Cooperative Multi-Cell Block Diagonalization with Per-Base-Station Power Constraints

Rui Zhang

Abstract

Block diagonalization (BD) is a practically favorable precoding technique that eliminates the interuser interference in downlink multiuser multiple-input multiple-output (MIMO) systems. In this paper, we apply BD to the downlink transmission in a cooperative multi-cell system, where the signals from different base stations (BSs) to all the mobile stations (MSs) are jointly designed with the perfect knowledge of the downlink channels and transmit messages. Specifically, this paper studies the BD precoder design to maximize the weighted sum-rate achievable for all the MSs. The associated optimization problem can be formulated in an auxiliary MIMO broadcast channel (BC) with a set of transmit power constraints equivalent to those for different BSs in the multi-cell system. Based on convex optimization techniques, this paper designs an efficient algorithm to solve this problem, and derives the structure of the corresponding optimal BD precoding matrix. Moreover, for the special case of single-antenna BSs and MSs, it is shown that the proposed solution leads to the optimal zero-forcing beamforming (ZF-BF) precoder design for the multiple-input single-output (MISO) BC with the per-antenna power constraints.

Index Terms

Block diagonalization, convex optimization, cooperative multi-cell system, MIMO broadcast channel, per-antenna power constraint, per-base-station power constraint, zero-forcing beamforming.

I. Introduction

The study of downlink beamforming and power control in cellular systems has been an active area of research for many years. Conventionally, most of the related works have focused on a single-cell setup, where the co-channel interferences experienced by the mobile stations (MSs) in a particular cell caused by the base stations (BSs) of the other cells are treated as additional noises at the receivers. For this setup, the downlink transmission in a single cell with a multi-antenna BS and multiple single-antenna/multi-antenna MSs can be modeled by a multiple-input single-output/multiple-output (MISO/MIMO) broadcast channel (BC). It is known that the dirty paper coding (DPC) technique achieves the capacity region for the Gaussian MISO/MIMO BC, which constitutes all the simultaneously achievable rates for all the MSs [1]. However, DPC requires complicated nonlinear encoding and decoding schemes and is thus difficult to implement in real-time systems. Consequently, linear transmit and receive beamforming

R. Zhang is with the Institute for Infocomm Research, A*STAR, Singapore (e-mail:rzhang@i2r.a-star.edu.sg).

schemes for the Gaussian MISO/MIMO BC have drawn a lot of attentions in the literature [2], [3], [4], [5], [6]. In particular, a simple linear precoding scheme for the MIMO BC is known as block diagonalization (BD) [7], [8], [9], [10]. With BD, the transmitted signal from the BS intended for each MS is multiplied by a precoding matrix, which is restricted to be orthogonal to the downlink channels associated with all the other MSs. Thereby, all the interuser interferences are eliminated and each MS perceives an interference-free MIMO channel. In the special case of MISO BC, BD reduces to the well-known zero-forcing beamforming (ZF-BF). Although BD is in general inferior in terms of achievable rates as compared with the DPC-based optimal nonlinear precoding scheme as well as the minimum-mean-squared-error (MMSE)-based optimal linear precoding scheme, it performs very well in the high signal-to-noise-ratio (SNR) regime and achieves the same degrees of freedom (DoF) for the MISO-/MIMO-BC sum-rate as the optimal precoding schemes [11]. Moreover, BD can be generalized to incorporate nonlinear DPC processing, which leads to a precoding scheme known as ZF-DPC [11].

Recently, there has been a rapidly growing interest in shifting the design paradigm from the conventional single-cell downlink transmission to the multi-cell cooperative downlink transmission [12], [13], [14], [15], [16]. In these works, it is assumed that multiple BSs in a cellular network are connected to a central processing unit (which can be a dedicated control station or a preselected BS), which has the knowledge of the transmit messages for all the MSs and the channels from different BSs to all the MSs. Thereby, the central processing unit is able to jointly design the downlink transmissions for all the BSs and provide them with the appropriate signals to transmit. As demonstrated in these works, by utilizing the co-channel interference across different cells for coherent transmission, the cooperative multi-cell downlink processing leads to enormous throughput gains as compared with the conventional single-cell processing with the co-channel interference treated as additional noise. Moreover, design of distributed multi-cell downlink beamforming with the use of belief prorogation and message passing between BSs has been studied in [17], without a central processing unit.

In this work, we will focus on the downlink transmit optimization in a fully cooperative multi-cell system assisted by a central processing unit equipped with all the required downlink channel and message knowledge. For this setup, the associated optimization problem can be formulated in an auxiliary MISO/MIMO BC with the number of transmitting antennas (receiving MSs) equal to the sum of those across all the cooperative BSs (cells). However, instead of adopting the conventional sum-power constraint for the auxiliary MISO/MIMO BC, we need to apply a set of transmit power constraints equivalent to those for individual BSs in the multi-cell system. Since most prior works on MISO-/MIMO-BC transmit optimization have dealt with the sum-power constraint, it remains unclear whether the developed results therein can be applied to the case with the equivalent per-BS power constraints. In this paper, we study the BD precoder design optimization for the cooperative multi-cell system to maximize the weighted sum-rate for all the MSs subject to the per-BS power constraints. To the author's best knowledge, this problem has not been completely solved yet in the literature.

The computation problem for the achievable rate region of the Gaussian MISO/MIMO BC subject to per-antenna power constraints has been studied in [18]. This work has been generalized in [19] to deal with arbitrary numbers of linear transmit power constraints. The results in [18], [19] can be directly applied in a cooperative multi-cell system to handle the per-BS power constraints if the DPC-based optimal nonlinear precoder or the MMSE-based optimal linear precoder is applied. On the other hand, the ZF-BF precoding technique has been studied in [20], [21], [22] for the MISO BC subject to per-antenna power constraints. In [20], the ZF-BF precoding matrix is taken as the pseudo inverse of the MISO-BC channel matrix and thereby decomposes the MISO BC into parallel interference-free subchannels for different MSs. Then, the power allocation over the subchannels is optimized under the per-antenna power constraints. However, it was pointed out in [21] that although the ZF-BF precoding matrix for the MISO BC based on the channel pseudo inverse is optimal for the sum-power constraint case, it is in general suboptimal for the per-antenna power constraint case. Thus, in [21] the authors proposed to apply the principle of generalized matrix inverse to design the ZF-BF precoding matrix with the per-antenna power constraints. The proposed scheme has been improved in [22] in terms of computational efficiency. However, these schemes cannot be applied to solve the BD precoder optimization for the MIMO BC with the equivalent per-BS power constraints.

In lieu of the above existing results, the main contributions of this paper are summarized as follows:

- We formulate the MIMO-BC transmit optimization problem with the BD precoding and the equivalent per-BS power constraints as a convex optimization problem. Based on convex optimization techniques, we design an efficient algorithm to solve this problem. We also derive the structure of the corresponding optimal transmit covariance matrix for the MIMO BC, which provides insights on how the optimal BD precoder design in the per-BS power constraint case differs from that in the conventional sum-power constraint case.
- For the special case of single-antenna BSs and MSs, we show that the proposed solution leads to the optimal ZF-BF precoder design for the MISO BC with per-antenna power constraints. Due to different problem formulations, it is shown that the developed method differs from those in prior works [21], [22].
- We present a heuristic suboptimal scheme for the studied problem, which bears a lower complexity than the optimal scheme. For this scheme, the conventional BD precoder design for the sum-power constraint case is

jointly deployed with the optimized power allocation subject to the per-BS power constraints. This scheme can be considered as the extension of that given in [20] for the MISO BC with the ZF-BF precoding and per-antenna power constraints to the MIMO BC with the BD precoding and per-BS power constraints.

The rest of this paper is organized as follows. Section II introduces the signal model for the downlink transmission in a cooperative multi-cell system, and presents the problem formulation for the weighted sum-rate maximization with the BD precoding and per-BS power constraints. Section III studies the solution of the formulated problem, and in particular discusses the developed solution for the special case of MISO BC with the per-antenna power constraints. Section IV develops a heuristic scheme for the studied problem. Section V provides numerical examples on the performance of the proposed optimal and suboptimal schemes. Finally, Section VI concludes this paper.

Notations: Scalars are denoted by lower-case letters, vectors denoted by bold-face lower-case letters, and matrices denoted by bold-face upper-case letters. $\boldsymbol{I}$ and $\mathbf{0}$ denote the identity matrix and the all-zero matrix, respectively, with appropriate dimensions. For a square matrix $\boldsymbol{S}$, $\mathrm{Tr}(\boldsymbol{S})$, $|\boldsymbol{S}|$, $\boldsymbol{S}^{-1}$, and $\boldsymbol{S}^{1/2}$ denote the trace, determinant, inverse (if $\boldsymbol{S}$ is full-rank), and square-root of $\boldsymbol{S}$, respectively; and $\boldsymbol{S} \succeq \mathbf{0}$ ($\boldsymbol{S} \preceq \mathbf{0}$) means that $\boldsymbol{S}$ is positive (negative) semi-definite. $\mathrm{Diag}(\boldsymbol{a})$ denotes a diagonal matrix with the main diagonal given by $\boldsymbol{a}$. For a matrix $\boldsymbol{M}$ of arbitrary size, $\boldsymbol{M}^H$, $\boldsymbol{M}^T$, $\mathrm{Rank}(\boldsymbol{M})$, and $\boldsymbol{M}^\dagger$ denote the conjugate transpose, transpose, rank, and pseudo inverse of $\boldsymbol{M}$, respectively. $\mathbb{E}[\cdot]$ denotes the statistical expectation. The distribution of a circular symmetric complex Gaussian (CSCG) random vector with the mean vector $\boldsymbol{x}$ and the covariance matrix $\boldsymbol{\Sigma}$ is denoted by $\mathcal{CN}(\boldsymbol{x}, \boldsymbol{\Sigma})$; and $\sim$ stands for "distributed as". $\mathbb{C}^{x \times y}$ denotes the space of $x \times y$ matrices with complex-valued elements. $\|\boldsymbol{x}\|$ denotes the Euclidean norm of a complex vector $\boldsymbol{x}$; and $|z|$ denotes the norm of a complex number z.

II. System Model and Problem Formulation

We consider a multi-cell system consisting of A cells, each of which has a BS to coordinate the transmission with K_a MSs, $K_a \geq 1$ and $a = 1, \cdots, A$. Denote the total number of MSs in the system as $K = \sum_{a=1}^{A} K_a$. For convenience, we assume that all the BSs are equipped with the same number of antennas, denoted by $M_B \geq 1$. Denote the total number of antennas across all the BSs as $M = M_B A$. We also assume that each of K MSs is equipped with N antennas, $N \geq 1$. Since we are interested in a fully cooperative multi-cell system, the jointly designed downlink transmission for all the BSs can be conveniently modeled as an auxiliary MIMO BC with M transmitting antennas and K MSs each having N receiving antennas. For convenience, we assign the indexes to the transmitting antennas in the auxiliary MIMO BC belonging to different BSs in the multi-cell system according to the BS index, i.e., the $((a-1)M_B + 1)$-th to (aM_B)-th antennas are taken as the M_B antennas from the ath

BS, $a = 1, \cdots, A$. Similarly, the indexes of MSs in the auxiliary MIMO BC are assigned in accordance with their belonged cell indexes in the multi-cell system, i.e., the $(\sum_{i=1}^{a-1} K_i + 1)$-th to $(\sum_{i=1}^{a} K_i)$-th MSs are taken as the K_a MSs from the ath cell, $a = 1, \cdots, A$. Accordingly, the discrete-time baseband signal for the auxiliary MIMO BC is given by

$$\boldsymbol{y}_k = \boldsymbol{H}_k \boldsymbol{x}_k + \sum_{j=1, j \neq k}^{K} \boldsymbol{H}_k \boldsymbol{x}_j + \boldsymbol{z}_k, \quad k = 1, \cdots, K \tag{1}$$

where $\boldsymbol{x}_k \in \mathbb{C}^{M \times 1}$ and $\boldsymbol{y}_k \in \mathbb{C}^{N \times 1}$ denote the transmitted and received signals for the kth MS, respectively; $\boldsymbol{H}_k \in \mathbb{C}^{N \times M}$ denotes the downlink channel from all the M cooperative transmitting antennas to the kth MS; and $\boldsymbol{z}_k \in \mathbb{C}^{N \times 1}$ denotes the received noise at the kth MS. For convenience, we assume that $\boldsymbol{z}_k \sim \mathcal{CN}(\boldsymbol{0}, \boldsymbol{I}), \forall k$.

We assume that $\boldsymbol{x}_k$'s are independent over k. It is further assumed that the Gaussian codebook is used at the transmitter and thus $\boldsymbol{x}_k \sim \mathcal{CN}(\boldsymbol{0}, \boldsymbol{S}_k), \forall k$, where $\boldsymbol{S}_k = \mathbb{E}[\boldsymbol{x}_k \boldsymbol{x}_k^H]$ denotes the transmit covariance matrix for the kth MS, $\boldsymbol{S}_k \in \mathbb{C}^{M \times M}$ and $\boldsymbol{S}_k \succeq \boldsymbol{0}$. The overall downlink transmit covariance matrix for the M cooperative transmitting antennas is then obtained as $\boldsymbol{S} = \sum_{k=1}^{K} \boldsymbol{S}_k$. Since these transmitting antennas come from A different BSs, they need to satisfy a set of per-BS power constraints expressed as

$$\mathrm{Tr}\left(\boldsymbol{B}_a \boldsymbol{S}\right) \leq P \text{ or } \sum_{k=1}^{K} \mathrm{Tr}\left(\boldsymbol{B}_a \boldsymbol{S}_k\right) \leq P, \quad a = 1, \cdots, A \tag{2}$$

where

$$\boldsymbol{B}_a \triangleq \mathrm{Diag}\Bigg(\underbrace{0, \cdots, 0}_{(a-1)M_B}, \underbrace{1, \cdots, 1}_{M_B}, \underbrace{0, \cdots, 0}_{(A-a)M_B}\Bigg) \tag{3}$$

and P denotes the per-BS power constraint, which is assumed identical for all the BSs. Note that in the special case of single-antenna BSs and MSs, i.e., $M_B = N = 1$, the per-BS power constraints in (2) reduce to the per-antenna power constraints in an equivalent MISO BC.

Next, consider the BD precoding scheme in the auxiliary MIMO BC, which eliminates the interuser interference, i.e., in (1) we have $\boldsymbol{H}_k \boldsymbol{x}_j = \boldsymbol{0}, \forall j \neq k$. It is easy to show that the above constraints for $k = 1, \cdots, K$ are equivalent to the following set of "ZF constraints":

$$\boldsymbol{H}_j \boldsymbol{S}_k \boldsymbol{H}_j^H = \boldsymbol{0}, \quad \forall j \neq k. \tag{4}$$

Assuming that the row vectors in all $\boldsymbol{H}_k$'s are linearly independent, from the constraints in (4) it follows that $NK \leq M$ needs to be true in order to have a set of feasible $\boldsymbol{S}_k$'s with $\mathrm{Rank}(\boldsymbol{S}_k) = N, \forall k$, i.e., all the MSs receive the same number of data streams equal to N. In practice, the total number of MSs in the system can be very large such that the above condition is not satisfied. In such scenarios, the transmissions to MSs can be scheduled into different time-slots, where in each time-slot the number of MSs scheduled for transmission satisfies the above condition. The interested readers may refer to [23], [24] for the detailed design of downlink transmission scheduling in the MISO/MIMO BC with the ZF-BF/BD precoding. For the rest of this paper, we assume that $NK \leq M$ for simplicity.

We are now ready to present the weighted sum-rate maximization problem for the downlink transmission in a cooperative multi-cell system with the BD precoding and per-BS power constraints as follows (P1):

$$\begin{aligned}
\max_{\boldsymbol{S}_1,\cdots,\boldsymbol{S}_K} \quad & \sum_{k=1}^{K} w_k \log\left|\boldsymbol{I} + \boldsymbol{H}_k \boldsymbol{S}_k \boldsymbol{H}_k^H\right| \\
\text{s.t.} \quad & \boldsymbol{H}_j \boldsymbol{S}_k \boldsymbol{H}_j^H = 0, \ \forall j \neq k \\
& \sum_{k=1}^{K} \operatorname{Tr}\left(\boldsymbol{B}_a \boldsymbol{S}_k\right) \leq P, \ \forall a \\
& \boldsymbol{S}_k \succeq \boldsymbol{0}, \ \forall k
\end{aligned}$$

where w_k is the given non-negative rate weight for the kth MS. For the purpose of exposition, we assume that $w_k > 0, \forall k$. It is easy to verify that (P1) is a convex optimization problem, since the objective function is concave over $\boldsymbol{S}_k$'s and all the constraints specify a convex set over $\boldsymbol{S}_k$'s. Thus, (P1) can be solved using standard convex optimization techniques, e.g., the interior-point method [25]. However, such an approach does not reveal any insight on the structure of the optimal downlink transmit covariance matrix $\boldsymbol{S}_k$'s with the BD precoding. Therefore, in this paper we take a different approach to solve (P1). This approach is based on the Lagrange duality [25] for convex optimization problems, and leads to the optimal BD precoder structure as will be shown in the next section.

Remark 2.1: It is noted that (P1) can be modified to incorporate additional per-antenna power constraints at all the BSs. Let $P^{(\mathrm{pa})}$ denote the given per-antenna power threshold. Then, a set of M per-antenna power constraints can be added into (P1) as

$$\sum_{k=1}^{K} \operatorname{Tr}\left(\boldsymbol{B}_i^{(\mathrm{pa})} \boldsymbol{S}_k\right) \leq P^{(\mathrm{pa})}, \quad i = 1, \cdots, M \tag{5}$$

where $\boldsymbol{B}_i^{(\mathrm{pa})}$ is a diagonal matrix with the ith diagonal element equal to one and all the others equal to zero. Since the resulting optimization problem bears the same structure as (P1), they can be similarly solved.

Remark 2.2: It is noted that (P1) can be modified to solve the weighted sum-rate maximization problem for the cooperative multi-cell downlink transmission with the ZF-DPC precoding [11]. With ZF-DPC, given a fixed encoding order for the transmitted signals to different MSs (without loss of generality, we assume that the encoding order is given by the MS index), the signal for a later encoded MS is designed with the non-causal knowledge of all the earlier encoded MS signals, of which the associated interferences can be precanceled by DPC. By extending the ZF-DPC scheme in [11] for the MISO BC to the case of MIMO BC, (P1) can be modified to obtain the optimal ZF-DPC precoder design by rewriting the set of ZF constraints in (P1) as

$$\boldsymbol{H}_j\boldsymbol{S}_k\boldsymbol{H}_j^H = 0, \quad \forall j > k. \tag{6}$$

Clearly, the resulting problem is of a similar structure as (P1) and thus can be solved similarly.

III. Proposed Solution

In this section, we first present a new method to solve (P1), which reveals the structure of the optimal BD precoding matrix for the general case with arbitrary numbers of antennas at the BS or MS. Then, we investigate the developed solution for the special case of single-antenna BSs and MSs (i.e., an equivalent MISO BC with the per-antenna power constraints), in comparison with other existing solutions in the literature [21], [22].

A. General Case

To solve (P1), as a first step, it is desirable to remove the set of ZF constraints, as discussed as follows. Define $\boldsymbol{G}_k = [\boldsymbol{H}_1^T, \cdots, \boldsymbol{H}_{k-1}^T, \boldsymbol{H}_{k+1}^T, \cdots, \boldsymbol{H}_k^T]^T, k = 1, \cdots, K$, where $\boldsymbol{G}_k \in \mathbb{C}^{L\times M}$ with $L = N(K-1)$. Let the singular value decomposition (SVD) of $\boldsymbol{G}_k$ be denoted as $\boldsymbol{G}_k = \boldsymbol{U}_k\boldsymbol{\Sigma}_k\boldsymbol{V}_k^H$, where $\boldsymbol{U}_k \in \mathbb{C}^{L\times L}$, $\boldsymbol{V}_k \in \mathbb{C}^{M\times L}$, and $\boldsymbol{\Sigma}_k$ is a $L\times L$ positive diagonal matrix. Note that $\mathrm{Rank}(\boldsymbol{G}_k) = L < M$ under the previous assumption that $NK \leq M$. Define the projection matrix $\boldsymbol{P}_k = (\boldsymbol{I} - \boldsymbol{V}_k\boldsymbol{V}_k^H)$. Without loss of generality, we can express $\boldsymbol{P}_k = \tilde{\boldsymbol{V}}_k\tilde{\boldsymbol{V}}_k^H$, where $\tilde{\boldsymbol{V}}_k \in \mathbb{C}^{M\times(M-L)}$ satisfies that $\boldsymbol{V}_k^H\tilde{\boldsymbol{V}}_k = \boldsymbol{0}$ and $\tilde{\boldsymbol{V}}_k^H\tilde{\boldsymbol{V}}_k = \boldsymbol{I}$. Note that $[\boldsymbol{V}_k, \tilde{\boldsymbol{V}}_k]$ forms a $M \times M$ unitary matrix. Then, we have the following lemma:

Lemma 3.1: The optimal solution of (P1) satisfies the following structure:

$$\boldsymbol{S}_k = \tilde{\boldsymbol{V}}_k\boldsymbol{Q}_k\tilde{\boldsymbol{V}}_k^H, \ k = 1, \cdots, K \tag{7}$$

where $\boldsymbol{Q}_k \in \mathbb{C}^{(M-L)\times(M-L)}$ and $\boldsymbol{Q}_k \succeq \boldsymbol{0}$.

Proof: Please refer to Appendix I. ■

Remark 3.1: In prior works [7], [8], [9], [10] on the design of BD precoder for the MIMO BC, it has been observed that the columns in the BD precoding matrix for the kth MS, $\boldsymbol{T}_k$, where $\boldsymbol{S}_k = \boldsymbol{T}_k\boldsymbol{T}_k^H$, should be linear combinations of those in $\tilde{\boldsymbol{V}}_k$ in order to satisfy the constraints: $\boldsymbol{H}_j\boldsymbol{T}_k = \boldsymbol{0}, \forall j \neq k$. Lemma 3.1 crystalizes this observation by providing an explicit structure of the corresponding optimal transmit covariance matrix $\boldsymbol{S}_k$.

With the optimal structures for $\boldsymbol{S}_k$'s given in Lemma 3.1, it can be verified that all the ZF constraints in (P1) are satisfied and thus can be removed. Thus, (P1) reduces to the following equivalent problem (P2):

$$\begin{aligned}\max_{\boldsymbol{Q}_1,\cdots,\boldsymbol{Q}_K} \quad & \sum_{k=1}^{K} w_k \log\left|\boldsymbol{I}+\boldsymbol{H}_k\tilde{\boldsymbol{V}}_k\boldsymbol{Q}_k\tilde{\boldsymbol{V}}_k^H\boldsymbol{H}_k^H\right| \\ \text{s.t.} \quad & \sum_{k=1}^{K} \mathrm{Tr}\left(\boldsymbol{B}_a\tilde{\boldsymbol{V}}_k\boldsymbol{Q}_k\tilde{\boldsymbol{V}}_k^H\right) \leq P, \ \forall a \\ & \boldsymbol{Q}_k \succeq \boldsymbol{0}, \ \forall k.\end{aligned}$$

Similar to (P1), it can be shown that (P2) is convex. Thus, (P2) is solvable by the Lagrange duality method as follows. By introducing a set of non-negative dual variables, $\mu_a, a = 1, \cdots, A$, associated with the set of per-BS power constraints in (P2), the Lagrangian function of (P2) can be written as

$$L(\{\boldsymbol{Q}_k\},\{\mu_a\}) = \sum_{k=1}^{K} w_k \log\left|\boldsymbol{I}+\boldsymbol{H}_k\tilde{\boldsymbol{V}}_k\boldsymbol{Q}_k\tilde{\boldsymbol{V}}_k^H\boldsymbol{H}_k^H\right| - \sum_{a=1}^{A}\mu_a\left(\sum_{k=1}^{K}\mathrm{Tr}\left(\boldsymbol{B}_a\tilde{\boldsymbol{V}}_k\boldsymbol{Q}_k\tilde{\boldsymbol{V}}_k^H\right) - P\right) \tag{8}$$

where $\{\boldsymbol{Q}_k\}$ and $\{\mu_a\}$ denote the set of $\boldsymbol{Q}_k$'s and the set of μ_a's, respectively. The Lagrange dual function for (P2) is then defined as

$$g(\{\mu_a\}) = \max_{\boldsymbol{Q}_k\succeq\boldsymbol{0},\forall k} L(\{\boldsymbol{Q}_k\},\{\mu_a\}). \tag{9}$$

Moreover, the dual problem of (P2) is defined as (P2-D):

$$\min_{\mu_a\geq 0,\forall a} \ g(\{\mu_a\}).$$

Since (P2) is convex and satisfies the Slater's condition [25], the duality gap between the optimal objective value of (P2) and that of (P2-D) is zero. Thus, (P2) can be solved equivalently by solving (P2-D). Moreover, (P2-D) is convex and can be solved by the subgradient-based method, e.g., the ellipsoid method [26], given the fact that the subgradient of function $g(\{\mu_a\})$ at a fixed set of μ_a's is $P - \sum_{k=1}^{K}\mathrm{Tr}\left(\boldsymbol{B}_a\tilde{\boldsymbol{V}}_k\boldsymbol{Q}_k^{\star}\tilde{\boldsymbol{V}}_k^H\right)$ for μ_a, $a = 1, \cdots, A$, where $\{\boldsymbol{Q}_k^{\star}\}$ is the optimal solution for the maximization problem in (9) with the given set of μ_a's.

Next, we focus on solving for $\{\boldsymbol{Q}_k^{\star}\}$ with a fixed set of μ_a's. It is important to observe that the maximization problem in (9) can be separated into K independent subproblems each involving only one of $\boldsymbol{Q}_k$'s. All of these subproblems have the same structure and thus can be solved by the same method. Specifically, for a given k, the corresponding subproblem can be expressed as (P3) (by discarding the irrelevant terms):

$$\max_{\boldsymbol{Q}_k\succeq\boldsymbol{0}} w_k \log\left|\boldsymbol{I}+\boldsymbol{H}_k\tilde{\boldsymbol{V}}_k\boldsymbol{Q}_k\tilde{\boldsymbol{V}}_k^H\boldsymbol{H}_k^H\right| - \mathrm{Tr}\left(\boldsymbol{B}_\mu\tilde{\boldsymbol{V}}_k\boldsymbol{Q}_k\tilde{\boldsymbol{V}}_k^H\right)$$

where $\boldsymbol{B}_\mu \triangleq \sum_{a=1}^{A}\mu_a\boldsymbol{B}_a$. Note that $\boldsymbol{B}_\mu$ is a diagonal matrix with the diagonal elements given by different μ_a's in the order of $a = 1, \cdots, A$. We then have the following lemma:

Lemma 3.2: Let A_μ denote the number of μ_a's in the main diagonal of $\boldsymbol{B}_\mu$, $a \in \{1, \cdots, A\}$, with $\mu_a > 0$. Then, for (P3) to have a bounded objective value, it needs that $A_\mu \geq \lceil \frac{M-N(K-1)}{M_B} \rceil$.

Proof: Please refer to Appendix II. ■

Remark 3.2: It is noted that by applying the Karash-Kuhn-Tucker (KKT) conditions [25] to (P2), the fact that $\mu_a > 0$ for a given $a \in \{1, \cdots, A\}$ implies that the corresponding power constraint must be tight with the optimal solution for $\{\boldsymbol{Q}_k\}$. Accordingly, in (P1) the optimal downlink transmit covariance matrix $\boldsymbol{S}_k$'s must make the ath per-BS power constraint tight. Therefore, Lemma 3.2 provides a lower bound on the number of BSs for which the corresponding transmit power constraints must be tight with the optimal $\boldsymbol{S}_k$'s for (P1).

With Lemma 3.2 and $L = N(K-1)$, we can assume without loss of generality that $M_B A_\mu \geq (M - L)$ since we are only interested in the case where the objective value of (P3) or that of (P1) is bounded. Accordingly, we have $\mathrm{Rank}(\tilde{\boldsymbol{V}}_k^H \boldsymbol{B}_\mu \tilde{\boldsymbol{V}}_k) = \min(M_B A_\mu, M - L) = M - L$. Thus, $\tilde{\boldsymbol{V}}_k^H \boldsymbol{B}_\mu \tilde{\boldsymbol{V}}_k \in \mathbb{C}^{(M-L)\times(M-L)}$ is a full-rank matrix and its inverse exists. Moreover, since $\mathrm{Tr}(\boldsymbol{XY}) = \mathrm{Tr}(\boldsymbol{YX})$, in (P3) we have $\mathrm{Tr}(\boldsymbol{B}_\mu \tilde{\boldsymbol{V}}_k \boldsymbol{Q}_k \tilde{\boldsymbol{V}}_k^H) = \mathrm{Tr}((\tilde{\boldsymbol{V}}_k^H \boldsymbol{B}_\mu \tilde{\boldsymbol{V}}_k)^{1/2} \boldsymbol{Q}_k (\tilde{\boldsymbol{V}}_k^H \boldsymbol{B}_\mu \tilde{\boldsymbol{V}}_k)^{1/2})$. We thus define

$$\tilde{\boldsymbol{Q}}_k = (\tilde{\boldsymbol{V}}_k^H \boldsymbol{B}_\mu \tilde{\boldsymbol{V}}_k)^{1/2} \boldsymbol{Q}_k (\tilde{\boldsymbol{V}}_k^H \boldsymbol{B}_\mu \tilde{\boldsymbol{V}}_k)^{1/2}. \tag{10}$$

Then, (P3) can be reformulated to maximize

$$w_k \log \left| \boldsymbol{I} + \boldsymbol{H}_k \tilde{\boldsymbol{V}}_k (\tilde{\boldsymbol{V}}_k^H \boldsymbol{B}_\mu \tilde{\boldsymbol{V}}_k)^{-1/2} \tilde{\boldsymbol{Q}}_k (\tilde{\boldsymbol{V}}_k^H \boldsymbol{B}_\mu \tilde{\boldsymbol{V}}_k)^{-1/2} \tilde{\boldsymbol{V}}_k^H \boldsymbol{H}_k^H \right| - \mathrm{Tr}\left(\tilde{\boldsymbol{Q}}_k\right) \tag{11}$$

subject to $\tilde{\boldsymbol{Q}}_k \succeq \boldsymbol{0}$. Note that $\mathrm{Rank}(\boldsymbol{H}_k \tilde{\boldsymbol{V}}_k (\tilde{\boldsymbol{V}}_k^H \boldsymbol{B}_\mu \tilde{\boldsymbol{V}}_k)^{-1/2}) = \min(N, M - L) = N$. Thus, the following SVD can be obtained as

$$\boldsymbol{H}_k \tilde{\boldsymbol{V}}_k (\tilde{\boldsymbol{V}}_k^H \boldsymbol{B}_\mu \tilde{\boldsymbol{V}}_k)^{-1/2} = \hat{\boldsymbol{U}}_k \hat{\boldsymbol{\Sigma}}_k \hat{\boldsymbol{V}}_k^H \tag{12}$$

where $\hat{\boldsymbol{U}}_k \in \mathbb{C}^{N\times N}$, $\hat{\boldsymbol{V}}_k \in \mathbb{C}^{(M-L)\times N}$, and $\hat{\boldsymbol{\Sigma}}_k = \mathrm{Diag}(\hat{\sigma}_{k,1}, \cdots, \hat{\sigma}_{k,N})$. Substituting the above SVD into (11) and applying the Hadamard's inequality (see, e.g., [27]) yields the following optimal solution for (11) as $\tilde{\boldsymbol{Q}}_k^\star = \hat{\boldsymbol{V}}_k \boldsymbol{\Lambda}_k \hat{\boldsymbol{V}}_k^H$, where $\boldsymbol{\Lambda}_k = \mathrm{Diag}(\lambda_{k,1}, \cdots, \lambda_{k,N})$, where $\lambda_{k,i}$, $i = 1, \cdots, N$, can be obtained by the standard water-filling algorithm [27] as

$$\lambda_{k,i} = \left(w_k - \frac{1}{\hat{\sigma}_{k,i}^2} \right)^+ \tag{13}$$

where $(x)^+ \triangleq \max(0, x)$. To summarize, the optimal solution of (P3) for a given set of μ_a's can be expressed as

$$\boldsymbol{Q}_k^\star = (\tilde{\boldsymbol{V}}_k^H \boldsymbol{B}_\mu \tilde{\boldsymbol{V}}_k)^{-1/2} \hat{\boldsymbol{V}}_k \boldsymbol{\Lambda}_k \hat{\boldsymbol{V}}_k^H (\tilde{\boldsymbol{V}}_k^H \boldsymbol{B}_\mu \tilde{\boldsymbol{V}}_k)^{-1/2}, \ k = 1, \cdots, K. \tag{14}$$

Note that when the optimal solution for $\{\mu_a\}$ in (P2-D) is obtained, the corresponding solution in (14) becomes optimal for (P2). By combining this result with Lemma 3.1, we obtain the following theorem:

Theorem 3.1: The optimal solution of (P1) has the following structure:

$$\boldsymbol{S}_k^{\star} = \tilde{\boldsymbol{V}}_k(\tilde{\boldsymbol{V}}_k^H\boldsymbol{B}_\mu\tilde{\boldsymbol{V}}_k)^{-1/2}\hat{\boldsymbol{V}}_k\boldsymbol{\Lambda}_k\hat{\boldsymbol{V}}_k^H(\tilde{\boldsymbol{V}}_k^H\boldsymbol{B}_\mu\tilde{\boldsymbol{V}}_k)^{-1/2}\tilde{\boldsymbol{V}}_k^H,\ k=1,\cdots,K. \tag{15}$$

Next, the following discussions are made regarding the optimal solution of (P1) given in Theorem 3.1:

- **Channel diagonalization:** For the convenience of transmitter design, a common practice is to extract the precoding matrix from the transmit covariance matrix. Let $\boldsymbol{S}_k^{\star} = \boldsymbol{T}_k^{\star}(\boldsymbol{T}_k^{\star})^H$, where $\boldsymbol{T}_k^{\star}$ denotes the optimal precoding matrix for the transmitted signal to the kth MS. Then, from (15) one possible choice of $\boldsymbol{T}_k^{\star}$ is

$$\boldsymbol{T}_k^{\star} = \tilde{\boldsymbol{V}}_k(\tilde{\boldsymbol{V}}_k^H\boldsymbol{B}_\mu\tilde{\boldsymbol{V}}_k)^{-1/2}\hat{\boldsymbol{V}}_k\boldsymbol{\Lambda}_k^{1/2}. \tag{16}$$

  A desirable property of linear precoding for a point-to-point MIMO channel is that the precoding matrix, when jointly deployed with a unitary decoding matrix at the receiver, is able to diagonalize the MIMO channel into parallel scalar channels, over which independent encoding and decoding can be applied to simplify the transceiver design. Therefore, it is important to check whether $\boldsymbol{T}_k^{\star}$ given in (16) satisfies this "channel diagonalization" property as follows:

$$\boldsymbol{H}_k\boldsymbol{T}_k^{\star} = \boldsymbol{H}_k\tilde{\boldsymbol{V}}_k(\tilde{\boldsymbol{V}}_k^H\boldsymbol{B}_\mu\tilde{\boldsymbol{V}}_k)^{-1/2}\hat{\boldsymbol{V}}_k\boldsymbol{\Lambda}_k^{1/2} \tag{17}$$

$$= \hat{\boldsymbol{U}}_k\hat{\boldsymbol{\Sigma}}_k\hat{\boldsymbol{V}}_k^H\hat{\boldsymbol{V}}_k\boldsymbol{\Lambda}_k^{1/2} \tag{18}$$

$$= \hat{\boldsymbol{U}}_k\hat{\boldsymbol{\Sigma}}_k\boldsymbol{\Lambda}_k^{1/2} \tag{19}$$

  where (18) is due to (12). Therefore, with a unitary decoding matrix $\hat{\boldsymbol{U}}_k^H$ applied at the kth MS receiver, the MIMO channel for the kth MS is diagonalized into N scalar channels with channel gains given by the main diagonal of the diagonal matrix, $\hat{\boldsymbol{\Sigma}}_k\boldsymbol{\Lambda}_k^{1/2}$. Furthermore, it can be easily verified that such channel diagonalization is capacity-lossless.

- **Comparison with the sum-power constraint case:** It is noted that (P1) can be modified to deal with the case where a single sum-power constraint over all the BSs (instead of a set of per-BS power constraints) is applied. This can be done via replacing the set of per-BS power constraints in (P1) by

$$\sum_{k=1}^{K}\mathrm{Tr}(\boldsymbol{S}_k) \le P^{(\mathrm{sum})} \tag{20}$$

  where $P^{(\mathrm{sum})}$ denotes the given sum-power constraint. Note also that (P1) in this case corresponds to the conventional BD precoder design problem for the MIMO BC with a sum-power constraint [7], [8], [9], [10]. It can be shown that the developed solution for (P1) can be applied to this case, while the corresponding matrix $\boldsymbol{B}_\mu$ should be modified as $\mu\boldsymbol{I}$ with μ denoting the dual variable associated with the sum-power constraint in (20). From (15), it follows that the optimal solution for this modified problem of (P1) is

$$\boldsymbol{S}_k^{\star\star} = \frac{1}{\mu}\tilde{\boldsymbol{V}}_k\hat{\boldsymbol{V}}_k\boldsymbol{\Lambda}_k\hat{\boldsymbol{V}}_k^H\tilde{\boldsymbol{V}}_k^H, \ k = 1, \cdots, K. \tag{21}$$

Moreover, from (12) with $\boldsymbol{B}_\mu = \mu\boldsymbol{I}$, it follows that $\hat{\boldsymbol{V}}_k$ is obtained from the SVD: $\frac{1}{\sqrt{\mu}}\boldsymbol{H}_k\tilde{\boldsymbol{V}}_k = \hat{\boldsymbol{U}}_k\hat{\boldsymbol{\Sigma}}_k\hat{\boldsymbol{V}}_k^H$ and is thus independent of μ. Accordingly, the optimal precoding matrix in the sum-power constraint case is $\boldsymbol{T}_k^{\star\star} = \frac{1}{\sqrt{\mu}}\tilde{\boldsymbol{V}}_k\hat{\boldsymbol{V}}_k\boldsymbol{\Lambda}_k^{1/2}$. Comparing $\boldsymbol{T}_k^{\star\star}$ with $\boldsymbol{T}_k^{\star}$ in (16) for the per-BS power constraint case, we see that $\boldsymbol{T}_k^{\star\star}$ consists of orthogonal columns (beamforming vectors) since $\hat{\boldsymbol{V}}_k^H\tilde{\boldsymbol{V}}_k^H\tilde{\boldsymbol{V}}_k\hat{\boldsymbol{V}}_k = \boldsymbol{I}$, while $\boldsymbol{T}_k^{\star}$ in general consists of non-orthogonal columns if $\boldsymbol{B}_\mu$ is a non-identity diagonal matrix (i.e., the optimal μ_a's are not all equal).

At last, the algorithm for solving (P1) is summarized as follows (A1):

- **Initialize** $\mu_a \geq 0, a = 1, \cdots, A$.
- **Repeat**
  1. Solve $\boldsymbol{Q}_k^\star, k = 1, \cdots, K$ using (14) with the given μ_a's;
  2. Compute the subgradient of $g(\{\mu_a\})$ as $P - \sum_{k=1}^K \mathrm{Tr}\left(\boldsymbol{B}_a\tilde{\boldsymbol{V}}_k\boldsymbol{Q}_k^\star\tilde{\boldsymbol{V}}_k^H\right), a = 1, \cdots, A$, and update μ_a's accordingly based on the ellipsoid method [26];
- **Until** all the μ_a's converge within a prescribed accuracy.
- **Set** $\boldsymbol{S}_k^\star = \tilde{\boldsymbol{V}}_k\boldsymbol{Q}_k^\star\tilde{\boldsymbol{V}}_k^H, k = 1, \cdots, K$.

B. Special Case: MISO BC with Per-Antenna Power Constraints

In this subsection, we investigate further the developed solution for the special case of $M_B = N = 1$, where the auxiliary MIMO BC with the per-BS power constraints reduces to an equivalent MISO BC with the corresponding per-antenna power constraints. With $N = 1$, $\boldsymbol{H}_k$ degrades to a row-vector $\boldsymbol{h}_k^H$, where $\boldsymbol{h}_k \in \mathbb{C}^{M\times 1}, k = 1, \cdots, K$. Accordingly, the SVD in (12) is rewritten as

$$\boldsymbol{h}_k^H\tilde{\boldsymbol{V}}_k(\tilde{\boldsymbol{V}}_k^H\boldsymbol{B}_\mu\tilde{\boldsymbol{V}}_k)^{-1/2} = \hat{\sigma}_k\hat{\boldsymbol{v}}_k^H \tag{22}$$

where $\hat{\sigma}_k > 0$ and $\hat{\boldsymbol{v}}_k \in \mathbb{C}^{(M-L)\times 1}$. From (13), (15), and (22), it follows that the optimal downlink transmit covariance matrix for the kth MS, $\boldsymbol{S}_k^\star$, in the case of $N = 1$ is expressed as

$$\boldsymbol{S}_k^\star = \lambda_k \tilde{\boldsymbol{V}}_k (\tilde{\boldsymbol{V}}_k^H \boldsymbol{B}_\mu \tilde{\boldsymbol{V}}_k)^{-1/2} \hat{\boldsymbol{v}}_k \hat{\boldsymbol{v}}_k^H (\tilde{\boldsymbol{V}}_k^H \boldsymbol{B}_\mu \tilde{\boldsymbol{V}}_k)^{-1/2} \tilde{\boldsymbol{V}}_k^H \tag{23}$$

$$= \lambda_k \hat{\sigma}_k^{-2} \tilde{\boldsymbol{V}}_k (\tilde{\boldsymbol{V}}_k^H \boldsymbol{B}_\mu \tilde{\boldsymbol{V}}_k)^{-1} \tilde{\boldsymbol{V}}_k^H \boldsymbol{h}_k \boldsymbol{h}_k^H \tilde{\boldsymbol{V}}_k (\tilde{\boldsymbol{V}}_k^H \boldsymbol{B}_\mu \tilde{\boldsymbol{V}}_k)^{-1} \tilde{\boldsymbol{V}}_k^H \tag{24}$$

where $\lambda_k = (w_k - 1/\hat{\sigma}_k^2)^+$. It is thus easy to observe that in this case $\mathrm{Rank}(\boldsymbol{S}_k^\star) \le 1$. Thus, the corresponding optimal precoding matrix reduces to a (beamforming) vector denoted by $\boldsymbol{t}_k^\star \in \mathbb{C}^{M\times 1}$, where $\boldsymbol{S}_k^\star = \boldsymbol{t}_k^\star (\boldsymbol{t}_k^\star)^H$ and

$$\boldsymbol{t}_k^\star = \lambda_k^{1/2} \hat{\sigma}_k^{-1} \tilde{\boldsymbol{V}}_k (\tilde{\boldsymbol{V}}_k^H \boldsymbol{B}_\mu \tilde{\boldsymbol{V}}_k)^{-1} \tilde{\boldsymbol{V}}_k^H \boldsymbol{h}_k. \tag{25}$$

Note that (25) holds regardless of M_B, while $M_B = 1$ corresponds to the per-antenna power constraint case in the equivalent MISO BC. Furthermore, the optimal beamforming vector for the kth MS in the conventional sum-power constraint case (with $\boldsymbol{B}_\mu = \mu \boldsymbol{I}$) is obtained from (25) as

$$\boldsymbol{t}_k^{\star\star} = \lambda_k^{1/2} \hat{\sigma}_k^{-1} \mu^{-1} \tilde{\boldsymbol{V}}_k \tilde{\boldsymbol{V}}_k^H \boldsymbol{h}_k. \tag{26}$$

<u>Remark</u> *3.3:* Denote $\boldsymbol{T} = [\boldsymbol{t}_1, \cdots, \boldsymbol{t}_K] \in \mathbb{C}^{M\times K}$ as the precoding matrix for a MISO BC with M transmitting antennas and K single-antenna receiving MSs. Then, for the sum-power constraint case with $\boldsymbol{t}_k = \boldsymbol{t}_k^{\star\star}$ given in (26), the corresponding optimal precoding matrix $\boldsymbol{T}^{\star\star}$ becomes the conventional ZF-BF design for the MISO BC based on the channel pseudo inverse [4], i.e., $\boldsymbol{T}^{\star\star}$ can be shown in the form of $\boldsymbol{T}^{\star\star} = \boldsymbol{H}^\dagger \hat{\boldsymbol{\Lambda}}$, where $\boldsymbol{H} = [\boldsymbol{h}_1, \cdots, \boldsymbol{h}_K]^H$ and $\hat{\boldsymbol{\Lambda}} = \mathrm{Diag}(\hat{\lambda}_1, \cdots, \hat{\lambda}_K)$, where $\hat{\lambda}_k = \lambda_k^{1/2} \hat{\sigma}_k, k = 1, \cdots, K$. However, it is observed that the ZF-BF design based on the channel pseudo inverse is in general suboptimal for the MISO BC with the per-antenna/per-BS power constraints, where the optimal precoding matrix $\boldsymbol{T}^\star$ is obtained with $\boldsymbol{t}_k = \boldsymbol{t}_k^\star$ given in (25). Note that $\boldsymbol{t}_k^\star$ becomes collinear with $\boldsymbol{t}_k^{\star\star}$ regardless of μ_a's when $M = K$. In this case, $\tilde{\boldsymbol{V}}_k$ degrades to $\tilde{\boldsymbol{v}}_k \in \mathbb{C}^{M\times 1}$, and $\boldsymbol{t}_k^\star$ and $\boldsymbol{t}_k^{\star\star}$ bear the same structure as $p_k \tilde{\boldsymbol{v}}_k$, with $p_k \ge 0$. Furthermore, it can be shown that this result holds regardless of the values of M_B provided that $N = 1$ and $M = M_B A = K$.

<u>Remark</u> *3.4:* In [21], the authors proposed a ZF-BF precoder design for the MISO BC with per-antenna power constraints in the form of the generalized inverse of $\boldsymbol{H}$. The corresponding precoding matrix is expressed as

$$\boldsymbol{T} = [\boldsymbol{g}_1 a_1, \cdots, \boldsymbol{g}_K a_K] + \boldsymbol{U}^\perp [\boldsymbol{b}_1, \cdots, \boldsymbol{b}_K] \tag{27}$$

where $\boldsymbol{g}_k$ is the normalized (to unit-norm) kth column in $\boldsymbol{H}^\dagger$, $k = 1, \cdots, K$; $\boldsymbol{U}^\perp \in \mathbb{C}^{M\times(M-K)}$ is a projection matrix onto the orthogonal complement of the space spanned by the row vectors in $\boldsymbol{H}$, $(\boldsymbol{U}^\perp)^H \boldsymbol{U}^\perp = \boldsymbol{I}$; a_k's and $\boldsymbol{b}_k$'s are design variables, $k = 1, \cdots, K$. In other words, each beamforming vector $\boldsymbol{t}_k$ in $\boldsymbol{T}$ given by (27) should be a linear combination of $\boldsymbol{g}_k$ and the columns in $\boldsymbol{U}^{\perp}$. We see that the beamforming vectors given in (27) are in accordance with the optimal $\boldsymbol{t}_k^{\star}$'s given in (25) due to the fact that for the MISO BC with $N = 1$ and thus $L = M - N(K-1) = M - K + 1$, the space spanned by the columns in $\tilde{\boldsymbol{V}}_k \in \mathbb{C}^{M \times L}$ is the same as that spanned by $\boldsymbol{g}_k$ and the columns in $\boldsymbol{U}^{\perp}$. Note that in [22], an algorithm is proposed to obtain the ZF-BF precoding matrix for the MISO BC with per-antenna power constraints by numerically searching over a_k's and $\boldsymbol{b}_k$'s in (27). In contrast, in this paper the closed-form structures for the optimal ZF-BF precoding vectors are obtained in (25).

Remark 3.5: It is worth comparing the proposed method for solving (P1) in the MISO BC case with the method presented in [21]. For the method in [21], a set of transmit beamforming vectors are applied in a MISO BC as $\boldsymbol{t}_1, \cdots, \boldsymbol{t}_K$. Thus, the weighted sum-rate maximization problem with the ZF-BF precoding and per-antenna power constraints can be formulated as (P4):

$$\begin{aligned} \max_{\boldsymbol{t}_1,\cdots,\boldsymbol{t}_K} \quad & \sum_{k=1}^{K} w_k \log\left(1 + |\boldsymbol{h}_k^H \boldsymbol{t}_k|^2\right) \\ \text{s.t.} \quad & \boldsymbol{h}_j^H \boldsymbol{t}_k = 0, \forall j \neq k \\ & \sum_{k=1}^{K} \mathrm{Tr}\left(\boldsymbol{B}_i \boldsymbol{t}_k \boldsymbol{t}_k^H\right) \leq P, \ \forall i \end{aligned}$$

where $\boldsymbol{B}_i \in \mathbb{C}^{M \times M}$ denotes a diagonal matrix with the ith diagonal element equal to one and all the others equal to zero, $i = 1, \cdots, M$; and P refers to the per-antenna power constraint. Note that (P4) is non-convex due to the fact that the objective function is not necessarily concave over $\boldsymbol{t}_k$'s. In [21], it is proposed to convert (P4) into an equivalent problem in terms of $\boldsymbol{S}_k \triangleq \boldsymbol{t}_k \boldsymbol{t}_k^H, k = 1, \cdots, K$, which is expressed as (P5):

$$\begin{aligned} \max_{\boldsymbol{S}_1,\cdots,\boldsymbol{S}_K} \quad & \sum_{k=1}^{K} w_k \log\left(1 + \boldsymbol{h}_k^H \boldsymbol{S}_k \boldsymbol{h}_k\right) \\ \text{s.t.} \quad & \boldsymbol{h}_j^H \boldsymbol{S}_k \boldsymbol{h}_j = 0, \ \forall j \neq k \\ & \sum_{k=1}^{K} \mathrm{Tr}\left(\boldsymbol{B}_i \boldsymbol{S}_k\right) \leq P, \ \forall i \\ & \boldsymbol{S}_k \succeq \boldsymbol{0}, \ \forall k \\ & \mathrm{Rank}(\boldsymbol{S}_k) = 1, \ \forall k. \end{aligned}$$

Note that (P5) can be treated as (P1) in the case of $N = 1$ and $M_B = 1$ (thus $M = M_B A = A$), and with an additional set of rank-one constraints for $\boldsymbol{S}_k$'s. However, these rank-one constraints are non-convex and thus render (P5) non-convex in general. As a special form of (P1), (P5) without the rank-one constraints is convex, and thus can be solved efficiently by, e.g., the interior-point method [25]. However, the obtained solution for $\boldsymbol{S}_k$ is not guarantied to be rank-one. In [21], it is proved that there always exists a set of rank-one solution $\boldsymbol{S}_k$'s for (P5), and a method is provided to reconstruct the rank-one solution of (P5) from the corresponding solution (with rank greater than one) of (P5) without the rank-one constraints. In contrast, the proposed method in this paper obtains the closed-form solution for (P5), which, as given in (24), is known to be rank-one.

IV. Suboptimal Solution

In this section, we propose a suboptimal solution for (P1), which bears a lower complexity as compared with the optimal solution obtained by (A1). First, we define the projected channel of $\boldsymbol{H}_k$ associated with the projection matrix $\boldsymbol{P}_k$ as $\boldsymbol{H}_k^{\perp} = \boldsymbol{H}_k\boldsymbol{P}_k = \boldsymbol{H}_k\tilde{\boldsymbol{V}}_k\tilde{\boldsymbol{V}}_k^H, k = 1, \cdots, K$, where $\boldsymbol{H}_k^{\perp} \in \mathbb{C}^{N\times M}$, and $\mathrm{Rank}(\boldsymbol{H}_k^{\perp}) = \min(N, M - L) = N$. Next, define the SVD of $\boldsymbol{H}_k^{\perp}$ as

$$\boldsymbol{H}_k^{\perp} = \boldsymbol{U}_k^{\perp}\boldsymbol{\Sigma}_k^{\perp}(\boldsymbol{V}_k^{\perp})^H \tag{28}$$

where $\boldsymbol{U}_k^{\perp} \in \mathbb{C}^{N\times N}$, $\boldsymbol{\Sigma}_k^{\perp} = \mathrm{Diag}(\sigma_{k,1}^{\perp}, \cdots, \sigma_{k,N}^{\perp})$, and $\boldsymbol{V}_k^{\perp} \in \mathbb{C}^{M\times N}$. Then, the proposed suboptimal solution for (P1) has the following structure:

$$\bar{\boldsymbol{S}}_k = \boldsymbol{V}_k^{\perp}\bar{\boldsymbol{\Lambda}}_k(\boldsymbol{V}_k^{\perp})^H \tag{29}$$

where $\bar{\boldsymbol{\Lambda}}_k = \mathrm{Diag}(\bar{\lambda}_{k,1}, \cdots, \bar{\lambda}_{k,N})$ denotes the power allocation for the kth MS. We see that the above structure for $\boldsymbol{S}_k$ is in general suboptimal for (P1) as compared with the optimal structure given in (15). Note that the structure in (29) is optimal for the case of sum-power constraint as discussed in Section III-A, since it can be shown that in (21) $\tilde{\boldsymbol{V}}_k\hat{\boldsymbol{V}}_k = \boldsymbol{V}_k^{\perp}$ with $\boldsymbol{B}_\mu = \mu\boldsymbol{I}$. With $\bar{\boldsymbol{S}}_k$'s given in (29), it can be shown that the ZF constraints in (P1) are satisfied and thus can be removed; furthermore, in the objective function of (P1), the following equalities hold:

$$\log\left|\boldsymbol{I} + \boldsymbol{H}_k\bar{\boldsymbol{S}}_k\boldsymbol{H}_k^H\right| \tag{30}$$

$$= \log\left|\boldsymbol{I} + (\boldsymbol{H}_k^{\perp} + \boldsymbol{H}_k\boldsymbol{V}_k\boldsymbol{V}_k^H)\bar{\boldsymbol{S}}_k(\boldsymbol{H}_k^{\perp} + \boldsymbol{H}_k\boldsymbol{V}_k\boldsymbol{V}_k^H)^H\right| \tag{31}$$

$$= \log\left|\boldsymbol{I} + \boldsymbol{H}_k^{\perp}\bar{\boldsymbol{S}}_k(\boldsymbol{H}_k^{\perp})^H\right| \tag{32}$$

$$= \log\left|\boldsymbol{I} + \boldsymbol{U}_k^{\perp}\boldsymbol{\Sigma}_k^{\perp}\bar{\boldsymbol{\Lambda}}_k\boldsymbol{\Sigma}_k^{\perp}(\boldsymbol{U}_k^{\perp})^H\right| \tag{33}$$

$$= \log\left|\boldsymbol{I} + (\boldsymbol{\Sigma}_k^{\perp})^2\bar{\boldsymbol{\Lambda}}_k\right| \tag{34}$$

where (31) is due to the fact that $\tilde{\boldsymbol{V}}_k\tilde{\boldsymbol{V}}_k^H + \boldsymbol{V}_k\boldsymbol{V}_k^H = \boldsymbol{I}$; (32) is due to the fact that $\bar{\boldsymbol{S}}_k\boldsymbol{V}_k = \boldsymbol{0}$ since $\tilde{\boldsymbol{V}}_k^H\boldsymbol{V}_k = \boldsymbol{0}$; (33) is due to (28) and (29); and (34) is due to the fact that $\log|\boldsymbol{I} + \boldsymbol{X}\boldsymbol{Y}| = \log|\boldsymbol{I} + \boldsymbol{Y}\boldsymbol{X}|$. From (34), we see that the MIMO channel for the kth MS is diagonalized into N scalar channels with channel gains given by

$\bar{\lambda}_{k,i}, i = 1, \cdots, N$. Accordingly, (P1) is reduced to the following problem (P6):

$$\begin{aligned}\max_{\{\bar{\lambda}_{k,i}\}}. \quad & \sum_{k=1}^{K} w_k \sum_{i=1}^{N} \log\left(1 + (\sigma_{k,i}^{\perp})^2 \bar{\lambda}_{k,i}\right) \\ \text{s.t.} \quad & \sum_{k=1}^{K}\sum_{i=1}^{N} \|\boldsymbol{v}_k^{\perp}[a,i]\|^2 \bar{\lambda}_{k,i} \leq P, \ \forall a \\ & \bar{\lambda}_{k,i} \geq 0, \ \forall k, i\end{aligned}$$

where $\{\bar{\lambda}_{k,i}\}$ denotes the set of $\bar{\lambda}_{k,i}$'s, $k = 1, \cdots, K$ and $i = 1, \cdots, N$, while $\boldsymbol{v}_k^{\perp}[a,i]$ denotes the vector consisting of the elements from the ith column and the $((a-1)M_B + 1)$-th to (aM_B)-th rows in $\boldsymbol{V}_k^{\perp}, a = 1, \cdots, A$ and $i = 1, \cdots, N$. It can be verified that (P6) is a convex optimization problem. Thus, similar to (P2), the Lagrange duality method can be applied to solve (P6) by introducing a set of dual variables, $\mu_a, a = 1, \cdots, A$, associated with the set of per-BS power constraints in (P6). For brevity, we omit here the details for derivations and present the optimal solution (power allocation) for $\{\bar{\lambda}_{k,i}\}$ as follows:

$$\bar{\lambda}_{k,i} = \left(\frac{w_k}{\sum_{a=1}^{A} \mu_a \|\boldsymbol{v}_k^{\perp}[a,i]\|^2} - \frac{1}{(\sigma_{k,i}^{\perp})^2}\right)^{+}. \tag{35}$$

Similar to (A1), the following algorithm (A2) can be used to obtain the proposed suboptimal solution for (P1):

- **Initialize** $\mu_a \geq 0, a = 1, \cdots, A$.
- **Compute** the SVDs: $\boldsymbol{H}_k \tilde{\boldsymbol{V}}_k \tilde{\boldsymbol{V}}_k^H = \boldsymbol{U}_k^{\perp} \boldsymbol{\Sigma}_k^{\perp} (\boldsymbol{V}_k^{\perp})^H, k = 1, \cdots, K$.
- **Repeat**
  1. Solve $\{\bar{\lambda}_{k,i}\}$ using (35) with the given μ_a's;
  2. Compute the subgradient of the dual function for (P6) as $P - \sum_{k=1}^{K}\sum_{i=1}^{N} \|\boldsymbol{v}_k^{\perp}[a,i]\|^2 \bar{\lambda}_{k,i}$, $a = 1, \cdots, A$, and update μ_a's accordingly based on the ellipsoid method [26];
- **Until** all the μ_a's converge within a prescribed accuracy.
- **Set** $\bar{\boldsymbol{S}}_k = \boldsymbol{V}_k^{\perp} \bar{\boldsymbol{\Lambda}}_k (\boldsymbol{V}_k^{\perp})^H, k = 1, \cdots, K$.

As compared with (A1), (A2) has a lower complexity due to the fact that for each loop in the "Repeat", only optimization for the power allocation is implemented, rather than the precoding matrix optimization in (A1). Due to the suboptimal structure of the downlink transmit covariance matrix as given in (29), (A2) in general leads to a suboptimal solution for (P1), unless in the special case of $N = 1$ and $M = K$ where the transmit covariance structure in (29) is known to be optimal (see Remark 3.3). In this special case, (A2) can be used as an alternative algorithm of (A1) to obtain the optimal solution for (P1).

Remark 4.1: It is worth noting that (A2) can be shown equivalent to the algorithm proposed in [20] for the special case of $M_B = N = 1$, i.e., the MISO BC with the ZF-BF precoding and the per-antenna power constraints. In this case, similar to Remark 3.3, the proposed suboptimal solution in (29) can be shown corresponding to a precoding matrix in the form of $\boldsymbol{T} = \boldsymbol{H}^{\dagger}\boldsymbol{\Theta}$, where $\boldsymbol{H}$ denotes the downlink MISO-BC channel, and $\boldsymbol{\Theta}$ is a diagonal matrix with the main diagonal to be optimized similarly as we have done for (P6). According to our previous discussions, this algorithm is indeed suboptimal for (P1) if $M > K$.

At last, as a counterpart of Lemma 3.2, we have the following lemma:

Lemma 4.1: Let A^* denote the number of active per-BS power constraints with the optimal solution for (P6). It then holds that $A^* \leq NK$.

Proof: Please refer to Appendix III. ■

Lemma 4.1 provides an upper bound on the number of active per-BS power constraints for the suboptimal solution of (P1) obtained by (A2). It thus follows that in the case of $(A/NK) \gg 1$, most of the BSs in the cooperative multi-cell system cannot transmit with their full powers with the BD precoder design by (A2).

V. Numerical Examples

In this section, we provide numerical examples to corroborate the proposed studies in this paper. For the purpose of exposition, we assume that the channel $\boldsymbol{H}_k$'s in (1) are independent over k, and all the elements in each channel matrix are independent CSCG random variables with zero mean and unit variance. Moreover, we consider the sum-rate maximization for the cooperative multi-cell downlink transmission, i.e., w_k's are all equal in (P1). The obtained numerical results along with discussions are presented in the following subsections.

A. Convergence Behavior

In Fig. 1, we show the convergence behavior of Algorithm (A1) for solving (P1). It is assumed that $A = 2, M_B = 4, K = 4$, and $N = 2$. The transmit power constraint P for each of the two BSs is set equal to 10. The initial values assigned to μ_a's in (A1) are $\mu_1 = \mu_2 = 0.2$. The achievable sum-rate and the consumed transmit powers by the two BSs are shown for different iterations in (A1), each with a pair of updated values for μ_1 and μ_2. As observed, the plotted rate and powers all converge to fixed values after around 30 iterations. The converged transmit powers for the two BSs are observed both equal to their given constraint value, which is 10. This result is in accordance with Remark 3.2, where we have shown that the number of active per-BS power constraints with the optimal solution for (P1) must be no smaller than $\lceil\frac{M-N(K-1)}{M_B}\rceil$, which can be shown equal to 1.5 with the assumed parameters; thus, the number of active per-BS power constraints should be two in this case. Generally speaking, the convergence speed of (A1) depends critically on the total number of per-BS power constraints, A, which is also the number of dual variable μ_a's to be searched. With the ellipsoid method, it is known that the complexity for searching μ_a's in (A1) is $\mathcal{O}(A^2)$ for large values of A [26]. Thus, the number of iterations for the convergence of (A1) grows asymptotically in the order of the square of the number of BSs in the system.

B. MISO BC with Per-Antenna Power Constraints

Next, we consider a special case of the cooperative multi-cell downlink transmission with $M_B = N = 1$, which is equivalent to a MISO BC with the corresponding per-antenna power constraints. The per-BS/per-antenna power constraint is assumed to be $P = 10$. In Fig. 2, we compare the achievable sum-rate with the optimal ZF-BF precoder by (A1) against that with the suboptimal precoder by (A2). The number of MSs is fixed as $K = 2$, while the total number of transmitting antennas M ranges from 2 to 10. It is observed that when $M = K = 2$, the achievable rates for both the optimal and suboptimal precoders are identical, which is in accordance with our discussions in Section IV. It is also observed that when $M > K$, the sum-rate gain of the optimal precoder solution over the suboptimal solution increases with M. In order to explain this observation, in Fig. 3 we show the histograms for the number of active per-antenna power constraints with the optimal and suboptimal solutions over 100 random MISO-BC realizations for the case of $M = 8$. It is observed that the number of active per-antenna power constraints with the optimal solution is always no less than $\lceil \frac{M-N(K-1)}{M_B} \rceil = 7$, while that with the suboptimal solution is always no larger than $NK = 2$, in accordance with Lemma 3.2 and 4.1, respectively. We thus see that when M becomes much larger than K for the MISO BC with the per-antenna power constraints, the optimal ZF-BF design can utilize the full transmit powers from the most number of antennas (equal to $M - K + 1$), while the suboptimal design can only have at most K antennas transmitting with their full powers. This explains why in Fig. 2 the rate gap between the optimal and suboptimal ZF-BF designs enlarges as M increases with a fixed K.

C. MIMO BC with Per-Antenna Power Constraints

At last, we consider the case of multi-antenna MS receivers. For the corresponding auxiliary MIMO BC, we assume that $A = 4, M_B = 1, K = 2$, and $N = 2$. Note that in this case although $M = NK$, i.e., the total number of transmitting antennas are equal to that of receiving antennas, (A2) in general leads to a suboptimal solution for (P1) due to the fact that $N > 1$. In Fig. 4, we show the achievable sum-rates for both the optimal and suboptimal BD precoders vs. the per-BS/per-antenna transmit power constraint P. It is observed that although the optimal precoder solution still performs better than the suboptimal one, their rate gap is not as large as that in Fig. 2 when $M > NK$. This is due to the fact that in the case of $M = NK$, the number of antennas transmitting with full powers for the suboptimal solution is no longer limited by NK according to Lemma 4.1. The practical rule of thump here is that when M is not substantially larger than NK, the low-complexity suboptimal BD precoder by (A2) can be applied to achieve the performance close to that of the optimal BD precoder by (A1).

VI. Conclusion

This paper studies the design of block diagonalization (BD) precoder for the cooperative multi-cell downlink transmission subject to individual power constraints for the bases stations (BSs). By applying convex optimization techniques, this paper derives the closed-form structure of the optimal BD precoder to maximize the weighted sum-rate of users. The optimal BD precoder is shown in general different from the conventional design with the sum-power constraint. An efficient algorithm is proposed to compute the optimal downlink transmit covariance matrix, and the maximum achievable sum-rate is compared with that of a heuristic method, which deploys the conventional BD precoder structure for the sum-power constraint case jointly with the optimized power allocation. Furthermore, the proposed solution leads to the optimal zero-forcing beamforming (ZF-BF) precoder design for the special case of MISO BC with the per-antenna power constraints, and the relationship between the proposed solution and the existing solution based on the theory of generalized matrix inverse is investigated.

References

[1] H. Weingarten, Y. Steinberg, and S. Shamai (Shitz), "The capacity region of the Gaussian multiple-input multiple-output broadcast channel," *IEEE Trans. Inf. Theory*, vol. 52, no. 9, pp. 3936-64, Sep. 2006.

[2] F. Rashid-Farrokhi, K. Liu, and L. Tassiulas, "Transmit beamforming and power control for cellular wireless systems," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 8, pp. 1437-1450, Oct. 1998.

[3] 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.

[4] C. B. Peel, B. Hochwald, and A. L. Swindlehurst, "A vector perturbation technique for near capacity multi-antenna multi-user communication – Part I: Channel inversion and regularization," *IEEE Trans. Commun.*, vol. 53, no. 1, pp. 195-202, Jan. 2005.

[5] A. Wiesel, Y. C. Eldar, and S. Shamai (Shitz), "Linear precoding via conic optimization for fixed MIMO receivers," *IEEE Trans. Sig. Process.*, vol. 54, no. 1, pp. 161-176, Jan. 2006.

[6] M. Stojnic, H. Vikalo, and B. Hassibi, "Maximizing the sum-rate of multi-antenna broadcast channels using linear preprocessing," *IEEE Trans. Wireless Commun.*, vol. 5, no. 9, pp. 2338-2342, Sep. 2006.

[7] Q. H. Spencer, A. L. Swindlehurst, and M. Haardt, "Zero-forcing methods for downlink spatial multiplexing in multiuser MIMO channels," *IEEE Trans. Sig. Proces.*, vol. 52, no. 2, pp. 461-471, Feb. 2004.

[8] K. K. Wong, R. D. Murch, and K. B. Letaief, "A joint-channel diagonalization for multiuser MIMO antenna systems," *IEEE Trans. Wireless Commun.*, vol. 2, no. 4, pp. 773-786, July 2003.

[9] L. U. Choi and R. D. Murch, "A transmit preprocessing technique for multiuser MIMO systems using a decomposition approach," *IEEE Trans. Wireless Commun.*, vol. 3, no. 1, pp. 20-24, Jan. 2004.

[10] Z. Pan, K. K. Wong, and T. S. Ng, "Generalized multiuser orthogonal space-division multiplexing," *IEEE Trans. Wireless Commun.*, vol. 3, no. 6, pp. 1969-1973, Nov. 2004.

[11] G. Carie and S. Shamai, "On the achievable throughput of a multiantenna Gaussian broadcast channel," *IEEE Trans. Inf. Theory*, vol. 49, no. 7, pp. 1691-1706, July 2003.

[12] S. Shamai (Shitz) and B. M. Zaidel, "Enhancing the cellular downlink capacity via co-processing at the transmitting end," in *Proc. IEEE Veh. Technol. Conf. (VTC)*, vol. 3, pp. 1745-1749, May 2001.

[13] H. Zhang and H. Dai, "Cochannel interference mitigation and cooperative processing in downlink multicell multiuser MIMO networks," *EURASIP J. Wireless Commun. Netw.*, no. 2, pp. 222-235, 2004.

[14] M. Karakayali, G. J. Foschini, and R. A. Valenzuela, "Network coordination for spectrally efficient communications in cellular systems," *IEEE Wireless Commun.*, vol. 13, no. 4, pp. 56-61, Aug. 2006.

[15] O. Somekh, B. Zaidel, and S. Shamai (Shitz), "Sum rate characterization of joint multiple cell-site processing," *IEEE Trans. Inf. Theory*, vol. 53, no. 12, pp. 4473-4497, Dec. 2007.

[16] S. Jing, D. Tse, J. Hou, J. Soriaga, J. Smee, and R. Padovani, "Downlink macro-diversity in cellular networks," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, pp. 24-29, June 2007.

[17] B. Ng, J. Evans, S. Hanly, and D. Aktas, "Distributed downlink beamforming with cooperative base stations," *IEEE Trans. Inf. Theory*, vol. 54, no. 12, pp. 5491-5499, Dec. 2008.

[18] W. Yu and T. Lan, "Transmitter optimization for the multi-antenna downlink with per-antenna power constraints," *IEEE Trans. Sig. Proces.*, vol. 55, no. 6, pp. 2646-2660, June 2007.

[19] L. Zhang, R. Zhang, Y.-C. Liang, Y. Xin, and H. V. Poor, "On the Gaussian MIMO BC-MAC duality with multiple transmit covariance constraints," submitted to *IEEE Trans. Inf. Theory*. Available [online] at arXiv:0809.4101.

[20] F. Boccardi and H. Huang, "Optimum power allocation for the MIMO-BC zero-forcing precoder with per-antenna power constraints," in *Proc. Conf. Inf. Sciences Systems (CISS)*, Mar. 2006.

[21] A. Wiesel, Y. Eldar, and S. Shamai, "Zero-forcing precoding and generalized inverses," *IEEE Trans. Sig. Proces.*, vol. 55, no. 9, pp. 4409-4418, Sep. 2008.

[22] H. Huh, H. Papadopoulos, and G. Caire, "MIMO broadcast channel optimization under general linear constraints," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, June 2009.

[23] T. Yoo and A. Goldsmith, "On the optimality of multi-antenna broadcast scheduling using zero-forcing beamforming," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 3, pp. 528-541, Mar. 2006.

[24] Z. Shen, R. Chen, J. G. Andrews, R. W. Heath, and B. L. Evans, "Low complexity user selection algorithms for multiuser MIMO systems with block diagonalization," *IEEE Trans. Sig. Proces.*, vol. 54, no. 9, pp. 3658-3663, Sep. 2006.

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

[26] R. G. Bland, D. Goldfarb, and M. J. Todd, "The ellipsoid method: a survey," *Operations Research*, vol. 29, no. 6, pp. 1039-1091, 1981.

[27] T. Cover and J. Thomas, *Elements of information theory*, New York: Wiley, 1991.

APPENDIX I
PROOF OF LEMMA 3.1

Let $\{\boldsymbol{S}_1^\star, \cdots, \boldsymbol{S}_K^\star\}$ denote the optimal solution of (P1). Without loss of generality, for any given $k \in \{1, \cdots, K\}$, we can express $\boldsymbol{S}_k^\star$ in the following form:

$$\boldsymbol{S}_k^\star = [\tilde{\boldsymbol{V}}_k, \boldsymbol{V}_k] \begin{bmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{B}^H & \boldsymbol{C} \end{bmatrix} [\tilde{\boldsymbol{V}}_k, \boldsymbol{V}_k]^H \tag{36}$$

$$= \tilde{\boldsymbol{V}}_k \boldsymbol{A} \tilde{\boldsymbol{V}}_k^H + \tilde{\boldsymbol{V}}_k \boldsymbol{B} \boldsymbol{V}_k^H + \boldsymbol{V}_k \boldsymbol{B}^H \tilde{\boldsymbol{V}}_k^H + \boldsymbol{V}_k \boldsymbol{C} \boldsymbol{V}_k^H \tag{37}$$

where $\boldsymbol{A} \in \mathbb{C}^{(M-L)\times(M-L)}$, $\boldsymbol{B} \in \mathbb{C}^{(M-L)\times L}$, and $\boldsymbol{C} \in \mathbb{C}^{L\times L}$. Note that $\boldsymbol{A} = \boldsymbol{A}^H$ and $\boldsymbol{C} = \boldsymbol{C}^H$. Since $\boldsymbol{S}_k^\star$ must satisfy the set of ZF constraints in (P1), it follows that

$$\boldsymbol{V}_k^H \boldsymbol{S}_k^\star \boldsymbol{V}_k = 0. \tag{38}$$

From (37) and (38), it follows that $\boldsymbol{C} = \boldsymbol{0}$. Furthermore, from the theory of Schur complement [25], it is known that $\boldsymbol{S}_k^\star \succeq \boldsymbol{0}$ iff the following conditions are satisfied:

$$\boldsymbol{A} \succeq \boldsymbol{0} \tag{39}$$

$$(\boldsymbol{I} - \boldsymbol{A}\boldsymbol{A}^\dagger)\boldsymbol{B} = \boldsymbol{0} \tag{40}$$

$$\boldsymbol{C} - \boldsymbol{B}^H\boldsymbol{A}^\dagger\boldsymbol{B} \succeq \boldsymbol{0}. \tag{41}$$

Since $\boldsymbol{A} \succeq \boldsymbol{0}$, it follows that $\boldsymbol{B}^H\boldsymbol{A}^\dagger\boldsymbol{B} \succeq \boldsymbol{0}$. Using this fact together with $\boldsymbol{C} = \boldsymbol{0}$, from (41) it follows that $\boldsymbol{B}^H\boldsymbol{A}^\dagger\boldsymbol{B} = \boldsymbol{0}$. Thus, from (40) it follows that $\boldsymbol{B} = \boldsymbol{0}$. With $\boldsymbol{B} = \boldsymbol{0}$ and $\boldsymbol{C} = \boldsymbol{0}$, from (37) it follows that $\boldsymbol{S}_k^\star = \tilde{\boldsymbol{V}}_k \boldsymbol{A} \tilde{\boldsymbol{V}}_k^H$. By letting $\boldsymbol{A} = \boldsymbol{Q}_k$, the proof of Lemma 3.1 thus follows.

APPENDIX II
PROOF OF LEMMA 3.2

We prove Lemma 3.2 by contradiction. Suppose that there exist a number of strictly positive μ_a's such that $A_\mu < \lceil \frac{M-N(K-1)}{M_B} \rceil$. Then, it follows that $A_\mu < \frac{M-N(K-1)}{M_B}$. Since $L = N(K-1)$, it thus follows that $M_B A_\mu < (M-L)$. Let $\mathcal{S}$ denote the set consisting of the indexes corresponding to all the non-zero diagonal elements in $\boldsymbol{B}_\mu$, i.e., if $\mu_a > 0$ for any $a \in \{1, \cdots, A\}$, then $(a-1)M_B + i \in \mathcal{S}, i = 1, \cdots, M_B$. Note that the size of $\mathcal{S}$ is denoted by $|\mathcal{S}| = M_B A_\mu$. Let $\boldsymbol{E}_k(\mathcal{S})$ and $\boldsymbol{F}_k(\mathcal{S}^c)$ denote the matrix consisting of the rows in $\tilde{\boldsymbol{V}}_k \in \mathbb{C}^{M\times(M-L)}$ with the row indexes given by the elements in $\mathcal{S}$ and $\mathcal{S}^c$, respectively, where $\mathcal{S}^c$ denotes the complement of $\mathcal{S}$. Note that $|\mathcal{S}| + |\mathcal{S}^c| = M$ and $|\mathcal{S}^c| > 0$ since $M_B A_\mu < (M-L) < M$. From $\boldsymbol{E}_k(\mathcal{S}) \in \mathbb{C}^{M_B A_\mu \times (M-L)}$ and $M_B A_\mu < (M-L)$, it follows that $\boldsymbol{E}_k(\mathcal{S})$ is not full row-rank. Thus, we could find a vector $\boldsymbol{q}_k \in \mathbb{C}^{(M-L)\times 1}$ with $\|\boldsymbol{q}_k\| = 1$ such that $\boldsymbol{E}_k(\mathcal{S})\boldsymbol{q}_k = \boldsymbol{0}$ and $\boldsymbol{F}_k(\mathcal{S}^c)\boldsymbol{q}_k \neq \boldsymbol{0}$. Accordingly, we have $\boldsymbol{B}_\mu \tilde{\boldsymbol{V}}_k \boldsymbol{q}_k = \boldsymbol{0}$ and $\tilde{\boldsymbol{V}}_k \boldsymbol{q}_k \neq \boldsymbol{0}$. Denote $\boldsymbol{w}_k = \tilde{\boldsymbol{V}}_k \boldsymbol{q}_k$. Note that the indexes of the non-zero elements in $\boldsymbol{w}_k$ belong to $\mathcal{S}^c$. Suppose that the solution of (P3) is taken as $\boldsymbol{Q}_k^\star = p(\boldsymbol{q}_k \boldsymbol{q}_k^H)$ with $p \geq 0$. Substituting the above solution into the objective function of (P3) yields that

$$w_k \log \left| \boldsymbol{I} + \boldsymbol{H}_k \tilde{\boldsymbol{V}}_k \boldsymbol{Q}_k^\star \tilde{\boldsymbol{V}}_k^H \boldsymbol{H}_k^H \right| - \mathrm{Tr}\left(\boldsymbol{B}_\mu \tilde{\boldsymbol{V}}_k \boldsymbol{Q}_k^\star \tilde{\boldsymbol{V}}_k^H \right) \tag{42}$$

$$= w_k \log \left| \boldsymbol{I} + p\boldsymbol{H}_k \boldsymbol{w}_k \boldsymbol{w}_k^H \boldsymbol{H}_k^H \right|. \tag{43}$$

Let $\boldsymbol{R}_k = \boldsymbol{H}_k \boldsymbol{w}_k$. Then, (43) can be further expressed as $w_k \log \left| \boldsymbol{I} + p \boldsymbol{R}_k \boldsymbol{R}_k^H \right|$, whose value becomes unbounded as $p \rightarrow \infty$ provided that $\boldsymbol{R}_k \boldsymbol{R}_k^H \neq \mathbf{0}$ (which holds with probability one with independent channel realizations). Therefore, we conclude that the presumption that $A_\mu < \lceil \frac{M-N(K-1)}{M_B} \rceil$ cannot be true. Lemma 3.2 thus follows.

APPENDIX III
PROOF OF LEMMA 4.1

We prove Lemma 4.1 by contradiction. Suppose that $A^* \geq (NK+1)$. Let $\mathcal{B}$ be a subset of $\{1, \cdots, A\}$ consisting of the indexes of the BSs for which the transmit power constraints are tight with the optimal solution for (P6). Note that $|\mathcal{B}| = A^*$. Let $\bar{\lambda}_{k,i}^\star$ denote the optimal solution for (P6), $k = 1, \cdots, K$ and $i = 1, \cdots, N$. Thus, we have the following equalities from (P6):

$$\sum_{k=1}^{K} \sum_{i=1}^{N} \|\boldsymbol{v}_k^\perp[a,i]\|^2 \bar{\lambda}_{k,i}^\star = P, \ \forall a \in \mathcal{B}. \tag{44}$$

Accordingly, $\bar{\lambda}_{k,i}^\star$'s are the solutions for a set of A^* linear independent (which holds with probability one due to independent channel realizations) equations. However, since $A^* \geq (NK+1)$, we see that the number of equations exceeds that of unknowns, which is equal to NK. Thus, given $P > 0$, there exist no feasible solutions for $\bar{\lambda}_{k,i}^\star$'s. We thus conclude that the presumption that $A^* \geq (NK+1)$ cannot be true. Lemma 4.1 thus follows.

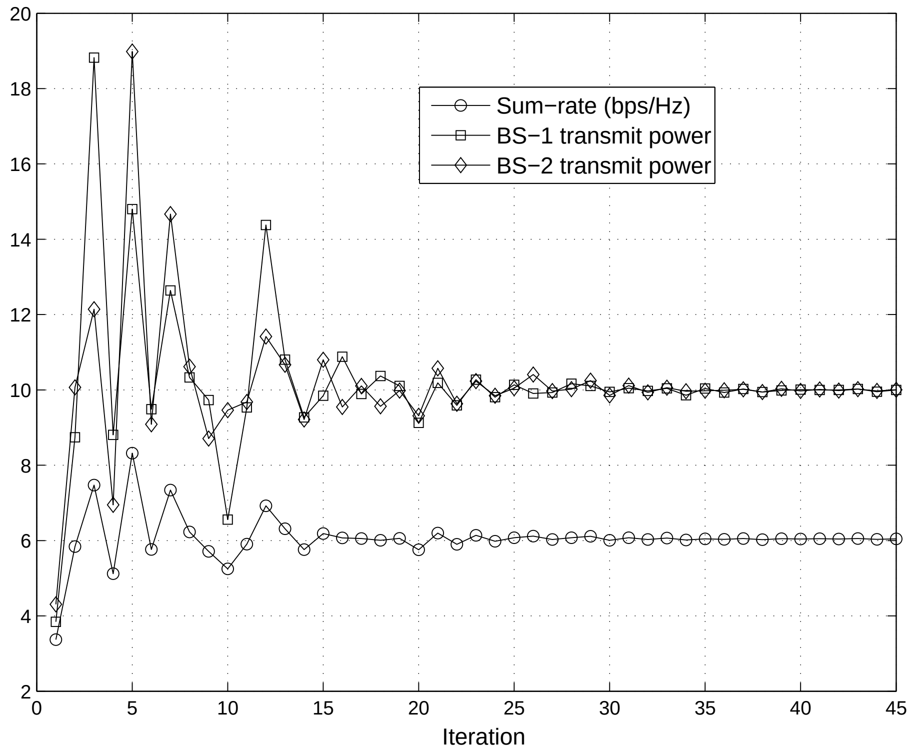

Fig. 1. Convergence behavior of Algorithm (A1).

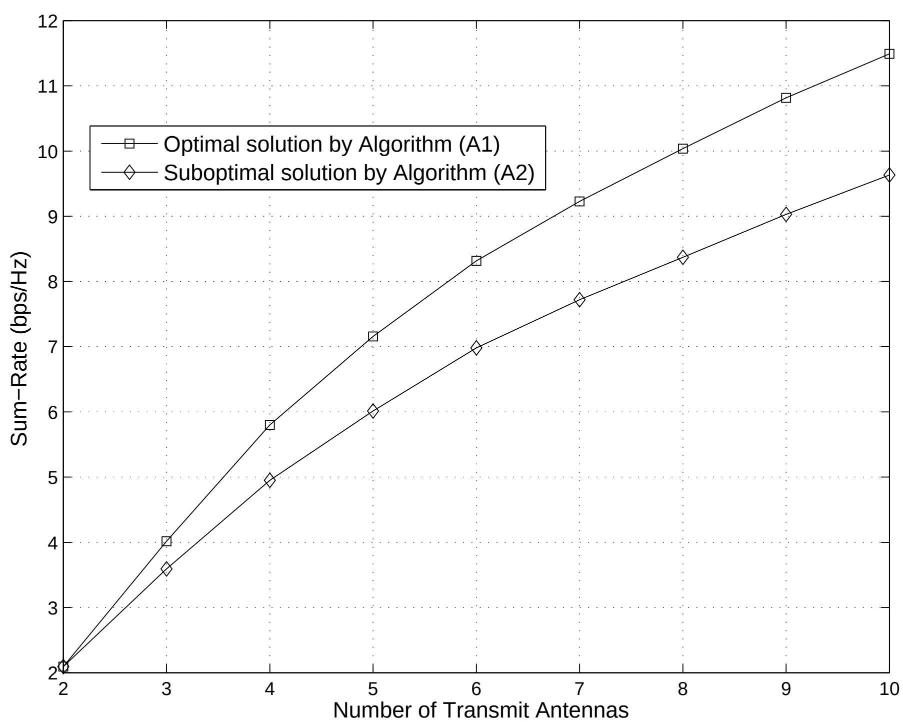

Fig. 2. Comparison of the sum-rate in the MISO BC with the ZF-BF precoding and the per-antenna power constraints.

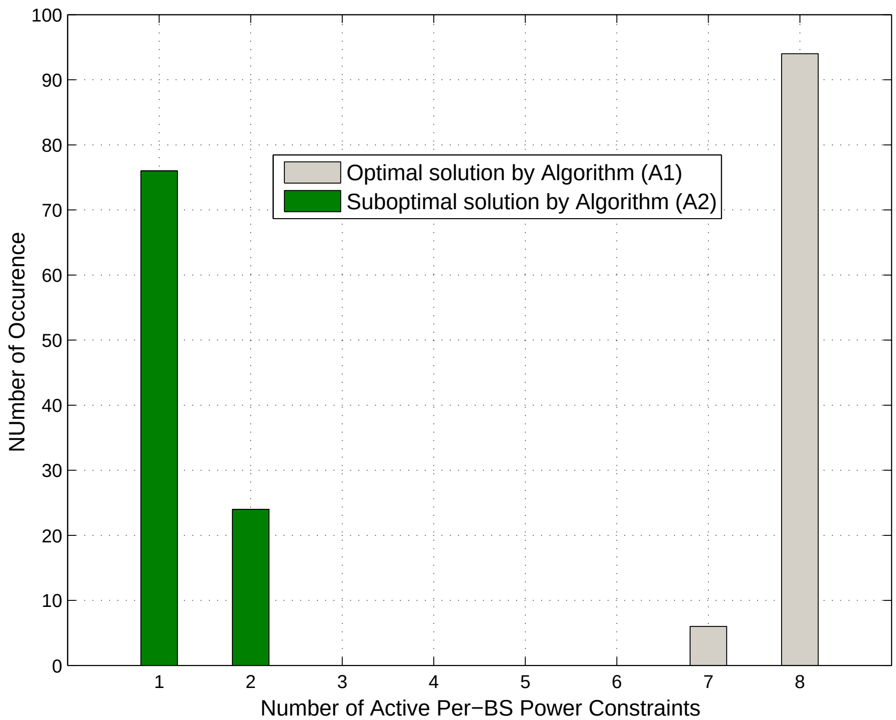

Fig. 3. Comparison of the number of active per-BS power constraints in the MISO BC with the ZF-BF precoding and the per-antenna power constraints.

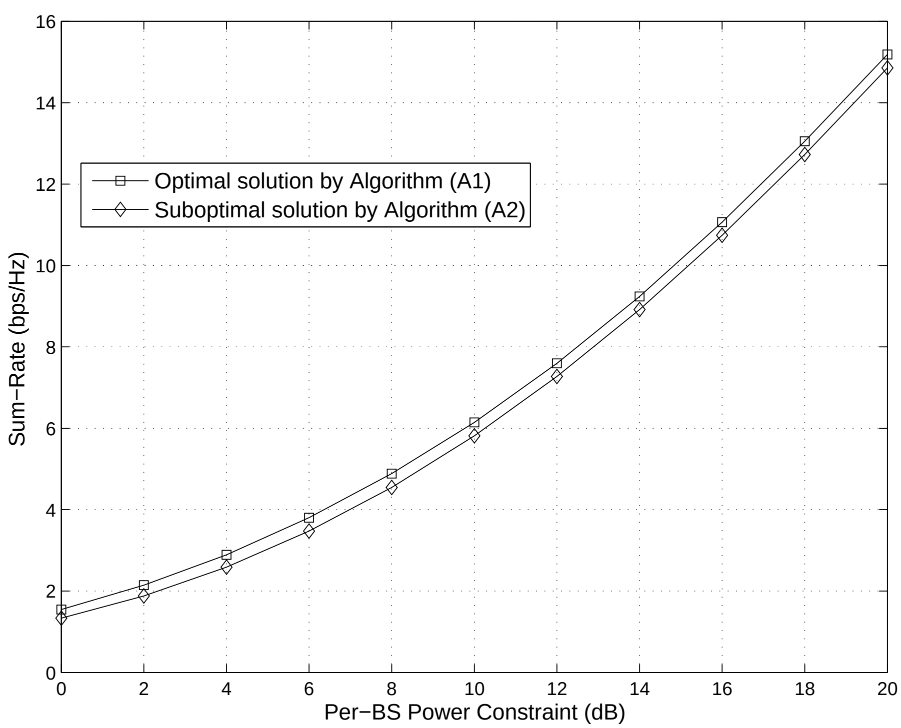

Fig. 4. Comparison of the sum-rate in the MIMO BC with the BD precoding and the per-antenna power constraints.