Block-Level Interference Exploitation Precoding for MU-MISO: An ADMM Approach

Yiran Wang, Yunsi Wen, Ang Li, *Senior Member, IEEE*, Xiaoyan Hu, *Member, IEEE*, and Christos Masouros, *Fellow, IEEE*

***Abstract*—We study constructive interference based block-level precoding (CI-BLP) in the downlink of multi-user multiple-input single-output (MU-MISO) systems. Specifically, our aim is to extend the analysis on CI-BLP to the case when the number of symbol slots in a given transmission block is smaller than the number of users. To this end, we mathematically prove the feasibility of using the pseudo-inverse to obtain a closed-form structure of the optimal CI-BLP precoding matrix. Similar to the case when the number of symbol slots in a given transmission block is not smaller than the number of users, we show that a quadratic programming (QP) optimization on simplex can be constructed. We also design a low-complexity algorithm based on the alternating direction method of multipliers (ADMM) framework, which can achieve a flexible trade-off between communication performance and execution time by modifying the maximum number of iterations. We further analyze the convergence and complexity of the proposed algorithm. Numerical results validate our analysis and the optimality of the QP optimization, and further show that the proposed ADMM algorithm can provide satisfactory results in dozens of iterations, which motivates the use of CI-BLP in practical wireless systems.**

***Index Terms*—MIMO, constructive interference (CI), block-level precoding (BLP), quadratic programming (QP) optimization, alternating direction method of multipliers (ADMM).**

I. Introduction

PRECODING has been widely studied in multiple-input multiple-output (MIMO) communication systems, which is able to support data transmission to multiple users simultaneously [1]. In the downlink, if the channel state information (CSI) is fully known to the base station, dirty paper coding (DPC) can achieve the best performance by pre-subtracting interference before transmission [2], but its prohibitive computational costs make it difficult to implement in practical systems. Therefore, low-complexity closed-form linear precoding schemes, represented by zero-forcing (ZF) [3] and regularized ZF (RZF) [4], are proposed to reduce the computational complexity in signal processing. At the same time, optimization-based precoding schemes are gaining more and more attention because they allow precoding to better meet various communication constraints and requirements in different scenarios. One popular example is the downlink signal-to-interference-plus-noise ratio (SINR) balancing approach, which aims to achieve a desired SINR for each user under transmit power constraints [5]. Another popular form is to minimize the transmit power under the SINR constraint of each user [6]. [7] proves that the SINR balancing and the power minimization problems are dual problem to each other, where an effective iterative algorithm is proposed by exploiting such duality to efficiently solve these two problems.

More recent research has shown that multi-user interference need not be completely eliminated. This is because interference can be utilized by interference exploitation precoding techniques to enhance the power of useful signals and benefit symbol detection, thus further improving the error-rate performance of MIMO communication systems. In [8], the concept of 'constructive interference' (CI) is introduced, and CI-based precoding has received increasing research attention [9]–[12]. In [13], multi-user interference is strictly aligned with the desired data symbol, and a CI-based maximum ratio transmission (MRT) precoding design is carried out to achieve improved performance. This approach was later shown to be sub-optimal and referred to as the 'strict phase-rotation' CI metric. The concept of 'constructive region' is introduced in [14] and [15], which shows that CI does not have to be strictly aligned with the desired data symbol, and as long as the interfered signals lie in the constructive region, the effect of interference is constructive. This observation alleviates the requirement that the interfering signals have to be strictly rotated to the direction of the intended data symbols, leading to further performance improvements. The CI metric introduced in [14] was later named the 'non-strict phase-rotation' CI metric and is widely adopted in the subsequent literature. Meanwhile,

Manuscript received December 19, 2023; revised June 11, 2024, and August 11, 2024; accepted October 05, 2024. The work of Ang Li was supported in part by the Young Elite Scientists Sponsorship Program by CIC (Grant No. 2021QNRC001), in part by the National Natural Science Foundation of China under Grant 62101422, 62371386, in part by the Science and Technology Program of Shaanxi Province under Grant 2024JC-JCQN-59, and in part by the open research fund of National Mobile Communications Research Laboratory, Southeast University (No. 2024D01). The work of Xiaoyan Hu was supported in part by the NSFC under Grant 62201449, in part by the Young Elite Scientists Sponsorship Program by CAST under Grant No.YESS20230611, and in part by the Key R&D Projects of Shaanxi Province under Grant 2023-YBGY-040. An earlier version of this paper was presented in part at the IEEE WCNC, Dubai, United Arab Emirates, April 2024. The associate editor coordinating the review of this paper and approving it for publication was Prof. Namyoon Lee. (Corresponding authors: Ang Li, Xiaoyan Hu.)

Y. Wang, Y. Wen and X. Hu are with the School of Information and Communications Engineering, Faculty of Electronic and Information Engineering, Xi'an Jiaotong University, Xi'an, Shaanxi 710049, China (e-mail:{yiranwang, yunsiwen}@stu.xjtu.edu.cn, xiaoyanhu@xjtu.edu.cn).

A. Li is with the School of Information and Communications Engineering, Faculty of Electronic and Information Engineering, Xi'an Jiaotong University, Xi'an, Shaanxi 710049, China, and is also with the National Mobile Communications Research Laboratory, Southeast University, China. (e-mail: ang.li.2020@xjtu.edu.cn)

C. Masouros is with the Department of Electronic and Electrical Engineering, University College London, Torrington Place, London, WC1E 7JE, UK (e-mail: c.masouros@ucl.ac.uk).

a relaxed CI metric based on a 'relaxed detection region' was introduced in [15], which expands the constructive region based on a phase margin that is related to the signal-to-noise ratio (SNR) target. The above CI-based precoding approaches are all designed for PSK modulation, while [12] was the first to extend the exploitation of CI to QAM modulation, where the CI effect can be exploited by the outer constellation points of a QAM constellation by employing the 'symbol-scaling' CI metric. Due to the significant advantages of CI, symbol-level precoding (SLP) based on CI has been applied to intelligent reflecting surface (IRS)-assisted communication [16], [17], 1-bit precoding [18]–[22], radar-communication coexistence [23], [24] and many other wireless communication scenarios. For high-order QAM modulation, [25] gives the expression of symbol error rate (SER) with respect to the transmitted signal and the rescaling factor based on the noise distribution, and constructs the optimization problem to minimize SER. This scheme can achieve better SER. However, the complexity of this scheme is high, especially when the number of transmit antennas is larger than the number of users.

CI-SLP as its name indicates, requires the base station to perform different precoding for each symbol slot, where different precoding optimization problems need to be solved per symbol. However, symbol-by-symbol optimization brings significant computational burden to the signal processing unit and requires high real-time processing capability. To alleviate the computational costs, several studies attempt to reduce the complexity of the CI-SLP optimization problem, including derivations of the optimal precoding structure of CI-SLP with efficient iterative algorithms [26], [27], sub-optimal solutions [28], [29], and deep learning-based methods [30]–[32]. Specifically, [26] and [27] derive the optimal precoding structure of CI-SLP for PSK and QAM modulation, respectively, and show that the CI-SLP optimization problem can equivalently be transformed into a quadratic programming (QP) optimization problem and solved using an iterative algorithm with a closed-form solution at each step. Building upon this, the work in [28] derives an exact closed-form but sub-optimal solution for the power minimization CI-SLP problem. Despite the above attempts to reduce the computational costs of solving the CI-SLP optimization problem for each symbol slot, these approaches still require solving precoding problems at the symbol level, i.e., the total number of CI-SLP optimization problems that needs to be solved in a channel coherence interval is not reduced. In order to further motivate the realization of CI-based precoding techniques in practical wireless communication systems, [33] proposed CI-based block-level precoding (CI-BLP) for multi-user multiple-input single-output (MU-MISO) communication system for the first time. Compared to CI-SLP approaches that optimize the precoding matrix (or the precoded signals) on a symbol level, CI-BLP scheme applies a constant precoding matrix to a block of symbol slots and the optimization only needs to be performed once per block of symbol slots, which can greatly reduce the update frequency of precoder and leave more computing resources for each optimization. Based on the Lagrange function and KKT conditions, a closed-form structure of the optimal CI-BLP precoding matrix is derived when the number of symbol slots in a block is not smaller than the number of users. By further studying the corresponding duality problem, the original optimization problem is transformed into a quadratic programming (QP) optimization on simplex.

The CI-BLP optimization problem has been fully studied in [33] when the number of symbol slots in the considered block is less than the number of users. We find that the CI-BLP scheme even outperforms CI-SLP when the number of symbol slots in a block is small, and in high-mobility scenarios, the channel coherence interval would become short, and the number of symbol slots in a block is limited to a small range [34]. These reasons encourage us to pay more attention to the CI-BLP optimization problem when the number of symbols in the considered block is smaller than the number of users. In fact, the supplement of this case completes the research of CI-BLP, which makes the application of CI-BLP more extensive. However, it is still unclear whether a similar QP problem exists for the case when the number of symbols in the considered block is smaller than the number of users. This is because the closed-form structure of the optimal CI-BLP precoding matrix in [33] cannot be directly applied to the above scenarios. Moreover, despite the fact that CI-BLP and the traditional CI-SLP method share a similar QP problem structure, it is still unclear whether the iterative closed-form algorithm proposed in CI-SLP [26] can offer complexity benefits when used to solve the QP problem for CI-BLP.

Therefore in this paper, we aim to extend the analysis on CI-BLP to the case where the number of symbol slots in the considered block is smaller than that of the users, and propose an efficient algorithm to obtain the optimal or sub-optimal CI-BLP precoding matrix. For clarity, we summarize the main contributions of the paper below:

1) We extend the analysis on CI-BLP to the case where the number of symbol slots in the considered block is smaller than that of the users, where the results in [33] are not directly applicable. Specifically, we mathematically prove the feasibility of using the pseudo-inverse to obtain a closed-form structure of the optimal CI-BLP precoding matrix as a function of the dual variables, which does not always hold in other cases. Building upon this, a similar QP optimization on simplex can also be constructed. This work thus generalizes and unifies the optimal precoding structure for CI-BLP.
2) To validate whether the existing iterative closed-form algorithm designed for CI-SLP can be directly applied to CI-BLP, we study the rank of the quadratic coefficient matrix of the formulated QP problem for CI-BLP mathematically. This work is important because the application of the iterative closed-form algorithm requires the quadratic coefficient matrix to be invertible. We show that the existing algorithm is applicable only when the number of symbol slots in a block is smaller than the number of users. However, the numerical convergence rate is not promising for CI-BLP.
3) To design an efficient CI-BLP algorithm with the considered symbol block of any length, we leverage the ADMM framework and obtain the closed-form solution of each subproblem within the ADMM iteration. In addition,

based on the equivalent transformation of the original QP problem, an improved ADMM algorithm is proposed, which can achieve a flexible trade-off between communication performance and execution time by modifying the maximum number of iterations. We analyze the convergence and complexity of the improved ADMM algorithm. Compared with the conventional ADMM algorithm, the updated variables can be obtained in a simpler way in the improved ADMM algorithm, and better performance can be obtained with fewer iterations.

Simulation results show that the QP optimization problem with the proposed precoding structure of CI-BLP solved by interior-point method (IPM) [35] offers the same result as the original optimization problem solved by CVX with the considered symbol block of any length. Without the need for sophisticated initialization and parameter optimization, the proposed ADMM algorithm can provide satisfactory results in dozens of iterations, which is much faster than IPM, which motivates the use of the block-level CI beamforming in practice.

The remainder of this paper is organized as follows. Section II introduces the system model and the problem fomulation of CI-BLP. Section III extend the analysis on CI-BLP to the case where the number of symbol slots in the considered block is smaller than that of the users. The discussion on whether the iterative closed-form algorithm in CI-SLP can be used directly in CI-BLP is given in Section IV, and the proposed ADMM algorithm is introduced in Section V. Numerical results are provided in Section VI, and Section VII concludes the paper.

Notations: Herein, lowercase, boldface lowercase and boldface uppercase letters denote scalars, vectors and matrices, respectively. $\mathbb{R}$ and $\mathbb{C}$ denote the set of real numbers and the set of complex numbers, respectively. Superscripts $^{\mathrm{T}}$ denotes the transpose. The operator $\|\cdot\|_2$ denotes the 2-norm of a vector. $\Re\{\cdot\}$ and $\Im\{\cdot\}$ extract the real and imaginary parts of the argument, respectively. $\mathrm{Diag}(\mathbf{x})$ represents a diagonal matrix with the elements of $\mathbf{x}$ as diagonal elements. We use $\Pi_{\mathbf{\Omega}}$ to represent the projection of the argument onto the set $\mathbf{\Omega}$. We define $\mathbf{I}_N$ as the $N \times N$ identity matrix, and $\mathbf{O}_N$ as the $N \times N$ all-zero matrix. Finally, $\mathbf{1}$ and $\mathbf{0}$ denote all-one vector and all-zero vector, respectively.

II. System Model and Preliminaries

A. System Model

We consider the generic multi-user multiple-input single-output (MU-MISO) downlink system, where a base station (BS) equipped with N_t antennas serves K single-antenna users simultaneously. For the transmission of a block of symbol slots, the data symbol vector in the n-th slot is denoted by $\mathbf{s}^n = [s_1^n, s_2^n, \cdots, s_K^n]^{\mathrm{T}} \in \mathbb{C}^K$, which is assumed to be drawn from a unit-norm $\mathcal{M}$-PSK constellation[1]. Accordingly, the received signal for user k in the n-th symbol slot can be expressed as

$$y_k^n = \mathbf{h}_k^{\mathrm{T}} \mathbf{W} \mathbf{s}^n + z_k^n, \tag{1}$$

where $\mathbf{h}_k \in \mathbb{C}^{N_t}$ is the channel between BS and user k, which is constant within the considered block, and $z_k^n \in \mathbb{C}$ is the additive noise of user k in the n-th symbol slot. $\mathbf{W} \in \mathbb{C}^{N_t \times K}$ is the precoding matrix that applies to all $\mathbf{s}^n$ in the considered block. $n \in \{n | n \leq N\}$, where N represents the length of the considered block which may be smaller than the channel coherence interval. An illustration of CI-BLP method is shown in Fig. 1.

B. Symbol-Scaling CI Metric

Traditionally, interference is usually viewed as a performance limiting factor in wireless communication systems. CI-based precoding is designed to exploit inter-user interference rather than eliminate it. The inter-user interference is used to enhance the power of useful signals through CI-based precoding design, thereby improving SER performance. To illustrate the symbol-scaling CI metric introduced in [36], below we depict one quarter of an 8PSK constellation in Fig. 2 as an example. Without loss of generality, we assume that $\overrightarrow{OA}$ is the nominal constellation point for user k in the n-th slot, i.e., $\overrightarrow{OA} = s_k^n$. $\overrightarrow{OB}$ represents the noiseless received signal with interference, where based on the geometry we obtain $\overrightarrow{OB} = \overrightarrow{OA} + \overrightarrow{AB} = \mathbf{h}_k^{\mathrm{T}} \mathbf{W} \mathbf{s}^n$, where $\overrightarrow{AB}$ can be regarded as the sum interference from other user streams.

Different from the common phase-rotation CI metric which uses phase relations, the symbol-scaling CI metric decomposes the signal along the decision boundaries and imposes scaling constraints on the decomposed components. In Fig. 2, $\overrightarrow{OA}$ is decomposed along the two decision boundaries for 8PSK modulation to obtain $\overrightarrow{OD}$ and $\overrightarrow{OE}$:

$$\overrightarrow{OA} = \overrightarrow{OD} + \overrightarrow{OE} = s_{k,\text{right}}^n + s_{k,\text{left}}^n. \tag{2}$$

Following a similar procedure, the received signal $\overrightarrow{OB}$ can also be decomposed along the two decision boundaries into

$$\overrightarrow{OB} = \overrightarrow{OF} + \overrightarrow{OG} = \alpha_{k,\text{right}}^n s_{k,\text{right}}^n + \alpha_{k,\text{left}}^n s_{k,\text{left}}^n. \tag{3}$$

[1]For the extension to QAM modulation, see [33].

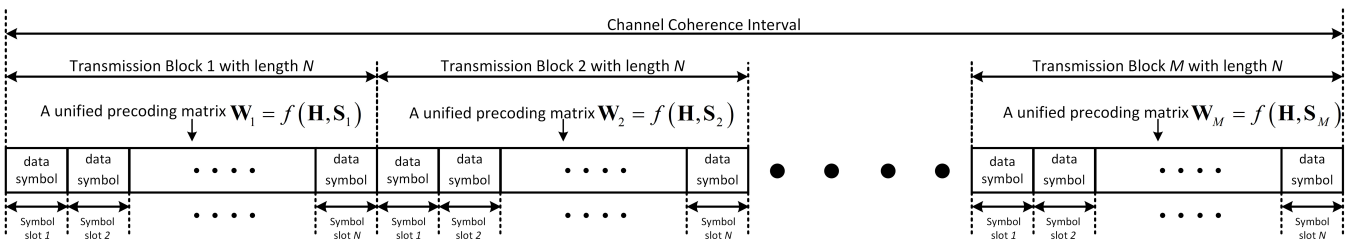

Fig. 1. An illustration of CI-BLP method.

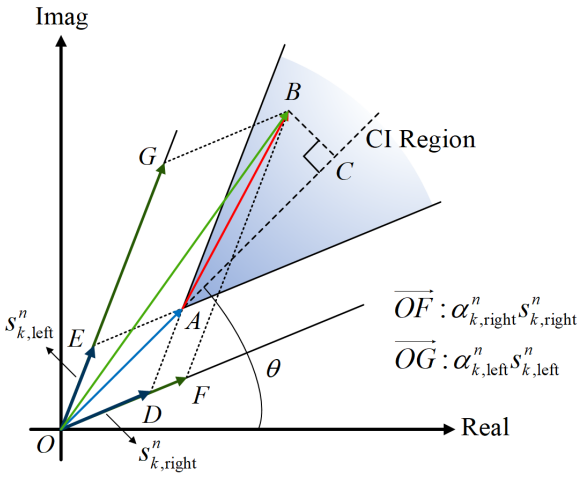

Fig. 2. Geometric diagram of the symbol-scaling CI metric for 8PSK.

where $\alpha^n_{k,\text{right}}$ and $\alpha^n_{k,\text{left}} > 0$ are non-negative scaling factors. By following the transformation in Section IV-A of [19], which we omit in this paper due to the limited space and also for brevity, we can construct a coefficient matrix $\mathbf{M}^n \in \mathbb{R}^{2K\times 2N_t}$ and obtain:

$$\boldsymbol{\alpha}^n_{\text{E}} = \mathbf{M}^n \mathbf{W}_{\text{E}} \mathbf{s}^n_{\text{E}}, \tag{4}$$

where $\boldsymbol{\alpha}^n_{\text{E}} = \left[\alpha^n_{1,\text{right}}, \cdots, \alpha^n_{K,\text{right}}, \alpha^n_{1,\text{left}}, \cdots, \alpha^n_{K,\text{left}}\right]^{\text{T}} \in \mathbb{R}^{2K}$, $\mathbf{W}_{\text{E}} \in \mathbb{R}^{2N_t\times 2K}$ and $\mathbf{s}^n_{\text{E}} \in \mathbb{R}^{2K}$ are defined as

$$\mathbf{W}_{\text{E}} = \begin{bmatrix} \Re(\mathbf{W}) & -\Im(\mathbf{W}) \\ \Im(\mathbf{W}) & \Re(\mathbf{W}) \end{bmatrix}, \ \mathbf{s}^n_{\text{E}} = \left[\Re\left(\mathbf{s}^n\right)^{\text{T}}, \Im\left(\mathbf{s}^n\right)^{\text{T}}\right]^{\text{T}}. \tag{5}$$

C. Problem Formulation for CI-BLP

Recalling Fig. 2, we can observe that the value of $\alpha^n_{k,\text{right}}$ or $\alpha^n_{k,\text{left}}$ represents the effect of inter-user interference, and a larger value of $\alpha^n_{k,\text{right}}$ or $\alpha^n_{k,\text{left}}$ means that the symbol s^n_k is pushed further away from one of its decision boundaries. Given the same noise at the receiver side, the signals that are located further away from the decision boundaries are more likely to be correctly demodulated. Accordingly, we can then construct the CI-BLP optimization problem that maximizes the minimum in $\boldsymbol{\alpha}^n_{\text{E}}$ for all the considered symbol slots within the block, so as to improve the average symbol error rate performance, given by

$$\begin{aligned} \mathcal{P}_0: \ & \max_{\mathbf{W}_{\text{E}}} \ \min_{k,n} \ \alpha^n_k \\ & \text{s.t.} \ \ \boldsymbol{\alpha}^n_{\text{E}} = \mathbf{M}^n \mathbf{W}_{\text{E}} \mathbf{s}^n_{\text{E}}, \ \forall n \le N, \\ & \quad \sum_{n=1}^{N} \|\mathbf{W}_{\text{E}} \mathbf{s}^n_{\text{E}}\|_2^2 \le N p_0, \end{aligned} \tag{6}$$

where α^n_k represents the k-th entry in $\boldsymbol{\alpha}^n_{\text{E}}$, and p_0 represents the transmit power budget per symbol slot.

D. QP Formulation for the Case of $N \ge K$

$\mathcal{P}_0$ is a joint optimization problem over all symbol slots within the considered block, and it is a convex problem that can be directly solved via optimization tools such as CVX. To motivate our extension to the case of $N < K$ in Section III, in this section we review the process of transforming $\mathcal{P}_0$ into a QP problem for $N \ge K$ [33].

We first introduce $\hat{\mathbf{W}} \in \mathbb{R}^{N_t\times 2K}$:

$$\hat{\mathbf{W}} = \left[\Re(\mathbf{W}) \ \ -\Im(\mathbf{W})\right], \tag{7}$$

based on which $\mathbf{W}_{\text{E}} = \mathbf{P}\hat{\mathbf{W}} + \mathbf{Q}\hat{\mathbf{W}}\mathbf{T}$, where $\mathbf{P} \in \mathbb{R}^{2N_t\times N_t}$, $\mathbf{Q} \in \mathbb{R}^{2N_t\times N_t}$ and $\mathbf{T} \in \mathbb{R}^{2K\times K}$ are defined as

$$\mathbf{P} = \begin{bmatrix} \mathbf{I}_{N_t} \\ \mathbf{O}_{N_t} \end{bmatrix}, \ \mathbf{Q} = \begin{bmatrix} \mathbf{O}_{N_t} \\ \mathbf{I}_{N_t} \end{bmatrix}, \ \mathbf{T} = \begin{bmatrix} \mathbf{O}_K & \mathbf{I}_K \\ -\mathbf{I}_K & \mathbf{O}_K \end{bmatrix}. \tag{8}$$

The expression for $\boldsymbol{\alpha}^n_{\text{E}}$ can be further transformed into:

$$\begin{aligned} \boldsymbol{\alpha}^n_{\text{E}} &= \mathbf{M}^n \mathbf{W}_{\text{E}} \mathbf{s}^n_{\text{E}} = \mathbf{M}^n \left(\mathbf{P}\hat{\mathbf{W}} + \mathbf{Q}\hat{\mathbf{W}}\mathbf{T}\right) \mathbf{s}^n_{\text{E}} \\ &= \mathbf{M}^n \mathbf{P}\hat{\mathbf{W}}\mathbf{s}^n_{\text{E}} + \mathbf{M}^n \mathbf{Q}\hat{\mathbf{W}}\mathbf{T}\mathbf{s}^n_{\text{E}} \\ &= \mathbf{A}^n \hat{\mathbf{W}}\mathbf{s}^n_{\text{E}} + \mathbf{B}^n \hat{\mathbf{W}}\mathbf{c}^n_{\text{E}}, \end{aligned} \tag{9}$$

where $\mathbf{A}^n \in \mathbb{R}^{2K\times N_t}$, $\mathbf{B}^n \in \mathbb{R}^{2K\times N_t}$, $\mathbf{c}^n_{\text{E}} \in \mathbb{R}^{2K}$ are defined as

$$\mathbf{A}^n = \mathbf{M}^n \mathbf{P}, \ \mathbf{B}^n = \mathbf{M}^n \mathbf{Q}, \ \mathbf{c}^n_{\text{E}} = \mathbf{T}\mathbf{s}^n_{\text{E}}. \tag{10}$$

Accordingly, the k-th entry of $\boldsymbol{\alpha}^n$ can be expressed as

$$\alpha^n_k = \left(\mathbf{a}^n_k\right)^{\text{T}} \hat{\mathbf{W}}\mathbf{s}^n_{\text{E}} + \left(\mathbf{b}^n_k\right)^{\text{T}} \hat{\mathbf{W}}\mathbf{c}^n_{\text{E}}, \tag{11}$$

where $\left(\mathbf{a}^n_k\right)^{\text{T}}$ and $\left(\mathbf{b}^n_k\right)^{\text{T}}$ represent the k-th row of $\mathbf{A}^n$ and $\mathbf{B}^n$, respectively. With the above expression, $\mathcal{P}_0$ can be expressed as a standard convex optimization problem:

$$\begin{aligned} \mathcal{P}_1: \ & \min_{\hat{\mathbf{W}}, t} \ -t \\ \text{s.t.} \ \ & t - \left(\mathbf{a}^n_k\right)^{\text{T}} \hat{\mathbf{W}}\mathbf{s}^n_{\text{E}} - \left(\mathbf{b}^n_k\right)^{\text{T}} \hat{\mathbf{W}}\mathbf{c}^n_{\text{E}} \le 0, \ \forall k \le 2K, \ n \le N, \\ & \sum_{n=1}^{N} \| \left(\mathbf{P}\hat{\mathbf{W}} + \mathbf{Q}\hat{\mathbf{W}}\mathbf{T}\right) \mathbf{s}^n_{\text{E}} \|_2^2 \le N p_0. \end{aligned} \tag{12}$$

The Lagrangian of $\mathcal{P}_1$ can be constructed as shown in (13) on the top of the next page, where $\boldsymbol{\delta}^n = [\delta^n_1, \delta^n_2, \cdots, \delta^n_{2K}]^{\text{T}}$ and μ are the non-negative dual variables associated with two inequality constraints of $\mathcal{P}_1$ respectively, and we note that

$$\mathbf{P}^{\text{T}}\mathbf{P} = \mathbf{Q}^{\text{T}}\mathbf{Q} = \mathbf{I}_{N_t}, \ \mathbf{P}^{\text{T}}\mathbf{Q} = \mathbf{Q}^{\text{T}}\mathbf{P} = \mathbf{0}. \tag{14}$$

Accordingly, the KKT conditions for the optimality of $\mathcal{P}_1$ can be formulated and shown in (15). Based on the KKT conditions, we first obtain that $\mu > 0$, otherwise $\boldsymbol{\delta}^n = \mathbf{0}$, $\forall n$, which contradicts with (15a). This means that the power constraint is active when the optimality is achieved, i.e.,

$$\sum_{n=1}^{N} \left(\mathbf{s}^n_{\text{E}}\right)^{\text{T}} \hat{\mathbf{W}}^{\text{T}}\hat{\mathbf{W}}\mathbf{s}^n_{\text{E}} + \sum_{n=1}^{N} \left(\mathbf{c}^n_{\text{E}}\right)^{\text{T}} \hat{\mathbf{W}}^{\text{T}}\hat{\mathbf{W}}\mathbf{c}^n_{\text{E}} = N p_0. \tag{16}$$

To proceed, we transform (15b) into

$$2\mu\hat{\mathbf{W}}\mathbf{D} = \sum_{n=1}^{N} \left[\left(\mathbf{A}^n\right)^{\text{T}} \boldsymbol{\delta}^n \left(\mathbf{s}^n_{\text{E}}\right)^{\text{T}} + \left(\mathbf{B}^n\right)^{\text{T}} \boldsymbol{\delta}^n \left(\mathbf{c}^n_{\text{E}}\right)^{\text{T}}\right], \tag{17}$$

where $\mathbf{D} \in \mathbb{R}^{2K\times 2K}$ is given by

$$\mathbf{D} = \left[\sum_{n=1}^{N} \mathbf{s}^n_{\text{E}} \left(\mathbf{s}^n_{\text{E}}\right)^{\text{T}} + \sum_{n=1}^{N} \mathbf{c}^n_{\text{E}} \left(\mathbf{c}^n_{\text{E}}\right)^{\text{T}}\right]. \tag{18}$$

$$
\begin{aligned}
&L\left(\hat{\mathbf{W}}, t, \delta_k^n, \mu\right) \\
&= -t + \sum_{n=1}^{N}\sum_{k=1}^{2K} \delta_k^n \left[t - \left(\mathbf{a}_k^n\right)^{\mathrm{T}} \hat{\mathbf{W}} \mathbf{s}_{\mathrm{E}}^n - \left(\mathbf{b}_k^n\right)^{\mathrm{T}} \hat{\mathbf{W}} \mathbf{c}_{\mathrm{E}}^n\right] + \mu \left[\sum_{n=1}^{N} \left(\mathbf{s}_{\mathrm{E}}^n\right)^{\mathrm{T}} \left(\mathbf{P}\hat{\mathbf{W}} + \mathbf{Q}\hat{\mathbf{W}}\mathbf{T}\right)^{\mathrm{T}} \left(\mathbf{P}\hat{\mathbf{W}} + \mathbf{Q}\hat{\mathbf{W}}\mathbf{T}\right) \mathbf{s}_{\mathrm{E}}^n - N p_0\right] \\
&= \left(\sum_{n=1}^{N} \mathbf{1}^{\mathrm{T}} \boldsymbol{\delta}^n - 1\right) t - \sum_{n=1}^{N} \left(\boldsymbol{\delta}^n\right)^{\mathrm{T}} \mathbf{A}^n \hat{\mathbf{W}} \mathbf{s}_{\mathrm{E}}^n - \sum_{n=1}^{N} \left(\boldsymbol{\delta}^n\right)^{\mathrm{T}} \mathbf{B}^n \hat{\mathbf{W}} \mathbf{c}_{\mathrm{E}}^n + \mu \sum_{n=1}^{N} \left(\mathbf{s}_{\mathrm{E}}^n\right)^{\mathrm{T}} \hat{\mathbf{W}}^{\mathrm{T}} \hat{\mathbf{W}} \mathbf{s}_{\mathrm{E}}^n + \mu \sum_{n=1}^{N} \left(\mathbf{c}_{\mathrm{E}}^n\right)^{\mathrm{T}} \hat{\mathbf{W}}^{\mathrm{T}} \hat{\mathbf{W}} \mathbf{c}_{\mathrm{E}}^n - \mu N p_0.
\end{aligned} \tag{13}
$$

$$
\frac{\partial L}{\partial t} = \sum_{n=1}^{N} \mathbf{1}^{\mathrm{T}} \boldsymbol{\delta}^n - 1 = 0, \tag{15a}
$$

$$
\frac{\partial L}{\partial \hat{\mathbf{W}}} = -\sum_{n=1}^{N} \left[\left(\boldsymbol{\delta}^n\right)^{\mathrm{T}} \mathbf{A}^n\right]^{\mathrm{T}} \left(\mathbf{s}_{\mathrm{E}}^n\right)^{\mathrm{T}} - \sum_{n=1}^{N} \left[\left(\boldsymbol{\delta}^n\right)^{\mathrm{T}} \mathbf{B}^n\right]^{\mathrm{T}} \left(\mathbf{c}_{\mathrm{E}}^n\right)^{\mathrm{T}} + 2\mu \hat{\mathbf{W}} \left[\sum_{n=1}^{N} \mathbf{s}_{\mathrm{E}}^n \left(\mathbf{s}_{\mathrm{E}}^n\right)^{\mathrm{T}} + \sum_{n=1}^{N} \mathbf{c}_{\mathrm{E}}^n \left(\mathbf{c}_{\mathrm{E}}^n\right)^{\mathrm{T}}\right] = \mathbf{0}, \tag{15b}
$$

$$
\delta_k^n \left[t - \left(\mathbf{a}_k^n\right)^{\mathrm{T}} \hat{\mathbf{W}} \mathbf{s}_{\mathrm{E}}^n - \left(\mathbf{b}_k^n\right)^{\mathrm{T}} \hat{\mathbf{W}} \mathbf{c}_{\mathrm{E}}^n\right] = 0,\ \delta_k^n \geq 0,\ \forall 1 \leq k \leq 2K, 1 \leq n \leq N \tag{15c}
$$

$$
\mu \left[\sum_{n=1}^{N} \left(\mathbf{s}_{\mathrm{E}}^n\right)^{\mathrm{T}} \hat{\mathbf{W}}^{\mathrm{T}} \hat{\mathbf{W}} \mathbf{s}_{\mathrm{E}}^n + \sum_{n=1}^{N} \left(\mathbf{c}_{\mathrm{E}}^n\right)^{\mathrm{T}} \hat{\mathbf{W}}^{\mathrm{T}} \hat{\mathbf{W}} \mathbf{c}_{\mathrm{E}}^n - N p_0\right] = 0,\ \mu \geq 0. \tag{15d}
$$

When $N \geq K$, $\mathbf{D}$ is thus full-rank and invertible. Accordingly, we can directly obtain a closed-form structure of the optimal CI-BLP precoding matrix $\hat{\mathbf{W}}$ as

$$
\hat{\mathbf{W}} = \frac{1}{2\mu} \sum_{n=1}^{N} \left[\left(\mathbf{A}^n\right)^{\mathrm{T}} \boldsymbol{\delta}^n \left(\mathbf{s}_{\mathrm{E}}^n\right)^{\mathrm{T}} + \left(\mathbf{B}^n\right)^{\mathrm{T}} \boldsymbol{\delta}^n \left(\mathbf{c}_{\mathrm{E}}^n\right)^{\mathrm{T}}\right] \mathbf{D}^{-1}. \tag{19}
$$

Since problem $\mathcal{P}_1$ satisfies Slater's condition, we can find the optimal solution of $\mathcal{P}_1$ by solving its dual problem. We can substitute the closed-form structure of $\hat{\mathbf{W}}$ into the objective function of the dual problem, and define $\boldsymbol{\delta}_{\mathrm{E}} \in \mathbb{R}^{2NK \times 1}$ and $\mathbf{U}_{m,n} \in \mathbb{R}^{2K \times 2K}$ as

$$
\boldsymbol{\delta}_{\mathrm{E}} = \left[\left(\boldsymbol{\delta}^1\right)^{\mathrm{T}}, \left(\boldsymbol{\delta}^2\right)^{\mathrm{T}}, \cdots, \left(\boldsymbol{\delta}^N\right)^{\mathrm{T}}\right]^{\mathrm{T}}, \tag{20}
$$

$$
\begin{aligned}
\mathbf{U}_{m,n} =& p_{m,n} \mathbf{A}^m \left(\mathbf{A}^n\right)^{\mathrm{T}} + f_{m,n} \mathbf{A}^m \left(\mathbf{B}^n\right)^{\mathrm{T}} \\
&+ g_{m,n} \mathbf{B}^m \left(\mathbf{A}^n\right)^{\mathrm{T}} + q_{m,n} \mathbf{B}^m \left(\mathbf{B}^n\right)^{\mathrm{T}},
\end{aligned} \tag{21}
$$

where $m \in \{1, \cdots, N\}$ and $n \in \{1, \cdots, N\}$. $p_{m,n}$, $f_{m,n}$, $g_{m,n}$ and $q_{m,n}$ are defined as

$$
\begin{aligned}
&p_{m,n} = \left(\mathbf{s}_{\mathrm{E}}^n\right)^{\mathrm{T}} \mathbf{D}^{-1} \mathbf{s}_{\mathrm{E}}^m,\ f_{m,n} = \left(\mathbf{c}_{\mathrm{E}}^n\right)^{\mathrm{T}} \mathbf{D}^{-1} \mathbf{s}_{\mathrm{E}}^m, \\
&g_{m,n} = \left(\mathbf{s}_{\mathrm{E}}^n\right)^{\mathrm{T}} \mathbf{D}^{-1} \mathbf{c}_{\mathrm{E}}^m,\ q_{m,n} = \left(\mathbf{c}_{\mathrm{E}}^n\right)^{\mathrm{T}} \mathbf{D}^{-1} \mathbf{c}_{\mathrm{E}}^m.
\end{aligned} \tag{22}
$$

The final dual problem of CI-BLP can be formulated as

$$
\begin{aligned}
\mathcal{P}_2:\ &\min_{\boldsymbol{\delta}_{\mathrm{E}}}\ \boldsymbol{\delta}_{\mathrm{E}}^{\mathrm{T}} \mathbf{U} \boldsymbol{\delta}_{\mathrm{E}} \\
&\text{s.t.}\ \ \mathbf{1}^{\mathrm{T}} \boldsymbol{\delta}_{\mathrm{E}} - 1 = 0, \\
&\qquad \delta_{\mathrm{E}}^m \geq 0,\ \forall m \in \{1, 2, \cdots, 2NK\}.
\end{aligned} \tag{23}
$$

Here, δ_{E}^m is the m-th entry of vector $\boldsymbol{\delta}_{\mathrm{E}}$, and $\mathbf{U} \in \mathbb{R}^{2NK \times 2NK}$ is a block matrix constructed as

$$
\mathbf{U} = \begin{bmatrix}
\mathbf{U}_{1,1} & \mathbf{U}_{1,2} & \cdots & \mathbf{U}_{1,N} \\
\mathbf{U}_{2,1} & \mathbf{U}_{2,2} & \cdots & \mathbf{U}_{2,N} \\
\vdots & \vdots & \ddots & \vdots \\
\mathbf{U}_{N,1} & \mathbf{U}_{N,2} & \cdots & \mathbf{U}_{N,N}
\end{bmatrix}. \tag{24}
$$

$\mathcal{P}_2$ is a QP optimization problem over a simplex, which can be more efficiently solved than $\mathcal{P}_1$ via the IPM. After solving $\mathcal{P}_2$ and obtaining $\hat{\mathbf{W}}$ via the closed-form structure in (19), the original complex precoding matrix $\mathbf{W}$ can be obtained via (7).

III. QP Formulation for the Case of $N < K$

The better performance of CI-BLP when $N < K$ and the demand for high-mobility scenarios encourage us to pay more attention to the CI-BLP optimization problem when $N < K$. However, the QP formulation derived in [33] is only applicable when $N \geq K$, because $\mathbf{D}$ is invertible only when $N \geq K$. It is not clear whether a similar QP problem exists when $N < K$. In this section, we extend the result in [33] to the case where the number of users is larger than the number of symbol slots within the considered transmission block, i.e., $N < K$.

When $N < K$, the direct inverse included in (19) becomes infeasible, as the matrix $\mathbf{D}$ is rank-deficient. In this case, we propose to employ the more general pseudo inverse instead of the direct matrix inverse. Accordingly, we express $\hat{\mathbf{W}}$ as

$$
\hat{\mathbf{W}} = \frac{1}{2\mu} \sum_{n=1}^{N} \left[\left(\mathbf{A}^n\right)^{\mathrm{T}} \boldsymbol{\delta}^n \left(\mathbf{s}_{\mathrm{E}}^n\right)^{\mathrm{T}} + \left(\mathbf{B}^n\right)^{\mathrm{T}} \boldsymbol{\delta}^n \left(\mathbf{c}_{\mathrm{E}}^n\right)^{\mathrm{T}}\right] \mathbf{D}^{+}. \tag{25}
$$

Then, one can easily follow a similar approach in Section II-D to obtain a QP optimization and the corresponding

solution. Although we note that a pseudo-inverse does not always guarantee that the original constraint is satisfied, our derivations in this section and the corresponding numerical results show that, in our problem the pseudo-inverse satisfies the original constraint and the closed-form structure of $\hat{\mathbf{W}}$ given by (25) is feasible.

In order to make the expression more concise, we can define a matrix $\mathbf{C} \in \mathbb{R}^{N_t \times 2K}$ as

$$\mathbf{C} = \frac{1}{2\mu} \sum_{n=1}^{N} \left[(\mathbf{A}^n)^{\mathrm{T}} \boldsymbol{\delta}^n (\mathbf{s}_{\mathrm{E}}^n)^{\mathrm{T}} + (\mathbf{B}^n)^{\mathrm{T}} \boldsymbol{\delta}^n (\mathbf{c}_{\mathrm{E}}^n)^{\mathrm{T}} \right]. \quad (26)$$

To prove that the closed-form structure of $\hat{\mathbf{W}}$ given by (25) is feasible, we need to prove that (25) can satisfy (17), that is

$$\hat{\mathbf{W}} = \mathbf{C}\mathbf{D}^{+} \Rightarrow \hat{\mathbf{W}}\mathbf{D} = \mathbf{C}. \quad (27)$$

In what follows, we further derive the equivalent condition of (27).

According to (18), $\mathbf{D}$ is a real symmetric matrix. Thus, by eigen decomposition, we can obtain

$$\begin{aligned} \mathbf{D} &= \mathbf{V}_{\mathrm{D}} \boldsymbol{\Sigma}_{\mathrm{D}} \mathbf{V}_{\mathrm{D}}^{\mathrm{T}} = \mathbf{V}_{\mathrm{D}} \mathrm{Diag}\left([\sigma_1, \cdots, \sigma_{r_d}, 0, \cdots, 0] \right) \mathbf{V}_{\mathrm{D}}^{\mathrm{T}}, \\ \mathbf{D}^{+} &= \mathbf{V}_{\mathrm{D}} \boldsymbol{\Sigma}_{\mathrm{D}}^{+} \mathbf{V}_{\mathrm{D}}^{\mathrm{T}} \\ &= \mathbf{V}_{\mathrm{D}} \mathrm{Diag}\left(\left[\frac{1}{\sigma_1}, \cdots, \frac{1}{\sigma_{r_d}}, 0, \cdots, 0 \right] \right) \mathbf{V}_{\mathrm{D}}^{\mathrm{T}}, \end{aligned} \quad (28)$$

where $\boldsymbol{\Sigma}_{\mathrm{D}}$ is a diagonal matrix whose diagonal elements are eigenvalues of $\mathbf{D}$, and r_d is the rank of $\mathbf{D}$, i.e., the number of non-zero eigenvalues of the matrix $\mathbf{D}$. $\mathbf{V}_{\mathrm{D}} \in \mathbb{R}^{2K \times 2K}$ is an orthogonal matrix composed of eigenvectors of $\mathbf{D}$. By substituting $\mathbf{D}$ and $\mathbf{D}^{+}$ in (28) into (27) and some simple algebraic operations, we can obtain

$$\mathbf{F} = \mathbf{E}\boldsymbol{\Sigma}_{\mathrm{D}}^{+} \Rightarrow \mathbf{F}\boldsymbol{\Sigma}_{\mathrm{D}} = \mathbf{E}, \quad (29)$$

where $\mathbf{E} = \mathbf{C}\mathbf{V}_{\mathrm{D}}$, $\mathbf{F} = \hat{\mathbf{W}}\mathbf{V}_{\mathrm{D}}$. In order to prove the above condition, we study the relationship between each rows for $\mathbf{E}$ and $\mathbf{F}$, where the i-th row can be expresseded by element as

$$F_{i,j} = E_{i,j} \cdot \frac{1}{\sigma_j} \Rightarrow F_{i,j} \cdot \sigma j = E_{i,j}, \ j \in \{1, \cdots, r_d\}, \quad (30)$$

$$F_{i,j} = E_{i,j} \cdot 0 \Rightarrow F_{i,j} \cdot 0 = E_{i,j}, \ j \in \{r_d + 1, \cdots, 2K\}, \quad (31)$$

where $i \in \{1, \cdots, N_t\}$. $E_{i,j}$ and $F_{i,j}$ represent the elements of the i-th row and j-th column of matrix $\mathbf{E}$ and matrix $\mathbf{F}$ respectively. (30) is naturally true, and (31) is equivalent to

$$E_{i,r_d+1} = \cdots = E_{i,2K} = 0. \quad (32)$$

By decomposing $\mathbf{E} = \left[\mathbf{e}^1, \mathbf{e}^2, \cdots, \mathbf{e}^{2K} \right]$, the equivalent condition of (27) can be rewritten as

$$\mathbf{e}^{r_D+1} = \cdots = \mathbf{e}^{2K} = \mathbf{0}. \quad (33)$$

Before we proceed, we present the following property:

Proposition 1: Suppose $\left[\mathbf{s}_{\mathrm{G}}^1, \mathbf{s}_{\mathrm{G}}^2, \cdots, \mathbf{s}_{\mathrm{G}}^{N_g} \right] \in \mathbb{R}^{M_g \times N_g}$. $\mathrm{rank}\left(\left[\mathbf{s}_{\mathrm{G}}^1, \mathbf{s}_{\mathrm{G}}^2, \cdots, \mathbf{s}_{\mathrm{G}}^{N_g} \right] \right) = r_g$, and $\mathbf{s}_{\mathrm{G}}^1, \mathbf{s}_{\mathrm{G}}^2, \cdots, \mathbf{s}_{\mathrm{G}}^{r_g}$ are linearly independent. Then for a matrix formulated as

$$\begin{aligned} \mathbf{G} &= \sum_{n=1}^{N_g} \mathbf{s}_{\mathrm{G}}^n \left(\mathbf{s}_{\mathrm{G}}^n \right)^{\mathrm{T}} = \mathbf{V}_{\mathrm{G}} \boldsymbol{\Sigma}_{\mathrm{G}} \mathbf{V}_{\mathrm{G}}^{\mathrm{T}} \\ &= \mathbf{V}_{\mathrm{G}} \mathrm{Diag}\left(\left[\tilde{\sigma}_1, \cdots, \tilde{\sigma}_{r_g}, 0, \cdots, 0 \right] \right) \mathbf{V}_{\mathrm{G}}^{\mathrm{T}}, \end{aligned} \quad (34)$$

where $\mathbf{V}_{\mathrm{G}} = \left[\mathbf{v}_{\mathrm{G}}^1, \mathbf{v}_{\mathrm{G}}^2, \cdots, \mathbf{v}_{\mathrm{G}}^{M_g} \right]$, the following condition must be satisfied for any $n \in \{1, 2, \cdots, N_g\}$, $m \in \{r_g + 1, r_g + 2, \cdots, M_g\}$:

$$\left(\mathbf{s}_{\mathrm{G}}^n \right)^{\mathrm{T}} \cdot \mathbf{v}_{\mathrm{G}}^m = 0. \quad (35)$$

Proof: See Appendix A. ∎

Proposition 1 shows that for the coefficient matrix $\mathbf{D}$ given in (18), the following conditions must be satisfied:

$$\left(\mathbf{s}_{\mathrm{E}}^n \right)^{\mathrm{T}} \cdot \mathbf{v}_{\mathrm{D}}^m = 0, \ \left(\mathbf{c}_{\mathrm{E}}^n \right)^{\mathrm{T}} \cdot \mathbf{v}_{\mathrm{D}}^m = 0, \quad (36)$$

where $n \in \{1, \cdots, N\}$, $m \in \{r_d + 1, \cdots, 2K\}$, and $\mathbf{v}_{\mathrm{D}}^m$ is the m-th column of matrix $\mathbf{V}_{\mathrm{D}}$. Accordingly, for any $m \in \{r_d + 1, \cdots, 2K\}$, we obtain that

$$\begin{aligned} \mathbf{e}^m &= \left(\frac{1}{2\mu} \sum_{n=1}^{N} \left[(\mathbf{A}^n)^{\mathrm{T}} \boldsymbol{\delta}^n (\mathbf{s}_{\mathrm{E}}^n)^{\mathrm{T}} + (\mathbf{B}^n)^{\mathrm{T}} \boldsymbol{\delta}^n (\mathbf{c}_{\mathrm{E}}^n)^{\mathrm{T}} \right] \right) \cdot \mathbf{v}_{\mathrm{D}}^m \\ &= \frac{1}{2\mu} \sum_{n=1}^{N} \left[(\mathbf{A}^n)^{\mathrm{T}} \boldsymbol{\delta}^n (\mathbf{s}_{\mathrm{E}}^n)^{\mathrm{T}} \cdot \mathbf{v}_{\mathrm{D}}^m + (\mathbf{B}^n)^{\mathrm{T}} \boldsymbol{\delta}^n (\mathbf{c}_{\mathrm{E}}^n)^{\mathrm{T}} \cdot \mathbf{v}_{\mathrm{D}}^m \right] \\ &= \frac{1}{2\mu} \sum_{n=1}^{N} \left[(\mathbf{A}^n)^{\mathrm{T}} \boldsymbol{\delta}^n \cdot 0 + (\mathbf{B}^n)^{\mathrm{T}} \boldsymbol{\delta}^n \cdot 0 \right] = \mathbf{0}, \end{aligned} \quad (37)$$

which means that the condition in (33) is satisfied and the closed-form structure of $\hat{\mathbf{W}}$ given by (25) is feasible.

We note that when $N \geq K$, $\mathbf{D}$ is full-rank and the closed-form structure of $\hat{\mathbf{W}}$ given by (25) is also feasible. To summarize, we can obtain a uniform closed-form structure of $\hat{\mathbf{W}}$ given by (25), and thus we can design optimal block-level precoding by solving $\mathcal{P}_2$ in (23) for the considered symbol block of any length.

IV. Discussion on the Iterative Closed-form Algorithm for CI-BLP

The QP problem of CI-SLP can be effectively solved by the iterative closed-form algorithm proposed in [26]. Compared with the common IPM, the iterative closed-form algorithm shows obvious advantages in solving the QP problem of CI-SLP. Before applying the iterative closed-form algorithm to the QP problem of CI-BLP, it is important to discuss the rank of $\mathbf{U}$ because the application of the iterative closed-form algorithm requires the quadratic coefficient matrix to be invertible.

Remark 1: *Before analyzing the rank of* $\mathbf{U}$*, we state that the variables introduced in the proof of each subsequent propositions in this section are only used to complete the proof and have nothing to do with other variables of the same name in this paper.*

According to (21) and (24), we can further get a new expression of $\mathbf{U}$. By defining $\hat{\mathbf{U}}_{m,n}$ as

$$\hat{\mathbf{U}}_{m,n} = \begin{bmatrix} p_{m,n}\mathbf{A}^m \left(\mathbf{A}^n\right)^{\mathrm{T}} & f_{m,n}\mathbf{A}^m \left(\mathbf{B}^n\right)^{\mathrm{T}} \\ g_{m,n}\mathbf{B}^m \left(\mathbf{A}^n\right)^{\mathrm{T}} & {}_{m,n}\mathbf{B}^m \left(\mathbf{B}^n\right)^{\mathrm{T}} \end{bmatrix}. \tag{38}$$

We can construct a block matrix as

$$\hat{\mathbf{U}} = \begin{bmatrix} \hat{\mathbf{U}}_{1,1} & \hat{\mathbf{U}}_{1,2} & \cdots & \hat{\mathbf{U}}_{1,N} \\ \hat{\mathbf{U}}_{2,1} & \hat{\mathbf{U}}_{2,2} & \cdots & \hat{\mathbf{U}}_{2,N} \\ \vdots & \vdots & \ddots & \vdots \\ \hat{\mathbf{U}}_{N,1} & \hat{\mathbf{U}}_{N,2} & \cdots & \hat{\mathbf{U}}_{N,N} \end{bmatrix}. \tag{39}$$

Then $\mathbf{U}$ can be written as

$$\mathbf{U} = \left(\mathbf{I}_N \otimes [\mathbf{I}_{2K}\ \mathbf{I}_{2K}]\right) \hat{\mathbf{U}} \left(\mathbf{I}_N \otimes [\mathbf{I}_{2K}\ \mathbf{I}_{2K}]\right)^{\mathrm{T}}. \tag{40}$$

We also define $\mathbf{\Delta}$ and $\mathbf{\Theta}$ as:

$$\mathbf{\Delta} = \begin{bmatrix} \mathbf{\Delta}_{1,1} & \mathbf{\Delta}_{1,2} & \cdots & \mathbf{\Delta}_{1,N} \\ \mathbf{\Delta}_{2,1} & \mathbf{\Delta}_{2,2} & \cdots & \mathbf{\Delta}_{2,N} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{\Delta}_{N,1} & \mathbf{\Delta}_{N,2} & \cdots & \mathbf{\Delta}_{N,N} \end{bmatrix}, \tag{41}$$

$$\mathbf{\Theta} = \begin{bmatrix} \mathbf{\Theta}_{1,1} & \mathbf{\Theta}_{1,2} & \cdots & \mathbf{\Theta}_{1,N} \\ \mathbf{\Theta}_{2,1} & \mathbf{\Theta}_{2,2} & \cdots & \mathbf{\Theta}_{2,N} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{\Theta}_{N,1} & \mathbf{\Theta}_{N,2} & \cdots & \mathbf{\Theta}_{N,N} \end{bmatrix}, \tag{42}$$

where

$$\begin{aligned} \mathbf{\Delta}_{m,n} &= \begin{bmatrix} p_{m,n} & f_{m,n} \\ g_{m,n} & q_{m,n} \end{bmatrix}, \\ \mathbf{\Theta}_{m,n} &= \begin{bmatrix} \mathbf{A}^m \left(\mathbf{A}^n\right)^{\mathrm{T}} & \mathbf{A}^m \left(\mathbf{B}^n\right)^{\mathrm{T}} \\ \mathbf{B}^m \left(\mathbf{A}^n\right)^{\mathrm{T}} & \mathbf{B}^m \left(\mathbf{B}^n\right)^{\mathrm{T}} \end{bmatrix}. \end{aligned} \tag{43}$$

Proposition 2: $\mathrm{rank}\left(\mathbf{D}\right) = \min\left\{2K, 2N\right\}$.
Proof: See Appendix B. ■

Proposition 3: $\mathrm{rank}\left(\mathbf{\Delta}\right) = \mathrm{rank}\left(\mathbf{D}\right)$.
Proof: See Appendix C. ■

Proposition 4: $\mathrm{rank}\left(\mathbf{\Theta}\right) = K$.
Proof: See Appendix D. ■

Proposition 5: $\mathrm{rank}\left(\hat{\mathbf{U}}\right) = \mathrm{rank}\left(\mathbf{\Delta}\right) \cdot \mathrm{rank}\left(\mathbf{\Theta}\right)$.
Proof: See Appendix E. ■

Proposition 6: $\mathrm{rank}\left(\mathbf{U}\right) = \mathrm{rank}\left(\hat{\mathbf{U}}\right) = \min\left\{2NK, 2K^2\right\}$.
Proof: See Appendix F. ■

Based on the above results, $\mathbf{U}$ is full rank when $N \leq K$, which means we can only use the iterative closed-form algorithm proposed in [26] when $N \leq K$. The iterative closed-form algorithm is thus not generic for CI-BLP problems.

Moreover, since the problem size of CI-BLP is N times that of CI-SLP, the iterative closed-form algorithm for CI-BLP requires a large number of iterations to converge. We simulate systems of different sizes using iterative closed-form algorithm, and obtain their corresponding cumulative distribution functions (CDF) of the required number of iterations, as shown in Fig. 3.

It is observed that the iterative closed-form solution is not ideal for CI-BLP. Therefore, we propose to leverage the ADMM framework to design an efficient CI-BLP algorithm with the considered symbol block of any length.

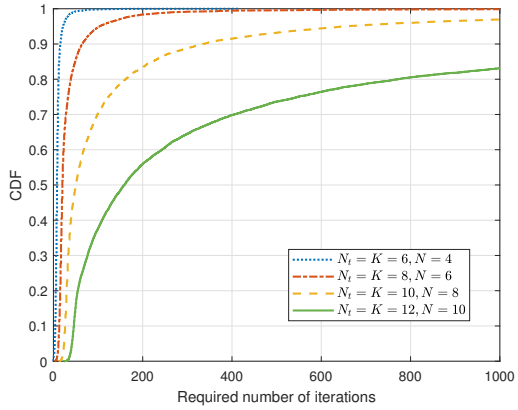

Fig. 3. CDF of the required number of iterations for CI-BLP, 8PSK.

V. The Proposed ADMM Algorithm

A. The Conventional ADMM Algorithm

To apply the ADMM framework, we introduce a variable $\boldsymbol{\omega}$ and a set $\Omega = \{\boldsymbol{\omega} | \omega_i \geq 0,\ \forall i \in \{1, 2, \ldots, 2NK\}\}$. $\mathcal{P}_2$ can be written as

$$\begin{aligned} \mathcal{P}_3: \min_{\boldsymbol{\delta}_{\mathrm{E}}, \boldsymbol{\omega}}\ & \boldsymbol{\delta}_{\mathrm{E}}^{\mathrm{T}} \mathbf{U} \boldsymbol{\delta}_{\mathrm{E}} \\ \text{s.t.}\ & \mathbf{1}^{\mathrm{T}} \boldsymbol{\delta}_{\mathrm{E}} - 1 = 0, \\ & \boldsymbol{\delta}_{\mathrm{E}} - \boldsymbol{\omega} = \mathbf{0}, \\ & \boldsymbol{\omega} \in \Omega. \end{aligned} \tag{44}$$

By defining an indicator function for $\boldsymbol{\omega}$:

$$I_{\Omega}\left(\boldsymbol{\omega}\right) = \begin{cases} 0, & \text{if } \boldsymbol{\omega} \in \Omega, \\ \infty, & \text{otherwise}, \end{cases} \tag{45}$$

the problem can be rewritten in an equivalent consensus form as follows [37]:

$$\begin{aligned} \mathcal{P}_4: \min_{\boldsymbol{\delta}_{\mathrm{E}} \in \{\boldsymbol{\delta}_{\mathrm{E}} | \mathbf{1}^{\mathrm{T}} \boldsymbol{\delta}_{\mathrm{E}} - 1 = 0\}, \boldsymbol{\omega}}\ & \boldsymbol{\delta}_{\mathrm{E}}^{\mathrm{T}} \mathbf{U} \boldsymbol{\delta}_{\mathrm{E}} + I_{\Omega}\left(\boldsymbol{\omega}\right) \\ \text{s.t.}\ & \boldsymbol{\delta}_{\mathrm{E}} - \boldsymbol{\omega} = \mathbf{0}. \end{aligned} \tag{46}$$

For the optimization problem $\mathcal{P}_4$, its augmented Lagrangian is expressed as

$$\begin{aligned} & L_{\rho}\left(\boldsymbol{\delta}_{\mathrm{E}}, \boldsymbol{\omega}, \boldsymbol{\lambda}\right) \\ &= \boldsymbol{\delta}_{\mathrm{E}}^{\mathrm{T}} \mathbf{U} \boldsymbol{\delta}_{\mathrm{E}} + I_{\Omega}\left(\boldsymbol{\omega}\right) - \left\langle \boldsymbol{\lambda}, \boldsymbol{\delta}_{\mathrm{E}} - \boldsymbol{\omega} \right\rangle + \frac{\rho}{2} \left\| \boldsymbol{\delta}_{\mathrm{E}} - \boldsymbol{\omega} \right\|_2^2 \\ &= \boldsymbol{\delta}_{\mathrm{E}}^{\mathrm{T}} \mathbf{U} \boldsymbol{\delta}_{\mathrm{E}} + I_{\Omega}\left(\boldsymbol{\omega}\right) + \frac{\rho}{2} \left\| \boldsymbol{\delta}_{\mathrm{E}} - \boldsymbol{\omega} - \frac{\boldsymbol{\lambda}}{\rho} \right\|_2^2 - \frac{\left\| \boldsymbol{\lambda} \right\|_2^2}{2\rho}, \end{aligned} \tag{47}$$

where $\boldsymbol{\lambda} \in \mathbb{R}^{2NK \times 1}$ is the dual vector and $\rho > 0$ is the penalty parameter. In the ADMM framework, the update of $\boldsymbol{\delta}_{\mathrm{E}}$, $\boldsymbol{\omega}$, and $\boldsymbol{\lambda}$ can be written as

$$\boldsymbol{\delta}_{\mathrm{E}}^{k+1} = \mathop{\arg\min}_{\boldsymbol{\delta}_{\mathrm{E}} \in \{\boldsymbol{\delta}_{\mathrm{E}} | \mathbf{1}^{\mathrm{T}} \boldsymbol{\delta}_{\mathrm{E}} - 1 = 0\}} L_{\rho}\left(\boldsymbol{\delta}_{\mathrm{E}}, \boldsymbol{\omega}^k, \boldsymbol{\lambda}^k\right), \tag{48}$$

$$\boldsymbol{\omega}^{k+1} = \mathop{\arg\min}_{\boldsymbol{\omega}} L_{\rho}\left(\boldsymbol{\delta}_{\mathrm{E}}^{k+1}, \boldsymbol{\omega}, \boldsymbol{\lambda}^k\right), \tag{49}$$

$$\boldsymbol{\lambda}^{k+1} = \boldsymbol{\lambda}^k - \rho\left(\boldsymbol{\delta}_{\mathrm{E}}^{k+1} - \boldsymbol{\omega}^{k+1}\right). \tag{50}$$

In the $\boldsymbol{\delta}_{\mathrm{E}}$-update, the subproblem for $\boldsymbol{\delta}_{\mathrm{E}}$ can be written as

$$\begin{aligned}\min_{\boldsymbol{\delta}_{\mathrm{E}}}\ & \boldsymbol{\delta}_{\mathrm{E}}^{\mathrm{T}}\mathbf{U}\boldsymbol{\delta}_{\mathrm{E}}+\frac{\rho}{2}\left\|\boldsymbol{\delta}_{\mathrm{E}}-\boldsymbol{\omega}^{k}-\frac{\boldsymbol{\lambda}^{k}}{\rho}\right\|_{2}^{2}\\ \text{s.t.}\ & \mathbf{1}^{\mathrm{T}}\boldsymbol{\delta}_{\mathrm{E}}-1=0.\end{aligned}\tag{51}$$

We can obtain the KKT condition of this subproblem:

$$\mathbf{1}^{\mathrm{T}}\boldsymbol{\delta}_{\mathrm{E}}-1=0,\tag{52}$$

$$2\mathbf{U}\boldsymbol{\delta}_{\mathrm{E}}+\rho\left(\boldsymbol{\delta}_{\mathrm{E}}-\boldsymbol{\omega}^{k}-\frac{\boldsymbol{\lambda}^{k}}{\rho}\right)+\nu\cdot\mathbf{1}=\mathbf{0},\tag{53}$$

where ν is the dual variable associated with the equality constraint. Then we can obtain

$$\begin{bmatrix}2\mathbf{U}+\rho & \mathbf{1}\\ \mathbf{1}^{\mathrm{T}} & 0\end{bmatrix}\begin{bmatrix}\boldsymbol{\delta}_{\mathrm{E}}\\ \nu\end{bmatrix}=\begin{bmatrix}\rho\boldsymbol{\omega}^{k}+\boldsymbol{\lambda}^{k}\\ 1\end{bmatrix}.\tag{54}$$

With the above formulation, we can obtain $\boldsymbol{\delta}_{\mathrm{E}}^{k+1}$ by an inverse operation

$$\begin{bmatrix}\boldsymbol{\delta}_{\mathrm{E}}^{k+1}\\ \nu\end{bmatrix}=\begin{bmatrix}2\mathbf{U}+\rho & \mathbf{1}\\ \mathbf{1}^{\mathrm{T}} & 0\end{bmatrix}^{-1}\begin{bmatrix}\rho\boldsymbol{\omega}^{k}+\boldsymbol{\lambda}^{k}\\ 1\end{bmatrix}.\tag{55}$$

In the $\boldsymbol{\omega}$-update, the subproblem for $\boldsymbol{\omega}$ can be written as

$$\min_{\boldsymbol{\omega}}\ I_{\Omega}\left(\boldsymbol{\omega}\right)+\frac{\rho}{2}\left\|\boldsymbol{\delta}_{\mathrm{E}}^{k+1}-\boldsymbol{\omega}-\frac{\boldsymbol{\lambda}^{k}}{\rho}\right\|_{2}^{2}.\tag{56}$$

Then $\boldsymbol{\omega}$ needs to satisfy

$$\rho\left(\boldsymbol{\delta}_{\mathrm{E}}^{k+1}-\boldsymbol{\omega}-\frac{\boldsymbol{\lambda}^{k}}{\rho}\right)\in\partial I_{\Omega}\left(\boldsymbol{\omega}\right),\tag{57}$$

and we can obtain

$$\boldsymbol{\omega}^{k+1}=\mathrm{m}ax\left\{\mathbf{0},\boldsymbol{\delta}_{\mathrm{E}}^{k+1}-\frac{\boldsymbol{\lambda}^{k}}{\rho}\right\}.\tag{58}$$

The corresponding algorithm is summarized in Algorithm 1.

Algorithm 1 The conventional ADMM algorithm

1: **Input**: $\mathbf{U}$
2: **Initialization**: $\boldsymbol{\delta}_{\mathrm{E}}^{1}=\boldsymbol{\omega}^{1}=\boldsymbol{\lambda}^{1}=\mathbf{0}$, ρ, maximum number of iterations K_{max}.
3: **for** $k=1,\cdots K_{max}$ **do**
4: Compute $\boldsymbol{\delta}_{\mathrm{E}}^{k+1}$ by (55);
5: Compute $\boldsymbol{\omega}^{k+1}$ by (58);
6: Update $\boldsymbol{\lambda}^{k+1}$ by (50);
7: **end for**
8: **Output**: $\boldsymbol{\delta}_{\mathrm{E}}=\boldsymbol{\delta}_{\mathrm{E}}^{K_{max}+1}$.

B. The Improved ADMM Algorithm

In addition to directly applying the ADMM framework in Section V-A, we also propose an improved ADMM algorithm to solve the QP problem of CI-BLP.

Proposition 7: $\mathcal{P}_2$ is equivalent to

$$\begin{aligned}\mathcal{P}_5:\ \min_{\boldsymbol{\delta}_{\mathrm{E}}}\ & \boldsymbol{\delta}_{\mathrm{E}}^{\mathrm{T}}\mathbf{U}\boldsymbol{\delta}_{\mathrm{E}}\\ \text{s.t.}\ & \mathbf{1}^{\mathrm{T}}\boldsymbol{\delta}_{\mathrm{E}}-1\geq 0,\\ & \delta_{\mathrm{E}}^{m}\geq 0,\ \forall m\in\{1,2,\cdots,2NK\}.\end{aligned}\tag{59}$$

Proof: Suppose $\boldsymbol{\delta}_{\mathrm{E}}^{*}$ is the optimal solution to $\mathcal{P}_5$, and it satisfies

$$\mathbf{1}^{\mathrm{T}}\boldsymbol{\delta}_{\mathrm{E}}^{*}-1>0,\quad(\delta_{\mathrm{E}}^{*})^{m}\geq 0,\ \forall m\in\{1,2,\cdots,2NK\}.\tag{60}$$

Then there must exist $\boldsymbol{\delta}_{\mathrm{E}}^{**}=\kappa\boldsymbol{\delta}_{\mathrm{E}}^{*}$ such that

$$\mathbf{1}^{\mathrm{T}}\boldsymbol{\delta}_{\mathrm{E}}^{**}-1=0,\quad(\delta_{\mathrm{E}}^{**})^{m}\geq 0,\ \forall m\in\{1,2,\cdots,2NK\},\tag{61}$$

where $\kappa=\frac{1}{\mathbf{1}^{\mathrm{T}}\boldsymbol{\delta}_{\mathrm{E}}^{*}}\in(0,1)$. Then we can obtain

$$\left(\boldsymbol{\delta}_{\mathrm{E}}^{**}\right)^{\mathrm{T}}\mathbf{U}\boldsymbol{\delta}_{\mathrm{E}}^{**}=\kappa^{2}\left(\boldsymbol{\delta}_{\mathrm{E}}^{*}\right)^{\mathrm{T}}\mathbf{U}\boldsymbol{\delta}_{\mathrm{E}}^{*}<\left(\boldsymbol{\delta}_{\mathrm{E}}^{*}\right)^{\mathrm{T}}\mathbf{U}\boldsymbol{\delta}_{\mathrm{E}}^{*},\tag{62}$$

which contradicts the assumption that $\boldsymbol{\delta}_{\mathrm{E}}^{*}$ is the optimal solution to $\mathcal{P}_5$. The optimal solution to $\mathcal{P}_5$ must satisfy $\mathbf{1}^{\mathrm{T}}\boldsymbol{\delta}_{\mathrm{E}}-1=0$, which means that the optimal solution to $\mathcal{P}_5$ is the optimal solution to $\mathcal{P}_2$, i.e., $\mathcal{P}_5$ is equivalent to $\mathcal{P}_2$. ■

Based on the above proposition, We can apply the ADMM framework to $\mathcal{P}_5$. We introduce a variable $\hat{\boldsymbol{\omega}}$ and a set $\hat{\Omega}=\{\hat{\boldsymbol{\omega}}|\hat{\omega}_i\geq 0,\ \forall i\in\{1,2,\ldots,2NK+1\}\}$. $\mathcal{P}_5$ is then equivalent to:

$$\begin{aligned}\mathcal{P}_6:\ \min_{\boldsymbol{\delta}_{\mathrm{E}},\hat{\boldsymbol{\omega}}}\ & \boldsymbol{\delta}_{\mathrm{E}}^{\mathrm{T}}\mathbf{U}\boldsymbol{\delta}_{\mathrm{E}}\\ \text{s.t.}\ & \begin{bmatrix}\mathbf{1}^{\mathrm{T}}\\ \mathbf{I}_{2NK\times 2NK}\end{bmatrix}\boldsymbol{\delta}_{\mathrm{E}}=\begin{bmatrix}1\\ \mathbf{0}\end{bmatrix}+\hat{\boldsymbol{\omega}},\\ & \hat{\boldsymbol{\omega}}\in\hat{\Omega}.\end{aligned}\tag{63}$$

By defining an indicator function for $\hat{\boldsymbol{\omega}}$:

$$I_{\hat{\Omega}}\left(\hat{\boldsymbol{\omega}}\right)=\begin{cases}0, & \text{if } \hat{\boldsymbol{\omega}}\in\hat{\Omega},\\ \infty, & \text{otherwise},\end{cases}\tag{64}$$

and

$$\boldsymbol{\Gamma}=\begin{bmatrix}\mathbf{1}^{\mathrm{T}}\\ \mathbf{I}_{2NK}\end{bmatrix}\in\mathbb{C}^{(2NK+1)\times(2NK)},\ \mathbf{c}=\begin{bmatrix}1\\ \mathbf{0}\end{bmatrix}\in\mathbb{C}^{(2NK+1)\times 1},\tag{65}$$

$\mathcal{P}_6$ can be written as

$$\begin{aligned}\mathcal{P}_7:\ \min_{\boldsymbol{\delta}_{\mathrm{E}},\hat{\boldsymbol{\omega}}}\ & \boldsymbol{\delta}_{\mathrm{E}}^{\mathrm{T}}\mathbf{U}\boldsymbol{\delta}_{\mathrm{E}}+I_{\hat{\Omega}}\left(\hat{\boldsymbol{\omega}}\right)\\ \text{s.t.}\ & \boldsymbol{\Gamma}\boldsymbol{\delta}_{\mathrm{E}}=\mathbf{c}+\hat{\boldsymbol{\omega}}.\end{aligned}\tag{66}$$

The corresponding augmented Lagrangian function for $\mathcal{P}_7$ is expressed as

$$\begin{aligned}\hat{L}_{\rho}\left(\boldsymbol{\delta}_{\mathrm{E}},\hat{\boldsymbol{\omega}},\hat{\boldsymbol{\lambda}}\right)=&\boldsymbol{\delta}_{\mathrm{E}}^{\mathrm{T}}\mathbf{U}\boldsymbol{\delta}_{\mathrm{E}}+I_{\hat{\Omega}}\left(\hat{\boldsymbol{\omega}}\right)+\hat{\boldsymbol{\lambda}}^{\mathrm{T}}\left(-\boldsymbol{\Gamma}\boldsymbol{\delta}_{\mathrm{E}}+\mathbf{c}+\hat{\boldsymbol{\omega}}\right)\\ &+\frac{\rho}{2}\left\|-\boldsymbol{\Gamma}\boldsymbol{\delta}_{\mathrm{E}}+\mathbf{c}+\hat{\boldsymbol{\omega}}\right\|_{2}^{2}\\ =&\boldsymbol{\delta}_{\mathrm{E}}^{\mathrm{T}}\mathbf{U}\boldsymbol{\delta}_{\mathrm{E}}+I_{\hat{\Omega}}\left(\hat{\boldsymbol{\omega}}\right)+\frac{\rho}{2}\left\|-\boldsymbol{\Gamma}\boldsymbol{\delta}_{\mathrm{E}}+\mathbf{c}+\hat{\boldsymbol{\omega}}+\frac{\hat{\boldsymbol{\lambda}}}{\rho}\right\|_{2}^{2}-\frac{\left\|\hat{\boldsymbol{\lambda}}\right\|_{2}^{2}}{2\rho},\end{aligned}\tag{67}$$

where $\hat{\boldsymbol{\lambda}}\in\mathbb{R}^{(2NK+1)\times 1}$ is the dual vector and $\rho>0$ is the penalty parameter. The update of $\boldsymbol{\delta}_{\mathrm{E}}$, $\hat{\boldsymbol{\omega}}$, and $\hat{\boldsymbol{\lambda}}$ can be written as

$$\boldsymbol{\delta}_{\mathrm{E}}^{k+1}=\arg\min_{\boldsymbol{\delta}_{\mathrm{E}}}\ \hat{L}_{\rho}\left(\boldsymbol{\delta}_{\mathrm{E}},\hat{\boldsymbol{\omega}}^{k},\hat{\boldsymbol{\lambda}}^{k}\right),\tag{68}$$

$$\hat{\boldsymbol{\omega}}^{k+1}=\arg\min_{\boldsymbol{\omega}}\ \hat{L}_{\rho}\left(\boldsymbol{\delta}_{\mathrm{E}}^{k+1},\hat{\boldsymbol{\omega}},\hat{\boldsymbol{\lambda}}^{k}\right),\tag{69}$$

$$\hat{\boldsymbol{\lambda}}^{k+1}=\hat{\boldsymbol{\lambda}}^{k}+\rho\left(-\boldsymbol{\Gamma}\boldsymbol{\delta}_{\mathrm{E}}^{k+1}+\mathbf{c}+\hat{\boldsymbol{\omega}}^{k+1}\right).\tag{70}$$

In the $\boldsymbol{\delta}_{\mathrm{E}}$-update, the subproblem for $\boldsymbol{\delta}_{\mathrm{E}}$ can be written as

$$\min_{\boldsymbol{\delta}_{\mathrm{E}}} \quad \boldsymbol{\delta}_{\mathrm{E}}^{\mathrm{T}}\mathbf{U}\boldsymbol{\delta}_{\mathrm{E}} + \frac{\rho}{2}\left\|-\boldsymbol{\Gamma}\boldsymbol{\delta}_{\mathrm{E}} + \mathbf{c} + \hat{\boldsymbol{\omega}}^{k} + \frac{\hat{\boldsymbol{\lambda}}^{k}}{\rho}\right\|_2^2. \tag{71}$$

The gradient of the objective function with respect to $\boldsymbol{\delta}_{\mathrm{E}}$ is 0, which leads to

$$\left(2\mathbf{U} + \rho\boldsymbol{\Gamma}^{\mathrm{T}}\boldsymbol{\Gamma}\right)\boldsymbol{\delta}_{\mathrm{E}} = \rho\boldsymbol{\Gamma}^{\mathrm{T}}\left(\mathbf{c} + \hat{\mathbf{w}}^{k} + \frac{\hat{\boldsymbol{\lambda}}^{k}}{\rho}\right). \tag{72}$$

Then we can get a closed-form solution:

$$\boldsymbol{\delta}_{\mathrm{E}}^{k+1} = \left(2\mathbf{U} + \rho\boldsymbol{\Gamma}^{\mathrm{T}}\boldsymbol{\Gamma}\right)^{-1}\rho\boldsymbol{\Gamma}^{\mathrm{T}}\left(\mathbf{c} + \hat{\mathbf{w}}^{k} + \frac{\hat{\boldsymbol{\lambda}}^{k}}{\rho}\right). \tag{73}$$

In the $\boldsymbol{\omega}$-update, the subproblem for $\boldsymbol{\omega}$ can be written as

$$\min_{\boldsymbol{\omega}} \quad I_{\Omega}(\boldsymbol{\omega}) + \frac{\rho}{2}\left\|-\boldsymbol{\Gamma}\boldsymbol{\delta}_{\mathrm{E}}^{k+1} + \mathbf{c} + \hat{\boldsymbol{\omega}} + \frac{\hat{\boldsymbol{\lambda}}^{k}}{\rho}\right\|_2^2. \tag{74}$$

Then $\boldsymbol{\omega}$ needs to satisfy

$$-\rho\left(-\boldsymbol{\Gamma}\boldsymbol{\delta}_{\mathrm{E}}^{k+1} + \mathbf{c} + \hat{\boldsymbol{\omega}} + \frac{\hat{\boldsymbol{\lambda}}^{k}}{\rho}\right) \in \partial I_{\Omega}(\boldsymbol{\omega}), \tag{75}$$

and we can obtain

$$\hat{\boldsymbol{\omega}}^{k+1} = \mathrm{m}ax\left\{\mathbf{0}, \boldsymbol{\Gamma}\boldsymbol{\delta}_{\mathrm{E}}^{k+1} - \mathbf{c} - \frac{\hat{\boldsymbol{\lambda}}^{k}}{\rho}\right\}. \tag{76}$$

The corresponding algorithm is summarized in Algorithm 2.

Algorithm 2 The improved ADMM algorithm

1: **Input**: $\mathbf{U}$
2: **Initialization**: $\boldsymbol{\delta}_{\mathrm{E}}^{1} = \mathbf{0}$, $\hat{\boldsymbol{\omega}}^{1} = \hat{\boldsymbol{\lambda}}^{1} = \mathbf{0}$, $\mathbf{c}$, $\boldsymbol{\Gamma}$, ρ, maximum number of iterations K_{max}.
3: **for** $k = 1, \cdots K_{max}$ **do**
4: Compute $\boldsymbol{\delta}_{\mathrm{E}}^{k+1}$ by (73);
5: Compute $\hat{\boldsymbol{\omega}}^{k+1}$ by (76);
6: Update $\hat{\boldsymbol{\lambda}}^{k+1}$ by (70);
7: **end for**
8: **Output**: $\boldsymbol{\delta}_{\mathrm{E}} = \boldsymbol{\delta}_{\mathrm{E}}^{K_{max}+1}$.

Remark 2: *In Algorithm 1, the ADMM framework is directly applied to optimization problem $\mathcal{P}_2$, and the update for$\boldsymbol{\delta}_{\mathrm{E}}$ is a subproblem with equality constraints. Although such a method is common, it affects the convergence behavior of the algorithm to some extent [37]. In Algorithm 2, the ADMM framework is applied to the equivalent optimization problem $\mathcal{P}_5$, and all updates are unconstrained optimization problems, which can guarantee the convergence of the algorithm.*

C. Convergence Analysis

First, the minimizer $\hat{\boldsymbol{\omega}}^{k+1}$ satisfies

$$\hat{L}_{\rho}\left(\boldsymbol{\delta}_{\mathrm{E}}^{k}, \hat{\boldsymbol{\omega}}^{k+1}, \hat{\boldsymbol{\lambda}}^{k}\right) \leq \hat{L}_{\rho}\left(\boldsymbol{\delta}_{\mathrm{E}}^{k}, \hat{\boldsymbol{\omega}}^{k}, \hat{\boldsymbol{\lambda}}^{k}\right). \tag{77}$$

It is easy to see that $\hat{L}_{\rho}\left(\boldsymbol{\delta}_{\mathrm{E}}, \hat{\boldsymbol{\omega}}^{k+1}, \hat{\boldsymbol{\lambda}}^{k}\right)$ is strongly convex with respect to $\boldsymbol{\delta}_{\mathrm{E}}$. Because of the property of strongly convex function:

$$\left(\partial_{\boldsymbol{\delta}_{\mathrm{E},1}}\hat{L}_{\rho}\left(\boldsymbol{\delta}_{\mathrm{E},1}, \hat{\boldsymbol{\omega}}^{k+1}, \hat{\boldsymbol{\lambda}}^{k}\right) - \partial_{\boldsymbol{\delta}_{\mathrm{E},2}}\hat{L}_{\rho}\left(\boldsymbol{\delta}_{\mathrm{E},2}, \hat{\boldsymbol{\omega}}^{k+1}, \hat{\boldsymbol{\lambda}}^{k}\right)\right)^{\mathrm{T}} \cdot \left(\boldsymbol{\delta}_{\mathrm{E},1} - \hat{\boldsymbol{\delta}}_{\mathrm{E},2}\right) \geq m\left\|\boldsymbol{\delta}_{\mathrm{E},1} - \hat{\boldsymbol{\delta}}_{\mathrm{E},2}\right\|_2^2, \tag{78}$$

where m is the strongly convex parameter of $\hat{L}_{\rho}\left(\boldsymbol{\delta}_{\mathrm{E}}, \hat{\boldsymbol{\omega}}^{k+1}, \hat{\boldsymbol{\lambda}}^{k}\right)$, we can obtain that

$$\begin{aligned}&\partial_{\boldsymbol{\delta}_{\mathrm{E},1}}\hat{L}_{\rho}\left(\boldsymbol{\delta}_{\mathrm{E},1}, \hat{\boldsymbol{\omega}}^{k+1}, \hat{\boldsymbol{\lambda}}^{k}\right) - \partial_{\boldsymbol{\delta}_{\mathrm{E},2}}\hat{L}_{\rho}\left(\boldsymbol{\delta}_{\mathrm{E},2}, \hat{\boldsymbol{\omega}}^{k+1}, \hat{\boldsymbol{\lambda}}^{k}\right)\\ &= \left(2\mathbf{U} + \rho\boldsymbol{\Gamma}^{\mathrm{T}}\boldsymbol{\Gamma}\right)\left(\boldsymbol{\delta}_{\mathrm{E},1} - \boldsymbol{\delta}_{\mathrm{E},2}\right).\end{aligned} \tag{79}$$

Since

$$\begin{aligned}\rho\boldsymbol{\Gamma}^{\mathrm{T}}\boldsymbol{\Gamma} &= \rho\begin{bmatrix}\mathbf{1} & \mathbf{I}_{2NK\times 2NK}\end{bmatrix}\begin{bmatrix}\mathbf{1}^{\mathrm{T}}\\ \mathbf{I}_{2NK\times 2NK}\end{bmatrix}\\ &= \rho\mathbf{1}\cdot\mathbf{1}^{\mathrm{T}} + \rho\mathbf{I},\end{aligned} \tag{80}$$

where $\rho\mathbf{1}\cdot\mathbf{1}^{\mathrm{T}}$ is a positive-semidefinite matrix and $\mathbf{U}$ is also a positive-semidefinite matrix, we can obtain

$$\begin{aligned}&\left(\partial_{\boldsymbol{\delta}_{\mathrm{E},1}}\hat{L}_{\rho}\left(\boldsymbol{\delta}_{\mathrm{E},1}, \hat{\boldsymbol{\omega}}^{k+1}, \hat{\boldsymbol{\lambda}}^{k}\right) - \partial_{\boldsymbol{\delta}_{\mathrm{E},2}}\hat{L}_{\rho}\left(\boldsymbol{\delta}_{\mathrm{E},2}, \hat{\boldsymbol{\omega}}^{k+1}, \hat{\boldsymbol{\lambda}}^{k}\right)\right)^{\mathrm{T}}\\ &\cdot\left(\boldsymbol{\delta}_{\mathrm{E},1} - \boldsymbol{\delta}_{\mathrm{E},2}\right)\\ &= \left(\boldsymbol{\delta}_{\mathrm{E},1} - \boldsymbol{\delta}_{\mathrm{E},2}\right)^{\mathrm{T}}\left(2\mathbf{U} + \rho\boldsymbol{\Gamma}^{\mathrm{T}}\boldsymbol{\Gamma}\right)\left(\boldsymbol{\delta}_{\mathrm{E},1} - \boldsymbol{\delta}_{\mathrm{E},2}\right)\\ &= \left(\boldsymbol{\delta}_{\mathrm{E},1} - \boldsymbol{\delta}_{\mathrm{E},2}\right)^{\mathrm{T}}\left(2\mathbf{U} + \rho\mathbf{1}\cdot\mathbf{1}^{\mathrm{T}} + \rho\mathbf{I}\right)\left(\boldsymbol{\delta}_{\mathrm{E},1} - \boldsymbol{\delta}_{\mathrm{E},2}\right)\\ &\geq \rho\left\|\boldsymbol{\delta}_{\mathrm{E},1} - \boldsymbol{\delta}_{\mathrm{E},2}\right\|_2^2.\end{aligned} \tag{81}$$

Thus, $m = \rho$ and $\hat{L}_{\rho}\left(\boldsymbol{\delta}_{\mathrm{E}}, \hat{\boldsymbol{\omega}}^{k+1}, \hat{\boldsymbol{\lambda}}^{k}\right)$ is ρ-strongly convex with respect to $\boldsymbol{\delta}_{\mathrm{E}}$.

Similarly, by definition of strongly convex function:

$$\begin{aligned}&\hat{L}_{\rho}\left(\boldsymbol{\delta}_{\mathrm{E},1}, \hat{\boldsymbol{\omega}}^{k+1}, \hat{\boldsymbol{\lambda}}^{k}\right) \geq \hat{L}_{\rho}\left(\boldsymbol{\delta}_{\mathrm{E},2}, \hat{\boldsymbol{\omega}}^{k+1}, \hat{\boldsymbol{\lambda}}^{k}\right)\\ &+ \left(\partial_{\boldsymbol{\delta}_{\mathrm{E},2}}\hat{L}_{\rho}\left(\boldsymbol{\delta}_{\mathrm{E},2}, \hat{\boldsymbol{\omega}}^{k+1}, \hat{\boldsymbol{\lambda}}^{k}\right)\right)^{\mathrm{T}}\left(\boldsymbol{\delta}_{\mathrm{E},1} - \boldsymbol{\delta}_{\mathrm{E},2}\right)\\ &+ \frac{\rho}{2}\left\|\boldsymbol{\delta}_{\mathrm{E},1} - \boldsymbol{\delta}_{\mathrm{E},2}\right\|_2^2.\end{aligned} \tag{82}$$

The minimizer $\boldsymbol{\delta}_{\mathrm{E}}^{k+1}$ satisfies

$$\partial_{\boldsymbol{\delta}_{\mathrm{E}}^{k+1}}\hat{L}_{\rho}\left(\boldsymbol{\delta}_{\mathrm{E}}^{k+1}, \hat{\boldsymbol{\omega}}^{k+1}, \hat{\boldsymbol{\lambda}}^{k}\right) = \mathbf{0}. \tag{83}$$

Thus, for any $\boldsymbol{\delta}_{\mathrm{E}}^{k}$, the minimizer $\boldsymbol{\delta}_{\mathrm{E}}^{k+1}$ also satisfies

$$\begin{aligned}&\hat{L}_{\rho}\left(\boldsymbol{\delta}_{\mathrm{E}}^{k+1}, \hat{\boldsymbol{\omega}}^{k+1}, \hat{\boldsymbol{\lambda}}^{k}\right)\\ &\leq \hat{L}_{\rho}\left(\boldsymbol{\delta}_{\mathrm{E}}^{k}, \hat{\boldsymbol{\omega}}^{k+1}, \hat{\boldsymbol{\lambda}}^{k}\right) - \frac{\rho}{2}\left\|\boldsymbol{\delta}_{\mathrm{E}}^{k+1} - \boldsymbol{\delta}_{\mathrm{E}}^{k}\right\|_2^2.\end{aligned} \tag{84}$$

Moreover, from the definition of $\hat{L}_{\rho}$ and with the use of (70), we have

$$\begin{aligned}&\hat{L}_{\rho}\left(\boldsymbol{\delta}_{\mathrm{E}}^{k+1}, \hat{\boldsymbol{\omega}}^{k+1}, \hat{\boldsymbol{\lambda}}^{k+1}\right) - \hat{L}_{\rho}\left(\boldsymbol{\delta}_{\mathrm{E}}^{k+1}, \hat{\boldsymbol{\omega}}^{k+1}, \hat{\boldsymbol{\lambda}}^{k}\right)\\ &= \frac{1}{\rho}\left\|\hat{\boldsymbol{\lambda}}^{k+1} - \hat{\boldsymbol{\lambda}}^{k}\right\|_2^2.\end{aligned} \tag{85}$$

Then, summing (77), (84) and (85) yields

$$\hat{L}_\rho\left(\boldsymbol{\delta}_{\mathrm{E}}^{k+1},\hat{\boldsymbol{\omega}}^{k+1},\hat{\boldsymbol{\lambda}}^{k+1}\right)-\hat{L}_\rho\left(\boldsymbol{\delta}_{\mathrm{E}}^{k},\hat{\boldsymbol{\omega}}^{k},\hat{\boldsymbol{\lambda}}^{k}\right)\leq\frac{1}{\rho}\left\|\hat{\boldsymbol{\lambda}}^{k+1}-\hat{\boldsymbol{\lambda}}^{k}\right\|_2^2-\frac{\rho}{2}\left\|\boldsymbol{\delta}_{\mathrm{E}}^{k+1}-\boldsymbol{\delta}_{\mathrm{E}}^{k}\right\|_2^2. \quad (86)$$

From (72), the minimizer $\boldsymbol{\delta}_{\mathrm{E}}^{k+1}$ satisfies

$$\left(2\mathbf{U}+\rho\boldsymbol{\Gamma}^{\mathrm{T}}\boldsymbol{\Gamma}\right)\boldsymbol{\delta}_{\mathrm{E}}^{k+1}-\rho\boldsymbol{\Gamma}^{\mathrm{T}}\left(\mathbf{c}+\hat{\mathbf{w}}^{k+1}+\frac{\hat{\boldsymbol{\lambda}}^{k}}{\rho}\right)=\mathbf{0}. \quad (87)$$

Substituting (70) into (87) we have

$$2\mathbf{U}\boldsymbol{\delta}_{\mathrm{E}}^{k+1}-\boldsymbol{\Gamma}^{\mathrm{T}}\hat{\boldsymbol{\lambda}}^{k+1}=\mathbf{0}, \quad (88)$$

which means

$$\left\|\boldsymbol{\Gamma}^{\mathrm{T}}\left(\hat{\boldsymbol{\lambda}}^{k+1}-\hat{\boldsymbol{\lambda}}^{k}\right)\right\|_2^2=\left\|2\mathbf{U}\left(\boldsymbol{\delta}_{\mathrm{E}}^{k+1}-\boldsymbol{\delta}_{\mathrm{E}}^{k}\right)\right\|_2^2. \quad (89)$$

Before we proceed, we present the following property:

Proposition 8:

$$\left\|\boldsymbol{\Gamma}^{\mathrm{T}}\left(\hat{\boldsymbol{\lambda}}^{k+1}-\hat{\boldsymbol{\lambda}}^{k}\right)\right\|_2^2>0. \quad (90)$$

Proof: When $N\leq K$, $\mathbf{U}$ is full rank according to Proposition 6, which means $\mathbf{U}^{\mathrm{T}}\mathbf{U}$ is a positive definite matrix because $\operatorname{rank}\left(\mathbf{U}^{\mathrm{T}}\mathbf{U}\right)=\operatorname{rank}\left(\mathbf{U}\right)=2NK$. Therefore,

$$\left\|2\mathbf{U}\left(\boldsymbol{\delta}_{\mathrm{E}}^{k+1}-\boldsymbol{\delta}_{\mathrm{E}}^{k}\right)\right\|_2^2=4\left(\boldsymbol{\delta}_{\mathrm{E}}^{k+1}-\boldsymbol{\delta}_{\mathrm{E}}^{k}\right)^{\mathrm{T}}\mathbf{U}^{\mathrm{T}}\mathbf{U}\left(\boldsymbol{\delta}_{\mathrm{E}}^{k+1}-\boldsymbol{\delta}_{\mathrm{E}}^{k}\right)>0. \quad (91)$$

Substituting (91) into (89) we can obtain (90).

When $N>K$, we can first obtain the matrix $\boldsymbol{\Gamma}\boldsymbol{\Gamma}^{\mathrm{T}}$ as

$$\boldsymbol{\Gamma}\boldsymbol{\Gamma}^{\mathrm{T}}=\begin{bmatrix}\mathbf{1}^{\mathrm{T}}\\ \mathbf{I}_{2NK\times 2NK}\end{bmatrix}\begin{bmatrix}\mathbf{1} & \mathbf{I}_{2NK\times 2NK}\end{bmatrix}=\begin{bmatrix}2NK & \mathbf{1}^{\mathrm{T}}\\ \mathbf{1} & \mathbf{I}_{2NK\times 2NK}\end{bmatrix}. \quad (92)$$

It is easy to see that the matrix $\boldsymbol{\Gamma}\boldsymbol{\Gamma}^{\mathrm{T}}$ has one zero eigenvalue and $2NK$ positive eigenvalues. The eigenvector corresponding to the zero eigenvalue of the matrix $\boldsymbol{\Gamma}\boldsymbol{\Gamma}^{\mathrm{T}}$ is $\left[-1,\mathbf{1}^{\mathrm{T}}\right]^{\mathrm{T}}$. This also means

$$\left\|\boldsymbol{\Gamma}^{\mathrm{T}}\left(\hat{\boldsymbol{\lambda}}^{k+1}-\hat{\boldsymbol{\lambda}}^{k}\right)\right\|_2^2=0 \quad (93)$$

only if $\left(\hat{\boldsymbol{\lambda}}^{k+1}-\hat{\boldsymbol{\lambda}}^{k}\right)$ is the eigenvector corresponding to the zero eigenvalue of the matrix $\boldsymbol{\Gamma}\boldsymbol{\Gamma}^{\mathrm{T}}$. Since each element of $\hat{\boldsymbol{\lambda}}^{k}$ is constantly changing with each iteration and the relationship between the individual elements of $\hat{\boldsymbol{\lambda}}^{k}$ is random, the probability that $\left(\hat{\boldsymbol{\lambda}}^{k+1}-\hat{\boldsymbol{\lambda}}^{k}\right)$ meets the above condition is almost zero, which is consistent with (90). Together with the case of $N\leq K$, the proof is completed. ■

According to Proposition 8, there must be a $\varsigma>0$ that satisfies

$$\left\|\boldsymbol{\Gamma}^{\mathrm{T}}\left(\hat{\boldsymbol{\lambda}}^{k+1}-\hat{\boldsymbol{\lambda}}^{k}\right)\right\|_2^2=\varsigma\cdot\left\|\hat{\boldsymbol{\lambda}}^{k+1}-\hat{\boldsymbol{\lambda}}^{k}\right\|_2^2. \quad (94)$$

Since

$$\left\|2\mathbf{U}\left(\boldsymbol{\delta}_{\mathrm{E}}^{k+1}-\boldsymbol{\delta}_{\mathrm{E}}^{k}\right)\right\|_2^2\leq 4\varphi^2\left\|\boldsymbol{\delta}_{\mathrm{E}}^{k+1}-\boldsymbol{\delta}_{\mathrm{E}}^{k}\right\|_2^2, \quad (95)$$

where $\varphi=\operatorname{eig}_{\max}\left(\mathbf{U}\right)$, we can obtain

$$\varsigma\cdot\left\|\hat{\boldsymbol{\lambda}}^{k+1}-\hat{\boldsymbol{\lambda}}^{k}\right\|_2^2\leq 4\varphi^2\left\|\boldsymbol{\delta}_{\mathrm{E}}^{k+1}-\boldsymbol{\delta}_{\mathrm{E}}^{k}\right\|_2^2. \quad (96)$$

Substituting (96) into (86) further results in

$$\hat{L}_\rho\left(\boldsymbol{\delta}_{\mathrm{E}}^{k+1},\hat{\boldsymbol{\omega}}^{k+1},\hat{\boldsymbol{\lambda}}^{k+1}\right)\leq\hat{L}_\rho\left(\boldsymbol{\delta}_{\mathrm{E}}^{k},\hat{\boldsymbol{\omega}}^{k},\hat{\boldsymbol{\lambda}}^{k}\right)-\left(\frac{\rho}{2}-\frac{4\varphi^2}{\varsigma\cdot\rho}\right)\left\|\boldsymbol{\delta}_{\mathrm{E}}^{k+1}-\boldsymbol{\delta}_{\mathrm{E}}^{k}\right\|_2^2, \quad (97)$$

which means if the condition

$$\rho>\frac{2\sqrt{2}\varphi}{\sqrt{\varsigma}} \quad (98)$$

holds, $\hat{L_\rho}$ is monotonously decreasing in the iteration procedure. This completes the proof for convergence.

D. Complexity Analysis

We analyze the computational complexity of the proposed algorithm in terms of the number of real multiplication operations.

Considering the special structure of $\boldsymbol{\Gamma}$ shown in (65) and $\boldsymbol{\Gamma}^{\mathrm{T}}\boldsymbol{\Gamma}$ shown in (80), the computational complexity can be greatly reduced. In Line 5 of Algorithm 2, the computation of $\left(2\mathbf{U}+\boldsymbol{\Gamma}^{\mathrm{T}}\boldsymbol{\Gamma}\right)^{-1}$ requires $\frac{1}{3}\left(2NK\right)^3$ real multiplications, which involve the cost of computing the Cholesky factorization. The computation of the matrix product of $\left(2\mathbf{U}+\boldsymbol{\Gamma}^{\mathrm{T}}\boldsymbol{\Gamma}\right)^{-1}\boldsymbol{\Gamma}^{\mathrm{T}}\left(\mathbf{c}+\hat{\mathbf{w}}^{k}+\hat{\boldsymbol{\lambda}}^{k}\right)$ requires $2NK\left(2NK+1\right)$ real multiplications. The total cost of $\boldsymbol{\delta}_{\mathrm{E}}$-update is $\left(\frac{1}{3}\left(2NK\right)^3+2NK\left(2NK+1\right)\right)$ real multiplications. The cost of projection $\hat{\boldsymbol{\omega}}$-update in Line 6 is negligible. And the $\hat{\boldsymbol{\lambda}}$-update in Line 7 does not require any real multiplication. Therefore, Algorithm 2 requires $\left(\frac{1}{3}\left(2NK\right)^3+2NK\left(2NK+1\right)\right)$ real multiplications when $k=1$.

Given that $\left(2\mathbf{U}+\boldsymbol{\Gamma}^{\mathrm{T}}\boldsymbol{\Gamma}\right)^{-1}$ does not change in each iteration, we can cache the result to perform the subsequent iterations efficiently. Accordingly, Algorithm 2 requires $2NK\left(2NK+1\right)$ real multiplications for each iteration when $k\geq 2$. The total number of real multiplications for Algorithm 2 is $\left[\frac{1}{3}\left(2NK\right)^3+T\cdot 2NK\left(2NK+1\right)\right]$, where T denotes the number of iterations.

VI. Simulation Result

We assume standard Rayleigh fading channel, random complex Gaussian distributed noise. The transmit power budget is $p_0=1$ per slot, and the SNR is defined as $\frac{1}{\sigma^2}$, where σ^2 is the noise power. The execution time results are obtained from a Windows 11 Desktop with i9-10900 and 16GB RAM.

For clarity, the following abbreviations are used throughout this section:

1) ZF: Traditional ZF precoding with block-level power normalization;
2) RZF: Traditional RZF precoding with block-level power normalization;

This article has been accepted for publication in IEEE Transactions on Communications. 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/TCOMM.2024.3480981

3) CI-SLP: Traditional CI-SLP method solved by IPM [35], $\mathcal{P}_8$ in [26];
4) CI-BLP-CVX: CI-BLP method solved by *CVX*, $\mathcal{P}_1$;
5) CI-BLP-IPM: CI-BLP method solved by IPM [35], $\mathcal{P}_2$;
6) CI-BLP-ADMM-P2-(K_{max}): CI-BLP method solved by the ADMM algorithm based on $\mathcal{P}_2$ with the maximum number of iterations K_{max};
7) CI-BLP-ADMM-P5-(K_{max}): CI-BLP method solved by the improved ADMM algorithm based on $\mathcal{P}_5$ with the maximum number of iterations K_{max}.

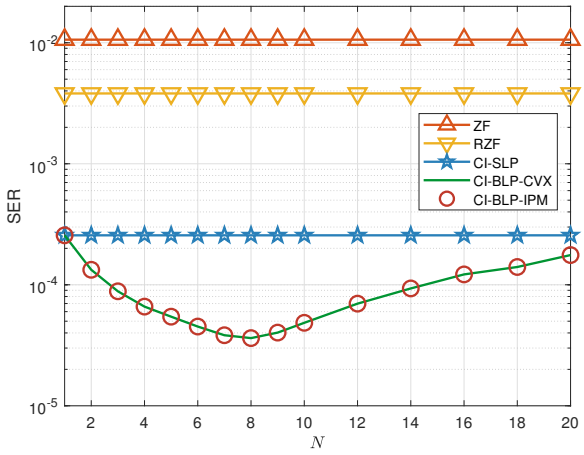

Fig. 4. SER performance of different schemes, $N_t = K = 10$, 8PSK, transmit SNR=35dB.

To verify the closed-form expression of the CI-BLP optimal precoding matrix for $N < K$, in Fig. 4 we depict the SER with respect to the block length N for a 10×10 MU-MISO system, where 8PSK modulation is employed at a transmit SNR of 35dB. As can be observed, CI-BLP returns the same SER performance as CI-SLP when $N = 1$, because CI-BLP reduces to CI-SLP when optimized independently for each symbol slot. As the block length N increases, we observe that the SER performance firstly improves since the benefit of the relaxed power constraint outweighs the loss due to using a fixed precoder over the block, while the SER performance becomes worse as N further increases, because the benefit of the relaxed power constraint cannot further compensate for the loss of the fixed precoder. In Fig. 4, CI-BLP-IPM returns the same SER performance as CI-BLP-CVX when $N < K$, which proves that our scenario expansion is feasible.

Fig. 5 to depicts the convergence results of the proposed ADMM algorithms for a 10×10 MU-MISO system with 8PSK modulation, where the length of the block is $N = 8$. In both algorithms, we take $\rho = 1$. Here, $\boldsymbol{\delta}_{\mathrm{E}}$ represents the solution obtained by the proposed ADMM algorithm, and $\boldsymbol{\delta}_{\mathrm{E}}^*$ represents the optimal solution obtained by IPM. It can be seen that there is no gap between the optimal solution and the solution respectively obtained by the two proposed ADMM algorithms with 600 iterations, which therefore means that both ADMM algorithms can reach the optimal solution when the iteration number is sufficient large.

Fig. 6 compares the SER performance of the conventional

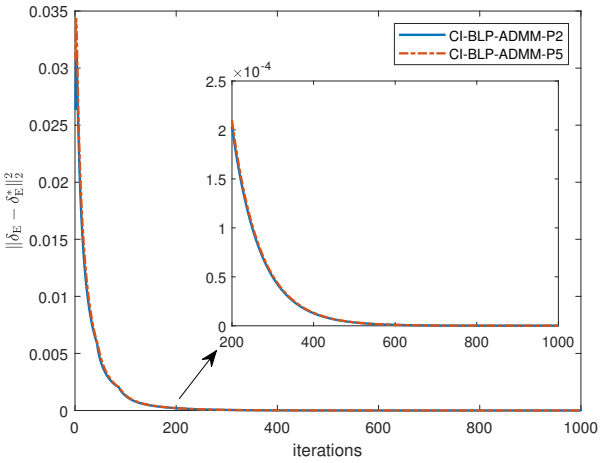

Fig. 5. Convergence behavior of the proposed ADMM algorithms, $N_t = K = 10$, $N = 8$, 8PSK.

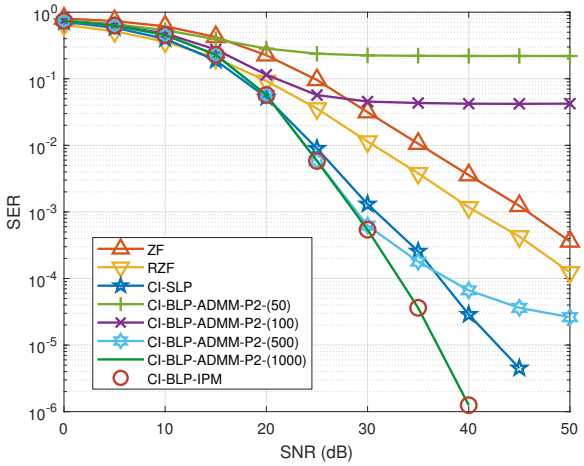

Fig. 6. SER performance of different schemes, $N_t = K = 10, N = 8$, 8PSK.

ADMM scheme based on $\mathcal{P}_2$ with other precoding schemes for a 10×10 MU-MISO system with 8PSK modulation, where the length of the block is $N = 8$. Although the conventional ADMM algorithm can obtain the optimal solution with 1000 iterations, the result with 500 iterations is still not ideal. This is because a direct application of the ADMM framework based on $\mathcal{P}_2$ inevitably results in the $\boldsymbol{\delta}_{\mathrm{E}}$-update being a constrained subproblem. Although such a method is common, it will affect the convergence process.

Fig. 7 compares the SER performance of the improved ADMM scheme based on $\mathcal{P}_5$ with other precoding schemes for a 10×10 MU-MISO system with 8PSK modulation. When $N = 8$, we observe that CI-BLP offers noticeable performance gains over traditional CI-SLP, owing to the relaxed power constraint over the entire block. The improved ADMM algorithm based on $\mathcal{P}_5$ can achieve satisfactory results after 50 or even 30 iterations, much faster than the conventional ADMM algorithm based on $\mathcal{P}_2$. When SNR$<$ 35dB, CI-BLP-ADMM returns the better SER performance than CI-SLP after only

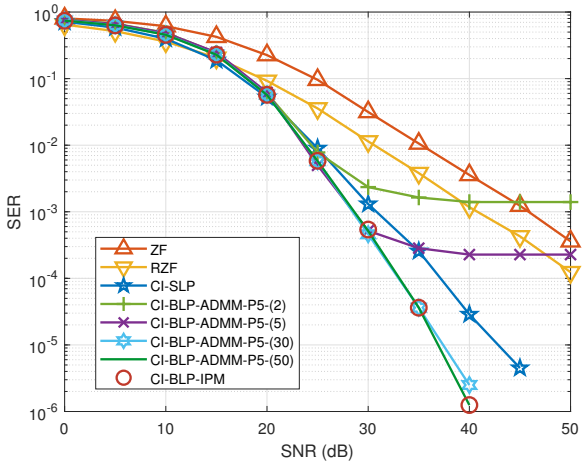

Fig. 7. SER performance of different schemes, $N_t = K = 10, N = 8$, 8PSK.

5 iterations. When the SNR of the practical system is high, the proposed ADMM design can obtain a finer solution by increasing the number of iterations, and thus obtain a more significant performance gain. Therefore, the proposed ADMM design can achieve a flexible trade-off between communication performance and execution time by modifying the maximum number of iterations.

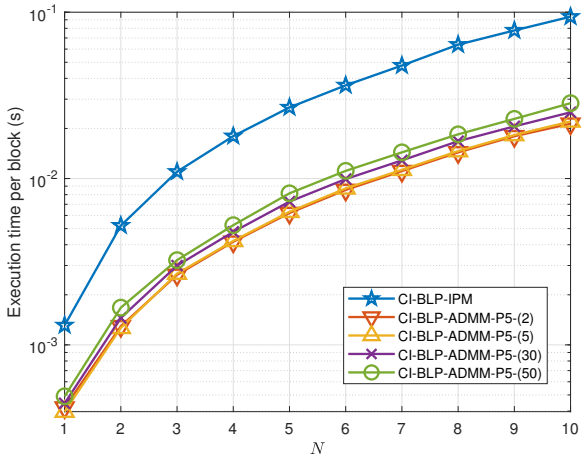

Fig. 8. Execution time per block of different schemes, $N_t = K = 10$, 8PSK, 30dB.

In Fig. 8 we compare the execution time required for each scheme as an indication to show the potential complexity benefits of the proposed ADMM algorithm. Since the size of the QP problem is independent of the modulation type, the modulation does not significantly affect the complexity, which is determined primarily by the number of users and transmit antennas. It is observed that the proposed ADMM algorithm is much faster than the IPM, which motivates the use of the CI-BLP method in practice.

VII. Conclusion

In this paper, we focus on extending the analysis on CI-BLP to the case where the number of symbol slots in a block is smaller than the number of users and obtain a QP optimization on simplex in this case. We study the possibility of applying the iterative closed-form algorithm for QP problem in CI-SLP to CI-BLP. We further propose an improved ADMM algorithm. All subproblems have simple closed-form solutions. The proposed algorithm is shown to achieve an identical performance to IPM with a reduced computational cost, which enables the use of CI-BLP method in practical wireless systems.

Appendix A
Proof for Proposition 1

By the definition of eigenvalues, for any $m \in \{r_g+1, r_g+2, \cdots, M_g\}$,

$$\mathbf{G} \cdot \mathbf{v}_{\mathrm{G}}^m = 0 \cdot \mathbf{v}_{\mathrm{G}}^m = \mathbf{0}. \tag{99}$$

By substituting (34) into (99), we can obtain

$$\mathbf{0} = \sum_{n=1}^{N_g} \mathbf{s}_{\mathrm{G}}^n \left(\mathbf{s}_{\mathrm{G}}^n\right)^{\mathrm{T}} \cdot \mathbf{v}_{\mathrm{G}}^m. \tag{100}$$

Since $\mathbf{s}_{\mathrm{G}}^1, \mathbf{s}_{\mathrm{G}}^2, \cdots, \mathbf{s}_{\mathrm{G}}^{r_g}$ is a maximal linearly independent system, we can get

$$\mathbf{s}_{\mathrm{G}}^n = \sum_{i=1}^{r_g} k_{n-r_g,i} \mathbf{s}_{\mathrm{G}}^i, \ n \in \{r_g+1, \cdots, N_g\}. \tag{101}$$

By substituting (101) into (100), we can obtain

$$\begin{aligned}
\mathbf{0} &= \sum_{n=1}^{r_g} \mathbf{s}_{\mathrm{G}}^n \left(\mathbf{s}_{\mathrm{G}}^n\right)^{\mathrm{T}} \cdot \mathbf{v}_{\mathrm{G}}^m \\
&\quad + \sum_{n=r_g+1}^{N_g} \left(\sum_{i=1}^{r_g} k_{n-r_g,i} \mathbf{s}_{\mathrm{G}}^i\right) \left(\sum_{j=1}^{r_g} k_{n-r_g,j} \mathbf{s}_{\mathrm{G}}^j\right)^{\mathrm{T}} \cdot \mathbf{v}_{\mathrm{G}}^m \\
&= \underbrace{l_{1,1}\mathbf{s}_{\mathrm{G}}^1}_{\hat{\mathbf{s}}_{\mathrm{G}}^1} \cdot \left(\mathbf{s}_{\mathrm{G}}^1\right)^{\mathrm{T}} \cdot \mathbf{v}_{\mathrm{G}}^m + \underbrace{\left(l_{1,2}\mathbf{s}_{\mathrm{G}}^1 + l_{2,2}\mathbf{s}_{\mathrm{G}}^2\right)}_{\hat{\mathbf{s}}_{\mathrm{G}}^2} \cdot \left(\mathbf{s}_{\mathrm{G}}^2\right)^{\mathrm{T}} \cdot \mathbf{v}_{\mathrm{G}}^m \\
&\quad + \cdots + \underbrace{\left(l_{1,r_g}\mathbf{s}_{\mathrm{G}}^1 + l_{2,r_g}\mathbf{s}_{\mathrm{G}}^2 + \cdots + l_{r_g,r_g}\mathbf{s}_{\mathrm{G}}^{r_g}\right)}_{\hat{\mathbf{s}}_{\mathrm{G}}^{r_g}} \cdot \left(\mathbf{s}_{\mathrm{G}}^{r_g}\right)^{\mathrm{T}} \cdot \mathbf{v}_{\mathrm{G}}^m,
\end{aligned} \tag{102}$$

where

$$l_{i,j} = \begin{cases} 1 + \sum_{n=1}^{N_g - rg} \left(k_{n,i}\right)^2 & \text{if } i = j, \\ 2\sum_{n=1}^{N_g - rg} k_{n,i} k_{n,j} & \text{if } i \neq j. \end{cases} \tag{103}$$

Based on (103), we can obtain that $l_{i,i} \neq 0$, which means that $\hat{\mathbf{s}}_{\mathrm{G}}^1, \hat{\mathbf{s}}_{\mathrm{G}}^2, \cdots, \hat{\mathbf{s}}_{\mathrm{G}}^{r_g}$ are linearly independent. Therefore,

$$\left(\mathbf{s}_{\mathrm{G}}^n\right)^{\mathrm{T}} \cdot \mathbf{v}_G^m = 0, \ n \in \{1, \cdots, r_g\}. \tag{104}$$

For $n \in \{r_g+1, \cdots, N_g\}$, combined with (101), we can further obtain

$$\begin{aligned}(\mathbf{s}_{\mathrm{G}}^n)^{\mathrm{T}} \cdot \mathbf{v}_{\mathrm{G}}^m &= \left(\sum_{i=1}^{r_g} k_{n-r_g,i}\mathbf{s}_{\mathrm{G}}^i\right)^{\mathrm{T}} \cdot \mathbf{v}_{\mathrm{G}}^m \\ &= \sum_{i=1}^{r_g} k_{n-r_g,i}\left(\mathbf{s}_{\mathrm{G}}^i\right)^{\mathrm{T}} \cdot \mathbf{v}_{\mathrm{G}}^m = 0,\end{aligned} \tag{105}$$

which completes the proof.

Appendix B
Proof for Proposition 2

We introduce $\mathbf{x}^n$ into this proof:

$$\mathbf{x}^n = \mathbf{s}_{\mathrm{E}}^n,\ \mathbf{x}^{N+n} = \mathbf{c}_{\mathrm{E}}^n,\ n \in \{1, \cdots, N\}. \tag{106}$$

According to (18), $\mathbf{D}$ can be expressed as

$$\mathbf{D} = \sum_{n=1}^{2N} \mathbf{x}^n \left(\mathbf{x}^n\right)^{\mathrm{T}} = \left[\mathbf{d}_1, \mathbf{d}_2, \cdots, \mathbf{d}_{2K}\right]^{\mathrm{T}}. \tag{107}$$

For any $m \in \{1, \cdots, 2K\}$,

$$\mathbf{d}_m = x_m^1 \cdot \mathbf{x}^1 + x_m^2 \cdot \mathbf{x}^2 + \cdots + x_m^{2N} \cdot \mathbf{x}^{2N}, \tag{108}$$

where x_i^n is the i-th entry of $\mathbf{x}^n$. Since $\mathbf{x}^1, \mathbf{x}^2, \cdots, \mathbf{x}^{2N}$ are linearly independent, the row rank of $\mathbf{D}$ is $\min\{2K, 2N\}$. In the same way we can get that the column rank of $\mathbf{D}$ is also $\min\{2K, 2N\}$, which completes the proof.

Appendix C
Proof for Proposition 3

According to (22) and (28), we introduce $\mathbf{x}^n$ and $\mathbf{y}^n$ into this proof and define them as

$$\mathbf{x}^n = \mathbf{V}_{\mathrm{D}}^{\mathrm{T}}\mathbf{s}_{\mathrm{E}}^n,\ \mathbf{y}^n = \mathbf{V}_{\mathrm{D}}^{\mathrm{T}}\mathbf{c}_{\mathrm{E}}^n,\ n \in \{1, \cdots, N\}. \tag{109}$$

Then, $p_{m,n}$, $f_{m,n}$, $g_{m,n}$, $q_{m,n}$ can be written as

$$\begin{aligned}p_{m,n} &= \left(\mathbf{s}_{\mathrm{E}}^n\right)^{\mathrm{T}} \mathbf{D}^{+}\mathbf{s}_{\mathrm{E}}^m = \left(\mathbf{x}^n\right)^{\mathrm{T}} \boldsymbol{\Sigma}_{\mathrm{D}}^{+}\mathbf{x}^m = \sum_{i=1}^{r_d} \frac{1}{\sigma_i} x_i^m x_i^n, \\ f_{m,n} &= \left(\mathbf{c}_{\mathrm{E}}^n\right)^{\mathrm{T}} \mathbf{D}^{+}\mathbf{s}_{\mathrm{E}}^m = \left(\mathbf{y}^n\right)^{\mathrm{T}} \boldsymbol{\Sigma}_{\mathrm{D}}^{+}\mathbf{x}^m = \sum_{i=1}^{r_d} \frac{1}{\sigma_i} x_i^m y_i^n, \\ g_{m,n} &= \left(\mathbf{s}_{\mathrm{E}}^n\right)^{\mathrm{T}} \mathbf{D}^{+}\mathbf{c}_{\mathrm{E}}^m = \left(\mathbf{x}^n\right)^{\mathrm{T}} \boldsymbol{\Sigma}_{\mathrm{D}}^{+}\mathbf{y}^m = \sum_{i=1}^{r_d} \frac{1}{\sigma_i} y_i^m x_i^n, \\ q_{m,n} &= \left(\mathbf{c}_{\mathrm{E}}^n\right)^{\mathrm{T}} \mathbf{D}^{+}\mathbf{c}_{\mathrm{E}}^m = \left(\mathbf{y}^n\right)^{\mathrm{T}} \boldsymbol{\Sigma}_{\mathrm{D}}^{+}\mathbf{y}^m = \sum_{i=1}^{r_d} \frac{1}{\sigma_i} y_i^m y_i^n,\end{aligned} \tag{110}$$

where σ_i is the i-th eigenvalues of $\mathbf{D}$. We further introduce $\mathbf{z}_i$ into this proof:

$$\mathbf{z}_i = \frac{1}{\sigma_i}\left[x_i^1, y_i^1, x_i^2, y_i^2, \cdots, x_i^N, y_i^N\right]. \tag{111}$$

where $i \in \{1, \cdots, r_d\}$. It is obvious that $\mathbf{z}_1, \mathbf{z}_2, \cdots, \mathbf{z}_{r_d}$ are linearly independent and each row of $\boldsymbol{\Delta}$ can be expressed as:

$$\begin{aligned}\boldsymbol{\Delta}_{2n-1} &= x_1^n\mathbf{z}_1 + x_2^n\mathbf{z}_2 + \cdots + x_{r_d}^n\mathbf{z}_{r_d}, \\ \boldsymbol{\Delta}_{2n} &= y_1^n\mathbf{z}_1 + y_2^n\mathbf{z}_2 + \cdots + y_{r_d}^n\mathbf{z}_{r_d},\end{aligned} \tag{112}$$

where $n \in \{1, \cdots, N\}$. Thus the row rank of $\boldsymbol{\Delta}$ is r_d. In the same way we can get that the column rank of $\boldsymbol{\Delta}$ is also r_d, which completes the proof.

Appendix D
Proof for Proposition 4

$\boldsymbol{\Theta}$ can be decomposed as

$$\boldsymbol{\Theta} = \hat{\mathbf{A}}\hat{\mathbf{B}}, \tag{113}$$

where $\hat{\mathbf{A}} \in \mathbb{R}^{4NK \times 2NK}$ and $\hat{\mathbf{B}} \in \mathbb{R}^{2NK \times 4NK}$ are defined as

$$\begin{aligned}\hat{\mathbf{A}} &= \mathbf{1}^{\mathrm{T}} \otimes \left[\left(\mathbf{A}^1\right)^{\mathrm{T}} \quad \left(\mathbf{B}^1\right)^{\mathrm{T}} \quad \cdots \quad \left(\mathbf{A}^N\right)^{\mathrm{T}} \quad \left(\mathbf{B}^N\right)^{\mathrm{T}}\right]^{\mathrm{T}}, \\ \hat{\mathbf{B}} &= \mathrm{Diag}\left(\left[\left(\mathbf{A}^1\right)^{\mathrm{T}} \quad \left(\mathbf{B}^1\right)^{\mathrm{T}} \quad \cdots \quad \left(\mathbf{A}^N\right)^{\mathrm{T}} \quad \left(\mathbf{B}^N\right)^{\mathrm{T}}\right]\right).\end{aligned} \tag{114}$$

It is obvious that the rank of $\hat{\mathbf{A}}$ is K and we can obtain

$$\mathrm{rank}\left(\boldsymbol{\Theta}\right) \le \mathrm{rank}\left(\hat{\mathbf{A}}\right) = K. \tag{115}$$

From another perspective, $\hat{\mathbf{A}}$ can be decomposed as

$$\hat{\mathbf{A}} = \boldsymbol{\Theta}\tilde{\mathbf{B}}, \tag{116}$$

where $\tilde{\mathbf{B}} \in \mathbb{R}^{4NK \times 2NK}$ is defined as (117). Then we can obtain

$$\mathrm{rank}\left(\boldsymbol{\Theta}\right) \ge \mathrm{rank}\left(\hat{\mathbf{A}}\right) = K. \tag{118}$$

With (115), we can get

$$\mathrm{rank}\left(\boldsymbol{\Theta}\right) = \mathrm{rank}\left(\hat{\mathbf{A}}\right) = K, \tag{119}$$

which completes the proof.

Appendix E
Proof for Proposition 5

First, we introduce $\mathbf{Y} \in \mathbb{R}^{3\times3}$ and $\mathbf{X} \in \mathbb{R}^{9\times9}$ into this proof and they are defined as

$$\mathbf{Y} = \begin{bmatrix}\mathbf{y}_1 \\ \mathbf{y}_2 \\ \mathbf{y}_3\end{bmatrix} = \begin{bmatrix}y_{1,1} & y_{1,2} & y_{1,3} \\ y_{2,1} & y_{2,2} & y_{2,3} \\ y_{3,1} & y_{3,2} & y_{3,3}\end{bmatrix}, \tag{120}$$

$$\mathbf{X} = \begin{bmatrix}\mathbf{x}_1 \\ \mathbf{x}_2 \\ \vdots \\ \mathbf{x}_9\end{bmatrix} = \begin{bmatrix}\mathbf{X}^{1,1} & \mathbf{X}^{1,2} & \mathbf{X}^{1,3} \\ \mathbf{X}^{2,1} & \mathbf{X}^{2,2} & \mathbf{X}^{2,3} \\ \mathbf{X}^{3,1} & \mathbf{X}^{3,2} & \mathbf{X}^{3,3}\end{bmatrix}, \tag{121}$$

where $\mathbf{X}^{m,n} \in \mathbb{R}^{3\times3}$ is defined as

$$\mathbf{X}^{m,n} = \begin{bmatrix}\mathbf{x}_1^{m,n} \\ \mathbf{x}_2^{m,n} \\ \mathbf{x}_3^{m,n}\end{bmatrix} = \begin{bmatrix}x_{1,1}^{m,n} & x_{1,2}^{m,n} & x_{1,3}^{m,n} \\ x_{2,1}^{m,n} & x_{2,2}^{m,n} & x_{2,3}^{m,n} \\ x_{3,1}^{m,n} & x_{3,2}^{m,n} & x_{3,3}^{m,n}\end{bmatrix}. \tag{122}$$

We suppose that $\mathbf{y}_1$ and $\mathbf{y}_2$ are linearly independent, and $\mathrm{rank}\left(\mathbf{Y}\right) = r_1 = 2$. We also suppose that $\mathbf{x}_1^{m,n}$ and $\mathbf{x}_2^{m,n}$ are linearly independent, and $\mathrm{rank}\left(\mathbf{X}\right) = \mathrm{rank}\left(\mathbf{X}^{m,n}\right) = r_2 = 2$,

$$\tilde{\mathbf{B}} = \mathrm{Diag}\left(\left[\mathbf{A}^1\left(\left(\mathbf{A}^1\right)^{\mathrm{T}}\mathbf{A}^1\right)^{-1} \quad \mathbf{B}^1\left(\left(\mathbf{B}^1\right)^{\mathrm{T}}\mathbf{B}^1\right)^{-1} \quad \cdots \quad \mathbf{A}^N\left(\left(\mathbf{A}^N\right)^{\mathrm{T}}\mathbf{A}^N\right)^{-1} \quad \mathbf{B}^N\left(\left(\mathbf{B}^N\right)^{\mathrm{T}}\mathbf{B}^N\right)^{-1}\right]\right). \tag{117}$$

which means $\mathbf{x}_1$ and $\mathbf{x}_2$ are linearly independent. Then we can get the following expression

$$\begin{aligned}\mathbf{y}_3 =& l^1\mathbf{y}_1 + l^2\mathbf{y}_2,\\ \mathbf{x}_i =& k_i^1\mathbf{x}_1 + k_i^2\mathbf{x}_2,\ i \in \{3,\cdots,9\}.\end{aligned} \tag{123}$$

We further introduce $\mathbf{Z} \in \mathbb{R}^{9\times 9}$ and $\tilde{\mathbf{V}} \in \mathbb{R}^{4\times 9}$, defined as

$$\mathbf{Z} = \begin{bmatrix}\mathbf{z}_1\\ \mathbf{z}_2\\ \vdots\\ \mathbf{z}_9\end{bmatrix} = \begin{bmatrix} y_{1,1}\mathbf{X}^{1,1} & y_{1,2}\mathbf{X}^{1,2} & y_{1,3}\mathbf{X}^{1,3}\\ y_{2,1}\mathbf{X}^{2,1} & y_{2,2}\mathbf{X}^{2,2} & y_{2,3}\mathbf{X}^{2,3}\\ y_{3,1}\mathbf{X}^{3,1} & y_{3,2}\mathbf{X}^{3,2} & y_{3,3}\mathbf{X}^{3,3}\end{bmatrix}, \tag{124}$$

$$\tilde{\mathbf{V}} = \begin{bmatrix}\tilde{\mathbf{v}}_1\\ \tilde{\mathbf{v}}_2\\ \tilde{\mathbf{v}}_3\\ \tilde{\mathbf{v}}_4\end{bmatrix} = \begin{bmatrix} y_{1,1}\mathbf{x}_1^{1,1} & y_{1,2}\mathbf{x}_1^{1,2} & y_{1,3}\mathbf{x}_1^{1,3}\\ y_{1,1}\mathbf{x}_2^{1,1} & y_{1,2}\mathbf{x}_2^{1,2} & y_{1,3}\mathbf{x}_2^{1,3}\\ y_{2,1}\mathbf{x}_1^{1,1} & y_{2,2}\mathbf{x}_1^{1,2} & y_{2,3}\mathbf{x}_1^{1,3}\\ y_{2,1}\mathbf{x}_2^{1,1} & y_{2,2}\mathbf{x}_2^{1,2} & y_{2,3}\mathbf{x}_2^{1,3}\end{bmatrix}. \tag{125}$$

According to (123), each row of $\mathbf{Z}$ can be expressed as a linear combination of $\tilde{\mathbf{v}}_1, \tilde{\mathbf{v}}_2, \tilde{\mathbf{v}}_3, \tilde{\mathbf{v}}_4$. Next we prove that $\tilde{\mathbf{v}}_1, \tilde{\mathbf{v}}_2, \tilde{\mathbf{v}}_3, \tilde{\mathbf{v}}_4$ are linearly independent.

Assuming that $\tilde{\mathbf{v}}_1, \tilde{\mathbf{v}}_2, \tilde{\mathbf{v}}_3, \tilde{\mathbf{v}}_4$ are linearly dependent, then there exist coefficients t_1, t_2, t_3, t_4 that are not all zero, such that

$$t_1\tilde{\mathbf{v}}_1 + t_2\tilde{\mathbf{v}}_2 + t_3\tilde{\mathbf{v}}_3 + t_4\tilde{\mathbf{v}}_4 = \mathbf{0}^{\mathrm{T}}, \tag{126}$$

that is,

$$\begin{aligned}\mathbf{0}^{\mathrm{T}} =& t_1y_{1,1}\mathbf{x}_1^{1,1} + t_2y_{1,1}\mathbf{x}_2^{1,1} + t_3y_{2,1}\mathbf{x}_1^{1,1} + t_4y_{2,1}\mathbf{x}_2^{1,1}\\ =& \left(t_1y_{1,1} + t_3y_{2,1}\right)\mathbf{x}_1^{1,1} + \left(t_2y_{1,1} + t_4y_{2,1}\right)\mathbf{x}_2^{1,1},\\ \mathbf{0}^{\mathrm{T}} =& t_1y_{1,2}\mathbf{x}_1^{1,2} + t_2y_{1,2}\mathbf{x}_2^{1,2} + t_3y_{2,2}\mathbf{x}_1^{1,2} + t_4y_{2,2}\mathbf{x}_2^{1,2}\\ =& \left(t_1y_{1,2} + t_3y_{2,2}\right)\mathbf{x}_1^{1,2} + \left(t_2y_{1,2} + t_4y_{2,2}\right)\mathbf{x}_2^{1,2},\\ \mathbf{0}^{\mathrm{T}} =& t_1y_{1,3}\mathbf{x}_1^{1,3} + t_2y_{1,3}\mathbf{x}_2^{1,3} + t_3y_{2,3}\mathbf{x}_1^{1,3} + t_4y_{2,3}\mathbf{x}_2^{1,3}\\ =& \left(t_1y_{1,3} + t_3y_{2,3}\right)\mathbf{x}_1^{1,3} + \left(t_2y_{1,3} + t_4y_{2,3}\right)\mathbf{x}_2^{1,3}.\end{aligned} \tag{127}$$

Since $\mathbf{x}_1^{m,n}$ and $\mathbf{x}_2^{m,n}$ are linearly independent, we can get

$$\begin{aligned}t_1y_{1,1} + t_3y_{2,1} = t_2y_{1,1} + t_4y_{2,1} = 0,\\ t_1y_{1,2} + t_3y_{2,2} = t_2y_{1,2} + t_4y_{2,2} = 0,\\ t_1y_{1,3} + t_3y_{2,3} = t_2y_{1,3} + t_4y_{2,3} = 0.\end{aligned} \tag{128}$$

Since t_1, t_2, t_3, t_4 are not all zero, the following equation must be satisfied

$$\frac{y_{1,1}}{y_{2,1}} = \frac{y_{1,2}}{y_{2,2}} = \frac{y_{1,3}}{y_{2,3}}, \tag{129}$$

which contradicts the fact that $\mathbf{y}_1$ and $\mathbf{y}_2$ are linearly independent. Therefore, the assumption does not hold. $\tilde{\mathbf{v}}_1, \tilde{\mathbf{v}}_2, \tilde{\mathbf{v}}_3, \tilde{\mathbf{v}}_4$ are linearly independent. And $\mathrm{rank}\left(\mathbf{Z}\right) = \mathrm{rank}\left(\tilde{\mathbf{V}}\right) = r_1 \cdot r_2$.

We extend this process by substituting $\boldsymbol{\Delta}$ into $\mathbf{Y}$ and $\boldsymbol{\Theta}$ into $\mathbf{X}$. And then we can obtain

$$\mathrm{rank}\left(\hat{\mathbf{U}}\right) = \mathrm{rank}\left(\boldsymbol{\Delta}\right) \cdot \mathrm{rank}\left(\boldsymbol{\Theta}\right), \tag{130}$$

which completes the proof.

Appendix F
Proof for Proposition 6

We introduce $\hat{\mathbf{X}} = [\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4, \mathbf{x}_5, \mathbf{x}_6]$ into this proof. We suppose that $\mathrm{rank}\left(\hat{\mathbf{X}}\right) = 3$ and $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_4$ are linearly independent, which means

$$\mathbf{x}_n = k_n^1\mathbf{x}_1 + k_n^2\mathbf{x}_2 + k_n^4\mathbf{x}_4,\ n \in \{3,5,6\}. \tag{131}$$

We further define

$$\mathbf{X} = \hat{\mathbf{X}} \cdot \begin{bmatrix}\mathbf{I}_3 & \mathbf{I}_3\end{bmatrix}^{\mathrm{T}} = [\mathbf{x}_1 + \mathbf{x}_4, \mathbf{x}_2 + \mathbf{x}_5, \mathbf{x}_3 + \mathbf{x}_6]. \tag{132}$$

It is obvious that $\mathrm{rank}\left(\mathbf{X}\right) \le \mathrm{rank}\left(\hat{\mathbf{X}}\right)$.

Assuming that $\mathrm{rank}\left(\mathbf{X}\right) < \mathrm{rank}\left(\hat{\mathbf{X}}\right)$. Then $\mathbf{x}_1 + \mathbf{x}_4, \mathbf{x}_2 + \mathbf{x}_5, \mathbf{x}_3 + \mathbf{x}_6$ are linearly dependent, which means there exist coefficients t_1, t_2, t_3 that are not all zero, such that

$$\begin{aligned}\mathbf{0}^{\mathrm{T}} =& t_1 \cdot (\mathbf{x}_1 + \mathbf{x}_4) + t_2 \cdot (\mathbf{x}_2 + \mathbf{x}_5) + t_3 \cdot (\mathbf{x}_3 + \mathbf{x}_6)\\ =& \left(t_1 + t_2k_5^1 + t_3\left(k_3^1 + k_6^1\right)\right) \cdot \mathbf{x}_1\\ &+ \left(t_2\left(1 + k_5^2\right) + t_3\left(k_3^2 + k_6^2\right)\right) \cdot \mathbf{x}_2\\ &+ \left(t_1 + t_2k_5^4 + t_3\left(k_3^4 + k_6^4\right)\right) \cdot \mathbf{x}_4.\end{aligned} \tag{133}$$

Since $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_4$ are linearly independent, we can get

$$\begin{aligned}t_1 + t_2k_5^1 + t_3\left(k_3^1 + k_6^1\right) = 0,\\ t_2\left(1 + k_5^2\right) + t_3\left(k_3^2 + k_6^2\right) = 0,\\ t_1 + t_2k_5^4 + t_3\left(k_3^4 + k_6^4\right) = 0.\end{aligned} \tag{134}$$

Since t_1, t_2, t_3 are not all zero, the following equation must be satisfied

$$\frac{\left(k_3^2 + k_6^2\right)k_5^1}{1 + k_5^2} + k_3^1 + k_6^1 = \frac{\left(k_3^2 + k_6^2\right)k_5^4}{1 + k_5^2} + k_3^4 + k_6^4. \tag{135}$$

The coefficients in (135) are determined by channels and symbols, and they are random data. Therefore, the left-hand side and right-hand side of (135) are strictly equal with probability 0, which means the assumption does not hold. Therefore, $\mathrm{rank}\left(\mathbf{X}\right) = \mathrm{rank}\left(\hat{\mathbf{X}}\right)$. We extend this process by substituting $\hat{\mathbf{U}}$ into $\hat{\mathbf{X}}$ and $\mathbf{U}$ into $\mathbf{X}$. And then we can obtain

$$\mathrm{rank}\left(\mathbf{U}\right) = \mathrm{rank}\left(\hat{\mathbf{U}}\right) = \min\left\{2NK, 2K^2\right\}, \tag{136}$$

which completes the proof.

References

[1] L. Zheng and D. N. C. Tse, "Diversity and Multiplexing: A Fundamental Tradeoff in Multiple-Antenna Channels," *IEEE Trans. Inf. Theory*, vol. 49, no. 5, pp. 1073–1096, May 2003.

[2] M. Costa, "Writing on Dirty Paper," *IEEE Trans. Inf. Theory*, vol. 29, no. 3, pp. 439–441, May 1983.

[3] T. Haustein, C. von Helmolt, E. Jorswieck, V. Jungnickel, and V. Pohl, "Performance of MIMO Systems with Channel Inversion," in *IEEE 55th Vehicular Technology Conference (VTC)*, Birmingham, AL, USA, 2002, pp. 35–39.

[4] C. Peel, B. Hochwald, and A. Swindlehurst, "A Vector-Perturbation Technique for Near-Capacity Multiantenna Multiuser Communication-part I: Channel Inversion and Regularization," *IEEE Trans. Commun.*, vol. 53, no. 1, pp. 195–202, Jan. 2005.

[5] A. Wiesel, Y. Eldar, and S. Shamai, "Linear Precoding via Conic Optimization for Fixed MIMO Receivers," *IEEE Trans. Sig. Process.*, vol. 54, no. 1, pp. 161–176, Jan. 2006.

[6] M. Bengtsson and B. Ottersten, "Optimal and Suboptimal Transmit Beamforming," *Handbook of Antennas in Wireless Communications*, Jan. 2001.

[7] M. Schubert and H. Boche, "Solution of the Multiuser Downlink Beamforming Problem with Individual SINR Constraints," *IEEE Trans. Veh. Tech.*, vol. 53, no. 1, pp. 18–28, Jan. 2004.

[8] C. Masouros, T. Ratnarajah, M. Sellathurai, C. B. Papadias, and A. K. Shukla, "Known Interference in the Cellular Downlink: A Performance Limiting Factor or a Source of Green Signal Power?" *IEEE Commun. Mag.*, vol. 51, no. 10, pp. 162–171, Oct. 2013.

[9] C. Masouros, M. Sellathurai, and T. Ratnarajah, "Vector Perturbation Based on Symbol Scaling for Limited Feedback MISO Downlinks," *IEEE Trans. Sig. Process.*, vol. 62, no. 3, pp. 562–571, Feb. 2014.

[10] C. Masouros and G. Zheng, "Exploiting Known Interference as Green Signal Power for Downlink Beamforming Optimization," *IEEE Trans. Sig. Process.*, vol. 63, no. 14, pp. 3628–3640, July 2015.

[11] M. Alodeh, S. Chatzinotas, and B. Ottersten, "Energy-Efficient Symbol-Level Precoding in Multiuser MISO Based on Relaxed Detection Region," *IEEE Trans. Wireless Commun.*, vol. 15, no. 5, pp. 3755–3767, May 2016.

[12] M. Alodeh, S. Chatzinotas, and B. Ottersten, "Symbol-Level Multiuser MISO Precoding for Multi-Level Adaptive Modulation," *IEEE Trans. Wireless Commun.*, vol. 16, no. 8, pp. 5511–5524, Aug. 2017.

[13] M. Alodeh, S. Chatzinotas, and B. Ottersten, "Constructive Multiuser Interference in Symbol Level Precoding for the MISO Downlink Channel," *IEEE Trans. Sig. Process.*, vol. 63, no. 9, pp. 2239–2252, May 2015.

[14] C. Masouros and G. Zheng, "Exploiting Known Interference as Green Signal Power for Downlink Beamforming Optimization," *IEEE Trans. Sig. Process.*, vol. 63, no. 14, pp. 3628–3640, July 2015.

[15] M. Alodeh, S. Chatzinotas, and B. Ottersten, "Energy-Efficient Symbol-Level Precoding in Multiuser MISO Based on Relaxed Detection Region," *IEEE Trans. Wireless Commun.*, vol. 15, no. 5, pp. 3755–3767, May 2016.

[16] R. Liu, M. Li, Q. Liu, and A. L. Swindlehurst, "Joint Symbol-Level Precoding and Reflecting Designs for IRS-Enhanced MU-MISO Systems," *IEEE Trans. Wireless Commun.*, vol. 20, no. 2, pp. 798–811, Feb. 2021.

[17] A. Li, L. Song, B. Vucetic, and Y. Li, "Interference Exploitation Precoding for Reconfigurable Intelligent Surface Aided Multi-User Communications With Direct Links," *IEEE Wireless Commun. Lett.*, vol. 9, no. 11, pp. 1937–1941, Nov. 2020.

[18] A. Li, C. Masouros, A. L. Swindlehurst, and W. Yu, "1-Bit Massive MIMO Transmission: Embracing Interference with Symbol-Level Precoding," *IEEE Commun. Mag.*, vol. 59, no. 5, pp. 121–127, May 2021.

[19] A. Li, C. Masouros, F. Liu, and A. L. Swindlehurst, "Massive MIMO 1-Bit DAC Transmission: A Low-Complexity Symbol Scaling Approach," *IEEE Trans. Wireless Commun.*, vol. 17, no. 11, pp. 7559–7575, Nov. 2018.

[20] A. Li, F. Liu, C. Masouros, Y. Li, and B. Vucetic, "Interference Exploitation 1-Bit Massive MIMO Precoding: A Partial Branch-and-Bound Solution With Near-Optimal Performance," *IEEE Trans. Wireless Commun.*, vol. 19, no. 5, pp. 3474–3489, May 2020.

[21] F. Sohrabi, Y.-F. Liu, and W. Yu, "One-Bit Precoding and Constellation Range Design for Massive MIMO With QAM Signaling," *IEEE J. Sel. Topics Sig. Process.*, vol. 12, no. 3, pp. 557–570, Jan. 2018.

[22] Y. Wang and A. Li, "ADMM Based Interference Exploitation Multi-User One-Bit Massive MIMO Precoding," *IEEE Trans. Veh. Tech.*, vol. 72, no. 7, pp. 9561–9566, 2023.

[23] F. Liu, C. Masouros, A. Li, T. Ratnarajah, and J. Zhou, "MIMO Radar and Cellular Coexistence: A Power-Efficient Approach Enabled by Interference Exploitation," *IEEE Trans. Sig. Process.*, vol. 66, no. 14, pp. 3681–3695, July 2018.

[24] R. Liu, M. Li, Q. Liu, and A. L. Swindlehurst, "Dual-Functional Radar-Communication Waveform Design: A Symbol-Level Precoding Approach," *IEEE J. Sel. Topics Sig. Process.*, vol. 15, no. 6, pp. 1316–1331, Nov. 2021.

[25] Y. Wang, H. Hou, W. Wang, and X. Yi, "Symbol-Level Precoding for Average SER Minimization in Multiuser MISO Systems," *IEEE Wireless Communications Letters*, vol. 13, no. 4, pp. 1103–1107, 2024.

[26] A. Li and C. Masouros, "Interference Exploitation Precoding Made Practical: Optimal Closed-Form Solutions for PSK Modulations," *IEEE Trans. Wireless Commun.*, vol. 17, no. 11, pp. 7661–7676, Nov. 2018.

[27] A. Li, C. Masouros, B. Vucetic, Y. Li, and A. L. Swindlehurst, "Interference Exploitation Precoding for Multi-Level Modulations: Closed-Form Solutions," *IEEE Trans. Commun.*, vol. 69, no. 1, pp. 291–308, Jan. 2021.

[28] A. Haqiqatnejad, F. Kayhan, and B. Ottersten, "Power Minimizer Symbol-Level Precoding: A Closed-Form Suboptimal Solution," *IEEE Sign. Process. Lett.*, vol. 25, no. 11, pp. 1730–1734, Nov. 2018.

[29] Y. Liu and W.-K. Ma, "Symbol-Level Precoding is Symbol-Perturbed zf When Energy Efficiency is Sought," in *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Calgary, Alberta, Canada, 2018, pp. 3869–3873.

[30] A. Mohammad, C. Masouros, and Y. Andreopoulos, "An Unsupervised Learning-Based Approach for Symbol-Level-Precoding," in *2021 IEEE Global Communications Conference (GLOBECOM)*, 2021, pp. 1–6.

[31] A. Mohammad, C. Masouros, and Y. Andreopoulos, "An Unsupervised Deep Unfolding Framework for Robust Symbol-Level Precoding," *IEEE Open Journal of the Communications Society*, vol. 4, pp. 1075–1090, Apr. 2023.

[32] A. Mohammad, C. Masouros, and Y. Andreopoulos, "A Memory-Efficient Learning Framework for SymbolLevel Precoding with Quantized NN Weights," *arXiv e-prints*, p. arXiv:2110.06542, Oct. 2021.

[33] A. Li, C. Shen, X. Liao, C. Masouros, and A. L. Swindlehurst, "Practical Interference Exploitation Precoding Without Symbol-by-Symbol Optimization: A Block-Level Approach," *IEEE Trans. Wireless Commun.*, vol. 22, no. 6, pp. 3982–3996, Nov. 2023.

[34] J. Z. C. S. Y. Yan, H. Zhao and Y. Ma, "MmWave Massive MIMO Hybrid Precoding Prediction in High Mobility Scenarios," in *2021 IEEE 93rd Vehicular Technology Conference (VTC2021-Spring)*, 2021, pp. 1–5.

[35] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.

[36] A. Li, D. Spano, J. Krivochiza, S. Domouchtsidis, C. G. Tsinos, C. Masouros, S. Chatzinotas, Y. Li, B. Vucetic, and B. Ottersten, "A Tutorial on Interference Exploitation via Symbol-Level Precoding: Overview, State-of-the-Art and Future Directions," *IEEE Commun. Sur. & Tut.*, vol. 22, no. 2, pp. 796–839, Mar. 2020.

[37] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, *Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers*, 2011.