

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/327122632>

Weighted Sum Rate Maximization for Zero-Forcing Methods with General Linear Covariance Constraints

Conference Paper · May 2018

DOI: 10.1109/ICC.2018.8422453

CITATIONS

0

READS

24

5 authors, including:



Ronan Farrell

National University of Ireland, Maynooth

188 PUBLICATIONS 603 CITATIONS

[SEE PROFILE](#)



Mark F. Flanagan

University College Dublin

136 PUBLICATIONS 655 CITATIONS

[SEE PROFILE](#)



Le-Nam Tran

University College Dublin

100 PUBLICATIONS 1,683 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



Crosstalk Cancellation via Digital Pre-Distortion (DPD) [View project](#)



Flexible Modulation Techniques for use in 5G communication systems [View project](#)

Weighted Sum Rate Maximization for Zero-forcing Methods with General Linear Covariance Constraints

Thuy M. Pham*, Ronan Farrell*, Holger Claussen[†], Mark Flanagan[‡], and Le-Nam Tran[‡]

*Department of Electronic Engineering, Maynooth University, Ireland. Email: {minhthuy.pham, ronan.farrell}@mu.ie

[†]Nokia Bell Labs, Ireland. Email: holger.claussen@nokia-bell-labs.com.

[‡]School of Electrical and Electronic Engineering, University College Dublin, Ireland. Email: {mark.flanagan,nam.tran}@ucd.ie

Abstract—In this paper, an efficient approach for weighted sum rate maximization (WSRMax) for zero-forcing (ZF) methods with general linear transmit covariance constraints (LTCCs) is proposed. This problem has been extensively studied separately for some special cases such as for sum power or per-antenna power constraints (PAPC). Due to some practical and regulatory requirements, these power constraints alone are not in general sufficient, which motivates the consideration of general LTCCs. On the other hand, the zero-forcing (ZF) is a simple linear precoding technique to mitigate inter-user interference. The problem of WSRMax for ZF methods with LTCCs was studied previously using a gradient descent algorithm with barrier functions, but this method was also shown to converge slowly. To derive an efficient solution to this problem, we first reformulate it as an equivalent minimax problem using Lagrangian duality. The obtained result in fact resembles BC-MAC duality but is specialized for ZF methods. We then combine alternating optimization and concave-convex procedure to efficiently compute a saddle point of the minimax problem. The proposed method is numerically shown to converge very fast and its complexity scales linearly with the number of users.

Index Terms—MIMO, zero-forcing, alternating optimization, linear transmit covariance constraints, minimax duality, closed-form, concave-convex procedure.

I. INTRODUCTION

It is known that the capacity of a Gaussian multiple-input multiple-output broadcast channel (BC) can be achieved by using the dirty paper coding (DPC) technique [1]. However, such a nonlinear coding strategy requires complex processing in both encoders and decoders, and thus is not attractive for practical applications. In this regard, linear precoding methods have received much more attention due to their capability to provide a good tradeoff between sum rate performance and implementation complexity. Among those, zero-forcing (ZF) is a promising alternative to DPC since it can attain a near-optimal performance but with lower complexity, for MIMO configurations with a large number of degrees of freedom.

Transmitter optimization is performed in accordance with various power constraints. Much of the early research on this topic was concerned with the sum power constraint (SPC) [1], [2]. To account for the fact that the power amplifier of each transmit antenna may have different nonlinear operation region, subsequent studies considered a per-antenna power

constraint (PAPC) [3]–[7]. MIMO capacity with joint SPC and PAPC was also investigated recently [8], [9]. In the context of cognitive networks, one may impose interference temperature constraints on a secondary user (SU) to limit the interference generated at a primary user (PU) [10]. These transmit power constraints are generally categorized as linear transmit covariance constraints (LTCCs).

The problem of efficient transmitter optimization for MIMO systems in general and for zero-forcing methods in particular has not yet been fully studied in the existing literature. To the best of the authors' knowledge, the most related work to that of the present paper was carried out in [10], where Huh *et al.* proposed a gradient descent algorithm with barrier functions to solve the problem of weighted sum rate maximization (WSRMax) for ZF under multiple generic LTCCs. However, the proposed method in [10] has two main drawbacks: (i) it is only applicable to MISO systems, and (ii) it converges very slowly.

To overcome the above shortcomings, we combine two powerful optimization tools, namely concave-convex procedure (CCP) and alternating optimization (AO), to derive a computationally efficient algorithm to solve the considered problem. Our major contributions can be summarized as follows:

- We propose for the first time an efficient solution to the WSRMax problem in multi-user MIMO channels with ZF methods, subject to multiple generic LTCCs. The proposed approach can also be extended to deal with other precoding schemes such as successive zero-forcing dirty paper coding (SZFDPC) [11], [12].
- We present the BC-MAC duality for multi-user MIMO channels with ZF methods under multiple LTCCs, resulting in a minimax problem for which a strong duality holds. Then, CCP and AO are applied to find the saddle point of the minimax problem. Each iteration of the proposed method can be solved efficiently by water-filling algorithms, leading to its fast convergence rate.
- We carry out numerical experiments to compare the proposed algorithm with the convex solver-based methods. The numerical results demonstrate that our proposed algorithm has low complexity, making it suitable to

characterize the performance of massive MIMO systems under generic LTCCs.

Notation: Standard notations are used in this paper. Bold lower and upper case letters represent vectors and matrices, respectively. \mathbf{I}_N defines an identity matrix of size N ; \mathbf{I} and $\mathbf{0}$ define identity and zero matrices respectively, of which the size can be easily inferred from the context. $\mathbb{C}^{M \times N}$ denotes the space of $M \times N$ complex matrices; \mathbf{H}^\dagger and \mathbf{H}^T are Hermitian and ordinary transpose of \mathbf{H} , respectively; $\mathbf{H}_{i,j}$ is the (i, j) th entry of \mathbf{H} ; $|\mathbf{H}|$ is the determinant of \mathbf{H} ; $\text{rank}(\mathbf{H})$ stands for the rank of \mathbf{H} . Furthermore, we denote the expected value of a random variable by $\mathbb{E}[\cdot]$, and $[x]_+ = \max(x, 0)$.

II. SYSTEM MODEL

Consider a K -user single-cell MIMO BC where the base station (BS) and each user have N and M_k antennas, respectively. Let $\mathbf{H}_k \in \mathbb{C}^{M_k \times N}$ be the channel matrix for user k . Then, the received signal at user k is given by

$$\mathbf{y}_k = \mathbf{H}_k \mathbf{x}_k + \sum_{j \neq k} \mathbf{H}_k \mathbf{x}_j + \mathbf{z}_k \quad (1)$$

where $\mathbf{x}_k \in \mathbb{C}^{N \times 1}$ is the downlink signal for the k th user and $\mathbf{z}_k \sim \mathcal{CN}(\mathbf{0}, \mathbf{I})$ is the background noise. In this paper, channel state information is perfectly known at both the BS and users. For linear precoding we can express \mathbf{x}_k as $\mathbf{x}_k = \mathbf{R}_k \mathbf{s}_k$, where $\mathbf{R}_k \in \mathbb{C}^{N \times M_k}$ and $\mathbf{s}_k \in \mathbb{C}^{M_k \times 1}$ denote the precoding matrix and information-bearing signal, respectively. We further assume that \mathbf{s}_k consists of independent zero-mean and unit energy symbols, i.e., $\mathbf{s}_k \sim \mathcal{CN}(\mathbf{0}, \mathbf{I})$. For the ZF technique, the inter-user interference to user k is suppressed by designing \mathbf{R}_k such that $\mathbf{H}_j \mathbf{R}_k = \mathbf{0}$ for all $j \neq k$. Thus, the WSRMax problem for ZF precoding with LTCCs is formulated as

$$\underset{\{\mathbf{S}_k \succeq \mathbf{0}\}}{\text{maximize}} \quad \sum_{k=1}^K w_k \log |\mathbf{I} + \mathbf{H}_k \mathbf{S}_k \mathbf{H}_k^\dagger| \quad (2a)$$

$$\text{subject to} \quad \mathbf{H}_j \mathbf{S}_k \mathbf{H}_j^\dagger = \mathbf{0}, \quad \forall j \neq k \quad (2b)$$

$$\sum_{k=1}^K \text{tr}(\mathbf{E}_{ik} \mathbf{S}_k) \leq P_i, \quad i = 1, 2, \dots, L \quad (2c)$$

where $\mathbf{S}_k = \mathbb{E}[\mathbf{x}_k \mathbf{x}_k^\dagger] = \mathbf{R}_k \mathbf{R}_k^\dagger$ is the input covariance matrix for user k , \mathbf{E}_{ik} is the i th positive semidefinite matrix of user k , P_i is the power constraint associated with $\{\mathbf{E}_{ik}\}_{k=1}^K$, and $w_k \geq 0$ is the weighting factor assigned to the k th user to maintain some degree of fairness. In fact we have omitted the constraint that $\text{rank}(\mathbf{S}_k) \leq \min(N, M_k)$ so that the precoder \mathbf{R}_k can be extracted from \mathbf{S}_k . However, it was proved in [13] that this relaxation does not affect the optimality. This property is also obvious from the BC-MAC duality presented later on. Note that (2c) is called the general LTCC since it includes several transmit power constraints as specific cases. For example, when $\mathbf{E}_{ik} = \mathbf{I}$, $L = 1$, (2) becomes WSRMax under SPC. If $\mathbf{E}_{ik} = \text{diag}(\mathbf{e}_i)$, $L = N$, where \mathbf{e}_i is a vector of all zeros except a value of one at the i th position, then the resulting problem is WSRMax under PAPC. Each of these two types of power constraints has been extensively studied.

Before proceeding further, let us first simplify (2) by eliminating the zero-interference constraints. Denote $\tilde{\mathbf{H}}_k$ to

be the channel matrix of all users, except for user k , i.e., $\tilde{\mathbf{H}}_k = [\mathbf{H}_1^\dagger, \dots, \mathbf{H}_{k-1}^\dagger, \mathbf{H}_{k+1}^\dagger, \dots, \mathbf{H}_K^\dagger]^\dagger$. For ZF precoding to be feasible, it should hold that $N \geq \sum_{k=1}^K M_k$, which is assumed in this paper [2]. Let \mathbf{V}_k be a basis of the null space of $\tilde{\mathbf{H}}_k$, then (2) reduces to the following maximization problem

$$\begin{aligned} & \underset{\{\tilde{\mathbf{S}}_k \succeq \mathbf{0}\}}{\text{maximize}} \quad \sum_{k=1}^K w_k \log |\mathbf{I} + \mathbf{H}_k \mathbf{V}_k \tilde{\mathbf{S}}_k \mathbf{V}_k^\dagger \mathbf{H}_k^\dagger| \\ & \text{subject to} \quad \sum_{k=1}^K \text{tr}(\mathbf{E}_{ik} \mathbf{V}_k \tilde{\mathbf{S}}_k \mathbf{V}_k^\dagger) \leq P_i, \quad i = 1, \dots, L. \end{aligned} \quad (3)$$

We note that for this general problem, (3) can be recast as a MAXDET program [14] and solved by a dedicated optimization package such as SDPT3 [15]. However, solving (3) by generic convex solvers is not computationally efficient for large-scale problems, nor does it provide useful insights into the structure of the optimal input covariance matrices. In particular, modern convex solvers are mostly based on interior-point methods whose complexity increases rapidly with the problem size, e.g., with the number of transmit antennas N and/or the number of users K in the considered context. For the special case of multi-user MISO systems, Huh *et al.* presented a gradient descent algorithm with barrier functions but it converges very slowly [10]. For large-scale MISO systems, Huh *et al.* also proposed a low-complexity solution but it cannot achieve the optimal performance. Therefore, an efficient algorithm for general MIMO systems for ZF precoding with LTCCs has remained an open problem. In the following, we propose a low-complexity method to solve this problem.

III. PROPOSED ALGORITHM

A. Algorithm Description

In this section, we first transform (3) into an equivalent problem in MAC, then apply AO and CCP to derive an efficient algorithm. Specifically, the proposed algorithm is based on the following theorem:

Theorem 1. *The equivalent problem in the dual MAC of the problem (3) in the BC is the following minimax problem*

$$\begin{aligned} & \min_{\lambda \geq 0} \max_{\{\tilde{\mathbf{S}}_k \succeq \mathbf{0}\}} \quad \sum_{k=1}^K w_k \log \frac{|\mathbf{V}_k^\dagger (\sum_{i=1}^L \lambda_i \mathbf{E}_{ik}) \mathbf{V}_k + \tilde{\mathbf{H}}_k^\dagger \tilde{\mathbf{S}}_k \tilde{\mathbf{H}}_k|}{|\mathbf{V}_k^\dagger (\sum_{i=1}^L \lambda_i \mathbf{E}_{ik}) \mathbf{V}_k|} \\ & \text{subject to} \quad \sum_{k=1}^K \text{tr}(\tilde{\mathbf{S}}_k) = P \\ & \quad \quad \quad \mathbf{p}^T \boldsymbol{\lambda} = P \end{aligned} \quad (4)$$

where $\tilde{\mathbf{H}}_k = \mathbf{H}_k \mathbf{V}_k$, $\boldsymbol{\lambda} = [\lambda_1, \lambda_2, \dots, \lambda_L]^T$ and $\mathbf{p} = [P_1, P_2, \dots, P_L]^T$. Let $(\boldsymbol{\lambda}^*, \{\tilde{\mathbf{S}}_k^*\})$ be the saddle point of (4). Then, the optimal solution \mathbf{S}_k^* to (3) is given by

$$\tilde{\mathbf{S}}_k^* = (\mathbf{V}_k^\dagger \boldsymbol{\Lambda}_k^* \mathbf{V}_k)^{-\frac{1}{2}} \mathbf{U}_k \mathbf{X}_k^\dagger \tilde{\mathbf{S}}_k^* \mathbf{X}_k \mathbf{U}_k^\dagger (\mathbf{V}_k^\dagger \boldsymbol{\Lambda}_k^* \mathbf{V}_k)^{-\frac{1}{2}} \quad (5)$$

where $\boldsymbol{\Lambda}_k^* = \sum_{i=1}^L \lambda_i^* \mathbf{E}_{ik}$, and $\mathbf{U}_k \boldsymbol{\Sigma}_k \mathbf{X}_k^\dagger$ is the compact singular value decomposition of $(\mathbf{V}_k^\dagger \boldsymbol{\Lambda}_k^* \mathbf{V}_k)^{-1/2} \tilde{\mathbf{H}}_k^\dagger$.

We remark that the above theorem is the generalization of the duality result in [16], [17] and its proof can be found in the Appendix.

Problem (4) is in fact a concave-convex program. Explicitly, the objective function in (4) is convex with respect to $\boldsymbol{\lambda}$

and concave with respect to $\{\bar{\mathbf{S}}_k\}$ [18]. Since the objective function is twice-differentiable, a saddle point exists. A naive approach to finding a saddle-point of (4) is to alternate between minimization and maximization. However, the convergence of such pure iterative alternating optimization is not guaranteed. Existing solutions to such a minimax problem as (4) are based on interior-point methods [6], [10]. Unfortunately, these methods do not scale well with the problem size, and thus are not suitable for large-scale MIMO systems. In the following, we show that (4) can be solved efficiently by combining AO and CCP in a novel way.

The proposed algorithm alternately optimizes $\{\bar{\mathbf{S}}_k\}$ and $\boldsymbol{\lambda}$ but a monotonic convergence is ensured. At the n th iteration of the proposed method, $\{\bar{\mathbf{S}}_k^n\}$ is found to be the solution to the following maximization problem:

$$\begin{aligned} \max \quad & \sum_{k=1}^K w_k \log |\mathbf{V}_k^\dagger \boldsymbol{\Lambda}_k^n \mathbf{V}_k + \tilde{\mathbf{H}}_k^\dagger \bar{\mathbf{S}}_k \tilde{\mathbf{H}}_k| \\ \text{s.t.} \quad & \sum_{k=1}^K \text{tr}(\bar{\mathbf{S}}_k) = P; \bar{\mathbf{S}}_k \succeq \mathbf{0}, k = 1, \dots, K \end{aligned} \quad (6)$$

where $\boldsymbol{\Lambda}_k^n = \sum_{i=1}^L \lambda_i^n \mathbf{E}_{ik}$. That is, to find $\{\bar{\mathbf{S}}_k^n\}$ we fix $\boldsymbol{\lambda}$. The above problem is actually the one of WSRMax under a single SPC, which admits a water-filling solution [2]. We skip the details for the sake of brevity and refer the interested reader to [2] for further details.

In the next step, we fix $\{\bar{\mathbf{S}}_k^n\}$ and solve for the optimal $\boldsymbol{\lambda}$ which minimizes the objective in (4). However, we do not optimize the objective directly as this does not guarantee convergence. Instead, we minimize an upper bound of the objective. In fact, this approach is motivated by the concept of CCP which generates a monotonic decrease of the objective. In this way, the concave part of the objective is linearized to obtain an upper bound. More explicitly, we apply the concavity property of the log-determinant function [18, p. 73] to the numerator of (4) to obtain the following inequality

$$\log |\mathbf{V}_k^\dagger \boldsymbol{\Lambda}_k \mathbf{V}_k + \tilde{\mathbf{H}}_k^\dagger \bar{\mathbf{S}}_k \tilde{\mathbf{H}}_k| \leq \log |\boldsymbol{\Phi}_k^n| + \text{tr}(\mathbf{V}_k \boldsymbol{\Phi}_k^{-n} \mathbf{V}_k^H (\boldsymbol{\Lambda}_k - \boldsymbol{\Lambda}_k^n)) \quad (7)$$

where $\boldsymbol{\Phi}_k^n \triangleq \mathbf{V}_k^\dagger \boldsymbol{\Lambda}_k \mathbf{V}_k + \tilde{\mathbf{H}}_k^\dagger \bar{\mathbf{S}}_k \tilde{\mathbf{H}}_k$, and $\boldsymbol{\Phi}_k^{-n} \triangleq (\boldsymbol{\Phi}_k^n)^{-1}$. Thus, at the n th iteration of the proposed algorithm, $\boldsymbol{\Lambda}_k^{n+1}$ is the solution to the following problem

$$\begin{aligned} \min \quad & \sum_{k=1}^K w_k (\text{tr}(\mathbf{V}_k \boldsymbol{\Phi}_k^{-n} \mathbf{V}_k^H \boldsymbol{\Lambda}_k) - \log |\mathbf{V}_k^\dagger \boldsymbol{\Lambda}_k \mathbf{V}_k|) \triangleq g(\boldsymbol{\lambda}) \\ \text{s.t.} \quad & \mathbf{p}^T \boldsymbol{\lambda} = P, \boldsymbol{\lambda} \succeq \mathbf{0} \end{aligned} \quad (8)$$

We remark that the feasible set of (8), denoted by $\Theta \triangleq \{\mathbf{p}^T \boldsymbol{\lambda} = P; \boldsymbol{\lambda} \succeq \mathbf{0}\}$, is a simplex. Since the projection on a simplex can be done efficiently [19], we can apply the gradient projection or conjugate gradient projection method to solve (8). The overall algorithms to solve the minimization (8) and WSRMax are summarized in Algorithm 1 and Algorithm 2, respectively. The convergence proof of Algorithm 2 is provided in the next subsection which in fact follows the similar arguments as those of [20], [21]. We also note that our proposed method can be easily modified to deal with the WSRMax problem for SZFDPC with multiple LTCCs. Specifically, we simply replace $\tilde{\mathbf{H}}_k = [\mathbf{H}_1^\dagger, \dots, \mathbf{H}_{k-1}^\dagger, \mathbf{H}_{k+1}^\dagger, \dots, \mathbf{H}_K^\dagger]^\dagger$

(defined immediately before (3)) by $\check{\mathbf{H}}_k = [\mathbf{H}_1^\dagger, \dots, \mathbf{H}_{k-1}^\dagger]^\dagger$ and the remaining steps are the same.

Algorithm 1: The gradient projection algorithm for solving (8).

Input: $\boldsymbol{\lambda}_0, \epsilon_1 > 0$.

- 1 Initialize $m := 0$, and $\tau_1 = 1 + \epsilon_1$.
 - 2 **while** $\tau_1 > \epsilon_1$ **do**
 - 3 Calculate the gradient $\tilde{\mathbf{u}}_m = -\nabla g(\boldsymbol{\lambda}_m)$.
 - 4 Choose an appropriate positive scalar ρ_m and create $\tilde{\boldsymbol{\lambda}}_m = \boldsymbol{\lambda}_m + \rho_m \tilde{\mathbf{u}}_m$.
 - 5 Project $\tilde{\boldsymbol{\lambda}}_m$ onto Θ to obtain $\bar{\boldsymbol{\lambda}}_m$.
 - 6 Choose appropriate step size ν_m using the Armijo rule [22] and set $\boldsymbol{\lambda}_{m+1} = \boldsymbol{\lambda}_m + \nu_m (\bar{\boldsymbol{\lambda}}_m - \boldsymbol{\lambda}_m)$.
 - 7 Compute $\tau_1 = |\nabla g(\boldsymbol{\lambda}_m)^T (\boldsymbol{\lambda}_{m+1} - \boldsymbol{\lambda}_m)|$.
 - 8 $m := m + 1$.
 - 9 **end**
- Output:** $\boldsymbol{\lambda}_m$ as the optimal solution to (8).
-

Algorithm 2: The proposed algorithm for solving (3).

Input: $\boldsymbol{\lambda}^0$ feasible to Θ , $\epsilon_2 > 0$.

- 1 Initialize $n := 0$, and $\tau_2 = 1 + \epsilon_2$.
 - 2 **while** $\tau_2 > \epsilon_2$ **do**
 - 3 Compute $\boldsymbol{\Lambda}_k^n = \sum_{i=1}^L \lambda_i^n \mathbf{E}_{ik}$
 - 4 Apply the water-filling algorithm to solve (6). Denote the optimal solution by $\{\bar{\mathbf{S}}_k^n\}$.
 - 5 For $n \geq 1$, compute $\tau_2 = |f(\boldsymbol{\lambda}^n, \bar{\mathbf{S}}^n) - f(\boldsymbol{\lambda}^{n-1}, \bar{\mathbf{S}}^{n-1})|$.
 - 6 For each k , set $\boldsymbol{\Phi}_k^n = (\mathbf{V}_k^\dagger \boldsymbol{\Lambda}_k^n \mathbf{V}_k + \tilde{\mathbf{H}}_k^\dagger \bar{\mathbf{S}}_k \tilde{\mathbf{H}}_k)$.
 - 7 Solve (8) to find $\boldsymbol{\lambda}^{n+1}$ using Algorithm 1.
 - 8 $n := n + 1$.
 - 9 **end**
- Output:** $\{\bar{\mathbf{S}}_k^n\}_{k=1}^K$ and apply the BC-MAC transformation to compute optimal $\{\tilde{\mathbf{S}}_k^n\}_{k=1}^K$.
-

B. Convergence Analysis

We will prove that Algorithm 2 produces a strictly decreasing objective following similar arguments to those in [20], [21]. Let $f(\boldsymbol{\lambda}, \{\bar{\mathbf{S}}_k\})$ be the objective of (4). Recall that $\boldsymbol{\Lambda}_k^n = \sum_{i=1}^L \lambda_i^n \mathbf{E}_{ik}$ is the optimal solution to the minimization (8), thus the following inequality holds:

$$\begin{aligned} f(\boldsymbol{\lambda}^n, \{\bar{\mathbf{S}}_k^n\}) &= \sum_{k=1}^K w_k \left(\log |\boldsymbol{\Phi}_k^n| - \log |\mathbf{V}_k^\dagger \boldsymbol{\Lambda}_k^n \mathbf{V}_k| \right) \geq \\ & \sum_{k=1}^K w_k \left(\log |\boldsymbol{\Phi}_k^n| + \text{tr}(\mathbf{V}_k \boldsymbol{\Phi}_k^{-n} \mathbf{V}_k^H (\boldsymbol{\Lambda}_k^{n+1} - \boldsymbol{\Lambda}_k^n)) \right. \\ & \quad \left. - \log |\mathbf{V}_k^\dagger \boldsymbol{\Lambda}_k^{n+1} \mathbf{V}_k| \right). \end{aligned} \quad (9)$$

Since $\log |\mathbf{V}_k^\dagger \boldsymbol{\Lambda}_k \mathbf{V}_k + \tilde{\mathbf{H}}_k^\dagger \bar{\mathbf{S}}_k \tilde{\mathbf{H}}_k|$ is jointly concave with $\boldsymbol{\Lambda}_k$ and $\{\bar{\mathbf{S}}_k\}$, and $\{\bar{\mathbf{S}}_k\}$ is the optimal solution to the maximization (6), it is easy to check that

$$\begin{aligned}
& \sum_{k=1}^K w_k \left(\log |\Phi_k^n| + \text{tr}(\mathbf{V}_k \Phi_k^{-n} \mathbf{V}_k^\dagger (\Lambda_k^{n+1} - \Lambda_k^n)) - \log |\mathbf{V}_k^\dagger \Lambda_k^{n+1} \mathbf{V}_k| \right) \\
& \geq \\
& \underbrace{\sum_{k=1}^K w_k \left(\log |\mathbf{V}_k^\dagger \Lambda_k^{n+1} \mathbf{V}_k + \tilde{\mathbf{H}}_k^\dagger \tilde{\mathbf{S}}_k^{n+1} \tilde{\mathbf{H}}_k| - \log |\mathbf{V}_k^\dagger \Lambda_k^{n+1} \mathbf{V}_k| \right)}_{f(\boldsymbol{\lambda}^{n+1}, \{\tilde{\mathbf{S}}_k^{n+1}\})}. \quad \text{CDF}
\end{aligned} \tag{10}$$

Combining these inequalities yields

$$\begin{aligned}
f(\boldsymbol{\lambda}^n, \{\tilde{\mathbf{S}}_k^n\}) & \geq \sum_{k=1}^K w_k \left(\log |\Phi_k^n| + \text{tr}(\mathbf{V}_k \Phi_k^{-n} \mathbf{V}_k^\dagger (\Lambda_k^{n+1} - \Lambda_k^n)) \right. \\
& \quad \left. - \log |\mathbf{V}_k^\dagger \Lambda_k^{n+1} \mathbf{V}_k| \right) \stackrel{(a)}{\geq} f(\boldsymbol{\lambda}^{n+1}, \{\tilde{\mathbf{S}}_k^{n+1}\}). \quad (11)
\end{aligned}$$

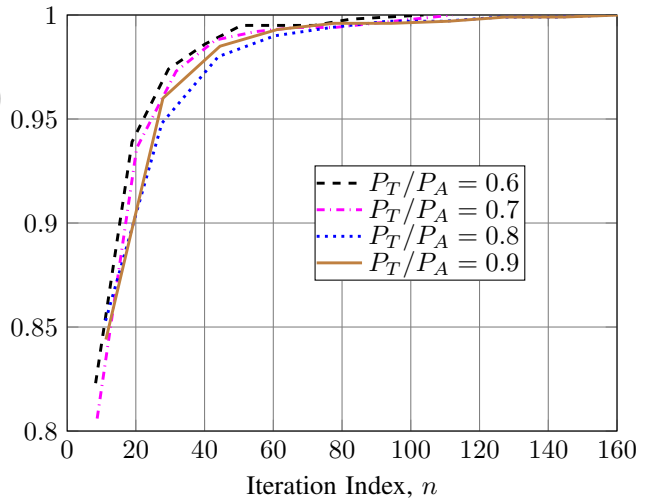
We notice that (a) is strict if $\{\Lambda_k^n\} \neq \{\Lambda_k^{n+1}\}$ or equivalently $\boldsymbol{\lambda}^n \neq \boldsymbol{\lambda}^{n+1}$. Thus, the sequence $\{f(\boldsymbol{\lambda}^n, \{\tilde{\mathbf{S}}_k^n\})\}$ is strictly decreasing unless it has converged. The convergence of the iterates to a limit point which is a saddle point can also be proved. However we omit the details due to space limitations.

C. Complexity Analysis

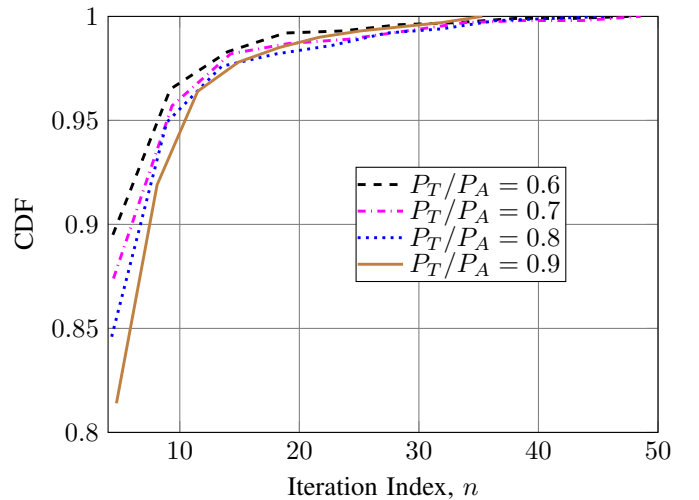
In this section, we analyze the complexity of the proposed algorithm by counting the number of flops [23], [24]. For simplicity, we assume that all the receivers have the same number of antennas i.e. $M_k = M$. Algorithm 2 performs the water-filling algorithm and eigenvalue decomposition to solve (6), which needs $K(N - (K - 1)M)^3 + K(4(N - (K - 1)M)^2 M - 8(N - (K - 1)M)M^2)$ flops [24]. In addition, the gradient computation in finding $\boldsymbol{\lambda}$ requires $(N - (K - 1)M)^3$ flops. The complexities of other steps are negligible and can be eliminated. Thus, the total per-iteration complexity of Algorithm 2 is $\mathcal{O}(KN^3)$ flops. Noticeably, the complexity of the proposed algorithm increases linearly with the number of users, which is desirable for dense networks.

IV. NUMERICAL RESULTS

In this section we numerically evaluate the performance of the proposed algorithm. In particular, we consider the case where both SPC and PAPC are considered at the BS. For notational convenience, we denote the SPC and the total power of PAPC as P_T and $P_A = \sum_{i=1}^N P_i$, respectively. Also, we are interested in the trivial case where $\min\{P_i\} < P_T < P_A$. In fact, if $\min\{P_i\} \geq P_T$, then the PAPC can be removed without loss of optimality. Similarly, if $P_T \geq P_A$, then the SPC can be eliminated. The power constraint for PAPC is set equally for all the transmit antennas, i.e., $P_i = P_A/N$, for $i = 1, \dots, N$. All users are equipped with the same number of receive antennas i.e., $M_k = M$. Unless stated otherwise, the number of receive antennas M and the error tolerances (i.e., ϵ_1, ϵ_2) are set to 2 and 10^{-6} respectively, for all simulations. Each result is obtained by averaging over 1000 i.i.d. channel



(a) The total power of per-antenna power constraint $P_A = 0$ dBW



(b) The total power of per-antenna power constraint $P_A = 10$ dBW

Fig. 1. Cumulative distribution function of the number of iterations to converge under joint SPC and PAPC. The number of transmit antennas $N = 4$, number of receive antennas $M = 2$ and number of users $K = 2$.

realizations. Other relevant simulation parameters are specified for each setup. System simulations were written in MATLAB and executed on a 64-bit desktop that supports 8 Gbyte RAM and Intel CORE i7.

In the first simulation, we plot the cumulative distribution functions (CDF) of the number of iterations taken by Algorithm 2 to converge. The low and high SNR scenarios as well as different power ratios i.e., P_T/P_A are studied. The CDF for each scenario is obtained over 1000 channel realizations. We can clearly see in Fig. 1 that 95% of the cases, Algorithm 2 terminates within 30 and 10 iterations for low and high power regions, respectively.

In Table I, we report the average run time for solving (3) by several approaches over 1000 channel realizations. As mentioned earlier, we can reformulate the considered problem as a MAXDET problem and then use a dedicated solver such as SDPT3 [15] to solve it optimally. For the

TABLE I

AVERAGE RUN TIME (SECONDS) COMPARISON FOR $P_T/P_A = 0.8$ AND $P_A = 10$ dBW, $M = 2$, $K = 4$. THE RUN TIME IS AVERAGED OVER 1000 CHANNEL REALIZATIONS.

No. of Tx. antennas N		16	32	64	128
Sum rate	Algorithm 2	0.290	1.298	9.645	97.183
	MOSEK	0.422	9.828	×	×
	SDPT3	1.427	13.992	×	×
WSRMax	Algorithm 2	0.406	1.794	10.383	122.073
	SDPT3	1.424	14.092	×	×

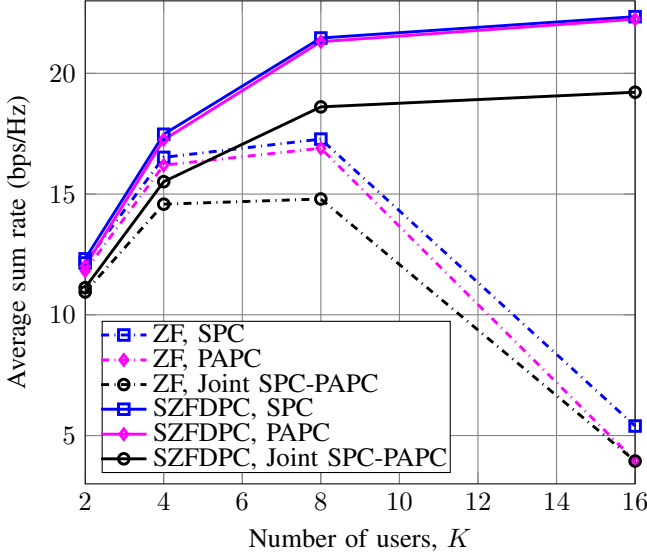


Fig. 2. Average sum rate of different precoding methods under different power constraints with the number of transmit antennas $N = 32$, the number of receive antennas $M = 2$; $P_T = 0$ dBW for SPC, $P_A = 0$ dBW for PAPC, $P_T/P_A = 0.8$ and $P_A = 0$ dBW for joint SPC and PAPC.

special case of the sum rate maximization problem, (3) can be further reformulated as a semidefinite program for which more powerful solvers e.g., MOSEK [25] can be invoked. Note that we execute these solvers through the parser YALMIP [26]. In Table I, ‘×’ stands for a case of extremely large computation time or a failure due to insufficient memory. Table I clearly shows that Algorithm 2 requires the lowest computation time among all approaches in comparison. In particular, the generic convex solvers i.e., SDPT3 and MOSEK only work effectively for small N . This result can be explained by the fact that interior-point-based solvers have a complexity that increases rapidly with the problem size.

Finally, we study the performance of the average sum rate of different precoding methods, including ZF, SZFDPC under various power constraint settings i.e., SPC, PAPC, and joint SPC and PAPC. For each of the precoding methods, the average sum rate with SPC is highest, followed by that with the PAPC. The sum rate with joint SPC and PAPC is lowest since adding more constraints to the transmitter optimization problem reduces the size of the feasible set. We can also see

that when the number of users increases, the performance of ZF methods decreases quickly. The reason is that the spatial multiplexing gains of ZF methods are greatly reduced when the number of users increases.

V. CONCLUSIONS

We have proposed an efficient algorithm for the WSRMax problem with ZF methods in multi-user MIMO channels subject to LTCCs. The considered problem has been reformulated as a minimax problem using the Lagrangian duality, and then AO and CCP are applied to obtain a semi-closed-form solution. We have benchmarked our proposed algorithm against generic optimization packages. The numerical results have demonstrated that the proposed algorithm is much faster and is able to deal with large-scale MIMO systems.

APPENDIX

PROOF OF THEOREM 1

Following the arguments in [12], we can prove the duality in Theorem 1. More explicitly, we first write the partial Lagrangian function of (3) as follows

$$\mathcal{L}(\{\tilde{\mathbf{S}}_k\}, \mathbf{a}) = \sum_{k=1}^K (w_k \log |\mathbf{I} + \tilde{\mathbf{H}}_k \tilde{\mathbf{S}}_k \tilde{\mathbf{H}}_k^\dagger| - \text{tr}(\mathbf{C}_k \tilde{\mathbf{S}})) + \mathbf{p}^T \mathbf{a} \quad (12)$$

where $\mathbf{C}_k = \mathbf{V}_k^\dagger \mathbf{A}_k \mathbf{V}_k$, $\mathbf{A}_k = \sum_i a_i \mathbf{E}_{ik}$, $\mathbf{a} = [a_1, a_2, \dots, a_L]^T$. Letting $\hat{\mathbf{S}}_k = \mathbf{C}_k^{1/2} \tilde{\mathbf{S}}_k \mathbf{C}_k^{1/2}$, we can rewrite $\mathcal{L}(\{\tilde{\mathbf{S}}_k\}, \{\mathbf{A}_k\})$ as:

$$\mathcal{L}(\{\hat{\mathbf{S}}_k\}, \mathbf{a}) = \sum_{k=1}^K (w_k \log |\mathbf{I} + \tilde{\mathbf{H}}_k \mathbf{C}_k^{-1/2} \hat{\mathbf{S}}_k \mathbf{C}_k^{-1/2} \tilde{\mathbf{H}}_k^\dagger| - \text{tr}(\hat{\mathbf{S}}_k)) + \mathbf{p}^T \mathbf{a}. \quad (13)$$

Denote $\mathbf{U}_k \boldsymbol{\Sigma}_k \mathbf{X}_k^\dagger$ to be the singular value decomposition of $\tilde{\mathbf{H}}_k \mathbf{C}_k^{-1/2}$, i.e., $\mathbf{U}_k \boldsymbol{\Sigma}_k \mathbf{X}_k^\dagger = \tilde{\mathbf{H}}_k \mathbf{C}_k^{-1/2}$, we can obtain the dual objective

$$\mathcal{D}(\{\mathbf{A}_k\}) = \max_{\hat{\mathbf{S}}_k \succeq 0} \sum_{k=1}^K (w_k \log |\mathbf{I} + \mathbf{C}_k^{-1/2} \tilde{\mathbf{H}}_k^\dagger \hat{\mathbf{S}}_k \tilde{\mathbf{H}}_k \mathbf{C}_k^{-1/2}| - \text{tr}(\hat{\mathbf{S}}_k)) + \mathbf{p}^T \mathbf{a} \quad (14)$$

where the relationship between $\hat{\mathbf{S}}_k$ and $\tilde{\mathbf{S}}_k$ is given by $\hat{\mathbf{S}}_k = \mathbf{X}_k \mathbf{U}_k^\dagger \tilde{\mathbf{S}}_k \mathbf{U}_k \mathbf{X}_k^\dagger$ [12]. Therefore, the dual problem is given as

$$\min_{\mathbf{a} \geq 0} \max_{\{\hat{\mathbf{S}}_k\} \succeq 0} \sum_{k=1}^K (w_k \log \frac{|\mathbf{V}_k^\dagger \mathbf{A}_k \mathbf{V}_k + \tilde{\mathbf{H}}_k^\dagger \hat{\mathbf{S}}_k \tilde{\mathbf{H}}_k|}{|\mathbf{V}_k^\dagger \mathbf{A}_k \mathbf{V}_k|} - \text{tr}(\hat{\mathbf{S}}_k)) + \mathbf{p}^T \mathbf{a}. \quad (15)$$

Now, by introducing a new optimization variable $\delta > 0$, (15) can be shown to be equivalent to

$$\begin{aligned} \min_{\mathbf{a} \geq 0, \delta > 0} \max_{\{\hat{\mathbf{S}}_k\} \succeq 0} & \sum_{k=1}^K w_k \log \frac{|\mathbf{V}_k^\dagger \mathbf{A}_k \mathbf{V}_k + \tilde{\mathbf{H}}_k^\dagger \hat{\mathbf{S}}_k \tilde{\mathbf{H}}_k|}{|\mathbf{V}_k^\dagger \mathbf{A}_k \mathbf{V}_k|} - \delta P \\ & + \mathbf{p}^T \mathbf{a} \\ \text{subject to} & \text{tr}(\hat{\mathbf{S}}) \leq \delta P. \end{aligned} \quad (16)$$

Defining

$$\bar{\mathbf{S}}_k = \frac{\dot{\mathbf{S}}_k}{\delta} \quad (17)$$

$$\boldsymbol{\lambda} = \frac{\mathbf{a}}{\delta}; \boldsymbol{\Lambda}_k = \sum_i \lambda_i \mathbf{E}_{ik} \quad (18)$$

and substituting these new optimization variables $\bar{\mathbf{S}}_k$ and $\boldsymbol{\lambda}$ into (16), we obtain

$$\begin{aligned} \min_{\lambda \geq 0, \delta > 0} \max_{\{\bar{\mathbf{S}}_k\}_{\geq 0}} & \sum_{k=1}^K w_k \log \frac{|\mathbf{V}_k^\dagger \boldsymbol{\Lambda}_k \mathbf{V}_k + \tilde{\mathbf{H}}_k^\dagger \bar{\mathbf{S}}_k \tilde{\mathbf{H}}_k|}{|\mathbf{V}_k^\dagger \boldsymbol{\Lambda}_k \mathbf{V}_k|} + \delta (\mathbf{p}^T \boldsymbol{\lambda} - P) \\ \text{subject to} & \sum_{k=1}^K \text{tr}(\bar{\mathbf{S}}_k) \leq P \end{aligned} \quad (19)$$

or, equivalently

$$\begin{aligned} \min_{\lambda \geq 0} \max_{\{\bar{\mathbf{S}}_k\}_{\geq 0}} & \sum_{k=1}^K w_k \log \frac{|\mathbf{V}_k^\dagger \boldsymbol{\Lambda}_k \mathbf{V}_k + \tilde{\mathbf{H}}_k^\dagger \bar{\mathbf{S}}_k \tilde{\mathbf{H}}_k|}{|\mathbf{V}_k^\dagger \boldsymbol{\Lambda}_k \mathbf{V}_k|} \\ \text{subject to} & \sum_{k=1}^K \text{tr}(\bar{\mathbf{S}}_k) \leq P \\ & \mathbf{p}^T \boldsymbol{\lambda} \leq P. \end{aligned} \quad (20)$$

Since the objective is decreasing with respect to each λ_i for a fixed $\bar{\mathbf{S}}_k$ and increasing with respect to each $\bar{\mathbf{S}}_k$, the inequalities of (20) can be made to be equalities without loss of optimality. As a result, we arrive at

$$\begin{aligned} \min_{\lambda \geq 0} \max_{\{\bar{\mathbf{S}}_k\}_{\geq 0}} & \sum_{k=1}^K w_k \log \frac{|\mathbf{V}_k^\dagger \boldsymbol{\Lambda}_k \mathbf{V}_k + \tilde{\mathbf{H}}_k^\dagger \bar{\mathbf{S}}_k \tilde{\mathbf{H}}_k|}{|\mathbf{V}_k^\dagger \boldsymbol{\Lambda}_k \mathbf{V}_k|} \\ \text{subject to} & \sum_{k=1}^K \text{tr}(\bar{\mathbf{S}}_k) = P \\ & \mathbf{p}^T \boldsymbol{\lambda} = P. \end{aligned} \quad (21)$$

which proves the theorem.

REFERENCES

- [1] H. Weingarten, Y. Steinberg, and S. S. Shamai, "The capacity region of the Gaussian multiple-input multiple-output broadcast channel," *IEEE Trans. Inf. Theory*, vol. 52, no. 9, pp. 3936–3964, Sep. 2006.
- [2] Q. Spencer, A. Swindlehurst, and M. Haardt, "Zero-forcing methods for downlink spatial multiplexing in multiuser MIMO channels," *IEEE Trans. Signal Process.*, vol. 52, no. 2, pp. 461–471, Feb. 2004.
- [3] K. Karakayali, R. Yates, G. Foschini, and R. Valenzuela, "Optimum zero-forcing beamforming with per-antenna power constraints," in *Proc. IEEE ISIT 2007*, Jun. 2007, pp. 101–105.
- [4] M. Vu, "MIMO capacity with per-antenna power constraint," in *Proc. IEEE GLOBECOM*, Dec. 2011, pp. 1–5.
- [5] S. Shi, M. Schubert, and H. Boche, "Per-antenna power constrained rate optimization for multiuser MIMO systems," in *Proc. WSA*, Feb. 2008, pp. 270–277.
- [6] W. Yu and T. Lan, "Transmitter optimization for the multi-antenna downlink with per-antenna power constraints," *IEEE Trans. Signal Process.*, vol. 55, no. 6, pp. 2646–2660, Jun. 2007.
- [7] A. Wiesel, Y. C. Eldar, and S. Shamai, "Zero-forcing precoding and generalized inverses," *IEEE Trans. Signal Process.*, vol. 56, no. 9, pp. 4409–4418, Sep. 2008.
- [8] P. L. Cao, T. J. Oechtering, R. F. Schaefer, and M. Skoglund, "Optimal transmit strategy for MISO channels with joint sum and per-antenna power constraints," *IEEE Trans. Signal Process.*, vol. 64, no. 16, pp. 4296–4306, August 2016.
- [9] S. Loyka, "The capacity of Gaussian MIMO channels under total and per-antenna power constraints," *IEEE Trans. Commun.*, vol. 65, no. 3, pp. 1035–1043, March 2017.
- [10] H. Huh, H. C. Papadopoulos, and G. Caire, "Multiuser MISO transmitter optimization for intercell interference mitigation," *IEEE Trans. Signal Process.*, vol. 58, no. 8, pp. 4272–4285, August 2010.
- [11] L.-N. Tran, M. Juntti, M. Bengtsson, and B. Ottersten, "Weighted sum rate maximization for MIMO broadcast channels using dirty paper coding and zero-forcing methods," *IEEE Trans. Commun.*, vol. 61, no. 6, pp. 2362–2373, Jun. 2013.

- [12] —, "Beamformer designs for MISO broadcast channels with zero-forcing dirty paper coding," *IEEE Trans. Wireless Commun.*, vol. 12, no. 3, pp. 1173–1185, Mar. 2013.
- [13] R. Zhang, "Cooperative multi-cell block diagonalization with per-base-station power constraints," *IEEE J. Sel. Areas Commun.*, vol. 28, no. 9, pp. 1435–1445, Dec. 2010.
- [14] L. Vandenberghe, S. Boyd, and S.-P. Wu, "Determinant maximization with linear matrix inequality constraints," *SIAM J. on Matrix Anal. and Appl.*, vol. 19, no. 2, pp. 499–533, 1998.
- [15] K. C. Toh, M. J. Todd, and R. Tutuncu, "SDPT3— a Matlab software package for semidefinite programming," *Optimization Methods and Software*, Nov. 1999.
- [16] W. Yu, "Uplink-downlink duality via minimax duality," *IEEE Trans. Inf. Theory*, vol. 52, no. 2, pp. 361–374, Feb. 2006.
- [17] L. Zhang, R. Zhang, Y. C. Liang, Y. Xin, and H. V. Poor, "On Gaussian MIMO BC-MAC duality with multiple transmit covariance constraints," *IEEE Trans. Inf. Theory*, vol. 58, no. 4, pp. 2064–2078, April 2012.
- [18] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.
- [19] L. Condat, "Fast projection onto the simplex and the ℓ_1 ball," *Mathematical Programming, Series A*, vol. 158, no. 1, pp. 575–585, Jul. 2016.
- [20] T. M. Pham, and R. Farrell, and L.-N. Tran, "Low-complexity approaches for MIMO capacity with per-antenna power constraint," in *Proc. IEEE VTC-Spring*, Jun. 2017.
- [21] —, "Alternating optimization for capacity region of Gaussian MIMO broadcast channels with per-antenna power constraint," in *Proc. IEEE VTC-Spring*, Jun. 2017.
- [22] L. Armijo, "Minimization of functions having Lipschitz continuous first partial derivatives," *Pacific J. Math.*, vol. 16, no. 1, pp. 1–3, 1966.
- [23] G. H. Golub and C. F. V. Loan, *Matrix Computations*, 3rd ed. The John Hopkins Univ. Press, 1996.
- [24] Z. Shen, R. Chen, J. Andrews, J. Heath, R.W., and B. Evans, "Low complexity user selection algorithms for multiuser MIMO systems with block diagonalization," *IEEE Trans. Signal Process.*, vol. 54, no. 9, pp. 3658–3663, Sep. 2006.
- [25] M. ApS, *The MOSEK optimization toolbox for MATLAB manual. Version 7.1 (Revision 28)*, 2015.
- [26] J. Löfberg, "YALMIP : A toolbox for modeling and optimization in MATLAB," in *Proc. the CACSD Conference*, 2004.