

Practical Interference Exploitation Precoding Without Symbol-by-Symbol Optimization: A Block-Level Approach

Ang Li[✉], *Senior Member, IEEE*, Chao Shen[✉], *Member, IEEE*, Xuewen Liao[✉], *Member, IEEE*, Christos Masouros[✉], *Senior Member, IEEE*, and A. Lee Swindlehurst[✉], *Fellow, IEEE*

Abstract—In this paper, we propose a constructive interference (CI)-based block-level precoding (CI-BLP) approach for the downlink of a multi-user multiple-input single-output (MU-MISO) communication system. Contrary to existing CI precoding approaches which have to be designed on a symbol-by-symbol level, here a constant precoding matrix is applied to a collection of symbols within a given transmission block, thus significantly reducing the computational costs over traditional CI-based symbol-level precoding (CI-SLP) as the CI-BLP optimization problem only needs to be solved once per block. For both PSK and QAM modulation, we formulate an optimization problem to maximize the minimum CI effect over the block subject to a block- rather than symbol-level power budget. We mathematically derive the optimal precoding matrix for CI-BLP as a function of the Lagrange multipliers in closed form. By formulating the dual problem, the original CI-BLP optimization problem is further shown to be equivalent to a quadratic programming (QP) optimization. Numerical results validate our derivations, and show that the proposed CI-BLP scheme achieves improved performance over the traditional CI-SLP method, thanks to the relaxed power constraint over the considered block of symbol slots.

Index Terms—MU-MISO, symbol-level precoding, constructive interference, optimization, Lagrangian.

Manuscript received 18 February 2022; revised 29 July 2022; accepted 13 November 2022. Date of publication 23 November 2022; date of current version 12 June 2023. This work was supported in part by the Young Elite Scientists Sponsorship Program by China Institute of Communications (CIC) under Grant 2021QNR001, in part by the National Natural Science Foundation of China under Grant 62101422, in part by the Science and Technology Program of Shaanxi Province under Grant 2021KWZ-01, in part by the Fundamental Research Funds for the Central Universities under Grant xzy012020007, in part by the Engineering and Physical Sciences Research Council (EPSRC) Project EP/S028455/1, and in part by the National Science Foundation under Grant CCF-2008714. The associate editor coordinating the review of this article and approving it for publication was M. Dong. (Corresponding author: Ang Li.)

Ang Li and Xuewen Liao are with the Faculty of Electronic and Information Engineering, School of Information and Communications Engineering, Xi'an Jiaotong University, Xi'an, Shaanxi 710049, China (e-mail: ang.li.2020@xjtu.edu.cn; yeplos@xjtu.edu.cn).

Chao Shen is with the Shenzhen Research Institute of Big Data, Shenzhen 518172, China (e-mail: chaoshen@sribd.cn).

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

A. Lee Swindlehurst is with the Center for Pervasive Communications and Computing, Henry Samueli School of Engineering, University of California at Irvine, Irvine, CA 92697 USA (e-mail: swindle@uci.edu).

Color versions of one or more figures in this article are available at <https://doi.org/10.1109/TWC.2022.3222780>.

Digital Object Identifier 10.1109/TWC.2022.3222780

I. INTRODUCTION

MULTIPLE-INPUT multiple-output (MIMO) technology has been widely adopted in current cellular communication systems and will be an indispensable part of future wireless communication systems because of its significant gains over single-antenna systems [1]. In the downlink transmission of a multi-user MIMO communication system, precoding is essential for realizing spatial multiplexing. When channel state information (CSI) is available at the base station (BS), dirty-paper coding (DPC) is able to achieve the channel capacity by pre-subtracting the interference prior to transmission [2]. Despite its optimal performance, it is difficult to employ DPC in practical wireless systems due to its prohibitive computational costs and unrealistic assumption of an infinite alphabet. To alleviate the requirements of DPC, precoding approaches such as Tomlinson-Harashima precoding (THP) [3] and vector perturbation (VP) precoding [4] have been proposed. To further reduce the signal processing complexity, a number of closed-form linear precoding schemes have been studied, among which the most representative examples include zero-forcing (ZF) [5] and regularized ZF (RZF) precoding [6]. On the other hand, optimization-based precoding methods have received increasing research attention recently because of their flexibility in optimizing certain performance metrics [7], [8], [9], [10], [11], [12]. One popular form is downlink multicast precoding that targets broadcasting common information to all users [7]. Another popular example is the downlink signal-to-interference-plus-noise ratio (SINR) balancing approach, which aims to achieve a desired SINR for each user subject to either a total transmit power [8] or a per-antenna power budget [9]. An alternative optimization-based design aims to minimize the required transmit power at the BS for a given received SINR target for each user [10]. It is further shown in the literature that the power minimization and the SINR balancing problems are duals of one another [8], [11], and uplink-downlink duality can be exploited to obtain efficient iterative algorithms. In addition, the weighted minimum mean-squared error (W-MMSE) precoder was proposed in [12] for weighted sum-rate maximization.

More recently, the concept of constructive interference (CI) has been introduced for downlink transmission in multi-user MIMO systems, and CI-based precoding has received

increasing research attention. CI precoding is able to achieve improved performance over the above linear precoding methods by exploiting both the CSI and the data symbol information [13], [14]. Unlike traditional practice where interference is treated as harmful to the system performance, the superiority of CI-based approaches lies in its recognition that, at the symbol level, the instantaneous interference can be categorized as either constructive or destructive interference (DI). This idea was first discussed in [15], and a similar concept referred to as ‘convex vector precoding’ was introduced in [16]. Based on this concept and by further exploiting the data symbol information in addition to CSI, a modified ZF precoder was designed in [17], where the beneficial CI is preserved while DI is eliminated by the ZF process. A more advanced correlation rotation method was further proposed in [18], in which a rotation matrix is applied to the precoder to modify all interference to be constructive.

As a step further, CI-based precoding has been implemented under various optimization criteria in order to achieve further performance improvement [19], [20], [21], [22], [23]. More specifically, [19] first proposed an optimization-based CI approach in the context of VP precoding by substituting the sophisticated sphere-search process with a linear scaling operation, leading to a quadratic programming (QP) formulation that reduces the computational complexity of traditional VP precoding. The work in [20] combines CI with maximum ratio transmission (MRT) precoding to improve the performance of the correlation-rotation CI precoding proposed in [18]. In addition, CI-based power minimization and SINR balancing problems are also studied. In [20], the interfering signals in both cases are optimized to be strictly aligned with the intended data symbols to achieve CI, an approach that was later shown to be sub-optimal and referred to as the ‘strict phase-rotation’ CI metric.

More advanced CI metrics are introduced in [21] and [22], where the concept of the ‘constructive region’ is introduced, within which all the interference is constructive to the intended data symbols. 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 [21] was later named the ‘non-strict phase-rotation’ CI metric and is widely adopted in the relevant literature. Meanwhile, a relaxed CI metric based on a ‘relaxed detection region’ was introduced in [22], 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 [17], [18], [19], [20], [21], [22] are all designed for PSK modulation, while [23] 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. Owing to the above benefits, the concept of CI has been applied to a number of wireless communication scenarios such as intelligent reflecting surface (IRS)-assisted communication [24], [25], mutual coupling exploitation [26], constant-envelope precoding [27], [28], 1-bit precoding [29]–[32], radar-communication coexistence [33], [34], physical-layer security [35], [36], etc. For a

more comprehensive literature review on CI and SLP, we refer the interested readers to [37] and [38].

It has to be mentioned that the benefits of CI exploited by the above approaches come at the cost of symbol-by-symbol signal processing operations, i.e., symbol-level precoding (SLP), in which the precoder must be optimized on a symbol-by-symbol basis. This poses a significant computational burden on multi-user MIMO communication systems, because the BS will need to solve a different CI-SLP optimization problem for each symbol slot. 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 [39], [40], sub-optimal solutions [41], [42], and deep learning-based methods [43], [44], [45]. Specifically, [39] and [40] 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 QP optimization problem and solved using an iterative algorithm with a closed-form solution at each step. Building upon this, the work in [41] derives an exact closed-form but sub-optimal solution for the power minimization CI-SLP problem. In [42] and [46], it is shown that SLP can be regarded as symbol-level ZF precoding with a perturbation vector applied to the data symbols, and a power minimization problem is further presented for a given symbol error rate (SER) target, which can be solved via a block-level optimization. Despite the above attempts to reduce the computational costs of solving the CI-SLP optimization problem for each symbol slot, most of the above approaches still require solving an optimization problem 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.

Therefore in this paper, for the first time in the literature we propose CI-based block-level precoding (CI-BLP) for both PSK and QAM modulation for the downlink of a MU-MISO system, which further motivates the use of CI-based precoding techniques in practical wireless communication systems. We summarize the main contributions of this paper below:

- 1) We propose CI-BLP that exploits CI on a block level for both PSK and QAM, where a constant precoding matrix is applied to a block of symbol slots within a transmission block. The corresponding optimization problem is formulated to maximize the minimum CI effect over all symbol slots employing the ‘symbol-scaling’ CI metric, subject to a block-level power budget.
- 2) For PSK modulation, based on the Lagrangian and Karush-Kuhn-Tucker (KKT) method, we derive the optimal precoding matrix for CI-BLP in closed form as a function of the Lagrange multipliers. By further studying the corresponding dual problem, the original CI-BLP optimization problem for PSK modulation is finally shown to be equivalent to a QP optimization over a simplex.
- 3) We further extend our mathematical analysis to QAM modulation, where we derive the optimal precoding matrix for CI-BLP and the dual problem formulation

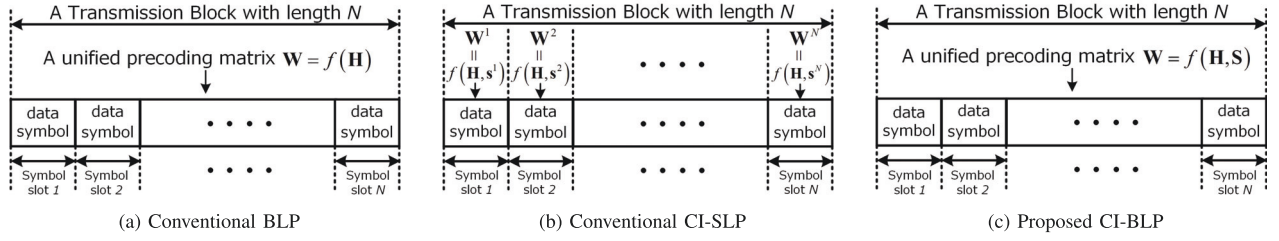


Fig. 1. A comparison between different precoding methodologies.

in a similar way. The original CI-BLP optimization problem for QAM modulation is also shown to be equivalent to a QP optimization, but not over a simplex any more. Another advantage of the proposed CI-BLP with QAM modulation is that it returns a constant power normalization factor over the considered block of symbol slots, thus reducing the signaling overhead for traditional CI-SLP with QAM.

- 4) Our above analyses for the PSK and QAM cases reveal that the proposed CI-BLP scheme shares a similar problem structure to the conventional CI-SLP method. While the problem size for the proposed CI-BLP approach is scaled up owing to the joint design over a block of symbol slots, the computational cost is reduced compared to traditional CI-SLP because the optimization problem only needs to be solved once per block.

Numerical results validate our derivations, and show that 1) the proposed CI-BLP approach can achieve an improved performance over the conventional CI-SLP scheme when the length of the block is short and the transmit power is sufficiently high, thanks to the relaxed block-level power budget; 2) the proposed CI-BLP method only exhibits a slight performance loss compared to conventional CI-SLP as the length of the block increases; 3) the proposed CI-BLP scheme offers reduced computational costs compared to traditional CI-SLP methods, as validated by the execution time result.

The remainder of this paper is organized as follows. Section II introduces the system model and briefly reviews CI. Section III studies the proposed CI-BLP scheme for PSK modulation, and Section IV studies the proposed CI-BLP for QAM modulation. Numerical results are shown in Section V, and Section VI concludes the paper.

Notation : a , $\mathbf{a}$, and $\mathbf{A}$ denote scalar, column vector and matrix, respectively. $(\cdot)^*$, $(\cdot)^T$, and $(\cdot)^H$ denote conjugate, transposition, and conjugate transposition, respectively. $\mathbf{A}(k, i)$ denotes the entry in the k -row and i -th column of $\mathbf{A}$. $\mathbb{C}^{n \times n}$ ($\mathbb{R}^{n \times n}$) represents an $n \times n$ matrix in the complex (real) set, and $\mathbf{I}_K$ denotes a $K \times K$ identity matrix. $\Re(\cdot)$ and $\Im(\cdot)$ extract the real and imaginary part of the argument, respectively. $\|\cdot\|_2$ denotes the ℓ_2 -norm, and j represents the imaginary unit. $\text{card}\{\cdot\}$ is the cardinality of a set.

II. SYSTEM MODEL AND CONSTRUCTIVE INTERFERENCE

A. System Model

We consider a multi-user MISO (MU-MISO) system during downlink transmission, where a BS equipped with N_T transmit

antennas is communicating with a total number of K single-antenna users in the same time-frequency resource, and where $K \leq N_T$. We introduce $\mathbf{S} = [\mathbf{s}^1, \mathbf{s}^2, \dots, \mathbf{s}^N] \in \mathbb{C}^{K \times N}$ as the data symbol matrix for the considered block of symbol slots, where N represents the length of the considered block which may be smaller than the channel coherence interval. The vector $\mathbf{s}^n = [s_1^n, s_2^n, \dots, s_K^n]^T \in \mathbb{C}^K$ contains the users' symbols for the n -th slot, drawn from normalized PSK or QAM constellations. Accordingly, the received signal for the k -th user in the n -th symbol slot is given by

$$y_k^n = \mathbf{h}_k^T \mathbf{W} \mathbf{s}^n + z_k^n, \quad (1)$$

where $\mathbf{h}_k \in \mathbb{C}^{N_T}$ represents the flat-fading channel vector between the BS and user k , which is constant within a channel coherence interval, and z_k^n is additive Gaussian noise with zero mean and variance σ^2 . The matrix $\mathbf{W} \in \mathbb{C}^{N_T \times K}$ is the block-level precoding matrix that is applied to all symbol slots in the block.

Fig. 1 depicts the difference between our proposed CI-BLP scheme and other standard precoding approaches, where $\mathbf{H}$ represents the CSI. Compared to existing CI-SLP approaches that optimize the precoding matrix (or the precoded signals) on a symbol level, our proposed CI-BLP scheme applies a constant precoding matrix $\mathbf{W}$ to all symbol slots, thus reducing the computational costs for the BS because the optimization only needs to be performed once per block of symbol slots. Moreover, our proposed CI-BLP approach is a linear precoding method, while CI-SLP has both linear and non-linear implementations. When compared to conventional BLP that also applies the same precoding matrix to all symbol slots based only on the CSI, our proposed CI-BLP scheme further exploits the data symbol information available at the BS for additional performance improvements, i.e., conventional BLP is linear data-independent precoding, while CI-BLP is linear data-dependent precoding. Since we focus on deriving the precoding structure for the proposed CI-BLP method, perfect CSI is assumed throughout the paper.

B. Constructive Interference

CI is defined as interference that is able to push the received signals away from the corresponding decision boundaries of the modulated symbol constellation, and the constructive region is defined as the area within which the received signals enjoy a larger distance to the decision boundaries compared to the corresponding nominal constellation point [37]. In this paper, the 'symbol-scaling' CI metric is employed for both PSK and QAM modulation. As an illustrative example,

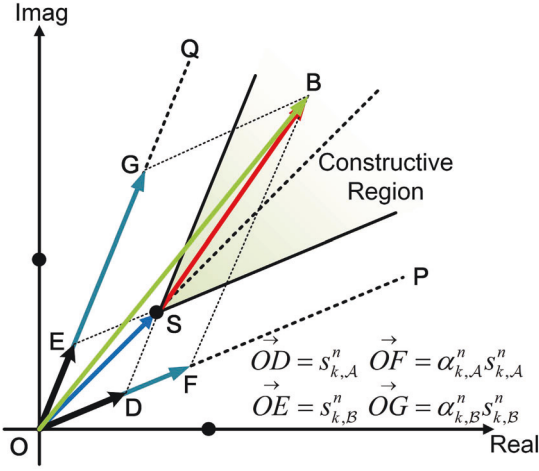


Fig. 2. An illustration for CI, 8PSK, 'symbol-scaling' metric.

Fig. 2 presents one quarter of a 8PSK constellation, where the green shaded area represents the constructive region corresponding to the 8PSK constellation point in the first quadrant. Without loss of generality, we assume $\vec{OS} = s_k^n$ is the data symbol of interest for user k in the n -th symbol slot, and we introduce $\vec{OB} = \mathbf{h}_k^T \mathbf{W} s_k^n$ as the corresponding received signal for user k excluding noise. From this perspective, the effect of interference is equivalent to a scaling and rotation operation on the transmitted data symbol $\vec{OS}$. Based on the geometry, if node 'B' is located in the constructive region, as depicted in Fig. 2, the distance between node 'B' and the two decision boundaries $\vec{OP}$ and $\vec{OQ}$ is larger than that between the nominal constellation point 'S' and the two decision boundaries, leading to an effective increase in the received SINR and a reduction in the probability of a detection error. Accordingly, CI is achieved in this case because the effect of interference contributes to the useful signal power. For a more detailed discussion on different CI metrics, we refer the interested readers to [37].

III. PROPOSED CI-BLP FOR PSK MODULATION

A. Problem Formulation

In this section, we study the proposed CI-BLP optimization for PSK modulation. Since it has already been shown in [40] that the 'non-strict phase-rotation' CI metric commonly adopted for PSK modulation is equivalent to the 'symbol-scaling' CI metric for PSK, we employ the 'symbol-scaling' CI metric in this section to be consistent with the mathematics of Section IV.

The 'symbol-scaling' CI metric decomposes both the modulated data symbols and the noiseless received signals along their corresponding decision boundaries, where two real-valued scaling coefficients are introduced to the noiseless received signals to jointly represent the scaling and rotation effect of the interference on the intended symbol of interest. Based on this criterion, to apply the 'symbol-scaling' CI metric to PSK modulation, each desired data symbol $\vec{OS} = s_k^n$ and the corresponding received signal excluding noise $\vec{OB} = \mathbf{h}_k^T \mathbf{W} s_k^n$ are decomposed along the two decision boundaries

corresponding to s_k^n , as depicted in Fig. 2. Mathematically, this decomposition can be expressed as

$$\begin{aligned} \vec{OS} &= \vec{OD} + \vec{OE} \Rightarrow s_k^n = s_{k,A}^n + s_{k,B}^n, \\ \vec{OB} &= \vec{OF} + \vec{OG} \Rightarrow \mathbf{h}_k^T \mathbf{W} s_k^n = \alpha_{k,A}^n s_{k,A}^n + \alpha_{k,B}^n s_{k,B}^n, \end{aligned} \quad (2)$$

where $\alpha_{k,A}^n \geq 0$ and $\alpha_{k,B}^n \geq 0$ are real-valued scaling coefficients. We observe that a larger value of $\min \{\alpha_{k,A}^n, \alpha_{k,B}^n\}$ represents a larger distance to the decision boundaries, and therefore a larger CI effect and a better SER performance. We define $\alpha_E^n \in \mathbb{R}^{2K}$ as

$$\alpha_E^n = [\alpha_{1,A}^n, \alpha_{2,A}^n, \dots, \alpha_{K,A}^n, \alpha_{1,B}^n, \alpha_{2,B}^n, \dots, \alpha_{K,B}^n]^T, \quad (3)$$

and by following the transformations in [30], α_E^n can be further expressed as

$$\alpha_E^n = \mathbf{M}^n \mathbf{W}_E s_E^n, \quad (4)$$

where the construction of $\mathbf{M}^n \in \mathbb{R}^{2K \times 2N_T}$ is shown in Appendix A, which directly follows Section IV-A of [30]. $\mathbf{W}_E \in \mathbb{R}^{2N_T \times 2K}$ and $s_E^n \in \mathbb{R}^{2K}$ in (4) are defined as

$$\mathbf{W}_E = \begin{bmatrix} \Re(\mathbf{W}) & -\Im(\mathbf{W}) \\ \Im(\mathbf{W}) & \Re(\mathbf{W}) \end{bmatrix}, \quad s_E^n = [\Re(s_k^n)^T, \Im(s_k^n)^T]^T. \quad (5)$$

The proposed CI-BLP optimization aims to maximize the minimum entry in α_E^n for all symbol slots within the considered transmission block, and the corresponding optimization problem can thus be constructed as:

$$\begin{aligned} \mathcal{P}_0^{\text{PSK}} : & \max_{\mathbf{W}_E} \min_{k,n} \alpha_k^n \\ \text{s.t. } & \text{C1} : \alpha_E^n = \mathbf{M}^n \mathbf{W}_E s_E^n, \quad \forall n \in \mathcal{N}, \\ & \text{C2} : \sum_{n=1}^N \|\mathbf{W}_E s_E^n\|_2^2 \leq N p_0, \end{aligned} \quad (6)$$

where α_k^n represents the k -th entry in α_E^n , p_0 represents the transmit power budget per symbol slot, and $\mathcal{N} = \{1, 2, \dots, N\}$. Compared to the traditional CI-SLP problem, $\mathcal{P}_0^{\text{PSK}}$ differs in that

- 1) a constant precoding matrix is applied to all symbols in the considered block, and it is optimized across all symbol slots, thus greatly reducing the computational costs over the traditional CI-SLP approach;
- 2) the power budget is enforced over the entire considered block instead of within each symbol slot, i.e., a relaxed power constraint is enforced compared to traditional CI-SLP.

The relaxed power constraint is an advantage for CI-BLP, but the use of a constant precoder for all symbols is a disadvantage compared with CI-SLP. For small N and a large enough transmit power, the benefit of the relaxed power constraint outweighs the loss due to using a fixed precoder, leading to a better performance for CI-BLP over CI-SLP.

$\mathcal{P}_0^{\text{PSK}}$ is a joint optimization over all symbol slots, but it is not trivial to directly solve it to obtain the complex-valued precoding matrix $\mathbf{W}$, because the variable $\mathbf{W}_E$ has to satisfy

the special structure shown in (5). To facilitate subsequent derivations, we introduce $\hat{\mathbf{W}}$:

$$\hat{\mathbf{W}} = [\Re(\mathbf{W}) \quad -\Im(\mathbf{W})] \in \mathbb{R}^{N_T \times 2K}, \quad (7)$$

based on which we can decompose $\mathbf{W}_E$ as

$$\mathbf{W}_E = \mathbf{P}\hat{\mathbf{W}} + \mathbf{Q}\hat{\mathbf{W}}\mathbf{T}, \quad (8)$$

where $\mathbf{P} \in \mathbb{R}^{2N_T \times N_T}$, $\mathbf{Q} \in \mathbb{R}^{2N_T \times N_T}$ and $\mathbf{T} \in \mathbb{R}^{2K \times 2K}$ are defined as

$$\mathbf{P} = \begin{bmatrix} \mathbf{I}_{N_T} \\ \mathbf{0} \end{bmatrix}, \quad \mathbf{Q} = \begin{bmatrix} \mathbf{0} \\ \mathbf{I}_{N_T} \end{bmatrix}, \quad \mathbf{T} = \begin{bmatrix} \mathbf{0} & \mathbf{I}_K \\ -\mathbf{I}_K & \mathbf{0} \end{bmatrix}. \quad (9)$$

Based on (8), the expression for α_E^n is further transformed into:

$$\begin{aligned} \alpha_E^n &= \mathbf{M}^n \mathbf{W}_E \mathbf{s}_E^n \\ &= \mathbf{M}^n (\mathbf{P}\hat{\mathbf{W}} + \mathbf{Q}\hat{\mathbf{W}}\mathbf{T}) \mathbf{s}_E^n \\ &= \mathbf{M}^n \mathbf{P}\hat{\mathbf{W}} \mathbf{s}_E^n + \mathbf{M}^n \mathbf{Q}\hat{\mathbf{W}}\mathbf{T} \mathbf{s}_E^n \\ &= \mathbf{A}^n \hat{\mathbf{W}} \mathbf{s}_E^n + \mathbf{B}^n \hat{\mathbf{W}} \mathbf{c}_E^n, \end{aligned} \quad (10)$$

where we introduce $\mathbf{A}^n \in \mathbb{R}^{2K \times N_T}$, $\mathbf{B}^n \in \mathbb{R}^{2K \times N_T}$ and $\mathbf{c}_E^n \in \mathbb{R}^{2K}$ as

$$\mathbf{A}^n = \mathbf{M}^n \mathbf{P}, \mathbf{B}^n = \mathbf{M}^n \mathbf{Q}, \mathbf{c}_E^n = \mathbf{T} \mathbf{s}_E^n. \quad (11)$$

With the expression for the k -th entry of α_E^n given by

$$\alpha_k^n = (\mathbf{a}_k^n)^T \hat{\mathbf{W}} \mathbf{s}_E^n + (\mathbf{b}_k^n)^T \hat{\mathbf{W}} \mathbf{c}_E^n, \quad (12)$$

where $(\mathbf{a}_k^n)^T$ and $(\mathbf{b}_k^n)^T$ denote the k -th row of $\mathbf{A}^n$ and $\mathbf{B}^n$ respectively, $\mathcal{P}_0^{\text{PSK}}$ can be expressed in the form of the standard convex optimization problem below:

$$\begin{aligned} \mathcal{P}_1^{\text{PSK}} : \min_{\hat{\mathbf{W}}, t} & -t \\ \text{s.t. } \mathbf{C1} : & t - (\mathbf{a}_k^n)^T \hat{\mathbf{W}} \mathbf{s}_E^n - (\mathbf{b}_k^n)^T \hat{\mathbf{W}} \mathbf{c}_E^n \leq 0, \\ & \forall k \in \mathcal{K}, n \in \mathcal{N}, \\ \mathbf{C2} : & \sum_{n=1}^N \left\| (\mathbf{P}\hat{\mathbf{W}} + \mathbf{Q}\hat{\mathbf{W}}\mathbf{T}) \mathbf{s}_E^n \right\|_2^2 - Np_0 \leq 0, \end{aligned} \quad (13)$$

where $\mathcal{K} = \{1, 2, \dots, 2K\}$.

B. Optimal Closed-Form Structure for $\hat{\mathbf{W}}$

We derive the optimal precoding matrix $\hat{\mathbf{W}}$ based on the Lagrangian and KKT conditions. By constructing the Lagrangian of $\mathcal{P}_1^{\text{PSK}}$ and applying the KKT conditions, the optimal precoding matrix $\hat{\mathbf{W}}$ as a function of the Lagrange multipliers can be obtained in closed form, given by the following proposition.

Proposition 1: When $N \geq K$, the optimal closed-form structure for $\hat{\mathbf{W}}$ in $\mathcal{P}_1^{\text{PSK}}$ is given by

$$\hat{\mathbf{W}} = \frac{1}{2\mu} \sum_{n=1}^N \left[(\mathbf{A}^n)^T \delta^n (\mathbf{s}_E^n)^T + (\mathbf{B}^n)^T \delta^n (\mathbf{c}_E^n)^T \right] \mathbf{D}^{-1}, \quad (14)$$

where μ is the non-negative Lagrange multiplier associated with the power constraint $\mathbf{C2}$ in $\mathcal{P}_1^{\text{PSK}}$, $\delta^n = [\delta_1^n, \delta_2^n, \dots,$

$\delta_{2K}^n]^T \in \mathbb{R}^{2K}$ is the non-negative Lagrange multiplier vector associated with $\mathbf{C1}$ in $\mathcal{P}_1^{\text{PSK}}$ and $\mathbf{D} \in \mathbb{R}^{2K \times 2K}$ is given by

$$\mathbf{D} = \left[\sum_{n=1}^N \mathbf{s}_E^n (\mathbf{s}_E^n)^T + \sum_{n=1}^N \mathbf{c}_E^n (\mathbf{c}_E^n)^T \right]. \quad (15)$$

Proof: See Appendix B.

We observe from (14) that the expression for the optimal precoding matrix $\hat{\mathbf{W}}$ includes all the data symbols transmitted within the considered block. Moreover, the optimization on $\hat{\mathbf{W}}$ is now transformed into an optimization on the Lagrange multipliers δ^n . In what follows, we consider the dual problem of $\mathcal{P}_1^{\text{PSK}}$ to further simplify the proposed CI-BLP optimization problem.

C. Dual Problem Formulation

For the convex CI-BLP optimization problem $\mathcal{P}_1^{\text{PSK}}$ in (13), it is easy to validate that Slater's condition is satisfied, meaning that the dual gap is zero [47]. Thus, we can solve $\mathcal{P}_1^{\text{PSK}}$ by solving its dual problem, given by

$$\mathcal{U}_1 = \max_{\{\delta^n\}, \mu} \min_{\hat{\mathbf{W}}, t} \mathcal{L}_1(\hat{\mathbf{W}}, t, \delta^n, \mu), \quad (16)$$

where the inner minimization is achieved with (42a), the active power constraint in (17) and the expression for $\hat{\mathbf{W}}$ in (14). By substituting (42a), (17) and (14) into $\mathcal{U}_1$ in (16), we arrive at the following proposition.

Proposition 2: The dual problem $\mathcal{U}_1$ is equivalent to the following dual problem that optimizes δ_E and μ :

$$\begin{aligned} \mathcal{P}_2^{\text{PSK}} : \min_{\delta_E, \mu} & \frac{1}{2\mu} \delta_E^T \mathbf{U} \delta_E \\ \text{s.t. } \mathbf{C1} : & \mathbf{1}^T \delta_E - 1 = 0, \\ \mathbf{C2} : & \delta_E^m \geq 0, \quad \forall m \in \{1, 2, \dots, 2NK\}, \\ \mathbf{C3} : & \frac{1}{4\mu^2} (\delta_E)^T (\mathbf{F} + \mathbf{G}) \delta_E = Np_0, \end{aligned} \quad (17)$$

where $\mathbf{F} = \sum_{l=1}^N \mathbf{F}^l \in \mathbb{R}^{2NK \times 2NK}$ can be obtained based on (53) and (55). $\mathbf{G} \in \mathbb{R}^{2NK \times 2NK}$ is similarly constructed based on (57) and (58). $\mathbf{U}$ is constructed based on (49) and (51).

Proof: See Appendix C.

At first glance, $\mathcal{P}_2^{\text{PSK}}$ seems difficult to handle because μ is present in the denominator of the criterion as well as the equality power constraint. Nevertheless, we show below that after some transformations, $\mathcal{P}_2^{\text{PSK}}$ is equivalent to a QP optimization problem over a simplex.

We begin by studying the relationship between $(\mathbf{F} + \mathbf{G})$ and $\mathbf{U}$ to further simplify $\mathcal{P}_2^{\text{PSK}}$, where the following proposition is obtained.

Proposition 3: $\mathbf{F}$, $\mathbf{G}$, and $\mathbf{U}$ satisfy the following condition:

$$\mathbf{F} + \mathbf{G} = \mathbf{U}. \quad (18)$$

Proof: See Appendix D.

According to **Proposition 3** and based on the block-level power constraint in (59), we can obtain the following expression for μ :

$$\frac{1}{4\mu^2} (\delta_E)^T \mathbf{U} \delta_E = Np_0 \Rightarrow \mu = \sqrt{\frac{(\delta_E)^T \mathbf{U} \delta_E}{4Np_0}}. \quad (19)$$

Substituting the above expression for μ into the objective function of the dual problem, $\mathcal{U}_1$ can be transformed into an optimization on δ_E only, given by

$$\begin{aligned}\mathcal{U}_1 &= \min_{\delta_E, \mu} \frac{1}{2\mu} \delta_E^T \mathbf{U} \delta_E \\ &= \min_{\delta_E} \frac{1}{2\sqrt{\frac{(\delta_E^T \mathbf{U} \delta_E)}{4Np_0}}} \delta_E^T \mathbf{U} \delta_E \\ &= \min_{\delta_E} \sqrt{Np_0 \delta_E^T \mathbf{U} \delta_E} \\ &= \min_{\delta_E} \delta_E^T \mathbf{U} \delta_E,\end{aligned}\quad (20)$$

where the last step is achieved because $y = \sqrt{x}$ is a monotonic function. Accordingly, the final dual problem of the proposed CI-BLP optimization for PSK modulation can be formulated as

$$\begin{aligned}\mathcal{P}_3^{\text{PSK}} : \quad & \min_{\delta_E} \delta_E^T \mathbf{U} \delta_E \\ \text{s.t. } & \mathbf{C1} : \mathbf{1}^T \delta_E - 1 = 0, \\ & \mathbf{C2} : \delta_E^m \geq 0, \quad \forall m \in \{1, 2, \dots, 2NK\},\end{aligned}\quad (21)$$

where we note that the block-level power constraint is no longer required because it is already inherently satisfied.

Comparing the QP formulation $\mathcal{P}_3^{\text{PSK}}$ for the proposed CI-BLP with that for traditional CI-SLP in the case of PSK modulation (shown as $\mathcal{P}_8$ in (50) of [39]), we observe that they share a similar problem formulation, with the only difference lying in the problem size. For traditional CI-SLP in [39], the number of variables to be optimized is $2K$, while for the proposed CI-BLP it is $2NK$. This is because the proposed CI-BLP finds a single precoding matrix for all symbol slots, and thus the data symbols for the entire block need to be jointly considered. Nevertheless, we note that for the proposed CI-BLP scheme, the optimization only needs to be performed once per block, while the traditional CI-SLP needs to solve N optimization problems for the considered block with length N , because traditional CI-SLP performs the optimization on a symbol-by-symbol basis. This will lead to a significant complexity reduction for the proposed CI-BLP over traditional CI-SLP, as will be shown in Section V.

$\mathcal{P}_3^{\text{PSK}}$ is a QP optimization problem over a simplex, which can be more efficiently solved than the original CI-BLP optimization problem $\mathcal{P}_1^{\text{PSK}}$ via the standard simplex method [48], [49] or the interior-point methods [50], which are not discussed in this paper for brevity. It is also worth mentioning that one advantage of the original problem formulation $\mathcal{P}_1^{\text{PSK}}$ over $\mathcal{P}_3^{\text{PSK}}$ is that the number of parameters to be optimized does not grow with the size N of the block. After solving $\mathcal{P}_3^{\text{PSK}}$ and obtaining $\hat{\mathbf{W}}$ via (14), the original complex precoding matrix $\mathbf{W}$ in (1) for PSK modulation can be obtained by

$$\mathbf{W} = \hat{\mathbf{W}}\hat{\mathbf{P}} - j\hat{\mathbf{W}}\hat{\mathbf{Q}}, \quad (22)$$

where the form of $\hat{\mathbf{P}}$ and $\hat{\mathbf{Q}}$ follows (9) but their dimension becomes $2K \times K$.

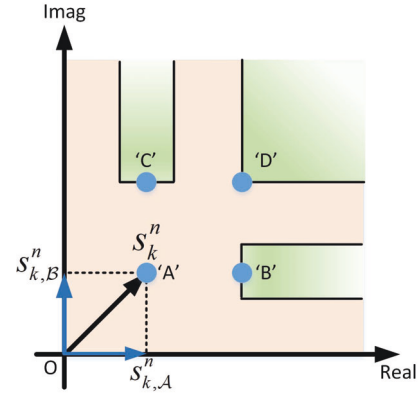


Fig. 3. An illustration for CI, 16QAM, 'symbol-scaling' metric.

IV. PROPOSED CI-BLP FOR QAM MODULATION

A. Problem Formulation

In this section, we extend the proposed CI-BLP approach to the case of QAM modulation, where we still employ the 'symbol-scaling' CI metric. Compared with PSK, one significant difference for QAM modulation is that not all QAM constellation points can exploit CI. To be more specific, the constellation points for QAM can be divided into 4 types, as shown in Fig. 3 where one quarter of a nominal 16QAM constellation is depicted as a representative example:

- 1) Type 'A': no CI can be exploited;
- 2) Type 'B': CI can be exploited for the real part;
- 3) Type 'C': CI can be exploited for the imaginary part;
- 4) Type 'D': CI can be exploited for both the real and imaginary part.

By following a similar procedure as that in Section III, we decompose each data symbol s_k^n and its corresponding received signal excluding noise as in (2), where for QAM modulation we have

$$s_{k,A}^n = \Re(s_k^n), s_{k,B}^n = \Im(s_k^n), \quad \forall k, n. \quad (23)$$

Introducing α_E^n as in (3) and following a similar formulation as in (4)-(5) and (7)-(12), the proposed CI-BLP optimization problem for QAM modulation can be formulated as:

$$\begin{aligned}\mathcal{P}_1^{\text{QAM}} : \quad & \min_{\hat{\mathbf{W}}, \hat{\mathbf{t}}} -\hat{\mathbf{t}} \\ \text{s.t. } & \mathbf{C1} : \hat{\mathbf{t}} - (\mathbf{a}_k^n)^T \hat{\mathbf{W}} \mathbf{s}_E^n - (\mathbf{b}_k^n)^T \hat{\mathbf{W}} \mathbf{c}_E^n \leq 0, \\ & \quad \forall k \in \mathcal{O}^n, n \in \mathcal{N}, \\ & \mathbf{C2} : \hat{\mathbf{t}} - (\mathbf{a}_k^n)^T \hat{\mathbf{W}} \mathbf{s}_E^n - (\mathbf{b}_k^n)^T \hat{\mathbf{W}} \mathbf{c}_E^n = 0, \\ & \quad \forall k \in \mathcal{I}^n, n \in \mathcal{N}, \\ & \mathbf{C3} : \sum_{n=1}^N \left\| (\mathbf{P}\hat{\mathbf{W}} + \mathbf{Q}\hat{\mathbf{W}}\mathbf{T}) \mathbf{s}_E^n \right\|_2^2 - Np_0 \leq 0.\end{aligned}\quad (24)$$

Recalling $\mathcal{K} = \{1, 2, \dots, 2K\}$, the set $\mathcal{O}^n$ consists of the indices in $\mathcal{K}$ that correspond to the real part of the data symbol belonging to constellation point type 'B', the imaginary part of the data symbol belonging to constellation point type 'C', and both the real and imaginary part of the data symbol belonging to constellation point type 'D', for which CI can be exploited.

$$\begin{aligned}
& \mathcal{L}_2(\hat{\mathbf{W}}, \hat{t}, \hat{\delta}_k^n, \hat{\vartheta}_k^n, \hat{\mu}) \\
&= -\hat{t} + \sum_{n=1}^N \sum_{k \in \text{card}\{\mathcal{O}^n\}} \hat{\delta}_k^n \left[\hat{t} - (\mathbf{a}_k^n)^T \hat{\mathbf{W}} \mathbf{s}_E^n - (\mathbf{b}_k^n)^T \hat{\mathbf{W}} \mathbf{c}_E^n \right] + \sum_{n=1}^N \sum_{k \in \text{card}\{\mathcal{I}^n\}} \hat{\vartheta}_k^n \left[\hat{t} - (\mathbf{a}_k^n)^T \hat{\mathbf{W}} \mathbf{s}_E^n - (\mathbf{b}_k^n)^T \hat{\mathbf{W}} \mathbf{c}_E^n \right] \\
&\quad + \hat{\mu} \left[\sum_{n=1}^N (\mathbf{s}_E^n)^T (\mathbf{P} \hat{\mathbf{W}} + \mathbf{Q} \hat{\mathbf{W}} \mathbf{T})^T (\mathbf{P} \hat{\mathbf{W}} + \mathbf{Q} \hat{\mathbf{W}} \mathbf{T}) \mathbf{s}_E^n - N p_0 \right] \\
&= \left(\sum_{n=1}^N \mathbf{1}^T \hat{\delta}^n + \sum_{n=1}^N \mathbf{1}^T \hat{\vartheta}^n - 1 \right) \hat{t} - \sum_{n=1}^N (\hat{\delta}^n)^T \mathbf{A}_{\mathcal{O}}^n \hat{\mathbf{W}} \mathbf{s}_E^n - \sum_{n=1}^N (\hat{\delta}^n)^T \mathbf{B}_{\mathcal{O}}^n \hat{\mathbf{W}} \mathbf{c}_E^n - \sum_{n=1}^N (\hat{\vartheta}^n)^T \mathbf{A}_{\mathcal{I}}^n \hat{\mathbf{W}} \mathbf{s}_E^n \\
&\quad - \sum_{n=1}^N (\hat{\vartheta}^n)^T \mathbf{B}_{\mathcal{I}}^n \hat{\mathbf{W}} \mathbf{c}_E^n + \hat{\mu} \left[\sum_{n=1}^N (\mathbf{s}_E^n)^T (\hat{\mathbf{W}}^T \mathbf{P}^T + \mathbf{T}^T \hat{\mathbf{W}}^T \mathbf{Q}^T) (\mathbf{P} \hat{\mathbf{W}} + \mathbf{Q} \hat{\mathbf{W}} \mathbf{T}) \mathbf{s}_E^n \right] - \hat{\mu} N p_0 \\
&= \left(\sum_{n=1}^N \mathbf{1}^T \beta^n - 1 \right) \hat{t} - \sum_{n=1}^N (\beta^n)^T \mathbf{A}^n \hat{\mathbf{W}} \mathbf{s}_E^n - \sum_{n=1}^N (\beta^n)^T \mathbf{B}^n \hat{\mathbf{W}} \mathbf{c}_E^n + \hat{\mu} \sum_{n=1}^N (\mathbf{s}_E^n)^T \hat{\mathbf{W}}^T \mathbf{P}^T \mathbf{P} \hat{\mathbf{W}} \mathbf{s}_E^n \\
&\quad + \underbrace{\hat{\mu} \sum_{n=1}^N (\mathbf{s}_E^n)^T \hat{\mathbf{W}}^T \mathbf{P}^T \mathbf{Q} \hat{\mathbf{W}} \mathbf{T} \mathbf{s}_E^n}_{=0} + \underbrace{\hat{\mu} \sum_{n=1}^N (\mathbf{s}_E^n)^T \mathbf{T}^T \hat{\mathbf{W}}^T \mathbf{Q}^T \mathbf{P} \hat{\mathbf{W}} \mathbf{s}_E^n}_{=0} + \hat{\mu} \sum_{n=1}^N \underbrace{(\mathbf{s}_E^n)^T \mathbf{T}^T \hat{\mathbf{W}}^T \mathbf{Q}^T \mathbf{Q} \hat{\mathbf{W}} \mathbf{T} \mathbf{s}_E^n}_{(\mathbf{c}_E^n)^T} - \hat{\mu} N p_0 \\
&= \left(\sum_{n=1}^N \mathbf{1}^T \beta^n - 1 \right) \hat{t} - \sum_{n=1}^N (\beta^n)^T \mathbf{A}^n \hat{\mathbf{W}} \mathbf{s}_E^n - \sum_{n=1}^N (\beta^n)^T \mathbf{B}^n \hat{\mathbf{W}} \mathbf{c}_E^n + \hat{\mu} \sum_{n=1}^N (\mathbf{s}_E^n)^T \hat{\mathbf{W}}^T \hat{\mathbf{W}} \mathbf{s}_E^n + \hat{\mu} \sum_{n=1}^N (\mathbf{c}_E^n)^T \hat{\mathbf{W}}^T \hat{\mathbf{W}} \mathbf{c}_E^n - \hat{\mu} N p_0
\end{aligned} \tag{26}$$

The set $\mathcal{I}^n$ consists of the indices in $\mathcal{K}$ that correspond to the real part of the data symbol belonging to constellation point type ‘C’, the imaginary part of the data symbol belonging to constellation point type ‘B’, and both the real and imaginary part of the data symbol belonging to constellation point type ‘A’, for which there exists no CI to be exploited. $\mathcal{O}^n$ and $\mathcal{I}^n$ satisfy:

$$\begin{aligned}
\mathcal{O}^n \cup \mathcal{I}^n &= \mathcal{K}, \mathcal{O}^n \cap \mathcal{I}^n = \emptyset, \\
\text{card}\{\mathcal{O}^n\} + \text{card}\{\mathcal{I}^n\} &= 2K, \quad \forall n \in \mathcal{N}.
\end{aligned} \tag{25}$$

Similar to the case for PSK, the proposed CI-BLP scheme for QAM offers complexity reduction over traditional CI-SLP, because only a single optimization problem needs to be solved per transmission block. More importantly, another advantage that is specific to QAM modulation over traditional CI-SLP is the reduction in signaling overhead. Since traditional CI-SLP is performed on a symbol level, the power normalization factor varies from symbol to symbol. Therefore, extra signaling overhead is required for CI-SLP since the power normalization factor needs to be broadcast to the users at the symbol rate for correct demodulation, while the proposed CI-BLP scheme returns a constant power normalization factor (equal to the optimal value t^* for $\mathcal{P}_1^{\text{QAM}}$) over the entire block, and thus it enjoys reduced signaling overhead compared with CI-SLP.

B. Optimal Closed-Form Structure for $\hat{\mathbf{W}}$

Similar to the case when PSK modulation is considered, we derive the optimal precoding matrix $\hat{\mathbf{W}}$ for $\mathcal{P}_1^{\text{QAM}}$ based on the Lagrangian and KKT conditions. To begin with, the Lagrangian of $\mathcal{P}_1^{\text{QAM}}$ can be constructed and is shown in (26),

at the top of the page, where $\hat{\delta}^n \in \mathbb{R}^{\text{card}\{\mathcal{O}^n\}}$ and $\hat{\vartheta}^n \in \mathbb{R}^{\text{card}\{\mathcal{I}^n\}}$ consist of the Lagrange multipliers that correspond to the inequality and equality constraints, respectively. For each symbol slot $n \in \mathcal{N}$, $\mathbf{A}_{\mathcal{O}}^n$ and $\mathbf{B}_{\mathcal{O}}^n$ consist of the rows of $\mathbf{A}^n$ and $\mathbf{B}^n$ respectively, whose indices belong to $\mathcal{O}^n$, while $\mathbf{A}_{\mathcal{I}}^n$ and $\mathbf{B}_{\mathcal{I}}^n$ consist of the rows of $\mathbf{A}^n$ and $\mathbf{B}^n$, whose indices belong to $\mathcal{I}^n$. $\beta^n = [\beta_1^n, \beta_2^n, \dots, \beta_K^n] \in \mathbb{R}^{2K}$, with the k -th entry defined as:

$$\beta_k^n = \begin{cases} \hat{\delta}_k^n, & \text{if } k \in \mathcal{O}^n \\ \hat{\vartheta}_k^n, & \text{if } k \in \mathcal{I}^n \end{cases}, \quad \forall k \in \mathcal{K}. \tag{27}$$

The main difference between the Lagrangian function $\mathcal{L}_2$ for QAM and the Lagrangian function $\mathcal{L}_1$ for PSK lies in the fact that the CI-BLP optimization for QAM includes equality constraints, as shown in **C2** for $\mathcal{P}_1^{\text{QAM}}$ in (24). This observation means that not all the entries in β^n need to be non-negative, which will mean that the final QP optimization for QAM modulation is not over a simplex any more. This is evident from the KKT conditions of the Lagrangian function $\mathcal{L}_2$ shown in (28), at the bottom of the next page. As can be observed in (28c) and (28d), the value of each $\hat{\delta}_k^n$ has to be non-negative, while the value of each $\hat{\vartheta}_k^n$ does not because it corresponds to the equality constraint.

Similar to the case for PSK modulation in Section III, when the optimal solution to $\mathcal{P}_1^{\text{QAM}}$ is obtained, the expression for the precoding matrix $\hat{\mathbf{W}}$ as a function of the Lagrange multipliers is found from (28b) to be

$$\hat{\mathbf{W}} = \frac{1}{2\hat{\mu}} \sum_{n=1}^N \left[(\mathbf{A}^n)^T \beta^n (\mathbf{s}_E^n)^T + (\mathbf{B}^n)^T \beta^n (\mathbf{c}_E^n)^T \right] \mathbf{D}^{-1}. \tag{29}$$

C. Dual Problem Formulation

The procedure for deriving the dual problem formulation for QAM is similar to that outlined in Section III-C for PSK, and therefore for brevity we directly present the final QP optimization problem for QAM modulation below:

$$\begin{aligned} \mathcal{P}_2^{\text{QAM}} : \min_{\beta_E} & \beta_E^T \mathbf{U} \beta_E \\ \text{s.t. } \mathbf{C1} : & \mathbf{1}^T \beta_E - 1 = 0, \\ \mathbf{C2} : & \beta_E^m \geq 0, \quad \forall m \in \mathcal{M}, \end{aligned} \quad (30)$$

where $\beta_E \in \mathbb{R}^{2NK}$ is defined as

$$\beta_E = [\beta^1, \beta^2, \dots, \beta^N]^T. \quad (31)$$

$\mathcal{M}$ contains the indices of the Lagrange multipliers corresponding to the inequality constraints for all symbol slots within the considered block, and is expressed mathematically as

$$\mathcal{M} = \left\{ m \mid m = 2(n-1)K + k, \text{ if } \beta_k^n = \hat{\delta}_k^n \right\}. \quad (32)$$

Compared to the QP formulation $\mathcal{P}_3^{\text{PSK}}$ for PSK in (21), the QP formulation $\mathcal{P}_2^{\text{QAM}}$ for QAM only requires a total number of card $\{\mathcal{M}\}$ entries in β_E to be non-negative. In this case, although the simplex method cannot be used for solving $\mathcal{P}_2^{\text{QAM}}$, the interior-point based methods [50] can still be employed to efficiently solve $\mathcal{P}_2^{\text{QAM}}$. After obtaining β_E , the precoding matrix $\hat{\mathbf{W}}$ can be obtained via (29), and the final complex precoding matrix $\mathbf{W}$ for QAM modulation can be obtained by (22).

V. NUMERICAL RESULTS

In this section, numerical results are presented based on Monte Carlo simulations. In each plot, we assume that the transmit power budget per symbol slot is $p_0 = 1$, leading to the total transmit power budget for the considered block of symbol slots as $P_{\text{total}} = Np_0 = N$. We compare our proposed CI-BLP schemes with closed-form ZF-based methods and traditional CI-SLP methods for both PSK and QAM modulation. Throughout this section, standard Rayleigh fading channels are employed, and perfect CSI is assumed. All the results are

generated based on 10000 independent channel realizations, and the execution time results are obtained from a Windows 11 Desktop with i9-12900k and 64GB RAM.

The following abbreviations are used throughout this section:

- 1) ‘ZF’: Traditional ZF precoding with block-level power normalization, where the precoding matrix is given by

$$\mathbf{W}_{\text{ZF}} = \frac{1}{f_{\text{ZF}}} \mathbf{H}^H (\mathbf{H} \mathbf{H}^H)^{-1}, \quad (33)$$

with block-level scaling factor

$$f_{\text{ZF}} = \sqrt{\frac{\mathbf{W}_{\text{ZF}}^H \mathbf{S}}{Np_0}}; \quad (34)$$

- 2) ‘RZF’: Traditional RZF precoding with block-level power normalization, where the precoding matrix is given by

$$\mathbf{W}_{\text{RZF}} = \frac{1}{f_{\text{RZF}}} \mathbf{H}^H \left(\mathbf{H} \mathbf{H}^H + \frac{K}{\rho} \mathbf{I} \right)^{-1}, \quad (35)$$

with block-level scaling factor

$$f_{\text{RZF}} = \sqrt{\frac{\mathbf{W}_{\text{RZF}}^H \mathbf{S}}{Np_0}}; \quad (36)$$

- 3) ‘SINR Balancing’: Block-level SINR balancing precoding for PSK and QAM modulation [11];
- 4) ‘CI-SLP-QP’: Traditional CI-SLP method with a symbol-level power constraint solved by QP, $\mathcal{P}_8$ in [39] for PSK and $\mathcal{P}_5$ in [40] for QAM;
- 5) ‘CI-SLP-Adaptive’: Traditional CI-SLP with a block-level power constraint, closed-form in-block power allocation [51];
- 6) ‘CI-BLP-CVX’: Original CI-BLP optimization $\mathcal{P}_1^{\text{PSK}}$ in (13) and $\mathcal{P}_1^{\text{QAM}}$ in (24);
- 7) ‘CI-BLP-QP’: Proposed QP solution for CI-BLP $\mathcal{P}_3^{\text{PSK}}$ in (21) and $\mathcal{P}_2^{\text{QAM}}$ in (30).

Fig. 4 depicts the SER of the proposed CI-BLP scheme when QPSK modulation is employed in a 12×12 MU-MISO system, where the length of the block is $N = 15$. As can be observed, both CI-based precoding approaches achieve an

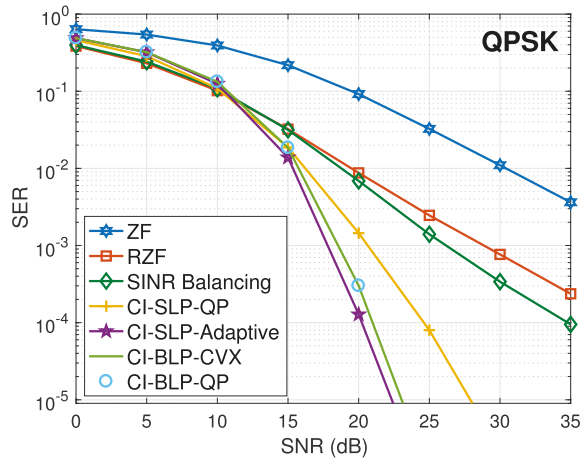
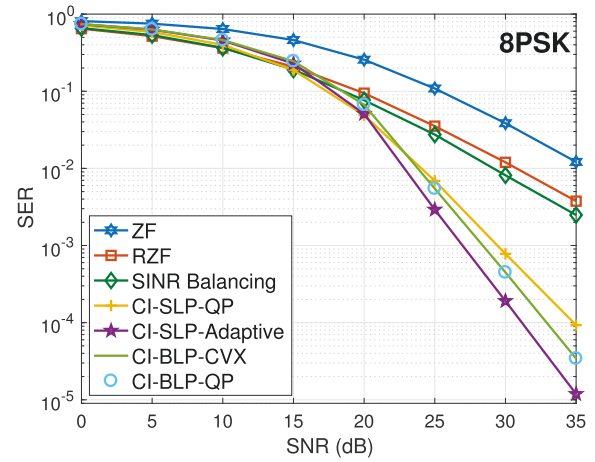
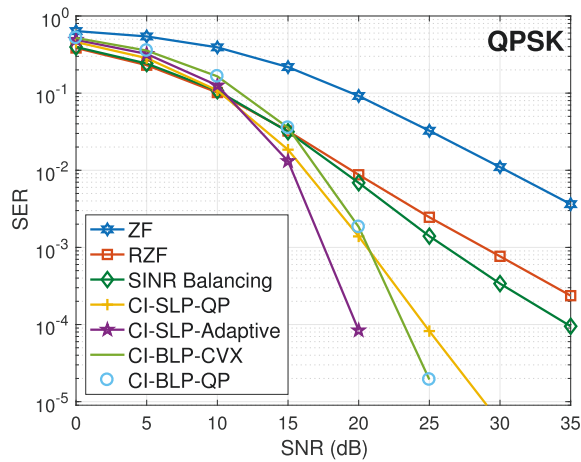
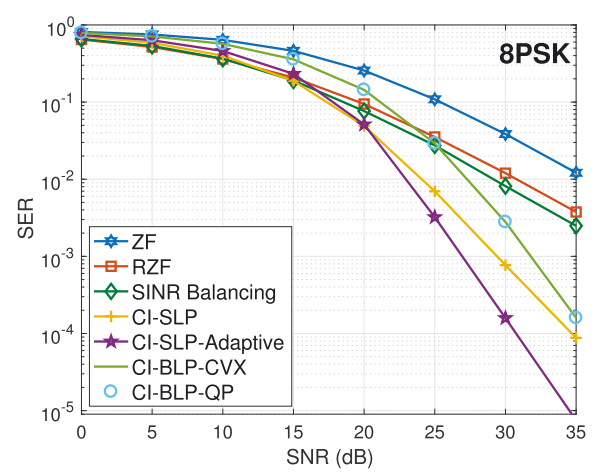
$$\frac{\partial \mathcal{L}_2}{\partial \hat{t}} = \sum_{n=1}^N \mathbf{1}^T \beta^n - 1 = 0 \quad (28a)$$

$$\begin{aligned} \frac{\partial \mathcal{L}_2}{\partial \hat{\mathbf{W}}} = & - \sum_{n=1}^N \left[(\beta^n)^T \mathbf{A}^n \right]^T (\mathbf{s}_E^n)^T - \sum_{n=1}^N \left[(\beta^n)^T \mathbf{B}^n \right]^T (\mathbf{c}_E^n)^T \\ & + 2\hat{\mu} \hat{\mathbf{W}} \left[\sum_{n=1}^N \mathbf{s}_E^n (\mathbf{s}_E^n)^T + \sum_{n=1}^N \mathbf{c}_E^n (\mathbf{c}_E^n)^T \right] = \mathbf{0} \end{aligned} \quad (28b)$$

$$\hat{\delta}_k^n \left[\hat{t} - (\mathbf{a}_k^n)^T \hat{\mathbf{W}} \mathbf{s}_E^n - (\mathbf{b}_k^n)^T \hat{\mathbf{W}} \mathbf{c}_E^n \right] = 0, \quad \hat{\delta}_k^n \geq 0, \quad \forall k \in \mathcal{O}^n, n \in \mathcal{N} \quad (28c)$$

$$\hat{t} - (\mathbf{a}_k^n)^T \hat{\mathbf{W}} \mathbf{s}_E^n - (\mathbf{b}_k^n)^T \hat{\mathbf{W}} \mathbf{c}_E^n = 0, \quad \forall k \in \mathcal{I}^n, n \in \mathcal{N} \quad (28d)$$

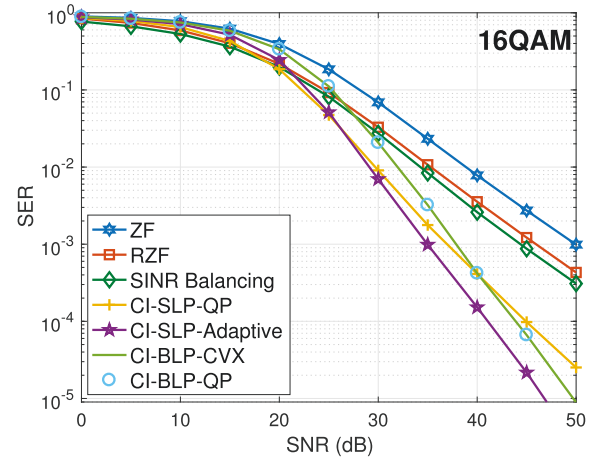
$$\hat{\mu} \left[\sum_{n=1}^N (\mathbf{s}_E^n)^T \hat{\mathbf{W}}^T \hat{\mathbf{W}} \mathbf{s}_E^n + \sum_{n=1}^N (\mathbf{c}_E^n)^T \hat{\mathbf{W}}^T \hat{\mathbf{W}} \mathbf{c}_E^n - Np_0 \right] = 0, \quad \hat{\mu} \geq 0 \quad (28e)$$

Fig. 4. SER v.s. SNR, QPSK, $N_T = K = 12$, $N = 15$.Fig. 6. SER v.s. SNR, 8PSK, $N_T = K = 12$, $N = 15$.Fig. 5. SER v.s. SNR, QPSK, $N_T = K = 12$, $N = 40$.Fig. 7. SER v.s. SNR, 8PSK, $N_T = K = 12$, $N = 40$.

improved performance over ZF precoding. When the length of the transmission block is short, we observe that the proposed CI-BLP offers noticeable performance gains over traditional CI-SLP that optimizes the precoding matrix on a symbol level, owing to the relaxed power constraint over the entire block, i.e., a power allocation among symbol slots is inherently performed for the proposed CI-BLP method. The result in Fig. 4 also validates the correctness of our derivations in the paper, as evidenced by the identical SER performance for ‘CI-BLP-CVX’ and ‘CI-BLP-QP’.

Fig. 5 compares the SER performance of different CI precoding approaches for a 12×12 MU-MISO system with QPSK modulation, where the length of the block is increased to $N = 40$. In this case, the proposed CI-BLP still offers significant performance gains over conventional ZF and RZF precoding. While we observe that the performance gain of the proposed CI-BLP becomes less significant when the block length increases, it still outperforms traditional CI-SLP when the transmit SNR goes above 20 dB. Compared to traditional CI-SLP methods that need to solve N optimization problems for the block, the proposed CI-BLP method only needs to solve the CI-BLP optimization problem once, thus further motivating the use of CI-based precoding in practical wireless systems.

Fig. 6 depicts the SER performance of different precoding methods when 8PSK modulation is employed, where

Fig. 8. SER v.s. SNR, 16QAM, $N_T = K = 12$, $N = 15$.

$K = N_T = 12$ and the block length is $N = 15$. Similar to the QPSK case, both CI-based precoding approaches outperform conventional ZF and RZF precoding. Again, as the transmit SNR increases, the proposed CI-BLP is able to achieve an improved SER over the traditional CI-SLP scheme, thanks to the inherent power allocation among different symbol slots.

Fig. 7 compares the SER performance of different precoding schemes when the length of the transmission block is increased

to $N = 40$ for 8PSK modulation, where $K = N_T = 12$. As N increases, the performance gain of the proposed CI-BLP method is reduced due to that fact that a larger number of constraints are simultaneously enforced in the corresponding CI-BLP optimization problem, but it is still able to outperform traditional ZF and RZF precoding when the transmit SNR is larger than 25dB and approach the traditional CI-SLP scheme. Again, the numerical results validate the correctness of our derivations for PSK modulation in Section III.

Fig. 8 depicts the SER performance of different precoding approaches when QAM modulation is employed for a 12×12 MU-MISO communication system, with a block length of $N = 15$. As can be observed, when we shift from PSK modulation to QAM modulation, the SER improvements for CI-based precoding become less significant, because only the outer QAM constellation points can exploit CI. The result in Fig. 8 validates the correctness of our derivations for the proposed CI-BLP with QAM modulation in Section IV, as evidence by the identical SER performance for ‘CI-BLP-CVX’ and ‘CI-BLP-QP’.

Fig. 9 depicts the complexity of the proposed CI-BLP method with traditional CI-SLP in terms of the execution time, where results for 6×6 , 12×12 and 18×18 MU-MISO systems are presented. For fairness of comparison, we only evaluate the time complexity of the ‘quadprog’ function in MATLAB that is used to solve the corresponding QP problem for both CI-BLP and CI-SLP, and avoid the run time consumed in constructing the required matrices/vectors. 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. From Fig. 9, we observe that the proposed CI-BLP approach offers a significant complexity gain over traditional

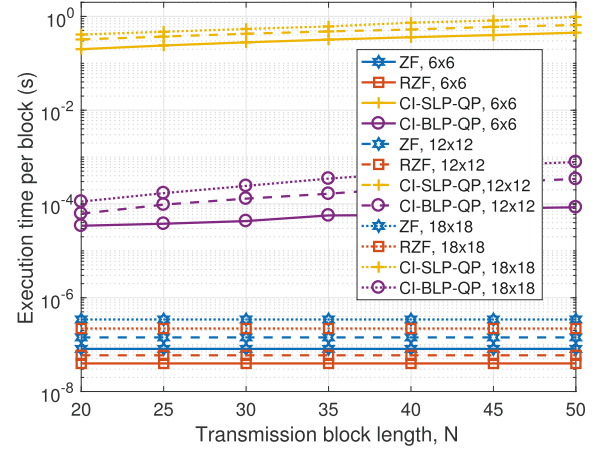


Fig. 9. Execution time v.s. block length N .

CI-SLP, and the complexity gains become more prominent as the block length N increases.

To illustrate the effect of block length N on the performance, in Fig. 10 we depict the SER with respect to the block length N , where 8PSK modulation is employed at a transmit SNR of 30dB. 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. Interestingly, 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. The results in Fig. 9 and Fig. 10 demonstrates that CI-BLP achieves an

$$\begin{aligned}
\mathcal{L}_1(\hat{\mathbf{W}}, t, \delta_k^n, \mu) &= -t + \sum_{n=1}^N \sum_{k=1}^{2K} \delta_k^n \left[t - (\mathbf{a}_k^n)^T \hat{\mathbf{W}} \mathbf{s}_E^n - (\mathbf{b}_k^n)^T \hat{\mathbf{W}} \mathbf{c}_E^n \right] + \mu \left[\sum_{n=1}^N (\mathbf{s}_E^n)^T (\mathbf{P} \hat{\mathbf{W}} + \mathbf{Q} \hat{\mathbf{W}} \mathbf{T})^T (\mathbf{P} \hat{\mathbf{W}} + \mathbf{Q} \hat{\mathbf{W}} \mathbf{T}) \mathbf{s}_E^n - N p_0 \right] \\
&= \left(\sum_{n=1}^N \mathbf{1}^T \delta^n - 1 \right) t - \sum_{n=1}^N (\delta^n)^T \mathbf{A}^n \hat{\mathbf{W}} \mathbf{s}_E^n - \sum_{n=1}^N (\delta^n)^T \mathbf{B}^n \hat{\mathbf{W}} \mathbf{c}_E^n \\
&\quad + \mu \left[\sum_{n=1}^N (\mathbf{s}_E^n)^T (\hat{\mathbf{W}}^T \mathbf{P}^T + \mathbf{T}^T \hat{\mathbf{W}}^T \mathbf{Q}^T) (\mathbf{P} \hat{\mathbf{W}} + \mathbf{Q} \hat{\mathbf{W}} \mathbf{T}) \mathbf{s}_E^n \right] - \mu N p_0 \\
&= \left(\sum_{n=1}^N \mathbf{1}^T \delta^n - 1 \right) t - \sum_{n=1}^N (\delta^n)^T \mathbf{A}^n \hat{\mathbf{W}} \mathbf{s}_E^n - \sum_{n=1}^N (\delta^n)^T \mathbf{B}^n \hat{\mathbf{W}} \mathbf{c}_E^n + \mu \sum_{n=1}^N (\mathbf{s}_E^n)^T \hat{\mathbf{W}}^T \mathbf{P}^T \mathbf{P} \hat{\mathbf{W}} \mathbf{s}_E^n \\
&\quad + \underbrace{\mu \sum_{n=1}^N (\mathbf{s}_E^n)^T \hat{\mathbf{W}}^T \mathbf{P}^T \mathbf{Q} \hat{\mathbf{W}} \mathbf{T} \mathbf{s}_E^n}_{=0} + \underbrace{\mu \sum_{n=1}^N (\mathbf{s}_E^n)^T \mathbf{T}^T \hat{\mathbf{W}}^T \mathbf{Q}^T \mathbf{P} \hat{\mathbf{W}} \mathbf{s}_E^n}_{=0} + \mu \sum_{n=1}^N \underbrace{(\mathbf{s}_E^n)^T \mathbf{T}^T}_{(\mathbf{c}_E^n)^T} \hat{\mathbf{W}}^T \mathbf{Q}^T \mathbf{Q} \hat{\mathbf{W}} \underbrace{\mathbf{T} \mathbf{s}_E^n}_{\mathbf{c}_E^n} - \mu N p_0 \\
&= \left(\sum_{n=1}^N \mathbf{1}^T \delta^n - 1 \right) t - \sum_{n=1}^N (\delta^n)^T \mathbf{A}^n \hat{\mathbf{W}} \mathbf{s}_E^n - \sum_{n=1}^N (\delta^n)^T \mathbf{B}^n \hat{\mathbf{W}} \mathbf{c}_E^n + \mu \sum_{n=1}^N (\mathbf{s}_E^n)^T \hat{\mathbf{W}}^T \hat{\mathbf{W}} \mathbf{s}_E^n + \mu \sum_{n=1}^N (\mathbf{c}_E^n)^T \hat{\mathbf{W}}^T \hat{\mathbf{W}} \mathbf{c}_E^n - \mu N p_0
\end{aligned} \tag{40}$$

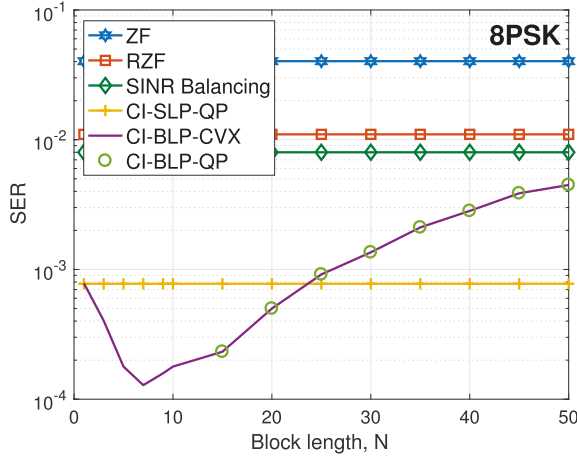


Fig. 10. SER v.s. block length N , 8PSK, $N_T = K = 12$, SNR = 30dB.

improved performance-complexity tradeoff when the value of the block length N is moderate.

VI. CONCLUSION

In this paper, a CI-based block-level precoding algorithm is proposed for the downlink of a MU-MISO communication system. As opposed to traditional CI-based precoding schemes that employ a symbol-level design, the proposed CI-BLP applies a constant precoding matrix to a block of symbol slots within the channel coherence interval, thus greatly reducing the number of optimization problems that need to be solved

for CI-based precoding. For both PSK and QAM modulation, the optimal precoding matrix for CI-BLP is derived by constructing the Lagrangian and formulating the KKT conditions. Further manipulations of the dual problem demonstrate that the CI-BLP problem is equivalent to a QP optimization. Numerical results show that owing to the relaxed block-level power constraint, the proposed CI-BLP approach offers an improved performance over the traditional CI-SLP scheme when the length of the considered block is short and the SNR is sufficiently high, and exhibits only a slight performance loss as the block length increases.

APPENDIX A

CONSTRUCTION OF $\mathbf{M}^n$ IN (4)

Based on Section IV-A of [30], $\mathbf{M}^n \in \mathbb{R}^{2K \times 2N_T}$ in (4) can be constructed based on the channel vector $\mathbf{h}_k$ and the data symbol s_k^n for each user, given by

$$\mathbf{M}^n = [\mathbf{j}_1 \quad \mathbf{j}_2 \quad \cdots \quad \mathbf{j}_K \quad \mathbf{l}_1 \quad \mathbf{l}_2 \quad \cdots \quad \mathbf{l}_K]^T, \quad (37)$$

where $\mathbf{j}_k \in \mathbb{R}^{2N_T}$ and $\mathbf{l}_k \in \mathbb{R}^{2N_T}$ are given by

$$\mathbf{j}_k = \begin{bmatrix} \frac{\Im(s_{k,B}^n) \Re(\mathbf{h}_k) - \Re(s_{k,B}^n) \Im(\mathbf{h}_k)}{\Re(s_{k,A}^n) \Im(s_{k,B}^n) - \Im(s_{k,A}^n) \Re(s_{k,B}^n)} \\ \frac{\Im(s_{k,B}^n) \Im(\mathbf{h}_k) + \Re(s_{k,B}^n) \Re(\mathbf{h}_k)}{\Re(s_{k,A}^n) \Im(s_{k,B}^n) - \Im(s_{k,A}^n) \Re(s_{k,B}^n)} \end{bmatrix}, \quad (38)$$

$$\frac{\partial \mathcal{L}_1}{\partial t} = \sum_{n=1}^N \mathbf{1}^T \boldsymbol{\delta}^n - 1 = 0 \quad (42a)$$

$$\begin{aligned} \frac{\partial \mathcal{L}_1}{\partial \hat{\mathbf{W}}} = & - \sum_{n=1}^N [(\boldsymbol{\delta}^n)^T \mathbf{A}^n]^T (\mathbf{s}_E^n)^T - \sum_{n=1}^N [(\boldsymbol{\delta}^n)^T \mathbf{B}^n]^T (\mathbf{c}_E^n)^T \\ & + 2u_0 \hat{\mathbf{W}} \left[\sum_{n=1}^N \mathbf{s}_E^n (\mathbf{s}_E^n)^T + \sum_{n=1}^N \mathbf{c}_E^n (\mathbf{c}_E^n)^T \right] = \mathbf{0} \end{aligned} \quad (42b)$$

$$\delta_k^n \left[t - (\mathbf{a}_k^n)^T \hat{\mathbf{W}} \mathbf{s}_E^n - (\mathbf{b}_k^n)^T \hat{\mathbf{W}} \mathbf{c}_E^n \right] = 0, \quad \delta_k^n \geq 0, \quad \forall k \in \mathcal{K}, \quad \forall n \in \mathcal{N} \quad (42c)$$

$$\mu \left[\sum_{n=1}^N (\mathbf{s}_E^n)^T \hat{\mathbf{W}}^T \hat{\mathbf{W}} \mathbf{s}_E^n + \sum_{n=1}^N (\mathbf{c}_E^n)^T \hat{\mathbf{W}}^T \hat{\mathbf{W}} \mathbf{c}_E^n - Np_0 \right] = 0, \quad \mu \geq 0 \quad (42d)$$

$$\begin{aligned} \mathcal{U}_1 = & \max_{\{\boldsymbol{\delta}^m\}, \mu} - \sum_{m=1}^N (\boldsymbol{\delta}^m)^T \mathbf{A}^m \hat{\mathbf{W}} \mathbf{s}_E^m - \sum_{m=1}^N (\boldsymbol{\delta}^m)^T \mathbf{B}^m \hat{\mathbf{W}} \mathbf{c}_E^m \\ = & \min_{\{\boldsymbol{\delta}^m\}, \mu} \sum_{m=1}^N (\boldsymbol{\delta}^m)^T \mathbf{A}^m \frac{1}{2\mu} \sum_{n=1}^N \left[(\mathbf{A}^n)^T \boldsymbol{\delta}^n (\mathbf{s}_E^n)^T + (\mathbf{B}^n)^T \boldsymbol{\delta}^n (\mathbf{c}_E^n)^T \right] \mathbf{D}^{-1} \mathbf{s}_E^m \\ & + \sum_{m=1}^N (\boldsymbol{\delta}^m)^T \mathbf{B}^m \frac{1}{2\mu} \sum_{n=1}^N \left[(\mathbf{A}^n)^T \boldsymbol{\delta}^n (\mathbf{s}_E^n)^T + (\mathbf{B}^n)^T \boldsymbol{\delta}^n (\mathbf{c}_E^n)^T \right] \mathbf{D}^{-1} \mathbf{c}_E^m \\ = & \min_{\{\boldsymbol{\delta}^m\}, \mu} \frac{1}{2\mu} \sum_{m=1}^N \sum_{n=1}^N (\boldsymbol{\delta}^m)^T \mathbf{A}^m (\mathbf{A}^n)^T \boldsymbol{\delta}^n (\mathbf{s}_E^n)^T \mathbf{D}^{-1} \mathbf{s}_E^m + \frac{1}{2\mu} \sum_{m=1}^N \sum_{n=1}^N (\boldsymbol{\delta}^m)^T \mathbf{A}^m (\mathbf{B}^n)^T \boldsymbol{\delta}^n (\mathbf{c}_E^n)^T \mathbf{D}^{-1} \mathbf{s}_E^m \\ & + \frac{1}{2\mu} \sum_{m=1}^N \sum_{n=1}^N (\boldsymbol{\delta}^m)^T \mathbf{B}^m (\mathbf{A}^n)^T \boldsymbol{\delta}^n (\mathbf{s}_E^n)^T \mathbf{D}^{-1} \mathbf{c}_E^m + \frac{1}{2\mu} \sum_{m=1}^N \sum_{n=1}^N (\boldsymbol{\delta}^m)^T \mathbf{B}^m (\mathbf{B}^n)^T \boldsymbol{\delta}^n (\mathbf{c}_E^n)^T \mathbf{D}^{-1} \mathbf{c}_E^m \end{aligned} \quad (45)$$

and

$$\mathbf{l}_k = \begin{bmatrix} \Re(s_{k,\mathcal{A}}^n) \Im(\mathbf{h}_k) - \Im(s_{k,\mathcal{A}}^n) \Re(\mathbf{h}_k) \\ \Re(s_{k,\mathcal{A}}^n) \Im(s_{k,\mathcal{B}}^n) - \Im(s_{k,\mathcal{A}}^n) \Re(s_{k,\mathcal{B}}^n) \\ \Re(s_{k,\mathcal{A}}^n) \Re(\mathbf{h}_k) + \Im(s_{k,\mathcal{A}}^n) \Im(\mathbf{h}_k) \\ \Re(s_{k,\mathcal{A}}^n) \Im(s_{k,\mathcal{B}}^n) - \Im(s_{k,\mathcal{A}}^n) \Re(s_{k,\mathcal{B}}^n) \end{bmatrix}. \quad (39)$$

APPENDIX B

PROOF FOR PROPOSITION 1

To begin with, the Lagrangian of $\mathcal{P}_1^{\text{PSK}}$ can be constructed as shown in (40), at the bottom of page 10, where $\boldsymbol{\delta}^n = [\delta_1^n, \delta_2^n, \dots, \delta_{2K}^n]^T \in \mathbb{R}^{2K}$ and μ are the non-negative

dual variables associated with the inequality constraints **C1** and **C2** respectively, and the last step is obtained by the fact that

$$\mathbf{P}^T \mathbf{P} = \mathbf{Q}^T \mathbf{Q} = \mathbf{I}_{N_T}, \quad \mathbf{P}^T \mathbf{Q} = \mathbf{Q}^T \mathbf{P} = \mathbf{0}. \quad (41)$$

Accordingly, the KKT conditions for optimality of $\mathcal{P}_1^{\text{PSK}}$ can be formulated and are shown in (42), at the bottom of the previous page. Based on the KKT conditions, we first obtain that $\mu > 0$, otherwise $\boldsymbol{\delta}^n = \mathbf{0}$, $\forall n$, which contradicts (42a), meaning that the block-level power constraint is active when optimality is achieved, i.e.,

$$\sum_{n=1}^N (\mathbf{s}_E^n)^T \hat{\mathbf{W}}^T \hat{\mathbf{W}} \mathbf{s}_E^n + \sum_{n=1}^N (\mathbf{c}_E^n)^T \hat{\mathbf{W}}^T \hat{\mathbf{W}} \mathbf{c}_E^n = N p_0. \quad (43)$$

$$\begin{aligned} \mathcal{U}_1 &= \min_{\{\boldsymbol{\delta}^m\}, \mu} \frac{1}{2\mu} \sum_{m=1}^N \sum_{n=1}^N (\boldsymbol{\delta}^m)^T \left[p_{m,n} \mathbf{A}^m (\mathbf{A}^n)^T \right] \boldsymbol{\delta}^n + \frac{1}{2\mu} \sum_{m=1}^N \sum_{n=1}^N (\boldsymbol{\delta}^m)^T \left[f_{m,n} \mathbf{A}^m (\mathbf{B}^n)^T \right] \boldsymbol{\delta}^n \\ &\quad + \frac{1}{2\mu} \sum_{m=1}^N \sum_{n=1}^N (\boldsymbol{\delta}^m)^T \left[g_{m,n} \mathbf{B}^m (\mathbf{A}^n)^T \right] \boldsymbol{\delta}^n + \frac{1}{2\mu} \sum_{m=1}^N \sum_{n=1}^N (\boldsymbol{\delta}^m)^T \left[q_{m,n} \mathbf{B}^m (\mathbf{B}^n)^T \right] \boldsymbol{\delta}^n \\ &= \min_{\{\boldsymbol{\delta}^m\}, \mu} \frac{1}{2\mu} \sum_{m=1}^N \sum_{n=1}^N (\boldsymbol{\delta}^m)^T \left[p_{m,n} \mathbf{A}^m (\mathbf{A}^n)^T + f_{m,n} \mathbf{A}^m (\mathbf{B}^n)^T + g_{m,n} \mathbf{B}^m (\mathbf{A}^n)^T + q_{m,n} \mathbf{B}^m (\mathbf{B}^n)^T \right] \boldsymbol{\delta}^n \end{aligned} \quad (47)$$

$$\begin{aligned} &\sum_{l=1}^N (\mathbf{s}_E^l)^T \hat{\mathbf{W}}^T \hat{\mathbf{W}} \mathbf{s}_E^l \\ &= \frac{1}{4\mu^2} \sum_{l=1}^N (\mathbf{s}_E^l)^T \left\{ \sum_{m=1}^N \left[(\mathbf{A}^m)^T \boldsymbol{\delta}^m (\mathbf{s}_E^m)^T + (\mathbf{B}^m)^T \boldsymbol{\delta}^m (\mathbf{c}_E^m)^T \right] \mathbf{D}^{-1} \right\}^T \left\{ \sum_{n=1}^N \left[(\mathbf{A}^n)^T \boldsymbol{\delta}^n (\mathbf{s}_E^n)^T + (\mathbf{B}^n)^T \boldsymbol{\delta}^n (\mathbf{c}_E^n)^T \right] \mathbf{D}^{-1} \right\} \mathbf{s}_E^l \\ &= \frac{1}{4\mu^2} \sum_{l=1}^N (\mathbf{s}_E^l)^T \mathbf{D}^{-1} \left\{ \sum_{m=1}^N \left[\mathbf{s}_E^m (\boldsymbol{\delta}^m)^T \mathbf{A}^m + \mathbf{c}_E^m (\boldsymbol{\delta}^m)^T \mathbf{B}^m \right] \right\} \left\{ \sum_{n=1}^N \left[(\mathbf{A}^n)^T \boldsymbol{\delta}^n (\mathbf{s}_E^n)^T + (\mathbf{B}^n)^T \boldsymbol{\delta}^n (\mathbf{c}_E^n)^T \right] \right\} \mathbf{D}^{-1} \mathbf{s}_E^l \\ &= \frac{1}{4\mu^2} \sum_{l=1}^N \sum_{m=1}^N \sum_{n=1}^N \left[\underbrace{(\mathbf{s}_E^l)^T \mathbf{D}^{-1} \mathbf{s}_E^m}_{p_{m,l}} (\boldsymbol{\delta}^m)^T \mathbf{A}^m (\mathbf{A}^n)^T \boldsymbol{\delta}^n \underbrace{(\mathbf{s}_E^n)^T \mathbf{D}^{-1} \mathbf{s}_E^l}_{p_{l,n}} \right] \\ &\quad + \frac{1}{4\mu^2} \sum_{l=1}^N \sum_{m=1}^N \sum_{n=1}^N \left[\underbrace{(\mathbf{s}_E^l)^T \mathbf{D}^{-1} \mathbf{s}_E^m}_{p_{m,l}} (\boldsymbol{\delta}^m)^T \mathbf{A}^m (\mathbf{B}^n)^T \boldsymbol{\delta}^n \underbrace{(\mathbf{c}_E^n)^T \mathbf{D}^{-1} \mathbf{s}_E^l}_{f_{l,n}} \right] \\ &\quad + \frac{1}{4\mu^2} \sum_{l=1}^N \sum_{m=1}^N \sum_{n=1}^N \left[\underbrace{(\mathbf{s}_E^l)^T \mathbf{D}^{-1} \mathbf{c}_E^m}_{g_{m,l}} (\boldsymbol{\delta}^m)^T \mathbf{B}^m (\mathbf{A}^n)^T \boldsymbol{\delta}^n \underbrace{(\mathbf{s}_E^n)^T \mathbf{D}^{-1} \mathbf{s}_E^l}_{p_{l,n}} \right] \\ &\quad + \frac{1}{4\mu^2} \sum_{l=1}^N \sum_{m=1}^N \sum_{n=1}^N \left[\underbrace{(\mathbf{s}_E^l)^T \mathbf{D}^{-1} \mathbf{c}_E^m}_{g_{m,l}} (\boldsymbol{\delta}^m)^T \mathbf{B}^m (\mathbf{B}^n)^T \boldsymbol{\delta}^n \underbrace{(\mathbf{c}_E^n)^T \mathbf{D}^{-1} \mathbf{s}_E^l}_{f_{l,n}} \right] \\ &= \frac{1}{4\mu^2} \sum_{l=1}^N \sum_{m=1}^N \sum_{n=1}^N \left\{ (\boldsymbol{\delta}^m)^T \left[p_{l,n} p_{m,l} \mathbf{A}^m (\mathbf{A}^n)^T \right] \boldsymbol{\delta}^n \right\} + \frac{1}{4\mu^2} \sum_{l=1}^N \sum_{m=1}^N \sum_{n=1}^N \left\{ (\boldsymbol{\delta}^m)^T \left[f_{l,n} p_{m,l} \mathbf{A}^m (\mathbf{B}^n)^T \right] \boldsymbol{\delta}^n \right\} \\ &\quad + \frac{1}{4\mu^2} \sum_{l=1}^N \sum_{m=1}^N \sum_{n=1}^N \left\{ (\boldsymbol{\delta}^m)^T \left[p_{l,n} g_{m,l} \mathbf{B}^m (\mathbf{A}^n)^T \right] \boldsymbol{\delta}^n \right\} + \frac{1}{4\mu^2} \sum_{l=1}^N \sum_{m=1}^N \sum_{n=1}^N \left\{ (\boldsymbol{\delta}^m)^T \left[f_{l,n} g_{m,l} \mathbf{B}^m (\mathbf{B}^n)^T \right] \boldsymbol{\delta}^n \right\} \\ &= \frac{1}{4\mu^2} \sum_{l=1}^N \sum_{m=1}^N \sum_{n=1}^N \left\{ (\boldsymbol{\delta}^m)^T \left[p_{l,n} p_{m,l} \mathbf{A}^m (\mathbf{A}^n)^T + f_{l,n} p_{m,l} \mathbf{A}^m (\mathbf{B}^n)^T + p_{l,n} g_{m,l} \mathbf{B}^m (\mathbf{A}^n)^T + f_{l,n} g_{m,l} \mathbf{B}^m (\mathbf{B}^n)^T \right] \boldsymbol{\delta}^n \right\} \end{aligned} \quad (52)$$

This is also intuitive because a linear scaling to $\hat{\mathbf{W}}$ directly leads to an increase in the objective value t . To proceed, we transform (42b) into

$$2\mu \hat{\mathbf{W}}\mathbf{D} = \sum_{n=1}^N \left[(\mathbf{A}^n)^T \delta^n (\mathbf{s}_E^n)^T + (\mathbf{B}^n)^T \delta^n (\mathbf{c}_E^n)^T \right]. \quad (44)$$

Based on the fact that the length of the transmission block N is in general larger than the number of users K , $\mathbf{D}$ is thus full-rank and invertible [52]. Accordingly, we can obtain the optimal precoding matrix $\hat{\mathbf{W}}$ as a function of the Lagrange multipliers δ^n by right-multiplying the inverse of $\mathbf{D}$ and obtain (14), which completes the proof. ■

APPENDIX C

PROOF FOR PROPOSITION 2

Based on (16), $\mathcal{U}_1$ can be further transformed and is shown in (45), at the bottom of page 11. By defining

$$p_{m,n} = (\mathbf{s}_E^n)^T \mathbf{D}^{-1} \mathbf{s}_E^m, q_{m,n} = (\mathbf{c}_E^n)^T \mathbf{D}^{-1} \mathbf{c}_E^m, \\ f_{m,n} = (\mathbf{c}_E^n)^T \mathbf{D}^{-1} \mathbf{s}_E^m, g_{m,n} = (\mathbf{s}_E^n)^T \mathbf{D}^{-1} \mathbf{c}_E^m, \quad (46)$$

the objective function of the dual problem $\mathcal{U}_1$ can be further simplified and is given by (47), shown at the bottom of the previous page. Further defining

$$\delta_E = \left[(\delta^1)^T, (\delta^2)^T, \dots, (\delta^N)^T \right]^T \in \mathbb{R}^{2NK} \quad (48)$$

and $\mathbf{U}_{m,n} \in \mathbb{R}^{2K \times 2K}$ given by

$$\mathbf{U}_{m,n} = p_{m,n} \mathbf{A}^m (\mathbf{A}^n)^T + f_{m,n} \mathbf{A}^m (\mathbf{B}^n)^T + g_{m,n} \mathbf{B}^m (\mathbf{A}^n)^T \\ + q_{m,n} \mathbf{B}^m (\mathbf{B}^n)^T, \quad (49)$$

the objective function of the dual problem $\mathcal{U}_1$ can finally be obtained as

$$\mathcal{U}_1 = \min_{\{\delta^m\}, \mu} \frac{1}{2\mu} \sum_{m=1}^N \sum_{n=1}^N (\delta^m)^T \mathbf{U}_{m,n} \delta^n \\ = \min_{\{\delta^m\}, \mu} \frac{1}{2\mu} (\delta_E)^T \mathbf{U} \delta_E, \quad (50)$$

where $\mathbf{U} \in \mathbb{R}^{2NK \times 2NK}$ is a block matrix constructed as

$$\mathbf{U} = \begin{bmatrix} \mathbf{U}_{1,1} & \cdots & \cdots & \cdots & \mathbf{U}_{1,N} \\ \vdots & \ddots & \mathbf{U}_{m,n} & \ddots & \vdots \\ \mathbf{U}_{N,1} & \cdots & \cdots & \cdots & \mathbf{U}_{N,N} \end{bmatrix}. \quad (51)$$

The result in (50) reveals that the objective function of the dual problem for the proposed CI-BLP scheme with PSK modulation is in a quadratic form.

After studying the objective function $\mathcal{U}_1$, we focus on the block-level power constraint in (43). By substituting the expression for $\hat{\mathbf{W}}$ in (14) into (43), the transformation and simplification of the first term on the left-hand side of (43) is shown in (52), at the bottom of the previous page. Based on the result in (52) and by introducing $\mathbf{F}_{m,n}^l \in \mathbb{R}^{2K \times 2K}$ as

$$\mathbf{F}_{m,n}^l = p_{l,n} p_{m,l} \mathbf{A}^m (\mathbf{A}^n)^T + f_{l,n} p_{m,l} \mathbf{A}^m (\mathbf{B}^n)^T \\ + p_{l,n} g_{m,l} \mathbf{B}^m (\mathbf{A}^n)^T + f_{l,n} g_{m,l} \mathbf{B}^m (\mathbf{B}^n)^T, \quad (53)$$

the first term on the left-hand side of (43) can finally be expressed in a compact form as

$$\sum_{l=1}^N (\mathbf{s}_E^l)^T \hat{\mathbf{W}}^T \hat{\mathbf{W}} \mathbf{s}_E^l = \frac{1}{4\mu^2} \sum_{l=1}^N (\delta_E)^T \mathbf{F}^l \delta_E \\ = \frac{1}{4\mu^2} (\delta_E)^T \mathbf{F} \delta_E, \quad (54)$$

where $\mathbf{F} = \sum_{l=1}^N \mathbf{F}^l \in \mathbb{R}^{2NK \times 2NK}$, with each $\mathbf{F}^l$ constructed as below:

$$\mathbf{F}^l = \begin{bmatrix} \mathbf{F}_{1,1}^l & \cdots & \cdots & \cdots & \mathbf{F}_{1,N}^l \\ \vdots & \ddots & \mathbf{F}_{m,n}^l & \ddots & \vdots \\ \mathbf{F}_{N,1}^l & \cdots & \cdots & \cdots & \mathbf{F}_{N,N}^l \end{bmatrix}. \quad (55)$$

Following the above procedure, the second term on the left-hand side of (43) can be similarly expressed in a compact form as:

$$\sum_{n=1}^N (\mathbf{c}_E^n)^T \hat{\mathbf{W}}^T \hat{\mathbf{W}} \mathbf{c}_E^n = \frac{1}{4\mu^2} (\delta_E)^T \mathbf{G} \delta_E, \quad (56)$$

where $\mathbf{G} = \sum_{l=1}^N \mathbf{G}^l \in \mathbb{R}^{2NK \times 2NK}$, and each $\mathbf{G}^l$ is constructed as below:

$$\mathbf{G}^l = \begin{bmatrix} \mathbf{G}_{1,1}^l & \cdots & \cdots & \cdots & \mathbf{G}_{1,N}^l \\ \vdots & \ddots & \mathbf{G}_{m,n}^l & \ddots & \vdots \\ \mathbf{G}_{N,1}^l & \cdots & \cdots & \cdots & \mathbf{G}_{N,N}^l \end{bmatrix}, \quad (57)$$

with each $\mathbf{G}_{m,n}^l \in \mathbb{R}^{2K \times 2K}$ given by

$$\mathbf{G}_{m,n}^l = g_{l,n} f_{m,l} \mathbf{A}^m (\mathbf{A}^n)^T + q_{l,n} f_{m,l} \mathbf{A}^m (\mathbf{B}^n)^T \\ + g_{l,n} q_{m,l} \mathbf{B}^m (\mathbf{A}^n)^T + q_{l,n} q_{m,l} \mathbf{B}^m (\mathbf{B}^n)^T. \quad (58)$$

Based on the above derivations, the block-level power constraint (43) is equivalent to:

$$\frac{1}{4\mu^2} (\delta_E)^T \mathbf{F} \delta_E + \frac{1}{4\mu^2} (\delta_E)^T \mathbf{G} \delta_E = N p_0 \\ \Rightarrow \frac{1}{4\mu^2} (\delta_E)^T (\mathbf{F} + \mathbf{G}) \delta_E = N p_0. \quad (59)$$

Together with the constraints on the Lagrange multiplier δ_E from the KKT conditions and (50), we arrive at the formulation of $\mathcal{P}_2^{\text{PSK}}$, which completes the proof. ■

APPENDIX D

PROOF FOR PROPOSITION 3

We begin by expressing the generic (m,n) -th block in $(\mathbf{F} + \mathbf{G})$, which is shown in (60), at the top of the next page, based on (53) and (58), where we obtain that

$$\mathbf{F}_{m,n} + \mathbf{G}_{m,n} = \mathbf{U}_{m,n}, \quad \forall m, n \in \mathcal{N}, \quad (61)$$

which completes the proof. ■

$$\begin{aligned}
& \mathbf{F}_{m,n} + \mathbf{G}_{m,n} \\
&= \sum_{l=1}^N \mathbf{F}_{m,n}^l + \sum_{l=1}^N \mathbf{G}_{m,n}^l \\
&= \sum_{l=1}^N (p_{l,n} p_{m,l} + g_{l,n} f_{m,l}) \mathbf{A}^m (\mathbf{A}^n)^T + \sum_{l=1}^N (p_{l,n} p_{m,l} + g_{l,n} f_{m,l}) \mathbf{A}^m (\mathbf{B}^n)^T + \sum_{l=1}^N (p_{l,n} g_{m,l} + g_{l,n} q_{m,l}) \mathbf{B}^m (\mathbf{A}^n)^T \\
&\quad + \sum_{l=1}^N (f_{l,n} g_{m,l} + q_{l,n} q_{m,l}) \mathbf{B}^m (\mathbf{B}^n)^T \\
&= \sum_{l=1}^N \left[(\mathbf{s}_E^n)^T \mathbf{D}^{-1} \mathbf{s}_E^l (\mathbf{s}_E^l)^T \mathbf{D}^{-1} \mathbf{s}_E^m + (\mathbf{s}_E^n)^T \mathbf{D}^{-1} \mathbf{c}_E^l (\mathbf{c}_E^l)^T \mathbf{D}^{-1} \mathbf{s}_E^m \right] \mathbf{A}^m (\mathbf{A}^n)^T \\
&\quad + \sum_{l=1}^N \left[(\mathbf{c}_E^n)^T \mathbf{D}^{-1} \mathbf{s}_E^l (\mathbf{s}_E^l)^T \mathbf{D}^{-1} \mathbf{s}_E^m + (\mathbf{c}_E^n)^T \mathbf{D}^{-1} \mathbf{c}_E^l (\mathbf{c}_E^l)^T \mathbf{D}^{-1} \mathbf{s}_E^m \right] \mathbf{A}^m (\mathbf{B}^n)^T \\
&\quad + \sum_{l=1}^N \left[(\mathbf{s}_E^n)^T \mathbf{D}^{-1} \mathbf{s}_E^l (\mathbf{s}_E^l)^T \mathbf{D}^{-1} \mathbf{c}_E^m + (\mathbf{s}_E^n)^T \mathbf{D}^{-1} \mathbf{c}_E^l (\mathbf{c}_E^l)^T \mathbf{D}^{-1} \mathbf{c}_E^m \right] \mathbf{B}^m (\mathbf{A}^n)^T \\
&\quad + \sum_{l=1}^N \left[(\mathbf{c}_E^n)^T \mathbf{D}^{-1} \mathbf{s}_E^l (\mathbf{s}_E^l)^T \mathbf{D}^{-1} \mathbf{c}_E^m + (\mathbf{c}_E^n)^T \mathbf{D}^{-1} \mathbf{c}_E^l (\mathbf{c}_E^l)^T \mathbf{D}^{-1} \mathbf{c}_E^m \right] \mathbf{B}^m (\mathbf{B}^n)^T \\
&= (\mathbf{s}_E^n)^T \mathbf{D}^{-1} \underbrace{\left\{ \sum_{l=1}^N \left[\mathbf{s}_E^l (\mathbf{s}_E^l)^T + \mathbf{c}_E^l (\mathbf{c}_E^l)^T \right] \right\}}_{\mathbf{D}} \mathbf{D}^{-1} \mathbf{s}_E^m \mathbf{A}^m (\mathbf{A}^n)^T \\
&\quad + (\mathbf{c}_E^n)^T \mathbf{D}^{-1} \underbrace{\left\{ \sum_{l=1}^N \left[\mathbf{s}_E^l (\mathbf{s}_E^l)^T + \mathbf{c}_E^l (\mathbf{c}_E^l)^T \right] \right\}}_{\mathbf{D}} \mathbf{D}^{-1} \mathbf{s}_E^m \mathbf{A}^m (\mathbf{B}^n)^T \\
&\quad + (\mathbf{s}_E^n)^T \mathbf{D}^{-1} \underbrace{\left\{ \sum_{l=1}^N \left[\mathbf{s}_E^l (\mathbf{s}_E^l)^T + \mathbf{c}_E^l (\mathbf{c}_E^l)^T \right] \right\}}_{\mathbf{D}} \mathbf{D}^{-1} \mathbf{c}_E^m \mathbf{B}^m (\mathbf{A}^n)^T \\
&\quad + (\mathbf{c}_E^n)^T \mathbf{D}^{-1} \underbrace{\left\{ \sum_{l=1}^N \left[\mathbf{s}_E^l (\mathbf{s}_E^l)^T + \mathbf{c}_E^l (\mathbf{c}_E^l)^T \right] \right\}}_{\mathbf{D}} \mathbf{D}^{-1} \mathbf{c}_E^m \mathbf{B}^m (\mathbf{B}^n)^T \\
&= (\mathbf{s}_E^n)^T \mathbf{D}^{-1} \mathbf{s}_E^m \mathbf{A}^m (\mathbf{A}^n)^T + (\mathbf{c}_E^n)^T \mathbf{D}^{-1} \mathbf{s}_E^m \mathbf{A}^m (\mathbf{B}^n)^T + (\mathbf{s}_E^n)^T \mathbf{D}^{-1} \mathbf{c}_E^m \mathbf{B}^m (\mathbf{A}^n)^T + (\mathbf{c}_E^n)^T \mathbf{D}^{-1} \mathbf{c}_E^m \mathbf{B}^m (\mathbf{B}^n)^T \\
&= p_{m,n} \mathbf{A}^m (\mathbf{A}^n)^T + f_{m,n} \mathbf{A}^m (\mathbf{B}^n)^T + g_{m,n} \mathbf{B}^m (\mathbf{A}^n)^T + q_{m,n} \mathbf{B}^m (\mathbf{B}^n)^T \\
&= \mathbf{U}_{m,n} \\
&= \mathbf{U}(m, n)
\end{aligned} \tag{60}$$

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. H. M. Costa, "Writing on dirty paper," *IEEE Trans. Inf. Theory*, vol. IT-29, no. 3, pp. 439–441, May 1983.
- [3] L. Sun and M. Lei, "Quantized CSI-based tomlinson-harashima precoding in multiuser MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 12, no. 3, pp. 1118–1126, Mar. 2013.
- [4] B. M. Hochwald, C. B. Peel, and A. L. Swindlehurst, "A vector-perturbation technique for near-capacity multiantenna multiuser communication—Part II: Perturbation," *IEEE Trans. Commun.*, vol. 53, no. 3, pp. 537–544, Mar. 2005.
- [5] T. Haustein, C. von Helmolt, E. Jorswieck, V. Jungnickel, and V. Pohl, "Performance of MIMO systems with channel inversion," in *Proc. IEEE 55th Veh. Technol. Conf. Veh. Technol. (VTC Spring)*, vol. 1, May 2002, pp. 35–39.
- [6] C. B. Peel, B. M. Hochwald, and A. L. 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.
- [7] N. D. Sidiropoulos, T. N. Davidson, and Z.-Q. Luo, "Transmit beamforming for physical-layer multicasting," *IEEE Trans. Signal Process.*, vol. 54, no. 6, pp. 2239–2251, Jun. 2006.
- [8] A. Wiesel, Y. C. Eldar, and S. Shamai (Shitz), "Linear precoding via conic optimization for fixed MIMO receivers," *IEEE Trans. Signal Process.*, vol. 54, no. 1, pp. 161–176, Jan. 2006.

- [9] F. Wang, X. Wang, and Y. Zhu, "Transmit beamforming for multiuser downlink with per-antenna power constraints," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Sydney, NSW, Australia, Jun. 2014, pp. 4692–4697.
- [10] M. Bengtsson and B. Ottersten, "Optimal and suboptimal transmit beamforming," in *Handbook of Antennas in Wireless Communications*. Boca Raton, FL, USA: CRC Press, Jan. 2001, doi: 10.1201/9781315220031.
- [11] M. Schubert and H. Boche, "Solution of the multiuser downlink beamforming problem with individual SINR constraints," *IEEE Trans. Veh. Technol.*, vol. 53, no. 1, pp. 18–28, Jan. 2004.
- [12] S. S. Christensen, R. Agarwal, E. De Carvalho, and J. M. Cioffi, "Weighted sum-rate maximization using weighted MMSE for MIMO-BC beamforming design," *IEEE Trans. Wireless Commun.*, vol. 7, no. 12, pp. 4792–4799, Dec. 2008.
- [13] C. Masouros, T. Ratnarajah, M. Sellathurai, C. Papadias, and A. 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.
- [14] G. Zheng, I. Krikidis, C. Masouros, S. Timotheou, D.-A. Tzoumpakis, and Z. Ding, "Rethinking the role of interference in wireless networks," *IEEE Commun. Mag.*, vol. 52, no. 11, pp. 152–158, Nov. 2014.
- [15] C. Masouros and E. Alsusa, "A novel transmitter-based selective precoding technique for DS/CDMA systems," *IEEE Signal Process. Lett.*, vol. 14, no. 9, pp. 637–640, Sep. 2007.
- [16] R. D. Miguel and R. R. Müller, "On convex vector precoding for multiuser MIMO broadcast channels," *IEEE Trans. Signal Process.*, vol. 57, no. 11, pp. 4497–4508, Nov. 2009.
- [17] C. Masouros and E. Alsusa, "Dynamic linear precoding for the exploitation of known interference in MIMO broadcast systems," *IEEE Trans. Wireless Commun.*, vol. 8, no. 3, pp. 1396–1401, Mar. 2009.
- [18] C. Masouros, "Correlation rotation linear precoding for MIMO broadcast communications," *IEEE Trans. Signal Process.*, vol. 59, no. 1, pp. 252–262, Jan. 2011.
- [19] C. Masouros, M. Sellathurai, and T. Ratnarajah, "Vector perturbation based on symbol scaling for limited feedback MISO downlinks," *IEEE Trans. Signal Process.*, vol. 62, no. 3, pp. 562–571, Feb. 2014.
- [20] M. Alodeh, S. Chatzinotas, and B. Ottersten, "Constructive multiuser interference in symbol level precoding for the MISO downlink channel," *IEEE Trans. Signal Process.*, vol. 63, no. 9, pp. 2239–2252, May 2015.
- [21] C. Masouros and G. Zheng, "Exploiting known interference as green signal power for downlink beamforming optimization," *IEEE Trans. Signal Process.*, vol. 63, no. 14, pp. 3628–3640, Jul. 2015.
- [22] 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.
- [23] M. Alodeh, S. Chatzinotas, and B. Ottersten, "Symbol-level multiuser MISO precoding for multi-level adaptive modulation," *IEEE Trans. Signal Process.*, vol. 16, no. 8, pp. 5511–5524, Aug. 2017.
- [24] 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.
- [25] A. Li, L. Song, B. Vucetic, and Y. Li, "Joint symbol-level precoding and reflecting designs for IRS-enhanced MU-MISO systems," *IEEE Trans. Veh. Tech.*, vol. 9, no. 11, pp. 1937–1941, Nov. 2020.
- [26] A. Li and C. Masouros, "Exploiting constructive mutual coupling in P2P MIMO by analog-digital phase alignment," *IEEE Trans. Wireless Commun.*, vol. 16, no. 3, pp. 1948–1962, Mar. 2017.
- [27] F. Liu, C. Masouros, P. V. Amadori, and H. Sun, "An efficient manifold algorithm for constructive interference based constant envelope precoding," *IEEE Signal Process. Lett.*, vol. 24, no. 10, pp. 1542–1546, Oct. 2017.
- [28] H. Jeddai, A. Mezghani, A. L. Swindlehurst, and J. A. Nossek, "Quantized constant envelope precoding with PSK and QAM signaling," *IEEE Trans. Wireless Commun.*, vol. 17, no. 12, pp. 8022–8034, Dec. 2018.
- [29] 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.
- [30] A. Li, C. Masouros, F. Liu, and A. Lee 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.
- [31] 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, Feb. 2020.
- [32] 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 Signal Process.*, vol. 12, no. 3, pp. 557–570, Jun. 2018.
- [33] F. Liu, C. Masouros, A. Li, T. Ratnarajah, and J. Zhou, "Interference exploitation for radar and cellular coexistence: The power-efficient approach," *IEEE Trans. Signal Process.*, vol. 66, no. 14, pp. 3681–3695, Jul. 2018.
- [34] 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 Signal Process.*, vol. 15, no. 6, pp. 1316–1331, Nov. 2021.
- [35] Y. Fan, A. Li, X. Liao, and V. C. M. Leung, "Secure interference exploitation precoding in MISO wiretap channel: Destructive region redefinition with efficient solutions," *IEEE Trans. Inf. Forensics Security*, vol. 16, pp. 402–417, 2021.
- [36] Q. Xu, P. Ren, and A. L. Swindlehurst, "Rethinking secure precoding via interference exploitation: A smart eavesdropper perspective," *IEEE Trans. Inf. Forensics Security*, vol. 16, pp. 585–600, 2021.
- [37] A. Li et al., "A tutorial on interference exploitation via symbol-level precoding: Overview, state-of-the-art and future directions," *IEEE Commun. Surveys Tuts.*, vol. 22, no. 2, pp. 796–839, 2nd Quart., 2020.
- [38] M. Alodeh et al., "Symbol-level and multicast precoding for multi-user multi-antenna downlink: A state-of-the-art, classification, and challenges," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 3, pp. 1733–1757, 3rd Quart., 2018.
- [39] 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, Sep. 2018.
- [40] 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.
- [41] A. Haqiqatnejad, F. Kayhan, and B. Ottersten, "Power minimizer symbol-level precoding: A closed-form suboptimal solution," *IEEE Signal Process. Lett.*, vol. 25, no. 11, pp. 1730–1734, Nov. 2018.
- [42] Y. Liu and W.-K. Ma, "Symbol-level precoding is symbol-perturbed ZF when energy efficiency is sought," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Apr. 2018, pp. 3869–3873.
- [43] A. Mohammad, C. Masouros, and Y. Andreopoulos, "An unsupervised learning-based approach for symbol-level precoding," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2021, pp. 1–6.
- [44] A. Mohammad, C. Masouros, and Y. Andreopoulos, "An unsupervised deep unfolding framework for robust symbol level precoding," 2021, *arXiv:2111.08129*.
- [45] A. Mohammad, C. Masouros, and Y. Andreopoulos, "A memory-efficient learning framework for symbol level precoding with quantized NN weights," 2021, *arXiv:2110.06542*.
- [46] Y. Liu, M. Shao, W.-K. Ma, and Q. Li, "Symbol-level precoding through the lens of zero forcing and vector perturbation," *IEEE Trans. Signal Process.*, vol. 70, pp. 1687–1703, 2022.
- [47] L. Vandenberghe and S. Boyd, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.
- [48] P. Wolfe, "The simplex method for quadratic programming," *Econometrica*, vol. 27, no. 3, pp. 382–398, Jul. 1959.
- [49] G. Cornuejols and R. Tutuncu, *Optimization Methods in Finance*. Cambridge, U.K.: Cambridge Univ. Press, Dec. 2006.
- [50] F. Alizadeh and D. Goldfarb, "Second-order cone programming," *Math. Program.*, vol. 95, no. 1, pp. 3–51, 2003.
- [51] A. Li, F. Liu, X. Liao, Y. Shen, and C. Masouros, "Symbol-level precoding made practical for multi-level modulations via block-level rescaling," in *Proc. IEEE 22nd Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, Sep. 2021, pp. 71–75.
- [52] *Technical Specification Group Services and System Aspects*, Release 16, 3GPP document TR 21.916, 2022, Version 16.1.0.