

Joint Phase Noise and Off-Grid Channel Estimation for AFDM Systems via Sparse Bayesian Learning

You Xu¹, Huaijin Zhang¹, Lixia Xiao¹, Guanghua Liu¹, and Zilong Liu²

¹School of Cyber Science and Engineering, Huazhong University of Science and Technology, Wuhan 430074, China

²School of Computer Science and Electronics Engineering, University of Essex, UK

Email: youxu@ieee.org, {huaijinzhang, lixiaxiao}@hust.edu.cn, guanghualiu@ieee.org, zilong.liu@essex.ac.uk

Abstract—In practical affine frequency division multiplexing (AFDM) systems, the intricate coupling of oscillator phase noise (PN) and off-grid fractional shifts traps conventional estimators in a severe high-SNR error floor. To address these challenges, we propose a joint PN and channel estimation method based on sparse Bayesian learning (JPNCE-SBL). Specifically, a reduced-rank subspace projection is first introduced to capture the dominant eigen-energy of the Wiener PN process. Concurrently, a dynamic grid evolution strategy is designed to iteratively eliminate off-grid errors without requiring computationally prohibitive global grid densification. Both components are integrated into a unified Expectation-Maximization (EM) framework, where the channel and PN estimates are jointly updated at each iteration to prevent error propagation. Simulation results demonstrate that JPNCE-SBL significantly outperforms existing benchmarks in both NMSE and BER, closely approaching the perfect channel state information case under practical PN conditions.

Index Terms—Affine frequency division multiplexing, doubly-dispersive channels, phase noise, sparse Bayesian learning, grid evolution, joint estimation.

I. INTRODUCTION

Emerging 6G applications demand ultra-reliable connectivity in high-mobility scenarios [1], [2]. However, extreme Doppler spreads in doubly-dispersive channels induce severe inter-carrier interference (ICI) in conventional orthogonal frequency division multiplexing (OFDM) systems, fundamentally limiting their reliability. To overcome this limitation, affine frequency division multiplexing (AFDM) has emerged as a promising next-generation waveform [3], [4]. By leveraging the discrete affine Fourier transform (DAFT), AFDM achieves a complete delay-Doppler channel representation with superior resilience against mobility-induced impairments [5], [6].

Despite these theoretical advantages, practical AFDM deployment at high carrier frequencies (e.g., mmWave) faces a key hardware impairment: oscillator phase noise (PN) [3], [7]. PN arises from local oscillator instability, creating random phase rotations that destroy subcarrier orthogonality. This dynamic phase error also couples with the continuous delay-Doppler multipath parameters, forming a complex bilinear inverse problem [8]. Consequently, traditional channel estimators become unreliable, degrading both channel acquisition and data recovery [9].

Accurate channel state information is fundamental to reliable AFDM signal detection. Sparse Bayesian learning (SBL) formulates doubly-dispersive channel recovery as a probabilistic inference problem, intrinsically capturing multipath

sparsity without requiring predefined sparsity levels [10], [11]. However, standard SBL maps continuous delay and Doppler shifts onto a fixed discrete dictionary, causing off-grid basis mismatches that degrade estimation accuracy [12]. First-order approximation methods suffer from truncation errors under high mobility, leading to premature performance saturation. Global grid densification can reduce mismatch but expands the dictionary dimension substantially, making it impractical for real-time use [13], [14].

Moreover, existing SBL estimators typically assume ideal hardware and treat channel estimation and PN tracking as separate steps. In practical high-frequency AFDM systems, off-grid errors and time-varying PN are tightly coupled, causing nonlinear distortions to the sensing dictionary. Decoupled methods ignore this interaction, which leads to error propagation between the two processes and an irreducible high-SNR error floor [15], [16]. A joint framework is therefore required to simultaneously handle nonlinear PN and recover continuous channel parameters at low complexity.

To address these challenges, we propose a joint PN and channel estimation method based on sparse Bayesian learning (JPNCE-SBL) for AFDM systems over doubly-dispersive channels. The joint estimation task is formulated as a bilinear sparse recovery problem, where an EM-based SBL framework is developed to iteratively decouple PN tracking from sparse channel recovery, thereby preventing error propagation. Specifically, a reduced-rank subspace projection is introduced to capture the dominant energy of the Wiener PN process, while a dynamic grid evolution strategy iteratively eliminates off-grid errors without global grid densification. Both components are integrated into a unified EM framework, where the channel and PN estimates are jointly updated at each iteration. Simulation results confirm that JPNCE-SBL significantly outperforms existing benchmarks in both NMSE and BER, closely approaching the perfect CSI bound under practical PN conditions.

II. SYSTEM MODEL

Consider an N -subcarrier AFDM system with duration T and bandwidth $B = N\Delta f$, where $\Delta f = 1/T$. The data symbols $\mathbf{x} = [x_0, \dots, x_{N-1}]^T \in \mathbb{A}^{N \times 1}$ in the discrete affine Fourier (DAF) domain are transformed into the time domain

via the inverse DAFT (IDAFT) [4]:

$$s_n = \frac{1}{\sqrt{N}} \sum_{m=0}^{N-1} e^{j2\pi(c_1 n^2 + c_2 m^2 + \frac{1}{N} nm)} x_m, \quad (1)$$

where c_1 and c_2 denote the chirp parameters. In vector form, the above expression yields $\mathbf{s} = \mathbf{A}_{\text{af}}^H \mathbf{x}$, where $\mathbf{A}_{\text{af}}^H = \mathbf{\Lambda}_{c_1}^H \mathbf{F}_N^H \mathbf{\Lambda}_{c_2}^H$ and $\mathbf{\Lambda}_{c_i} \triangleq \text{diag}[e^{-j2\pi c_i 0^2}, \dots, e^{-j2\pi c_i (N-1)^2}]$. A chirp-periodic prefix (CPP) of length N_{cpp} is appended to \mathbf{s} to mitigate inter-symbol interference from multipath propagation.

After propagating through a time-varying doubly-dispersive channel and being corrupted by receiver-side phase noise [9], the received time-domain signal is modeled as

$$r_n = e^{j\phi_n} \sum_{p=1}^P h_p e^{j2\pi\nu_p n T_s} s_{[n-\ell_p]_N} + w_n, \quad (2)$$

where $e^{j\phi_n}$ denotes the n -th PN sample, $w_n \sim \mathcal{CN}(0, \sigma_w^2)$ is the additive white Gaussian noise (AWGN), and $\{h_p, \ell_p, \nu_p\}$ denote the complex gain, normalized delay, and normalized Doppler shift of the p -th propagation path, respectively. The corresponding vector-form received signal is expressed as

$$\mathbf{r} = \mathbf{P} \sum_{p=1}^P h_p \mathbf{V}^{\nu_p} \mathbf{\Pi}^{\ell_p} \mathbf{s} + \mathbf{w}, \quad (3)$$

where $\mathbf{P} = \text{diag}(e^{j\phi_0}, \dots, e^{j\phi_{N-1}})$, while \mathbf{V}^{ν_p} and $\mathbf{\Pi}^{\ell_p}$ denote the diagonal Doppler shift matrix and forward cyclic shift matrix, respectively. \mathbf{w} is the noise vector.

By performing the DAFT $\mathbf{y}_{\text{af}} = \mathbf{A}_{\text{af}} \mathbf{r}$ at the receiver, the DAF-domain received symbol $y_{\text{af}}[\tilde{m}]$ can be expressed as

$$y_{\text{af}}[\tilde{m}] = \sum_{m=0}^{N-1} H_{\text{pn}}[\tilde{m}, m] x_m + \tilde{w}_{\tilde{m}}, \quad (4)$$

where the equivalent channel kernel $H_{\text{pn}}[\tilde{m}, m]$ is given by

$$H_{\text{pn}}[\tilde{m}, m] = \frac{1}{N} \sum_{p=1}^P h_p e^{j2\pi(c_1 \ell_p^2 - \frac{m\ell_p}{N} + c_2(m^2 - \tilde{m}^2))} \mathcal{G}_p(\tilde{m}, m). \quad (5)$$

Here, the interaction function $\mathcal{G}_p(\tilde{m}, m) = \sum_{n=0}^{N-1} e^{j\phi_n} e^{-j\frac{2\pi}{N}(\tilde{m}-m+2Nc_1\ell_p+\nu_p)n}$ encapsulates the joint impact of PN and doubly-dispersive fading. It is worth noting that the PN term $e^{j\phi_n}$ disrupts the Dirichlet kernel structure inherent to the AFDM system, thereby inducing severe inter-carrier interference [8]. This disruption necessitates a robust framework to jointly estimate the sparse channel parameters and the PN parameter.

In practical oscillators, the time-domain phase noise (PN) is widely described as a non-stationary Wiener process [17]. The phase evolution is given by $\phi_n = \phi_{n-1} + \delta_n$, where the innovation $\delta_n \sim \mathcal{N}(0, \sigma_{PN}^2)$. Assuming $\phi_{-1} = 0$, the PN phase vector $\boldsymbol{\phi} = [\phi_0, \dots, \phi_{N-1}]^T$ follows a multivariate Gaussian distribution, i.e., $\boldsymbol{\phi} \sim \mathcal{N}(\mathbf{0}, \mathbf{R}_{\boldsymbol{\phi}})$. The elements of the covariance matrix $\mathbf{R}_{\boldsymbol{\phi}} \in \mathbb{R}^{N \times N}$ are compactly defined as $[\mathbf{R}_{\boldsymbol{\phi}}]_{m,n} = \sigma_{PN}^2(\min(m, n) + 1)$.

Furthermore, to alleviate the high-dimensional estimation burden, $\boldsymbol{\phi}$ is projected onto a lower-dimensional subspace via

the singular value decomposition $\mathbf{R}_{\boldsymbol{\phi}} = \mathbf{U}\mathbf{D}\mathbf{U}^T$, yielding a basis expansion model:

$$\boldsymbol{\phi} \approx \mathbf{B}\boldsymbol{\eta}, \quad (6)$$

where $\mathbf{B} = \mathbf{U}_{1:L} \mathbf{D}_{1:L}^{1/2} \in \mathbb{R}^{N \times L}$ comprises the L dominant basis vectors with $L \ll N$, and $\boldsymbol{\eta} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_L)$ denotes the reduced-rank PN coefficient vector. The parameterized time-domain PN matrix is then given by $\mathbf{P}(\boldsymbol{\eta}) = \text{diag}(e^{j\mathbf{B}\boldsymbol{\eta}})$. In the proposed JPNCE-SBL framework, tracking the low-dimensional $\boldsymbol{\eta}$ instead of $\boldsymbol{\phi}$ significantly reduces computational overhead while ensuring numerical stability.

Next, we consider the joint channel estimation problem. To exploit the inherent sparsity of the doubly-dispersive channel, the estimation task is formulated as a sparse recovery problem by establishing a virtual sampling grid [10], [12]. The delay and Doppler domains are discretized within the ranges $[0, \ell_{\text{max}}]$ and $[-k_{\text{max}} - 1, k_{\text{max}} + 1]$ with resolutions $r_\tau = \frac{\ell_{\text{max}}}{M_\tau - 1}$ and $r_\nu = \frac{2k_{\text{max}} + 2}{M_\nu - 1}$, respectively. The resulting virtual sampling delay and Doppler vectors are $[\tilde{\ell}_0, \dots, \tilde{\ell}_{M_\tau - 1}]^T$ and $[\tilde{k}_0, \dots, \tilde{k}_{M_\nu - 1}]^T$, where $\tilde{\ell}_a = ar_\tau$ and $\tilde{k}_b = br_\nu - k_{\text{max}} - 1$. The sampling grid is then defined as $\bar{\mathbf{G}} = \{\bar{g}_m\}_{m=0}^{M_S - 1} = \{\tilde{\ell}_m, \tilde{k}_m\}_{m=0}^{M_S - 1}$ with $M_S = M_\tau M_\nu$, comprising all combinations of virtual parameters where $\tilde{\ell}_m = \ell_a$ and $\tilde{k}_m = \tilde{k}_b$ for $a = \lfloor m/M_\nu \rfloor$ and $b = m \bmod M_\nu$. Then, by mapping the physical parameters onto this grid, the received signal can be reformulated as

$$\mathbf{r} = \mathbf{P}(\boldsymbol{\eta}) \boldsymbol{\Phi}(\bar{\mathbf{G}}) \bar{\mathbf{h}} + \mathbf{w}, \quad (7)$$

where $\bar{\mathbf{h}} \in \mathbb{C}^{M_S \times 1}$ is the sparse channel vector to be estimated, the dictionary matrix $\boldsymbol{\Phi}(\bar{\mathbf{G}}) \in \mathbb{C}^{N \times M_S}$ consists of the steering vectors $\boldsymbol{\phi}(\bar{g}_m) = \mathbf{V}^{\tilde{\nu}_m} \mathbf{\Pi}^{\tilde{\ell}_m} \mathbf{s}$ for all grid points $m = 0, \dots, M_S - 1$, and the diagonal matrix $\mathbf{P}(\boldsymbol{\eta}) \in \mathbb{C}^{N \times N}$ is the PN error parameterized by the coefficient vector $\boldsymbol{\eta}$. The joint estimation of $\bar{\mathbf{h}}$ and $\boldsymbol{\eta}$ is challenging due to their bilinear coupling in (7), the non-convexity induced by $\mathbf{P}(\boldsymbol{\eta}) = \text{diag}(e^{j\mathbf{B}\boldsymbol{\eta}})$, and the dictionary mismatch caused by off-grid effects. The goal is therefore to jointly recover the sparse $\bar{\mathbf{h}}$, the PN coefficients $\boldsymbol{\eta}$, and the refined grid $\bar{\mathbf{G}}$ from (7), which we address via the JPNCE-SBL in the following section.

III. THE PROPOSED JPNCE-SBL METHOD FOR AFDM

In this section, we develop a joint channel and PN estimation algorithm based on the SBL framework [14]. Unlike [14], which applies grid evolution solely to doubly fractional OTFS channel estimation without accounting for phase noise, the proposed JPNCE-SBL extends this strategy to AFDM systems by jointly handling time-varying PN and off-grid channel parameters within a unified EM framework. To exploit the inherent sparsity of the channel vector $\bar{\mathbf{h}}$, a hierarchical Gaussian-Gamma prior is adopted [12], where the first layer assumes that the entries of $\bar{\mathbf{h}}$ follow a conditional complex Gaussian distribution:

$$p(\bar{\mathbf{h}}|\boldsymbol{\gamma}) = \mathcal{CN}(\bar{\mathbf{h}}|\mathbf{0}, \boldsymbol{\Gamma}), \quad (8)$$

where $\boldsymbol{\Gamma} = \text{diag}(\boldsymbol{\gamma})$ and $\boldsymbol{\gamma} = [\gamma_0, \dots, \gamma_{M_S - 1}]^T$. In the second

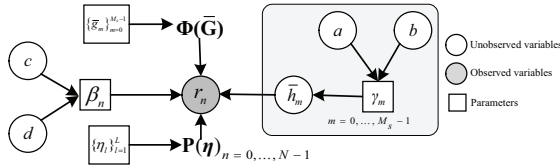


Fig. 1: The graphical model for the JPNCE-SBL method.

layer, to ensure non-informative priors, each hyperparameter γ_m and the noise precision $\beta = \sigma_w^{-2}$ are independently governed by Gamma distributions with small constants:

$$p(\gamma|a, b) = \prod_{m=0}^{M_S-1} \Gamma(\gamma_m|a, b), p(\mathbf{w}|\beta) = \mathcal{CN}(\mathbf{w}|\mathbf{0}, \beta^{-1}\mathbf{I}_N). \quad (9)$$

The corresponding graphical model for the joint estimation process is illustrated in Fig. 1. It is worth noting that marginalizing over γ naturally yields a sparsity-enforcing Student- t marginal prior for each entry of $\bar{\mathbf{h}}$ [10]. Based on the signal model established in Section II, the conditional likelihood of the observation is given by

$$p(\mathbf{r}|\bar{\mathbf{h}}; \beta, \bar{\mathbf{G}}, \boldsymbol{\eta}) = \mathcal{CN}(\mathbf{r}|\mathbf{P}(\boldsymbol{\eta})\Phi(\bar{\mathbf{G}})\bar{\mathbf{h}}, \beta^{-1}\mathbf{I}_N), \quad (10)$$

Thus, let the parameter set $\boldsymbol{\theta} = \{\gamma, \beta, \bar{\mathbf{G}}, \boldsymbol{\eta}\}$, the conditional posterior follows $p(\bar{\mathbf{h}}|\mathbf{r}; \boldsymbol{\theta}) \propto p(\mathbf{r}|\bar{\mathbf{h}}; \beta, \bar{\mathbf{G}}, \boldsymbol{\eta})p(\bar{\mathbf{h}}|\gamma) = \mathcal{CN}(\bar{\mathbf{h}}|\boldsymbol{\mu}, \boldsymbol{\Sigma})$, where the posterior statistics $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ [14]:

$$\boldsymbol{\mu} = \beta\boldsymbol{\Sigma}\Psi^H\mathbf{r}, \quad \boldsymbol{\Sigma} = \boldsymbol{\Gamma} - \boldsymbol{\Gamma}\Psi^H \left(\beta^{-1}\mathbf{I}_N + \Psi\boldsymbol{\Gamma}\Psi^H \right)^{-1} \Psi\boldsymbol{\Gamma}, \quad (11)$$

where $\Psi = \mathbf{P}(\boldsymbol{\eta})\Phi(\bar{\mathbf{G}})$ denotes the effective sensing matrix. According to the maximum a posteriori criterion, the sparse channel estimate is directly obtained as $\hat{\bar{\mathbf{h}}} = \boldsymbol{\mu}$.

A. The Proposed JPNCE-SBL

Since the posterior statistics depend heavily on the unknown parameters in $\boldsymbol{\theta}$, the Expectation-Maximization (EM) algorithm is employed for the iterative estimation of $\boldsymbol{\theta}$. Specifically, the marginal log-likelihood is maximized via the evidence lower bound [18]. During the E-step of the $(i+1)$ -th iteration, by taking the expectation of the joint log-likelihood $\ln p(\mathbf{r}, \bar{\mathbf{h}}; \boldsymbol{\theta})$ with respect to $p(\bar{\mathbf{h}}|\mathbf{r}; \boldsymbol{\theta}^{(i)})$, the objective log-likelihood function $\mathcal{Q}(\boldsymbol{\theta}, \boldsymbol{\theta}^{(i)})$ can be derived as:

$$\begin{aligned} \mathcal{Q}(\boldsymbol{\theta}, \boldsymbol{\theta}^{(i)}) &= \langle \ln p(\mathbf{r}, \bar{\mathbf{h}}; \boldsymbol{\theta}) \rangle_{p(\bar{\mathbf{h}}|\mathbf{r}; \boldsymbol{\theta}^{(i)})} \\ &\propto N \ln \beta - \beta \left\| \mathbf{r} - \Psi \boldsymbol{\mu}^{(i)} \right\|_2^2 - \beta \text{Tr} \left(\Psi \boldsymbol{\Sigma}^{(i)} \Psi^H \right) \\ &\quad + \sum_{m=0}^{M_S-1} \left(\ln \gamma_m^{-1} - \gamma_m^{-1} \Xi_m^{(i)} \right), \end{aligned} \quad (12)$$

where $\Xi_m^{(i)} = \sum_{m,m}^{(i)} + |\mu_m^{(i)}|^2$, and the second equality exploits the expectation trace property and $\langle \bar{\mathbf{h}}\bar{\mathbf{h}}^H \rangle_{p(\bar{\mathbf{h}}|\mathbf{r}; \boldsymbol{\theta}^{(i)})} = \boldsymbol{\Sigma}^{(i)} + \boldsymbol{\mu}^{(i)}(\boldsymbol{\mu}^{(i)})^H$. Subsequently, in the M-step, the parameters are updated via $\boldsymbol{\theta}^{(i+1)} = \arg \max_{\boldsymbol{\theta}} \mathcal{Q}(\boldsymbol{\theta}, \boldsymbol{\theta}^{(i)})$ [19]. By exploiting the separable structure of \mathcal{Q} with respect to γ , β , $\bar{\mathbf{G}}$, and $\boldsymbol{\eta}$, the optimization process is effectively decoupled and executed across three sequential stages.

1) *Stage 1: Updating Hyperparameters $\gamma^{(i+1)}$ and $\beta^{(i+1)}$:* When given the conditional posterior $p(\bar{\mathbf{h}}|\mathbf{r}; \boldsymbol{\theta}^{(i)})$, the hyperparameters $\gamma^{(i+1)}$ and $\beta^{(i+1)}$ are updated by maximizing the objective function $\mathcal{Q}(\boldsymbol{\theta}, \boldsymbol{\theta}^{(i)})$ with respect to γ and β , i.e., $(\gamma^{(i+1)}, \beta^{(i+1)}) = \arg \max_{\gamma, \beta} \mathcal{Q}(\boldsymbol{\theta}, \boldsymbol{\theta}^{(i)})$. Then, by setting the partial derivatives with respect to γ_m and β to zero, respectively, the closed-form update rules can be obtained as

$$\beta^{(i+1)} = \frac{N + c - 1}{d + \mathcal{C}_{\beta}^{(i)}}, \gamma_m^{(i+1)} = \frac{\sqrt{4\rho \left(|\mu_m^{(i)}|^2 + \sum_{m,m}^{(i)} + 1 \right)} - 1}{2\rho}, \quad (13)$$

where the residual term $\mathcal{C}_{\beta}^{(i)} = \|\mathbf{r} - \Psi^{(i)}\boldsymbol{\mu}^{(i)}\|_2^2 + (\beta^{(i)})^{-1} \sum_{m=0}^{M_S-1} (1 - \sum_{m,m}^{(i)} / \gamma_m^{(i)})$. When updating $\gamma^{(i+1)}$, we extract the \mathcal{P} indices with the largest magnitudes to form the active support set \mathcal{P} . The corresponding delay and Doppler components are preliminarily estimated as $\hat{g}_p^{(i)} = \{\bar{\ell}_p^{(i)}, \bar{\nu}_p^{(i)}\}$ for $p \in \mathcal{P}$. It is noted that conventional SBL methods cannot handle PN-induced dictionary mismatches and off-grid effects. While Off Grid SBL method uses first-order linear approximations, it remains sensitive to modeling errors. To address grid limitations, we resolve these issues in Stage 3 by evolving the grid $\bar{\mathbf{G}}$ to match the continuous parameters, which also ensure precise PN estimation.

2) *Stage 2: Updating PN Coefficient Vector $\boldsymbol{\eta}^{(i+1)}$:* In this stage, we update the reduced-rank PN coefficients $\boldsymbol{\eta}^{(i+1)}$ using the dictionary $\Phi(\bar{\mathbf{G}}^{(i)})$ and the channel mean $\boldsymbol{\mu}^{(i)}$ from the previous iteration. The objective is to minimize the cost function $\mathcal{L}(\boldsymbol{\eta})$ while fixing the current sampling grid $\bar{\mathbf{G}}^{(i)}$:

$$\hat{\boldsymbol{\eta}}^{(i+1)} = \arg \min_{\boldsymbol{\eta}} \mathcal{L}(\boldsymbol{\eta}^{(i)}), \quad (14)$$

where $\mathcal{L}(\boldsymbol{\eta}) = \log \det(\boldsymbol{\Sigma}) + (\mathbf{r} - \boldsymbol{\mu}_r(\boldsymbol{\eta}, \bar{\mathbf{G}}^{(i)}))^H \boldsymbol{\Sigma}^{-1} (\mathbf{r} - \boldsymbol{\mu}_r(\boldsymbol{\eta}, \bar{\mathbf{G}}^{(i)})) + 1/2\boldsymbol{\eta}^T \boldsymbol{\eta}$, and the reconstructed signal is $\boldsymbol{\mu}_r(\boldsymbol{\eta}, \bar{\mathbf{G}}^{(i)}) = \mathbf{P}(\boldsymbol{\eta})\Phi(\bar{\mathbf{G}}^{(i)})\boldsymbol{\mu}^{(i)}$. To obtain a closed-form solution for this non-convex problem, we perform two steps. First, we freeze the noise covariance $\boldsymbol{\Sigma}$ using the estimate from the i -th iteration. Second, since the PN innovation variance of practical oscillators is usually small, we can apply the small-angle approximation like Taylor series expansion to the PN matrix $\mathbf{P}(\boldsymbol{\eta})$, i.e.,

$$\mathbf{P}(\boldsymbol{\eta}) = \text{diag}(e^{j\mathbf{B}\boldsymbol{\eta}}) \approx \mathbf{I}_N + j \text{diag}(\mathbf{B}\boldsymbol{\eta}). \quad (15)$$

Then, the object function $\mathcal{L}(\boldsymbol{\eta})$ can be reformulated into a regularized quadratic form:

$$\begin{aligned} \mathcal{L}(\boldsymbol{\eta}) &\approx (\bar{\mathbf{r}}^{(i)} - \mathbf{V}^{(i)}\boldsymbol{\eta})^H \boldsymbol{\Sigma}^{-1} (\bar{\mathbf{r}}^{(i)} - \mathbf{V}^{(i)}\boldsymbol{\eta}) + \frac{1}{2}\boldsymbol{\eta}^T \boldsymbol{\eta}, \\ &\approx \boldsymbol{\eta}^T \Re(\mathbf{V}^{(i)H} \boldsymbol{\Sigma}^{-1} \mathbf{V}^{(i)})\boldsymbol{\eta} - 2\Re(\bar{\mathbf{r}}^{(i)H} \boldsymbol{\Sigma}^{-1} \mathbf{V}^{(i)})\boldsymbol{\eta} + \frac{1}{2}\boldsymbol{\eta}^T \boldsymbol{\eta}, \end{aligned} \quad (16)$$

where $\bar{\mathbf{r}}^{(i)}$ is the residual between the observation and the PN-free reconstructed signal, and $\mathbf{V}^{(i)}$ is the linearized coupling matrix that maps the PN subspace coefficients to the received signal domain, given by

$$\begin{aligned} \bar{\mathbf{r}}^{(i)} &= \mathbf{r} - \Phi(\bar{\mathbf{G}}^{(i)})\boldsymbol{\mu}^{(i)}, \\ \mathbf{V}^{(i)} &= j \text{diag}(\Phi(\bar{\mathbf{G}}^{(i)})\boldsymbol{\mu}^{(i)})\mathbf{B}. \end{aligned} \quad (17)$$

Finally, we set the gradient of (16) with respect to $\boldsymbol{\eta}$ to zero. This leads to the closed-form update rule for the PN coefficients, $\boldsymbol{\eta}^{(i+1)}$, which can be given as

$$\boldsymbol{\eta}^{(i+1)} = [\Re((\mathbf{V}^{(i)})^H \boldsymbol{\Sigma}^{-1} \mathbf{V}^{(i)}) + \frac{1}{2} \mathbf{I}_L]^{-1} \Re((\mathbf{V}^{(i)})^H \boldsymbol{\Sigma}^{-1} \bar{\mathbf{r}}^{(i)}). \quad (18)$$

Here, $\Re(\cdot)$ arises because $\boldsymbol{\eta}$ is real-valued, and the $\frac{1}{2} \mathbf{I}_L$ term reflects the Gaussian prior $\boldsymbol{\eta} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_L)$, regularizing the estimate when PN energy is low. After computing $\boldsymbol{\eta}^{(i+1)}$, we reconstruct the full-dimensional PN vector as $\hat{\boldsymbol{\phi}}_{\gamma}^{(i+1)} = \mathbf{B} \boldsymbol{\eta}^{(i+1)}$. The PN matrix $\mathbf{P}(\boldsymbol{\eta}^{(i+1)})$ is then refreshed for the subsequent channel estimation and grid refinement stages.

3) Stage 3: Updating Doppler Grid via Grid Evolution:

We develop a dynamic grid evolution (GE) strategy to mitigate discretization errors of fixed grids. In this work, delay parameters are assumed to lie on integer grid points, so GE is applied solely to the Doppler dimension; the analogous delay refinement follows the same derivation and can be straightforwardly incorporated as an extension. This method iteratively refines the virtual Doppler grid points \bar{k}_m without expanding the measurement matrix. Thus, the refined points converge toward the true continuous Doppler parameters to improve estimation accuracy. Specifically, we approximate the dictionary matrix $\Phi(\bar{\mathbf{G}})$ via a first-order linear expansion with respect to \bar{k}_m around the current grid points:

$$\hat{\Phi}(\bar{\mathbf{G}}) \approx \Phi(\bar{\mathbf{G}}) + \Omega(\bar{\mathbf{G}}) \text{diag}(\boldsymbol{\xi}), \quad (19)$$

where $\boldsymbol{\xi} = [\xi_0, \dots, \xi_{M_S-1}]^T$ is the off-grid Doppler residual vector, and the derivative matrix $\Omega(\bar{\mathbf{G}})$ is constructed with columns defined as $\omega(\bar{g}_m) = \partial \phi(\bar{g}_m) / \partial \bar{k}_m$.

In the M-step, we update the grid perturbation $\boldsymbol{\xi}$ by minimizing the expected residual error. We utilize the posterior mean $\boldsymbol{\mu}^{(i)}$ and covariance $\boldsymbol{\Sigma}^{(i)}$ from the E-step to perform this update. Then, the optimization problem can be shown as:

$$\begin{aligned} \boldsymbol{\xi}^{(i+1)} &= \arg \min_{\boldsymbol{\xi}} \mathcal{Q}^{(i)}(\boldsymbol{\theta}, \boldsymbol{\theta}^{(i)}) \\ &= \arg \min_{\boldsymbol{\xi}} \{ \|\mathbf{r} - \mathbf{P}(\boldsymbol{\eta}^{(i+1)}) \hat{\Phi}(\bar{\mathbf{G}}^{(i)}) \boldsymbol{\mu}^{(i)}\|_2^2 \\ &\quad + \text{Tr} [\mathbf{P}(\boldsymbol{\eta}^{(i+1)}) \hat{\Phi}(\bar{\mathbf{G}}^{(i)}) \boldsymbol{\Sigma}^{(i)} (\mathbf{P}(\boldsymbol{\eta}^{(i+1)}) \hat{\Phi}(\bar{\mathbf{G}}^{(i)}))^H] \}. \end{aligned} \quad (20)$$

Then, substituting (19) into (20) and applying the identity $\text{Tr}(\mathbf{A} \text{diag}(\boldsymbol{\xi}) \mathbf{B} \text{diag}(\boldsymbol{\xi})) = \boldsymbol{\xi}^T (\mathbf{A}^* \odot \mathbf{B}) \boldsymbol{\xi}$, the cross terms between Φ and Ω are handled via the Hadamard product. This reduces the objective to a standard quadratic form:

$$\mathcal{L}(\boldsymbol{\xi}) \approx \boldsymbol{\xi}^T \mathbf{Q}^{(i)} \boldsymbol{\xi} - 2 \mathbf{d}^{(i)T} \boldsymbol{\xi} + C, \quad (21)$$

where C is a constant term independent of $\boldsymbol{\xi}$, and the matrix $\mathbf{Q}^{(i)}$ and the vector $\mathbf{d}^{(i)}$ can be defined as

$$\begin{aligned} \mathbf{Q}^{(i)} &= \Re \left\{ ((\Omega^{(i)})^H \Omega^{(i)})^* \odot (\boldsymbol{\mu}^{(i)} (\boldsymbol{\mu}^{(i)})^H + \boldsymbol{\Sigma}^{(i)}) \right\}, \\ \mathbf{d}^{(i)} &= \Re \left\{ \text{diag}(\boldsymbol{\mu}^{(i)*}) (\Omega^{(i)})^H (\mathbf{P}^{(i+1)})^H (\mathbf{r} - \mathbf{P}^{(i+1)} \Phi^{(i)} \boldsymbol{\mu}^{(i)}) \right. \\ &\quad \left. - \text{diag}((\Omega^{(i)})^H \Phi^{(i)} \boldsymbol{\Sigma}^{(i)}) \right\}. \end{aligned} \quad (22)$$

To reduce complexity, we update the grid only on the active set \mathcal{P} from Stage 1. Specifically, let $\boldsymbol{\xi}_{\mathcal{P}}$ be the doppler offset

vector for \mathcal{P} . We extract the sub-matrix $\mathbf{Q}_{\mathcal{P}}^{(i)}$ and sub-vector $\mathbf{d}_{\mathcal{P}}^{(i)}$. Then, we can obtain the closed-form solution:

$$\boldsymbol{\xi}_{\mathcal{P}}^{(i+1)} = \left(\mathbf{Q}_{\mathcal{P}}^{(i)} \right)^{-1} \mathbf{d}_{\mathcal{P}}^{(i)}. \quad (23)$$

Moreover, to maintain grid integrity and avoid overlapping, we constrain the offsets to the resolution interval, and the doppler grid offset can be given as

$$\xi_p^{(i+1)} = \text{clip}(\xi_p^{(i+1)}, -r_\nu/2, r_\nu/2), \quad \forall p \in \mathcal{P}. \quad (24)$$

Finally, we update the virtual Doppler grid for the next iteration as $\bar{k}_p^{(i+1)} = \bar{k}_p^{(i)} + \xi_p^{(i+1)}$. Note that the JPNCE-SBL algorithm stops when the parameter change satisfies $\|\boldsymbol{\gamma}^{(i+1)} - \boldsymbol{\gamma}^{(i)}\|_2 / \|\boldsymbol{\gamma}^{(i)}\|_2 < \epsilon$ or reaches the max iterations n_{iter} . After stopping, we find the active set $\hat{\mathcal{P}}$ by keeping indices with $\gamma_p^{(i+1)} > \epsilon$, where ϵ is a small threshold. Then, we count the number of paths as $\hat{P} = |\hat{\mathcal{P}}|$. Finally, we get the channel gains, phase noise, the final delays, and the Doppler parameters. The complete JPNCE-SBL framework is summarized in Algorithm 1.

Algorithm 1 Proposed JPNCE-SBL Algorithm

Input: $\mathbf{r}, \bar{\mathbf{G}}^{(0)}, \mathbf{B}, \epsilon, n_{\text{iter}}$
Initialize: $i = 0, \boldsymbol{\eta}^{(0)} = \mathbf{0}, \boldsymbol{\gamma}^{(0)} = \mathbf{1}, \beta^{(0)} = 1, \boldsymbol{\xi}^{(0)} = \mathbf{0}$.
 Compute initial $\boldsymbol{\mu}^{(0)}$ and $\boldsymbol{\Sigma}^{(0)}$.
 1: **while** $\|\boldsymbol{\gamma}^{(i)} - \boldsymbol{\gamma}^{(i-1)}\|_2 / \|\boldsymbol{\gamma}^{(i-1)}\|_2 \geq \epsilon$ **and** $i < n_{\text{iter}}$ **do**
 2: $i \leftarrow i + 1$
 3: Update hyperparameters $\boldsymbol{\gamma}^{(i)}$ and $\beta^{(i)}$ via (13) using $\boldsymbol{\mu}^{(i-1)}$ and $\boldsymbol{\Sigma}^{(i-1)}$, then extract the active set \mathcal{P} .
 4: Update the reduced-rank PN coefficients $\boldsymbol{\eta}^{(i)}$ via (18) and reconstruct $\mathbf{P}(\boldsymbol{\eta}^{(i)})$.
 5: Compute and clip offsets $\boldsymbol{\xi}_{\mathcal{P}}^{(i)}$ via (23), then update the virtual grid $\bar{k}_p^{(i)} = \bar{k}_p^{(i-1)} + \xi_p^{(i)}, \forall p \in \mathcal{P}$.
 6: Reconstruct the dictionary $\Phi(\bar{\mathbf{G}}^{(i)})$ and evaluate the new posterior statistics $\boldsymbol{\mu}^{(i)}, \boldsymbol{\Sigma}^{(i)}$.
 7: **end while**
 8: **return** $\hat{P} = |\hat{\mathcal{P}}|, \hat{\mathbf{h}}, \hat{\boldsymbol{\eta}}, \hat{\boldsymbol{\ell}}, \hat{\boldsymbol{\nu}}$.

B. Computational Complexity Analysis

The overall complexity of JPNCE-SBL is asymptotically bounded by the E-step. The intermediate operations require much less computation. Specifically, hyperparameter updates, dictionary construction, and grid evolution only contribute lower-order terms such as $\mathcal{O}(NL^2)$ and $\mathcal{O}(|\mathcal{P}|^3)$ [16], [18]. In the E-step, updating the posterior statistics inherently requires inverting an $M_S \times M_S$ matrix. Since the observation dimension is typically smaller than the high-resolution grid size ($N < M_S$), we invoke the Woodbury matrix identity to fundamentally reduce the inversion dimension from M_S to N , lowering the E-step complexity to $\mathcal{O}(N^3 + N^2 M_S)$. Both the phase noise parameters and the active paths exhibit high sparsity ($L \ll N$ and $|\mathcal{P}| \ll M_S$), so the lower-order terms are negligible and the total complexity per iteration reduces to $\mathcal{O}(N^3 + N^2 M_S)$, confirming that PN tracking and grid evolution introduce almost no extra overhead.

IV. SIMULATION RESULTS

The proposed JPNCE-SBL algorithm is evaluated in an AFDM system comprising $N = 64$ subcarriers, with carrier frequency $f_c = 30$ GHz, subcarrier spacing $\Delta f = 15$ kHz, and 4-QAM modulations. The doubly-dispersive channel consists of $P = 3$ independent paths, following the Jakes model with a maximum normalized delay $\ell_{\max} = 12$ and Doppler shift $k_{\max} = 3$. For reliable channel acquisition, five pilot symbols are multiplexed with a 30 dB power boost. The receiver-side PN is modeled as a discrete-time Wiener process [17]. For the SBL configuration, grid resolutions are $r_\tau = r_\nu = 1$, initialized with $\rho = 10^{-2}$, $c = d = 10^{-6}$, and 100 maximum iterations. For data recovery, an MMSE detector is employed. The proposed scheme is benchmarked against Newton-OMP, OffGrid-SBL [12], and Grid Fission SBL [20] under two scenarios: an ideal oscillator (Without PN) and practical PN impairments (With PN). Performance is quantified via normalized mean square error (NMSE) and bit error rate (BER).

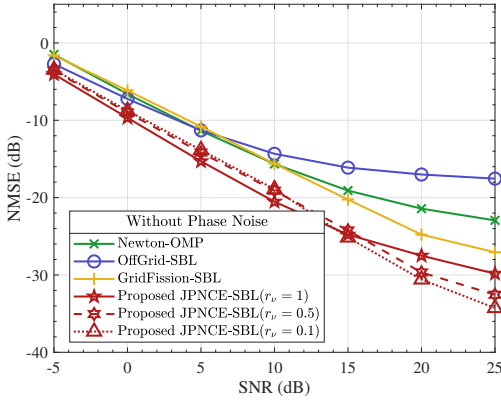


Fig. 2: NMSE performance comparison of different estimators versus SNR under an ideal oscillator scenario.

To evaluate the performance gains of the proposed grid evolution strategy, Fig. 2 presents the NMSE results under an ideal oscillator scenario. It can be observed that Newton-OMP and OffGrid-SBL suffer from severe high-SNR error floors due to discretization errors on fixed grids, while GridFission-SBL achieves the best baseline performance by reducing grid mismatch but still saturates at high SNR. In contrast, the proposed JPNCE-SBL consistently achieves superior accuracy by continuously refining the virtual grid, exhibiting no error floor across the entire SNR range. Furthermore, the NMSE performance of JPNCE-SBL remains stable across different grid resolutions $r_\nu \in \{0.1, 0.5, 1\}$, with only small performance loss as r_ν increases, demonstrating that the proposed grid evolution strategy effectively handles coarse grid initializations. This confirms the robustness of JPNCE-SBL against grid resolution variations and validates the effectiveness of the proposed dynamic grid update mechanism.

Fig. 3 and Fig. 4 evaluate the proposed subspace projection method through the PN tracking MSE and temporal trajectory, respectively. From Fig. 3, an irreducible error floor emerges at high SNRs due to the truncation loss of the finite-rank

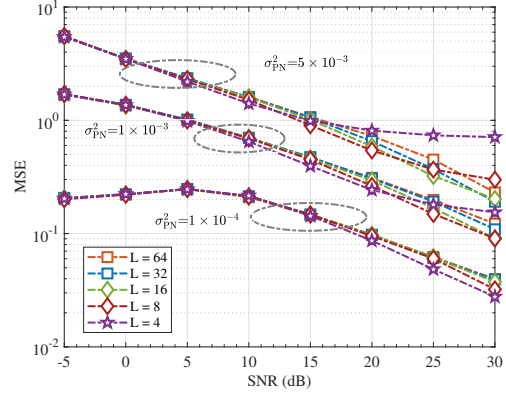


Fig. 3: MSE performance of the proposed subspace projection method for PN estimation under different subspace dimensions L and PN innovation variances.

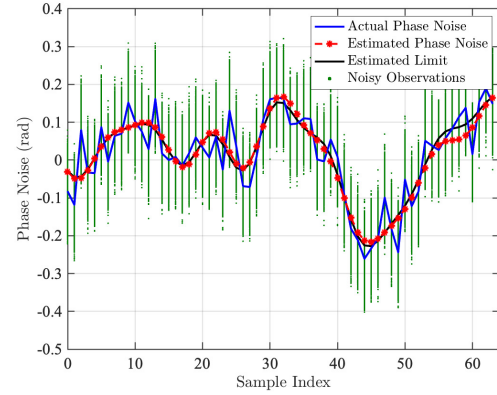


Fig. 4: PN trajectory tracking performance of the proposed subspace projection method.

approximation. It can be observed that a compact subspace of $L = 16$ is sufficient under mild PN conditions, while severe PN requires an expanded dimension to suppress premature saturation, indicating a fundamental trade-off between accuracy and complexity. Furthermore, Fig. 4 demonstrates that the estimated trajectory closely tracks the actual random evolution with high fidelity. This is attributed to the subspace projection eliminating broadband noise while preserving the correlated Wiener structure, enabling precise phase reconstruction even under rapid PN fluctuations.

From Fig. 5, it can be observed that when PN is ignored, all estimators suffer from severe error floors at low SNR (around 5 dB), as the PN-induced phase rotations corrupt the delay-Doppler representation and cause a model mismatch that dominates the thermal noise. When basic PN estimation is introduced, baseline methods such as OffGrid-SBL and GridFission-SBL show only marginal improvement, as they fail to decouple PN variations from the multipath components, leading to early saturation. In contrast, the proposed JPNCE-SBL effectively overcomes this limitation by jointly updating the virtual grid and the PN trajectory, achieving a continuously decreasing NMSE that closely approaches the ideal bound across the entire SNR range.

From Fig. 6, it can be observed that under perfect CSI,

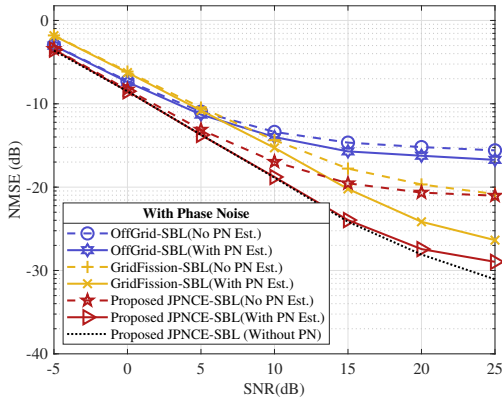


Fig. 5: NMSE performance comparison of different estimators versus SNR in the presence of transceiver phase noise.

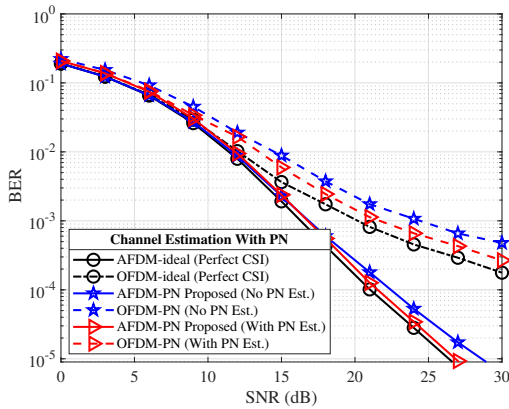


Fig. 6: BER performance comparison of different estimators versus SNR in the presence of transceiver phase noise.

AFDM outperforms OFDM by fully exploiting the delay-Doppler diversity in doubly-dispersive channels. However, when PN is uncompensated, both waveforms suffer from severe BER saturation, erasing the performance gain of AFDM. In contrast, the proposed JPNCE-SBL recovers this gain by providing high-precision CSI, closely approaching the perfect-CSI bound. This confirms that JPNCE-SBL not only effectively compensates for PN impairments but also maintains the superior performance of AFDM in high-mobility scenarios.

V. CONCLUSION

In this paper, we proposed the JPNCE-SBL framework for joint channel estimation and PN tracking in AFDM systems over doubly-dispersive channels. By incorporating a dynamic grid evolution strategy and a reduced-rank subspace projection mechanism, the proposed framework iteratively refines the continuous delay-Doppler parameters while tracking the time-varying PN trajectory with low computational overhead. Within a unified EM-based probabilistic model, the sparse channel hyperparameters, the PN subspace coefficients, and the virtual Doppler grid are simultaneously updated, thereby effectively preventing error propagation between the two estimation processes. Simulation results demonstrate that the proposed JPNCE-SBL achieves superior NMSE and BER perfor-

mance compared to existing benchmarks, closely approaching the perfect-CSI bound under practical PN impairments.

REFERENCES

- [1] W. Saad, M. Bennis, and M. Chen, "A vision of 6G wireless systems: Applications, trends, technologies, and open research problems," *IEEE Network*, vol. 34, no. 3, pp. 134–142, 2020.
- [2] Y. Zhang, Q. Yi, L. Musavian, T. Xu, and Z. Liu, "Non-orthogonal AFDM: A promising spectrum-efficient waveform for 6G high-mobility communications," in *Proc. IEEE 36th Int. Symp. Pers., Indoor Mobile Radio Commun. (PIMRC)*, Istanbul, Turkiye, 2025, pp. 1–6.
- [3] Z. Sui, Z. Liu, L. Musavian, L.-L. Yang, and L. Hanzo, "Generalized spatial modulation aided affine frequency division multiplexing," *IEEE Trans. Wireless Commun.*, early access, 2025, doi: 10.1109/TWC.2025.3613062.
- [4] X. Zhang, H. Yin, Y. Tang, Y. Ge, Y. Zeng, M. Wen, Z. Liu, *et al.*, "A unified multicarrier waveform framework for next-generation wireless networks: Principles, performance, and challenges," *IEEE Commun. Surveys Tuts.*, early access, 2026, doi: 10.1109/COMST.2026.3672602.
- [5] H. S. Rou *et al.*, "From orthogonal time-frequency space to affine frequency-division multiplexing: A comparative study of next-generation waveforms for integrated sensing and communications in doubly dispersive channels," *IEEE Signal Process. Mag.*, vol. 41, no. 5, pp. 71–86, Sep. 2024.
- [6] Y. Tao *et al.*, "Affine frequency division multiplexing with index modulation: Full diversity condition, performance analysis, and low-complexity detection," *IEEE J. Sel. Areas Commun.*, vol. 43, no. 4, pp. 1041–1055, Apr. 2025.
- [7] Z. Sui *et al.*, "MIMO-AFDM outperforms MIMO-OFDM in the face of hardware impairments," *arXiv preprint arXiv:2601.00502*, 2026.
- [8] S. Wu and Y. Bar-Ness, "OFDM systems in the presence of phase noise: Consequences and solutions," *IEEE Trans. Commun.*, vol. 52, no. 11, pp. 1988–1996, Nov. 2004.
- [9] R. Wang, H. Mehrpouyan, M. Tao, and Y. Hua, "Channel estimation and carrier recovery in the presence of phase noise in OFDM relay systems," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Austin, TX, USA, Dec. 2015, pp. 1–6.
- [10] X. Li, Z. Liu, Z. Zhou, and P. Fan, "Matched filtering-based channel estimation for AFDM systems in doubly selective channels," *IEEE Trans. Wireless Commun.*, 2026. [Online]. Available: <https://arxiv.org/abs/2507.0926>
- [11] K. Zheng, M. Wen, T. Mao, L. Xiao, and Z. Wang, "Channel estimation for AFDM with superimposed pilots," *IEEE Trans. Veh. Technol.*, vol. 74, no. 2, pp. 3389–3394, Feb. 2025.
- [12] Z. Wei *et al.*, "Off-grid channel estimation with sparse Bayesian learning for OTFS systems," *IEEE Trans. Wireless Commun.*, vol. 21, no. 9, pp. 7407–7426, Sep. 2022.
- [13] Q. Luo, J. Zhu, Z. Liu, Y. Tang, P. Xiao, G. Chen, and J. Shi, "Joint sparse graph for enhanced MIMO-AFDM receiver design," *IEEE Trans. Wireless Commun.*, vol. 25, pp. 3272–3286, 2026.
- [14] X. Li, P. Fan, Q. Wang, and Z. Liu, "Grid evolution for doubly fractional channel estimation in OTFS systems," *IEEE Trans. Veh. Technol.*, vol. 74, no. 2, pp. 3486–3490, Feb. 2025.
- [15] D. D. Lin and T. J. Lim, "The variational inference approach to joint data detection and phase noise estimation in OFDM," *IEEE Trans. Signal Process.*, vol. 55, no. 5, pp. 1862–1874, May 2007.
- [16] H. Mehrpouyan *et al.*, "Joint estimation of channel and oscillator phase noise in MIMO systems," *IEEE Trans. Signal Process.*, vol. 60, no. 9, pp. 4790–4807, Sep. 2012.
- [17] P. Mathecken, S. Werner, T. Riihonen, and R. Wichman, "Subspace-based phase noise estimation in OFDM receivers," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Apr. 2015, pp. 3227–3231.
- [18] L. Zhao, W.-J. Gao, and W. Guo, "Sparse Bayesian learning of delay-Doppler channel for OTFS system," *IEEE Commun. Lett.*, vol. 24, no. 12, pp. 2766–2769, Dec. 2020.
- [19] Y. Xu *et al.*, "Bayesian compressive sensing for NLOS mmWave imaging under imprecisely multiangle surfaces," *IEEE Signal Process. Lett.*, vol. 32, pp. 2075–2079, 2025.
- [20] R. R. Pote and B. D. Rao, "Maximum likelihood-based gridless DoA estimation using structured covariance matrix recovery and SBL with grid refinement," *IEEE Trans. Signal Process.*, vol. 71, pp. 802–815, Mar. 2023.