Parameter Design for Channel Knowledge Map Assisted Channel Estimation

Muhammad Awais and Yun Hee Kim
Department of Electronics and Information Convergence Engineering, Kyung Hee University, Korea
Emails: mawais@khu.ac.kr, yheekim@khu.ac.kr

***Abstract*—Efficient channel estimation in massive multiple-input multiple-output (mMIMO) with limited pilot resources remains a key challenge. Recently, the concept of a channel knowledge map (CKM) has been proposed to enable location and environment-aware estimation. For a CKM referred to as channel angle-distance map (CADM) proposed for near-field channels, we investigate the sensitivity of the CADM-assisted channel estimation to the number of stored multipath components for a given location. We explore how variations in the number of multipath components in CADM affect the mean squared error (MSE) performance in different signal-to-noise ratio (SNR) conditions. Simulation results reveal that a moderate overestimation of multipath components in CADM can improve the MSE performance in high SNR regimes, whereas excessive multipath data degrades the performance in low SNR regimes.**

***Index Terms*—Channel knowledge map (CKM), channel estimation, massive MIMO, near-field channels**

I. Introduction

Massive multiple-input multiple-output (mMIMO) systems, along with their advanced counterpart, extremely large MIMO (XL-MIMO) have been recognized as key enabling technologies for beyond-fifth generation (B5G) and sixth-generation (6G) mobile communication systems, owing to their ability to enhance the spectral efficiency through highly directional and sharpened beamforming. However, the large-scale antenna arrays significantly increase the complexity of channel estimation, as they demand a proportionally larger number of pilot signals. This results in increased training overhead and reduced spectral efficiency. Consequently, developing efficient channel estimation strategies with reduced pilot overhead has emerged as a critical research direction [1].

To address this challenge, the channel knowledge map (CKM) has recently emerged as a promising approach that leverages learned environmental and location-specific knowledge to assist channel estimation and beamforming [2]. By storing location-dependent information such as angles of arrival, propagation distances, beam indices, and large-scale fading parameters, CKM allows for pilot-efficient estimation by exploiting the fact that large-scale parameters tend to vary slowly with user movement. Building upon this concept, a variant termed the channel angle-distance map (CADM) was recently proposed by the authors to support both near-field and far-field scenarios [3]. CADM stores angle and distance information of dominant multipath components relative to the user or scatterer location in the polar-domain, enabling channel estimation with significantly reduced pilot overhead.

In constructing CADM, the number of stored multipath components per location is a critical design parameter that directly affects both computational complexity and estimation performance. Since the number of multipaths is not known in advance and cannot be perfectly estimated under noisy channels, the choice of the number of multipaths estimated and stored in CADM is non-trivial and its effect on the channel estimation performance has not been examined in prior studies. In this work, we investigate how the number of dominant multipaths stored in CADM influences online channel estimation when leveraging CADM data.

II. System Model

We consider an uplink orthogonal frequency division multiplexing (OFDM) transmission in an mMIMO system, where the base station (BS) is equipped with M antennas and each user is equipped with a single antenna. The uniform spherical wave model is adopted to characterize near-field propagation, with the signal models and notations following [3]. At the BS, the received signals are combined using an analog beamforming matrix $\mathbf{W}_p \in \mathbb{C}^{M_{RF}\times M}$, where M_{RF} denotes the number of RF chains and $|[\mathbf{W}_p]_{ij}| = 1/\sqrt{M}$. The corresponding baseband signal at the s-th subcarrier during the p-th symbol duration is expressed as

$$\mathbf{y}_{s,p} = \mathbf{W}_p \mathbf{h}_s x_{s,p} + \mathbf{W}_p \mathbf{n}_{s,p} \in \mathbb{C}^{M_{RF}\times 1}, \tag{1}$$

where $x_{s,p}$ and $\mathbf{n}_{s,p} \sim \mathcal{CN}(0, \sigma^2 \mathbf{I}_M)$ denote the known pilot symbol and the complex Gaussian noise vector, respectively, for $s = 1, 2, \cdots, S$ and $p = 1, 2, \cdots, P$. Here, S represents the total number of subcarriers in the system, and P denotes the number of pilot symbols used for channel estimation.

Under the assumption of a quasi-static channel, the channel response at the s-th subcarrier of frequency f_s is given by

$$\mathbf{h}_s = \sqrt{\frac{M}{L}} \sum_{l=1}^{L} e^{-j\frac{2\pi f_s}{c} r_l} g_l \mathbf{b}(\theta_l, r_l), \;\; s = 1, 2, \cdots, S, \tag{2}$$

where g_l, θ_l, and r_l denote the channel fading gain, angle of arrival (AoA), and propagation distance of the lth path, respectively, and L is the number of multipaths. The vector $\mathbf{b}(\theta_l, r_l)$ is the near-field steering vector with the mth entry

$$[\mathbf{b}(\theta_l, r_l)]_m = \frac{1}{\sqrt{M}} e^{-j\frac{2\pi}{\lambda}(\|\mathbf{q}_l - \mathbf{c}_m\| - r_l)} \tag{3}$$

where λ is the wavelength, $\mathbf{q}_l$ is the position vector of the scatterer corresponding to the l-th path, and $\mathbf{c}_m$ is the position vector of the m-th BS antenna element.

III. Stored Multipath Selection in CADM

The CADM associates each user location with its long-term multipath parameters, AoA, and distance information to support both near-field and far-field channels. For a given location $\mathbf{q}$ of a user, the ground-truth path information

$$\mathbf{\Omega} = \{(\theta_1, r_1), (\theta_2, r_2), \cdots, (\theta_L, r_L)\} \tag{4}$$

is estimated by a near-field channel estimation method [4] as

$$\tilde{\mathbf{\Omega}} \triangleq \left\{(\tilde{\theta}_l, \tilde{r}_l) \mid l = 1, 2, \cdots, \tilde{L}\right\} \tag{5}$$

to construct the CADM. Here, $\tilde{L}$ is the number of stored multipaths in CADM, which is an important design parameter that reduces the CADM construction time and memory at the cost of information loss for $\tilde{L} < L$. A larger $\tilde{L} > L$ contains the full path information, but includes unnecessary noise components and induces increased computation complexity.

For channel estimation, we express

$$\mathbf{h}_s = \mathbf{B}(\boldsymbol{\theta}, \mathbf{r})\mathbf{\Phi}_s(\mathbf{r})\mathbf{g}, \tag{6}$$

where $\boldsymbol{\theta} = [\theta_1, \theta_2, \cdots, \theta_L]^T$, $\mathbf{r} = [r_1, r_2, \cdots, r_L]^T$, $\mathbf{g} = [g_1, g_2, \cdots, g_L]^T$, and the others are defined in [3]. Then, the maximum likelihood (ML) estimation is formulated as

$$\min_{\boldsymbol{\theta}, \mathbf{r}, \mathbf{g}} \sum_{p=1}^{P} \sum_{s=1}^{S} \|\mathbf{y}_{s,p} - \mathbf{W}_p \mathbf{B}(\boldsymbol{\theta}, \mathbf{r})\mathbf{\Phi}_s(\mathbf{r})\mathbf{g}x_{s,p}\|^2. \tag{7}$$

If $(\boldsymbol{\theta}, \mathbf{r})$ is known perfectly, the ML estimate of $\mathbf{g}$ is given by the least squares (LS) solution. Even if the CADM provides an imperfect information $(\tilde{\boldsymbol{\theta}}, \tilde{\mathbf{r}})$, the LS estimate of $\mathbf{g}$ with imperfect information can be obtained, which is referred to as CADM-LS. To further improve the channel estimation, we can find an improved solution $(\tilde{\boldsymbol{\theta}}^{opt}, \tilde{\mathbf{r}}^{opt})$ utilizing a gradient descent-based algorithm as proposed in [3] using the CADM data as an initial point, which is called CADM-Optimized.

IV. Simulation Results

For performance evaluation, the true sparse channel is constructed with a fixed number of paths $L = 6$, where g_l is Rayleigh fading, $\theta_l \sim \mathcal{U}(0, 2\pi)$, and $r_l \sim \mathcal{U}(5, 10)$ meters. The polar-domain simultaneous orthogonal matching pursuit (P-SOMP) using 64 pilot symbols was utilized for CADM construction with different $\tilde{L}$ values. The other simulation parameters follow [3].

We provide the normalized minimum mean square error (NMSE) performance of the CADM-assisted channel estimation methods using 32 pilot symbols in Fig. 1. The performance is compared for $\tilde{L} \in \{3, 6, 12\}$ as the SNR increases. CADM-Optimized with $\tilde{L} = L = 6$ provides the best performance, while the CADM-LS counterpart exhibits an error floor in high SNR regimes due to the CADM data errors. Both CADM-based methods with $\tilde{L} > L$ exhibit a better performance than those with $\tilde{L} < L$ in high SNR regimes,

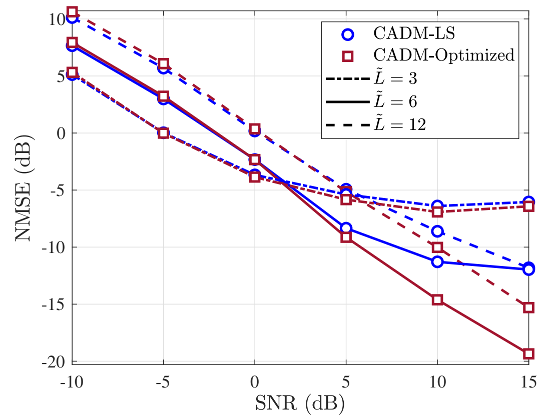

Fig. 1: NMSE vs SNR having $\tilde{L} \in \{3, 6, 12\}$ when $P = 8$.

while the vice versa is observed in low SNR regimes. In low SNR regimes, the performance degradation due to the noise components is more critical, so channel estimation with $\tilde{L} = 3$ outperforms that with $\tilde{L} = 6$. In contrast, in high SNR regimes, channel estimation with $\tilde{L} = 12$ outperforms $\tilde{L} = 3$ by containing all multipaths, where inclusion of extra multipaths without signal components does not significantly deteriorate performance. Therefore, it is desirable to construct the CADM with a larger $\tilde{L}$ and allow online CADM-based channel estimation methods to adapt $\tilde{L}$ values according to the SNR level.

V. Concluding Remarks

This paper explored the effect of the number of dominant multipaths stored in CADM on the CADM-assisted channel estimation methods. The results demonstrated a trade-off between robustness and accuracy, governed by the number of estimated multipaths. The results show that careful or adaptive selection of $\tilde{L}$ is vital for achieving optimal estimation performance in practical deployments, especially under constraints on pilot length or in dynamic SNR conditions.

VI. Acknowledgment

This work was supported by the National Research Foundation of Korea (NRF) under Grant RS-2025-16-067576 and by the Institute for Information & Communications Technology Planning & Evaluation (IITP) under the Information Technology Research Center (ITRC) support program (IITP-2025-RS-2021-II212046), funded by the Ministry of Science and ICT (MSIT), Korea.

References

[1] H. Lu, Y. Zeng, C. You, Y. Han, J. Zhang, Z. Wang, Z. Dong, S. Jin, C.-X. Wang, T. Jiang, X. You, and R. Zhang, “A tutorial on near-field XL-MIMO communications toward 6G,” *IEEE Commun. Surveys Tuts.*, vol. 26, no. 4, pp. 2213–2257, 2024.

[2] Y. Zeng, J. Chen, J. Xu, D. Wu, X. Xu, S. Jin, X. Gao, D. Gesbert, S. Cui, and R. Zhang, “A tutorial on environment-aware communications via channel knowledge map for 6G,” *IEEE Commun. Surveys Tuts.*, vol. 26, no. 3, pp. 1478–1519, Feb. 2024.

[3] M. Awais and Y. H. Kim, “Channel knowledge map assisted channel estimation for an extremely large antenna array system,” *IEEE Access*, vol. 13, pp. 158353–158365, Sept. 2025.

[4] M. Cui and L. Dai, “Channel estimation for extremely large-scale MIMO: Far-field or near-field?,” *IEEE Trans. Commun.*, vol. 70, no. 4, pp. 2663–2677, Jan. 2022.