. Iimori, T. Takahashi, K. Ishibashi, G. T. F. de Abreu, D. González G., and O. Gonsa, "Joint activity and channel estimation for extra-large MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 21, no. 9, pp. 7253–7270, 2022.

[39] K. Ito, T. Takahashi, K. Ishibashi, K. Igarashi, and S. Ibi, "Joint channel, CFO, and data estimation via bayesian inference for multi-user MIMO-OFDM systems," *IEEE Trans. Wireless Commun.*, vol. 24, no. 3, pp. 1898–1915, 2025.

[40] K. R. R. Ranasinghe, H. Seok Rou, G. Thadeu Freitas de Abreu, T. Takahashi, and K. Ito, "Joint channel, data, and radar parameter estimation for AFDM systems in doubly-dispersive channels," *IEEE Trans. Wireless Commun.*, vol. 24, no. 2, pp. 1602–1619, 2025.

[41] Y. Xiang, K. Xu, B. Xia, and X. Cheng, "Bayesian joint channel-and-data estimation for quantized OFDM over doubly selective channels," *IEEE Trans. Wireless Commun.*, vol. 22, no. 3, pp. 1523–1536, 2023.

[42] T. P. Minka, "Expectation propagation for approximate Bayesian inference," in *Proc. the Seventeenth Conf. on Uncertainty in Artificial Intelligence*, 2001, pp. 362–369.

[43] X. Meng, S. Wu, L. Kuang, and J. Lu, "An expectation propagation perspective on approximate message passing," *IEEE Signal Process. Lett.*, vol. 22, no. 8, pp. 1194–1197, 2015.

[44] Q. Zou, H. Zhang, C.-K. Wen, S. Jin, and R. Yu, "Concise derivation for generalized approximate message passing using expectation propagation," *IEEE Signal Process. Lett.*, vol. 25, no. 12, pp. 1835–1839, 2018.

[45] S. Rangan, P. Schniter, and A. K. Fletcher, "Vector approximate message passing," *IEEE Trans. Inf. Theory*, vol. 65, no. 10, pp. 6664–6684, 2019.

[46] K. Takeuchi, "Rigorous dynamics of expectation-propagation-based signal recovery from unitarily invariant measurements," *IEEE Tran. Inf. Theory*, vol. 66, no. 1, pp. 368–386, 2020.

[47] T. Takahashi, A. Tölli, S. Ibi, and S. Sampei, "Low-complexity large MIMO detection via layered belief propagation in beam domain," *IEEE Trans. Wireless Commun.*, vol. 21, no. 1, pp. 234–249, 2022.

[48] K. Furuta, T. Takahashi, and H. Ochiai, "Sparse bayesian learning using complex t-prior for beam-domain massive MIMO channel estimation," *IEEE Open J. the Commun. Society*, vol. 5, pp. 5905–5920, 2024.

[49] D. Donoho, "Compressed sensing," *IEEE Trans. Info. Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.

[50] K. Ito, T. Takahashi, K. Igarashi, S. Ibi, and S. Sampei, "AoA estimation-aided Bayesian receiver design via bilinear inference for mmWave massive MIMO," in *Proc. IEEE ICC*, 2023, pp. 6474–6479.

[51] Z. Gao, Z. Wan, D. Zheng, S. Tan, C. Masouros, D. W. K. Ng, and S. Chen, "Integrated sensing and communication with mmWave massive MIMO: A compressed sampling perspective," *IEEE Trans. Wireless Commun.*, vol. 22, no. 3, pp. 1745–1762, 2023.

[52] T. Takahashi, S. Ibi, and S. Sampei, "Design of adaptively scaled belief in multi-dimensional signal detection for higher-order modulation," *IEEE Trans. Commun.*, vol. 67, no. 3, pp. 1986–2001, Mar. 2019.

[53] A. Chockalingam and B. S. Rajan, *Large MIMO Systems*. Cambridge University Press, 2014.

[54] S. Rangan, "Generalized approximate message passing for estimation with random linear mixing," in *2011 IEEE Int. Symp. Inf. Theory Proceedings*, July 2011, pp. 2168–2172.

[55] F. Bellili, F. Sohrabi, and W. Yu, "Generalized approximate message passing for massive MIMO mmWave channel estimation with Laplacian prior," *IEEE Trans. Commun.*, vol. 67, no. 5, pp. 3205–3219, 2019.

[56] C. M. Bishop, *Pattern Recognition and Machine Learning*. Berlin, Heidelberg: Springer-Verlag, 2006.

[57] J. P. Vila and P. Schniter, "Expectation-maximization Gaussian-mixture approximate message passing," *IEEE Trans. Signal Process.*, vol. 61, no. 19, pp. 4658–4672, 2013.

[58] J. Max, "Quantizing for minimum distortion," *IRE Trans. Inf. Theory*, vol. 6, no. 1, pp. 7–12, 1960.

[59] N. J. Myers and R. W. Heath, "Message passing-based joint CFO and channel estimation in mmWave systems with one-bit ADCs," *IEEE Trans. Wireless Commun.*, vol. 18, no. 6, pp. 3064–3077, 2019.

[60] Q. Wan, J. Fang, H. Duan, Z. Chen, and H. Li, "Generalized Bussgang LMMSE channel estimation for one-bit massive MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 19, no. 6, pp. 4234–4246, 2020.

[61] X. Cheng, B. Xia, K. Xu, and S. Li, "Bayesian channel estimation and data detection in oversampled OFDM receiver with low-resolution ADC," *IEEE Trans. Wireless Commun.*, vol. 20, no. 9, pp. 5558–5571, 2021.

[62] M. Ma, N. T. Nguyen, I. Atzeni, and M. Juntti, "Analysis of oversampling in uplink massive MIMO-OFDM with low-resolution ADCs," in *Proc. IEEE SPAWC*, 2023, pp. 626–630.