Signal Processing 179 (2021) 107828

Contents lists available at ScienceDirect

Signal Processing

journal homepage: www.elsevier.com/locate/sigpro

Short communication

MISO NOMA downlink beamforming optimization with per-antenna power constraints

Yongwei Huang[a,*], Longtao Zhou[b]

[a] School of Information Engineering, Guangdong University of Technology, University Town, Guangzhou, Guangdong 510006, China
[b] Guangzhou Yangchengtong Public Transportation Card Co. Ltd., Yuexiu District, Guangzhou 510080, China

ARTICLE INFO

Article history:
Received 5 August 2020
Accepted 27 September 2020
Available online 7 October 2020

Keywords:
Downlink beamforming
Non-orthogonal multiple access
Per-antenna power constraint
SOCP based approximate algorithm

ABSTRACT

Consider a multiuser downlink beamforming optimization problem for the non-orthogonal multiple access (NOMA) transmission in a multiple-input single-output (MISO) system with a general number of users. The total transmission power minimization problem is formulated subject to both quality-of-service (QoS) constraints under the NOMA principal and per-antenna power constraints, which are more realistic since each antenna has its own power amplifier and is limited individually by the linearity of the power amplifier. The problem is a nonconvex quadratically constrained quadratic program, and the conventional semidefinite program (SDP) relaxation is not tight. In order to tackle the non-convex problem, we construct second-order cone program (SOCP) approximation for each signal-to-interference-plus-noise ratio (SINR) constraint and form an iterative algorithm, in which a sequence of SOCPs are solved. The optimal values of SOCPs in the sequence are proved to be non-increasing, and converging to a locally optimal value. In particular, we show that the SDP relaxation is tight for two-user case if one of the SINR constraints is strict (non-binding) at the optimality. Detailed simulation results are presented to demonstrate the performance gains of the NOMA downlink beamforming with per-antenna power constraints through the proposed algorithm, which is compared with some state-of-the-art methods.

1. Introduction

In a multiuser system, the effective utilization of non-orthogonal multiple access (NOMA) beamforming has been more and more popular, due to its features of improvement of spectral efficiency, user fairness, the system throughput and so on. The NOMA beamforming allows the multi-antenna base station (BS) to apply superposition coding (SC) using the spatial degree of freedom, and the receivers to conduct successive interference cancellation (SIC) with manageable costs, often assuming perfect channel state information known at the BS. Nowaday, the NOMA beamforming has been an important radio access technique, which is suitable for the fifth generation (5G) wireless networks [1,2].

In this work, we deal with an NOMA beamforming optimization problem with per-antenna power constraints in setting of a multiple-input single-output (MISO) multiuser downlink transmission system with a general number of users. In contrast to the sum-power constraint across all transmit antennas, the per-antenna power constraints are more realistic in practical implementations, since every antenna in the transmit array is equipped with its own power amplifier. As a consequence, the antenna is limited individually by the linearity of the power amplifier (see e.g. [3]). In this regard, we formulate a total transmission power minimization problem subject to both per-antenna power constraints and quality-of-service (QoS) constraints under the NOMA scheme, and propose a low per-iteration complexity algorithm based on second-order cone program (SOCP) approximation.

In order to proceed, let us mention some related works. In [4], multicast beamforming with SC for multiresolution broadcast with two users is considered, and the optimal beamforming vectors are obtained by minimizing the total transmission power for given target rates in a broadcast system with multiresolution transmissions where the interference is mitigated at a user using SIC. As well, the proposed multicast beamforming with SC is applied to NOMA systems as a two-stage beamforming method. A sum rate maximization problem in a MISO downlink system is studied in Hanif et al. [5] relying on NOMA principles, and a concave-convex procedure (CCP) based iterative algorithm is developed to solve the NOMA sum rate maximization problem. In [6], the application of the concept of quasi-degradation to a multi-user MISO downlink is considered, by employing the idea of user pairing. In particular, a QoS optimization problem with two users is formulated into

* Corresponding author.
E-mail addresses: ywhuang@gdut.edu.cn (Y. Huang), zhoult@gzyct.com (L. Zhou).

https://doi.org/10.1016/j.sigpro.2020.107828

minimization of the total transmit power constrained by the target individual rate, and closed-form expressions for different precoding algorithms including dirty-paper coding (DPC) and zero-forcing beamforming (ZFBF) are obtained. Then the hybrid NOMA precoding algorithm is presented, and it is combined with the proposed user pairing algorithm to yield practical transmission schemes. In [7], several efficient optimization algorithms for multiuser SC beamforming (SCBF) is studied to solve power minimization and rate maximization problems, and combination of SCBF with ZFBF through user grouping is proposed. Optimal results are guaranteed to be found for the two-user case, and it is experimentally shown that a nearly optimal performance can be achieved for the three user case with practical problem setups.

In this paper, we propose an SOCP based approximate algorithm for the MISO NOMA downlink beamforming optimization problem. The problem is formulated into the total transmission power minimization subject to the per-antenna power constraints and NOMA signal-to-interference-plus-noise rate (SINR) constraints. In the optimization problem, the number of the SINR constraints is bigger than that of those constraints in the traditional downlink beamforming problem in Bengtsson and Ottersten [8], and unlike the results in Yu and Lan [3], the semidefinite program (SDP) relaxation is not tight any more [9]. Nevertheless, we take into account an SOCP approximation for each SINR constraint for any given initial point, and the new approximation problem is an SOCP, which can be solved efficiently. The optimal solution of the SOCP is treated as a given point for the next iterative step. In the way, we obtain a sequence of SOCPs with the property that the optimal values are non-increasing. Therefore, the sequence of optimal solutions converges to a locally optimal solution. Our extensive simulations show that as long as the number of users is not high, the locally optimal value is very close to the optimal value of the corresponding SDP relaxation problem, which is our benchmark. In particular, when the two-user NOMA downlink system is considered, we show that the SDP relaxation is tight if one of SINR inequality constraints is strict at the optimality. In addition, simulation examples are presented in order to demonstrate the performance gains of the proposed algorithm, which is compared with state-of-the-art methods.

The paper is organized as follows. In Section 2, we introduce the system model and formulate the MISO NOMA downlink beamforming problem with per-antenna power constraints. In Section 3, we propose the an SOCP based approximate algorithm to solve the NOMA downlink beamforming problem. In Section 4, we present numerical examples showing the performance of the proposed algorithm in difference scenarios. Finally, the paper is concluded in Section 5.

2. System model and problem formulation

Consider a wireless NOMA downlink system where a K-antenna BS serves M single-antenna users. The transmitted signal by the BS is expressed as

$$\boldsymbol{x} = \sum_{m=1}^{M} s_m \boldsymbol{w}_m,$$

where $\boldsymbol{w}_m \in \mathbb{C}^K$ and s_m are the downlink beamforming vector and the information symbol for user m (with zero mean and unit variance), respectively. The received signal by user m is given by

$$y_m = \boldsymbol{h}_m^H \boldsymbol{x} + n_m,$$

where $\boldsymbol{h}_m$ is the channel vector between the BS and user m, and n_m is a zero-mean complex additive white Gaussian noise with variance σ_m^2 at user m.

In order to conduct SIC at the users, we need to build a decoding sequence, which is associated with power level of the users. We suppose that $\boldsymbol{h}_m$, $m = 1, \ldots, M$, follow the Rician channel model [10]:

$$\boldsymbol{h}_m = \sqrt{\beta_m}\left(\sqrt{\frac{\zeta}{1+\zeta}}\boldsymbol{a}(\theta_m) + \sqrt{\frac{1}{1+\zeta}}\boldsymbol{u}_m\right), \tag{1}$$

where $\boldsymbol{u}_m$ is a complex-valued Guassian random vector with zero mean and $\frac{1}{K}\boldsymbol{I}$ covariance, and

$$\boldsymbol{a}(\theta) = \frac{1}{\sqrt{K}}[1; e^{-j2\pi(d/\lambda)\sin\theta}; \ldots; e^{-j(K-1)2\pi(d/\lambda)\sin\theta}]$$

(a column vector) is the array respond vector (steering vector) for a uniform linear array (ULA) of half-wavelength spacing. Here θ_m is the angle of departure (AoD) to user m, and the large scaling fading factor β_m is given by $1/(d_m)^\eta$, where d_m is distance between the BS and receiver m, and η is the path loss exponent (a non-negative number). Clearly, if $\zeta = 0$, then $\boldsymbol{h}_m = \sqrt{\beta_m}\boldsymbol{u}_m$ which is a Rayleigh fading channel model (e.g., see [7]). If $\zeta = +\infty$, then $\boldsymbol{h}_m = \sqrt{\beta_m}\boldsymbol{a}(\theta_m)$, which is the scenario that the BS is equipped with a ULA.

Now suppose that the user set $\mathcal{S} = \{u_1, u_2, \ldots, u_M\}$ is an ordered set[1], and a user with stronger channel condition (the distance between the BS and the user is smaller) has a bigger index. Therefore, u_m can decode the information for u_n with $n \le m$, under the QoS conditions

$$\min_{n\le m\le M}\{\mathrm{SINR}_m^n\} \ge \gamma_n, \ 1 \le n \le M$$

where the SINR_m^n is given by

$$\mathrm{SINR}_m^n = \frac{|\boldsymbol{h}_m^H \boldsymbol{w}_n|^2}{\sum_{i=n+1}^{M} |\boldsymbol{h}_m^H \boldsymbol{w}_i|^2 + \sigma_m^2}, \ n \le m, \tag{2}$$

and

$$\gamma_n = 2^{R_n} - 1$$

with R_n being the target rate for user u_n, associated with its minimum QoS requirement.

Therefore, we consider the following NOMA downlink beamforming with per-antenna power constraints:

$$\underset{\{\boldsymbol{w}_m\}}{\text{minimize}} \quad \sum_{m=1}^{M} \boldsymbol{w}_m^H \boldsymbol{w}_m \tag{3a}$$

$$\text{subject to} \quad \boldsymbol{e}_k^H\left(\sum_{m=1}^{M} \boldsymbol{w}_m \boldsymbol{w}_m^H\right)\boldsymbol{e}_k \le P_k, \ k = 1, \ldots, K, \tag{3b}$$

$$\min_{n\le m\le M}\{\mathrm{SINR}_m^n\} \ge \gamma_n, \ 1 \le n \le M, \tag{3c}$$

where $\boldsymbol{e}_k$ is the kth column of the identity matrix of dimension K. Evidently, (3b) includes the per-antenna power constraints and P_k is a given power upper bound of each per-antenna constraint. Such a per-antenna basis power constraint is more realistic than a sum power constraint, since each antenna in the array has its own power amplifier in its analog front-end, and is limited individually by the linearity of the power amplifier (cf. [3]). The constraints in (3c) are the user QoS conditions for a NOMA downlink system. Alternatively, we note that the following downlink beamforming problem with per-antenna power constraints can be

[1] As for how to determine an optimal decoding sequence for a general number of users is an open problem [11]. Here we assume that the order is associated with the distance between the BS and a user due to the assumption that the channel vectors $\boldsymbol{h}_m$ follow the Rician channel model (see (1)).

considered:

$$\underset{\{\boldsymbol{w}_m\},\alpha}{\text{minimize}} \quad \alpha \tag{4a}$$

$$\text{subject to} \quad \boldsymbol{e}_k^H\left(\sum_{m=1}^{M}\boldsymbol{w}_m\boldsymbol{w}_m^H\right)\boldsymbol{e}_k \le \alpha P_k,\ k=1,\dots,K, \tag{4b}$$

$$\min_{n\le m\le M}\{\text{SINR}_m^n\} \ge \gamma_n,\ 1\le n\le M, \tag{4c}$$

This is an extension of the problem studied in Yu and Lan [3]. We will first focus on problem (3) and design an efficient approximate algorithm, and then show that the algorithm is applicable to (4). Before proceeding, let us mention some relations between (3) and (4). Clearly, if there is a feasible solution $(\{\boldsymbol{w}_m\},\ \alpha)$ for (4) with $\alpha\ge 1$, then the feasible set of (4) (fixing this α) contains that of (3). Suppose that $(\{\boldsymbol{w}_m^\star\},\alpha^\star)$ is an optimal solution for (4). If $\alpha^\star\le 1$ (i.e. $\alpha^\star P_k\le P_k$), then the optimal solution $\{\boldsymbol{w}_m^\star\}$ for problem (4) is a feasible point for problem (3). As a result, the minimal total transmission power obtained by (3) is less than or equal to the total transmission power obtained by (4).

3. An SOCP based approximate algorithm for the NOMA downlink beamforming problem

In this section, we establish an SOCP based approximate algorithm for (3). To this end, invoking the SINR expressions (2), NOMA downlink beamforming problem (3) is rewritten into a separable quadratically constrained quadratic program (QCQP):

$$\underset{\{\boldsymbol{w}_m\}}{\text{minimize}} \quad \sum_{m=1}^{M}\boldsymbol{w}_m^H\boldsymbol{w}_m \tag{5a}$$

$$\text{subject to} \quad \left(\sum_{m=1}^{M}|\boldsymbol{e}_k^H\boldsymbol{w}_m|^2\right)\le P_k,\ k=1,\dots,K, \tag{5b}$$

$$|\boldsymbol{h}_m^H\boldsymbol{w}_n|^2 \ge \gamma_n\sum_{i=n+1}^{M}|\boldsymbol{h}_m^H\boldsymbol{w}_i|^2+\gamma_n\sigma_m^2, \tag{5c}$$

$$n\le m\le M,\ 1\le n\le M.$$

Observe that the per-antenna power constraints are convex but the NOMA SINR constraints are non-convex. Furthermore, the total number of SINR constraints is $M(M+1)/2$, instead of M as in Yu and Lan [3]. Therefore, the traditional SDP relaxation for (5) is not tight any longer (e.g. see [9]), which indicates that we need to establish a new method without involving the SDP relaxation.

3.1. A general number of users case

In this subsection, we study the scenario of a general number of users in the NOMA downlink system.

Introducing auxiliary variables $\{t_{mn}\}$, problem (5) can be further recast into

$$\underset{\{\boldsymbol{w}_m\},\{t_{mn}\}}{\text{minimize}} \quad \sum_{m=1}^{M}\boldsymbol{w}_m^H\boldsymbol{w}_m \tag{6a}$$

$$\text{subject to} \quad (5b)\ \text{satisfied}, \tag{6b}$$

$$|\boldsymbol{h}_m^H\boldsymbol{w}_n|\ge t_{mn} \tag{6c}$$

$$t_{mn}\ge\sqrt{\gamma_n\sum_{i=n+1}^{M}|\boldsymbol{h}_m^H\boldsymbol{w}_i|^2+\gamma_n\sigma_m^2}, \tag{6d}$$

$$n\le m\le M,\ 1\le n\le M.$$

It is seen that beamforming problem (6) includes the nonconvex constraints (all others are convex):

$$|\boldsymbol{h}_m^H\boldsymbol{w}_n|\ge t_{mn},\ n\le m\le M,\ 1\le n\le M.$$

Therefore, it is hard to get a globally optimal solution for (6), and as a compromise we consider to develop an approximate algorithm for it (in order to obtain an approximate/suboptimal solution). Toward the end, we suppose that $\{\boldsymbol{w}_1^0,\dots,\boldsymbol{w}_M^0\}$ is any given initial point (to be updated). Observe that

$$|\boldsymbol{h}_m^H\boldsymbol{w}_n|\ge\frac{|\boldsymbol{w}_n^H\boldsymbol{h}_m\boldsymbol{h}_m^H\boldsymbol{w}_n^0|}{|\boldsymbol{h}_m^H\boldsymbol{w}_n^0|}\ge\frac{\Re(\boldsymbol{w}_n^H\boldsymbol{h}_m\boldsymbol{h}_m^H\boldsymbol{w}_n^0)}{|\boldsymbol{h}_m^H\boldsymbol{w}_n^0|}. \tag{7}$$

It follows that if

$$\frac{\Re(\boldsymbol{w}_n^H\boldsymbol{h}_m\boldsymbol{h}_m^H\boldsymbol{w}_n^0)}{|\boldsymbol{h}_m^H\boldsymbol{w}_n^0|}\ge t_{mn},$$

then one has

$$|\boldsymbol{h}_m^H\boldsymbol{w}_n|\ge t_{mn}.$$

Thereby, let us consider the following beamforming problem in the SOCP form:

$$\underset{\{\boldsymbol{w}_m\},\{t_{mn}\}}{\text{minimize}} \quad \sum_{m=1}^{M}\boldsymbol{w}_m^H\boldsymbol{w}_m \tag{8a}$$

$$\text{subject to} \quad (5b),\ (6d)\ \text{satisfied}, \tag{8b}$$

$$\frac{\Re(\boldsymbol{w}_n^H\boldsymbol{h}_m\boldsymbol{h}_m^H\boldsymbol{w}_n^0)}{|\boldsymbol{h}_m^H\boldsymbol{w}_n^0|}\ge t_{mn},\ n\le m\le M,\ 1\le n\le M. \tag{8c}$$

Observe that for the given (initial) point $(\boldsymbol{w}_1^0,\dots,\boldsymbol{w}_M^0)$, problem (8) is always a convex SOCP restriction problem of the orginal NOMA downlink beamforming problem (5) (in other words, the feasible set of (8) is contained in that of (5)), rather than an SOCP relaxation problem (see e.g. [12] for a reference).

To further proceed the analysis, let $l:=1$. Solve the SOCP (8), finding a solution $(\boldsymbol{w}_1^l,\dots,\boldsymbol{w}_M^l;\{t_{mn}^l\})$. Again observe that

$$|\boldsymbol{h}_m^H\boldsymbol{w}_n|\ge\frac{|\boldsymbol{w}_n^H\boldsymbol{h}_m\boldsymbol{h}_m^H\boldsymbol{w}_n^l|}{|\boldsymbol{h}_m^H\boldsymbol{w}_n^l|}\ge\frac{\Re(\boldsymbol{w}_n^H\boldsymbol{h}_m\boldsymbol{h}_m^H\boldsymbol{w}_n^l)}{|\boldsymbol{h}_m^H\boldsymbol{w}_n^l|},$$

and that

$$\frac{\Re(\boldsymbol{w}_n^H\boldsymbol{h}_m\boldsymbol{h}_m^H\boldsymbol{w}_n^l)}{|\boldsymbol{h}_m^H\boldsymbol{w}_n^l|}\ge t_{mn}\ \text{implies}\ |\boldsymbol{h}_m^H\boldsymbol{w}_n|\ge t_{mn}.$$

Then, similarly construct another convex restriction of (5) based on $(\boldsymbol{w}_1^l,\dots,\boldsymbol{w}_M^l)$:

$$\underset{\{\boldsymbol{w}_m\},\{t_{mn}\}}{\text{minimize}} \quad \sum_{m=1}^{M}\boldsymbol{w}_m^H\boldsymbol{w}_m \tag{9a}$$

$$\text{subject to} \quad (5b),\ (6d)\ \text{satisfied}, \tag{9b}$$

$$\frac{\Re(\boldsymbol{w}_n^H\boldsymbol{h}_m\boldsymbol{h}_m^H\boldsymbol{w}_n^l)}{|\boldsymbol{h}_m^H\boldsymbol{w}_n^l|}\ge t_{mn},\ n\le m\le M,\ 1\le n\le M. \tag{9c}$$

Then, solve (9) getting $(\boldsymbol{w}_1^{l+1},\dots,\boldsymbol{w}_M^{l+1};\{t_{mn}^{l+1}\})$. Set $l:=l+1$ and solve (9), and in the way an iterative procedure is formed. Let

$$v_l=\sum_{m=1}^{M}\|\boldsymbol{w}_m^l\|^2,\ l=1,\dots,$$

stand for the optimal values. We are about to show the important property that $\{v_l\}$ is a non-increasing sequence.

Proposition 3.1. *It holds that $v_l\ge v_{l+1}$ for $l\ge 1$.*

Proof. The idea is to prove that the optimal solution $(\boldsymbol{w}_1^l,\ldots,\boldsymbol{w}_M^l;\{t_{mn}^l\})$ for problem (8) with $\boldsymbol{w}_n^0$ replaced with $\boldsymbol{w}_n^{l-1}$, is feasible for (9). If that is the case, the optimal solution $(\boldsymbol{w}_1^{l+1},\ldots,\boldsymbol{w}_M^{l+1};\{t_{mn}^{l+1}\})$ for (9) simply has the property:

$$v_l=\sum_{m=1}^M(\boldsymbol{w}_m^l)^H\boldsymbol{w}_m^l\geq\sum_{m=1}^M(\boldsymbol{w}_m^{l+1})^H\boldsymbol{w}_m^{l+1}=v_{l+1},$$

for $l\geq 1$.

We have an immediate check whether $(\boldsymbol{w}_1^l,\ldots,\boldsymbol{w}_M^l;\{t_{mn}^l\})$ is feasible for (9). Since it is optimal for problem (8) with $\boldsymbol{w}_n^0$ replaced with $\boldsymbol{w}_n^{l-1}$, then it is feasible too and thus, (8c) are satisfied:

$$\frac{\Re((\boldsymbol{w}_n^l)^H\boldsymbol{h}_m\boldsymbol{h}_m^H\boldsymbol{w}_n^{l-1})}{|\boldsymbol{h}_m^H\boldsymbol{w}_n^{l-1}|}\geq t_{mn}^l,$$

which imply that

$$|\boldsymbol{h}_m^H\boldsymbol{w}_n^l|\geq t_{mn}^l.$$

Accordingly, we have

$$\frac{\Re((\boldsymbol{w}_n^l)^H\boldsymbol{h}_m\boldsymbol{h}_m^H\boldsymbol{w}_n^l)}{|\boldsymbol{h}_m^H\boldsymbol{w}_n^l|}=|\boldsymbol{h}_m^H\boldsymbol{w}_n^l|\geq t_{mn}^l, \tag{10}$$

It is evident that (5b) and (6d) are also fulfilled:

$$\left(\sum_{m=1}^M|\boldsymbol{e}_k^H\boldsymbol{w}_m^l|^2\right)\leq P_k,\ k=1,\ldots,K, \tag{11}$$

and

$$t_{mn}^l\geq\sqrt{\gamma_n\sum_{i=n+1}^M|\boldsymbol{h}_m^H\boldsymbol{w}_i^l|^2+\gamma_n\sigma_m^2}, \tag{12}$$

for $n\leq m\leq M, 1\leq n\leq M$, respectively. It follows from (10)–(12) that $(\boldsymbol{w}_1^l,\ldots,\boldsymbol{w}_M^l;\{t_{mn}^l\})$ is feasible for (9). □

Remark that problems (9) for $l=1,2,\ldots,$ are convex restrictions of (5) and the optimal values are non-increasing, namely, $v_1\geq v_2\geq\ldots$. Therefore the sequence of $(\boldsymbol{w}_1^l,\ldots,\boldsymbol{w}_M^l)$ converges to a locally minimal point for (5) (our extensive numerical simulations show that it goes to the globally minimal point when the number of users is not too high).

Note that we can remove the auxiliary variables, and reformulate (9) equivalently into the compact form

$$\underset{\{\boldsymbol{w}_m\}}{\text{minimize}}\ \sum_{m=1}^M\boldsymbol{w}_m^H\boldsymbol{w}_m \tag{13a}$$

$$\text{subject to}\ \ (5b)\ \text{satisfied}, \tag{13b}$$

$$\frac{\Re(\boldsymbol{w}_n^H\boldsymbol{h}_m\boldsymbol{h}_m^H\boldsymbol{w}_n^l)}{|\boldsymbol{h}_m^H\boldsymbol{w}_n^l|}\geq\sqrt{\gamma_n\sum_{i=n+1}^M|\boldsymbol{h}_m^H\boldsymbol{w}_i|^2+\gamma_n\sigma_m^2}\quad n\leq m\leq M,\ 1\leq n\leq M. \tag{13c}$$

Therefore, we summarize an approximate algorithm based on the SOCP restrictions as in Algorithm 1.

As for the computational complexity, it mainly includes solving SOCP (13) in every iteration. Let us check the worst-case computational complexity for (13) based on [13, page 335]. There are $K+1+M(M+1)/2$ constraints in the SOCP, the length of the optimization (real-valued) variables are $2NM+1$, and $(K+1)M^2+\sum_{i=1}^{M(M+1)/2}k_i^2$ is equal to $(K+1)M^2+\sum_{m=1}^M M^3=(K+1)M^2+(M(M+1)/2)^2$. Therefore, the total complexity is of order $O((M(M+1)/2+K+2)^{1/2}(2NM+1)((2NM+1)^2+(K+1)(M^2+1)+(M(M+1))/2+(M^2(M+1)^2)/4)$, which is approximately equal to $O(\sqrt{K+M^2}NM^3(8N^2+2K+M^2/2))$. Recall that the complexity is the worst-cast theoretical value, and it is not real computational cost.

Algorithm 1 An approximate algorithm for (5)

`Input:` $\{\boldsymbol{h}_m\}$, $\{\sigma_m\}$, $\{\gamma_n\}$, $\{P_k\}$, M, K, ζ, η, ξ;
`Output:` A solution $\{\boldsymbol{w}_m^\star\}$ for problem (5);
1: Suppose that $\{\boldsymbol{w}_m^0\}$ is an initial point; set $l=0$ and $v_0=\sum_{m=1}^M\|\boldsymbol{w}_m^0\|^2$ (a large number);
2: do
3: solve SOCP (13) with the given point $\{\boldsymbol{w}_m^l\}$, obtaining optimal solution $\{\boldsymbol{w}_m^{l+1}\}$ and optimal value v_{l+1};
4: $l:=l+1$;
5: **until** $v_{l-1}-v_l\leq\xi$
6: output $\{\boldsymbol{w}_m^l\}$.

We remark that from optimization point of view, problem (5) with $\boldsymbol{h}_m\boldsymbol{h}_m^H$ replaced with $\boldsymbol{H}_m$ (general rank) can be solved similarly by using the approximate algorithm.

As for how to solve (4), we construct the following SOCP restriction problem (similar to (13):

$$\underset{\{\boldsymbol{w}_m\},\alpha}{\text{minimize}}\ \ \alpha \tag{14a}$$

$$\text{subject to}\ \ (4b),(13c)\ \text{satisfied}, \tag{14b}$$

for $l=0,1,\ldots$. Like the proof of Proposition 3.1, it is not hard to show that the feasible set of SOCP (14) in the lth step is always contained in that of the SOCP in the $(l+1)$th step, for $l=1,2,\ldots$. Therefore, we have $\alpha_l\geq\alpha_{l+1}$ for all $l\geq 1$. In other words, Algorithm 1 is applicable to solve (4), simply by changing SOCP (13) in step 3 into (14).

Now, we would like to mention the difference between our method and the known CCP, which can be found, for example, in Lipp and Boyd [14]. In order to approximate SINR constraints (5c), the CCP includes a step of linearizing the left hand side $|\boldsymbol{h}_m^H\boldsymbol{w}_n|^2$ of (5c), such that the SINR constraints are turned into second-order cone (SOC) constraints. Appendix A.1 contains the procedure about how to construct an SOCP problem to be solved in each iteration in the CCP.

In contrast to (A.5), to approximate the original SINR constraints (i.e. (5c)), we utilize the SOC constraints (13c), which are variations of the SOC:

$$\left\{\begin{bmatrix}t\\ \boldsymbol{x}\end{bmatrix}\in\mathbb{R}\times\mathbb{R}^{2(M-n)+1}\ \middle|\ t\geq\|\boldsymbol{x}\|\right\} \tag{15}$$

and then problem (13) is solved in each iteration in our method. Comparing the CCP with our method, we see that the expression (13c) appears simpler than (A.3). In other words, the dimension of the SOC (15) is lower than that of (A.4), which means that the computational cost will be less.

In addition to the CCP, another SOCP (restriction) approximate method, called feasible point pursuit successive convex approximation (FPP-SCA) algorithm, has been proposed in Mehanna et al. [15]. Let us see the difference between the FPP-SCA algorithm and our proposed method. Appendix A.2 includes the procedure about how the FPP-SCA is applied to solve (5) and the SOCP restriction problem is constructed.

Comparing (A.10) with (13), we observe that (i) the SOC constraint (13c) appears simpler than (A.8); that is, the dimension of the SOC (see (15)) in our method is less than that of the SOC (A.8) (see (A.9)), which means that the computational cost of our method will be less; (ii) the slack variables shall be added to (A.10), and a parameter λ has to be selected properly, which makes the FPP-SCA algorithm more complicate.

More recently, the consensus alternating direction method of multipliers (consensus-ADMM, see [16]) has been proposed to solve a QCQP. However, how to apply the consensus-ADMM to a separable QCQP (like (5)) has not been mentioned therein. Therefore, we need to make some manipulations in order to implement the consensus-ADMM to solve (5). Fortunately, the iteration steps of the consensus-ADMM can be adapted according to Section 3.1 and 3.2 in Huang and Sidiropoulos [16], and we omit the details due to the limited paper room. In the simulation part, we will compare our proposed algorithm with the consensus-ADMM, in terms of the total transmission power and the cpu-time.

3.2. Two-user case

If there are two users in the NOMA downlink transmission system, then problem (5) is specified to

$$\underset{\{\boldsymbol{w}_1,\boldsymbol{w}_2\}}{\text{minimize}} \;\; \boldsymbol{w}_1^H\boldsymbol{w}_1+\boldsymbol{w}_2^H\boldsymbol{w}_2 \tag{16a}$$

$$\text{subject to} \;\; |\boldsymbol{e}_k^H\boldsymbol{w}_1|^2+|\boldsymbol{e}_k^H\boldsymbol{w}_2|^2\le P_k,\; k=1,\dots,K \tag{16b}$$

$$\boldsymbol{w}_1^H\boldsymbol{h}_1\boldsymbol{h}_1^H\boldsymbol{w}_1-\gamma_1\boldsymbol{w}_2^H\boldsymbol{h}_1\boldsymbol{h}_1^H\boldsymbol{w}_2\ge\gamma_1\sigma_1^2 \tag{16c}$$

$$\boldsymbol{w}_1^H\boldsymbol{h}_2\boldsymbol{h}_2^H\boldsymbol{w}_1-\gamma_1\boldsymbol{w}_2^H\boldsymbol{h}_2\boldsymbol{h}_2^H\boldsymbol{w}_2\ge\gamma_1\sigma_2^2 \tag{16d}$$

$$\boldsymbol{w}_2^H\boldsymbol{h}_2\boldsymbol{h}_2^H\boldsymbol{w}_2\ge\gamma_2\sigma_2^2, \tag{16e}$$

for which the traditional SDP relaxation is

$$\underset{\{\boldsymbol{W}_1,\boldsymbol{W}_2\}}{\text{minimize}} \;\; \text{tr}\,\boldsymbol{W}_1+\text{tr}\,\boldsymbol{W}_2 \tag{17a}$$

$$\text{subject to} \;\; \text{tr}\,(\boldsymbol{e}_k\boldsymbol{e}_k^H(\boldsymbol{W}_1+\boldsymbol{W}_2))\le P_k,\; k=1,\dots,K \tag{17b}$$

$$\text{tr}\,(\boldsymbol{h}_1\boldsymbol{h}_1^H(\boldsymbol{W}_1-\gamma_1\boldsymbol{W}_2^H))\ge\gamma_1\sigma_1^2 \tag{17c}$$

$$\text{tr}\,(\boldsymbol{h}_2\boldsymbol{h}_2^H(\boldsymbol{W}_1-\gamma_1\boldsymbol{W}_2^H))\ge\gamma_1\sigma_2^2 \tag{17d}$$

$$\text{tr}\,(\boldsymbol{h}_2\boldsymbol{h}_2^H\boldsymbol{W}_2)\ge\gamma_2\sigma_2^2 \tag{17e}$$

$$\boldsymbol{W}_1\succeq\boldsymbol{0},\;\boldsymbol{W}_2\succeq\boldsymbol{0}. \tag{17f}$$

The dual is the following maximization problem:

$$\underset{\{y_m\},\{x_k\}}{\text{maximize}} \;\; y_1\gamma_1\sigma_1^2+y_2\gamma_1\sigma_2^2+y_3\gamma_2\sigma_2^2-\sum_{k=1}^K x_kP_k \tag{18a}$$

$$\text{subject to} \;\; \boldsymbol{I}-y_1\boldsymbol{h}_1\boldsymbol{h}_1^H-y_2\boldsymbol{h}_2\boldsymbol{h}_2^H+\sum_{k=1}^K x_k\boldsymbol{e}_k\boldsymbol{e}_k^H\succeq\boldsymbol{0} \tag{18b}$$

$$\boldsymbol{I}+y_1\gamma_1\boldsymbol{h}_1\boldsymbol{h}_1^H+y_2\gamma_1\boldsymbol{h}_2\boldsymbol{h}_2^H-y_3\boldsymbol{h}_2\boldsymbol{h}_2^H+\sum_{k=1}^K x_k\boldsymbol{e}_k\boldsymbol{e}_k^H\succeq\boldsymbol{0} \tag{18c}$$

$$y_1\ge 0, y_2\ge 0, y_3\ge 0, x_k\ge 0, k=1,\dots,K. \tag{18d}$$

Theorem 3.2. *Suppose that $(\boldsymbol{W}_1^\star,\boldsymbol{W}_2^\star)$ is an optimal solution for SDP (17) and $\boldsymbol{W}_i^\star\neq\boldsymbol{0}$ for $i=1,2$. Then $\boldsymbol{W}_2^\star$ is of rank one. If either $\text{tr}\,(\boldsymbol{h}_1\boldsymbol{h}_1^H(\boldsymbol{W}_1^\star-\gamma_1\boldsymbol{W}_2^{\star H}))>\gamma_1\sigma_1^2$ or $\text{tr}\,(\boldsymbol{h}_2\boldsymbol{h}_2^H(\boldsymbol{W}_1^\star-\gamma_1\boldsymbol{W}_2^{\star H}))>\gamma_1\sigma_2^2$, then $\boldsymbol{W}_1^\star$ is a rank-one matrix.*

The proof is immediate if one verifies the complementary conditions (i.e., the complementary slackness in the KKT optimality conditions). We hence omit it.

Observe a property that at an optimal solution $(\boldsymbol{W}_1^\star,\boldsymbol{W}_2^\star)$, at least one of two inequality constraints (17c)–(17d) holds with equality, and the inequality constraint (17e) becomes equality (active or binding).

4. Simulations results

In this section, we provide simulation results demonstrating the performance of the proposed approximate algorithm solving NOMA downlink beamforming optimization problems with per-antenna power constraints.

Example 1. Let us consider a NOMA downlink system where the BS is equipped with $K=8$ antennas, and the receiver noise power $\sigma_m^2=1$ W (Watt) for all m. Suppose that the BS serves $M=4$ users (as in problem (5)). Assume that the Rayleigh fading channel $\boldsymbol{h}_m$ follows $\mathcal{N}(\boldsymbol{0},1/((d_m)^\eta K)\boldsymbol{I})$ (i.e. (1) with $\zeta=0$), where d_m is the distance between the BS and user m, and η is the path loss exponent. We set $d_M=d_4=2$ m (meter), $d_3=5$ m, $d_2=10$ m, $d_1=15$ m, and $\eta=2$. The power upper bound for each antenna is $P_k=6$ W, $k=1,\dots,K$. The SINR targets $\gamma_4=1$, $\gamma_3=0.5$, $\gamma_2=0.1$, and $\gamma_1=0.05$.

In Fig. 1, we plot the optimal transmission power on antenna 1 in Fig. 1(a) for a set of 100 channel realizations (where the feasible sets of problem (5) without the per-antenna power constraints are nonempty), and in Fig. 1(b) for another set of 100 channel realizations (where the feasible sets of problem (5) are nonempty). Both problem (5) without the per-antenna power constraints and problem (5) are solved by Algorithm 1 (rather than by solving its SDP relaxation) for every feasible channel realization. We observe in Fig. 1(a) that without the per-antenna power constraints, there are many realizations for which the optimal transmission power on antenna 1 is over per-antenna power upper bound P_1, while in Fig. 1(b) that in all realizations the transmission power on antenna 1 is equal to or below P_1, which is in line with the condition that the per-antenna power constraints are enforced.

Example 2. We consider a scenario that up to four users are served. We assume that the channel vector $\boldsymbol{h}_m$ follows the model (1) with $\zeta=10$ and $\eta=1$, and the SINR thresholds are equal to each other, i.e. $\gamma_m=\gamma$, $\forall m$. The number of transmit antennas $K=20$. The distances between the BS and the users are $(d_1,d_2,d_3,d_4)=(6,5,4,3)$, and the AoDs in (1) are $(\theta_1,\theta_2,\theta_3,\theta_4)=(30^\circ,40^\circ,50^\circ,50^\circ)$. When saying three (two) users served by the BS, it means that $\{u_1,u_2,u_3\}$ ($\{u_1,u_2\}$) are selected. The power in dB is the ratio of the power in W over the noise power in W, so that 1 corresponds to 0 dB. Herein, the power upper bound $P_k=10\alpha^\star$ dB for each antenna, where $\alpha^\star$ is the optimal value for (4) with the same settings as stated in this example. Every data point in the figure is the average over 100 channel realizations.

Fig. 2 examines how different the optimal values obtained by the SOCP based approximation and by the SDP relaxation are. Here, the SDP relaxation refers to the conventional SDP relaxation problem for QCQP problem (5), namely the following problem:

$$\underset{\{\boldsymbol{W}_m\}}{\text{minimize}} \;\; \sum_{m=1}^M \text{tr}\,\boldsymbol{W}_m \tag{19a}$$

$$\text{subject to} \;\; \left(\sum_{m=1}^M \text{tr}\,(\boldsymbol{W}_m\boldsymbol{e}_k\boldsymbol{e}_k^H)\right)\le P_k,\; k=1,\dots,K, \tag{19b}$$

$$\text{tr}\,(\boldsymbol{W}_n\boldsymbol{h}_m\boldsymbol{h}_m^H)\ge\gamma_n\sum_{i=n+1}^M\text{tr}\,(\boldsymbol{W}_i\boldsymbol{h}_m\boldsymbol{h}_m^H)+\gamma_n\sigma_m^2, \tag{19c}$$

$$n\le m\le M,\; 1\le n\le M, \tag{19d}$$

$$\boldsymbol{W}_m\succeq\boldsymbol{0},\; m=1,\dots,M. \tag{19e}$$

Fig. 2 displays the total transmission power versus the common SINR threshold γ. We observe that the difference between the two

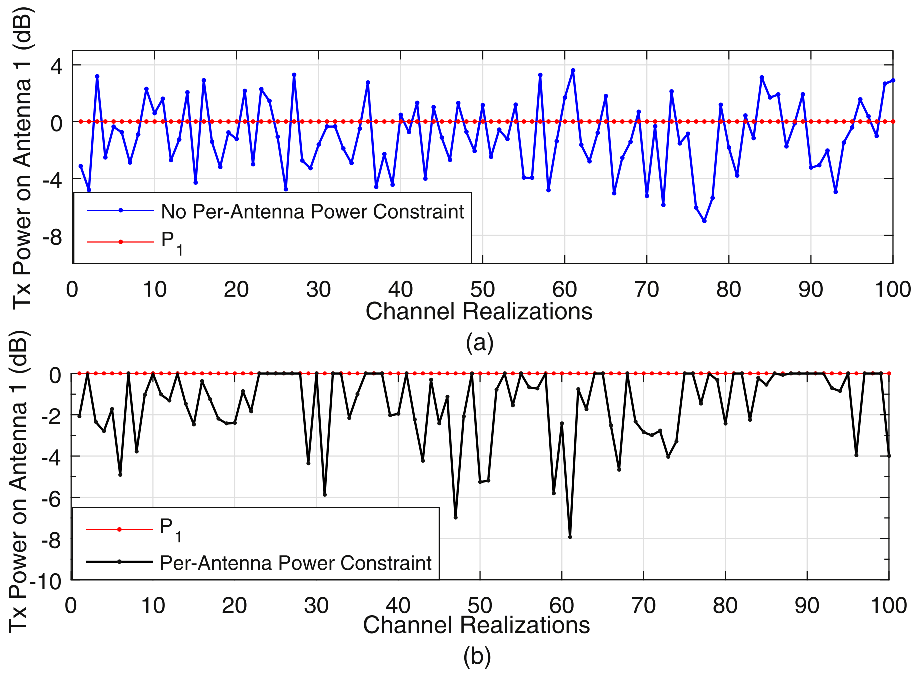

Fig. 1. The optimal transmission power on antenna 1 versus channel realizations; (a) NOMA beamforming problem (5) excluding the per-antenna power constraints; (b) NOMA beamforming problem (5). The transmission power on antenna 1 is divided by $P_1 = 6$ W so that 1 corresponds to 0 dB (thus the power upper bound P_1 of antenna 1 is 0 dB).

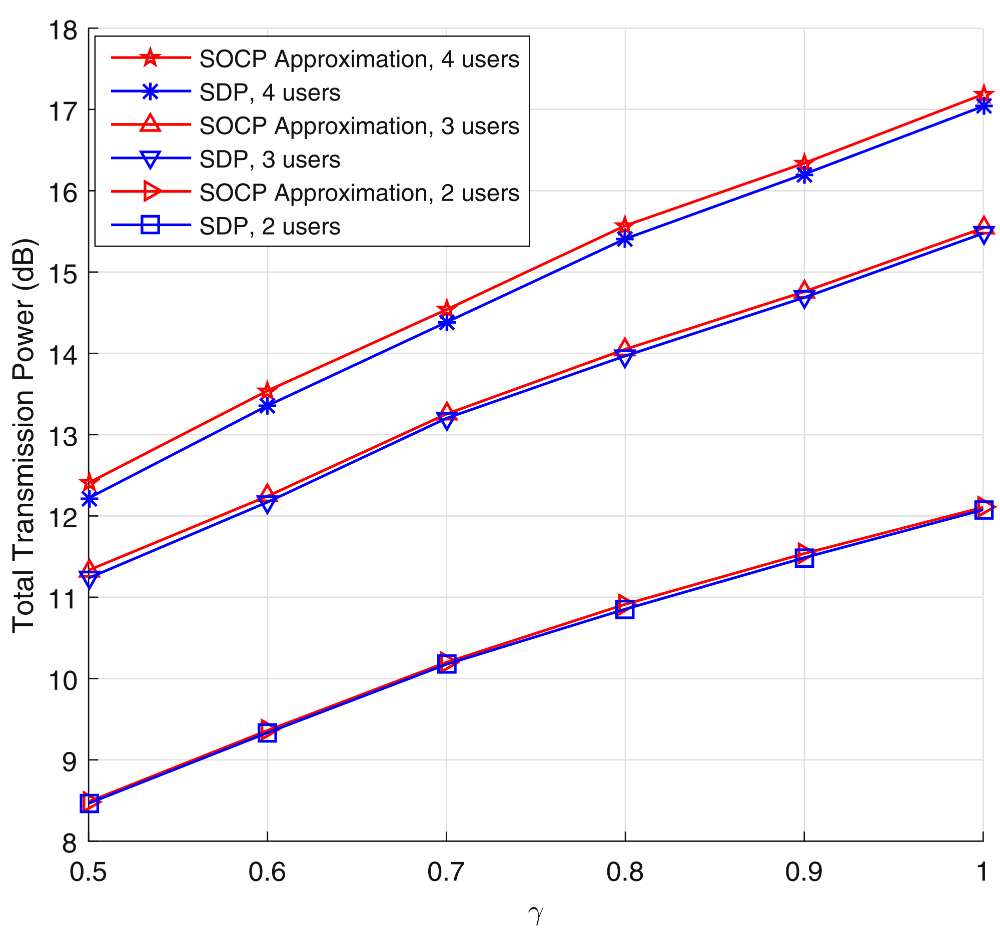

Fig. 2. The optimal total transmission power versus SINR threshold γ, with $\zeta = 10$, $\eta = 1$, and $K = 20$.

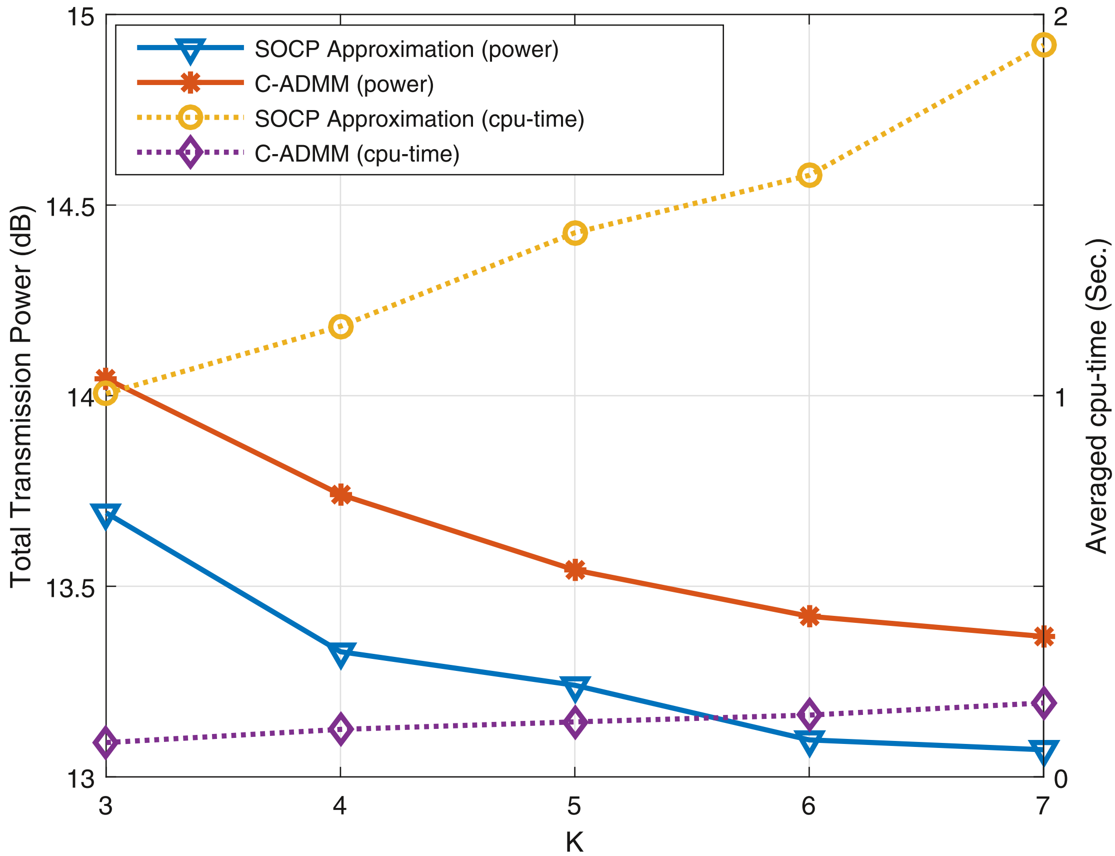

Fig. 3. The optimal total transmission power and the cpu running time versus number of antennas K, with $\zeta = 10$, $\eta = 1$, and $\gamma = 1$ dB.

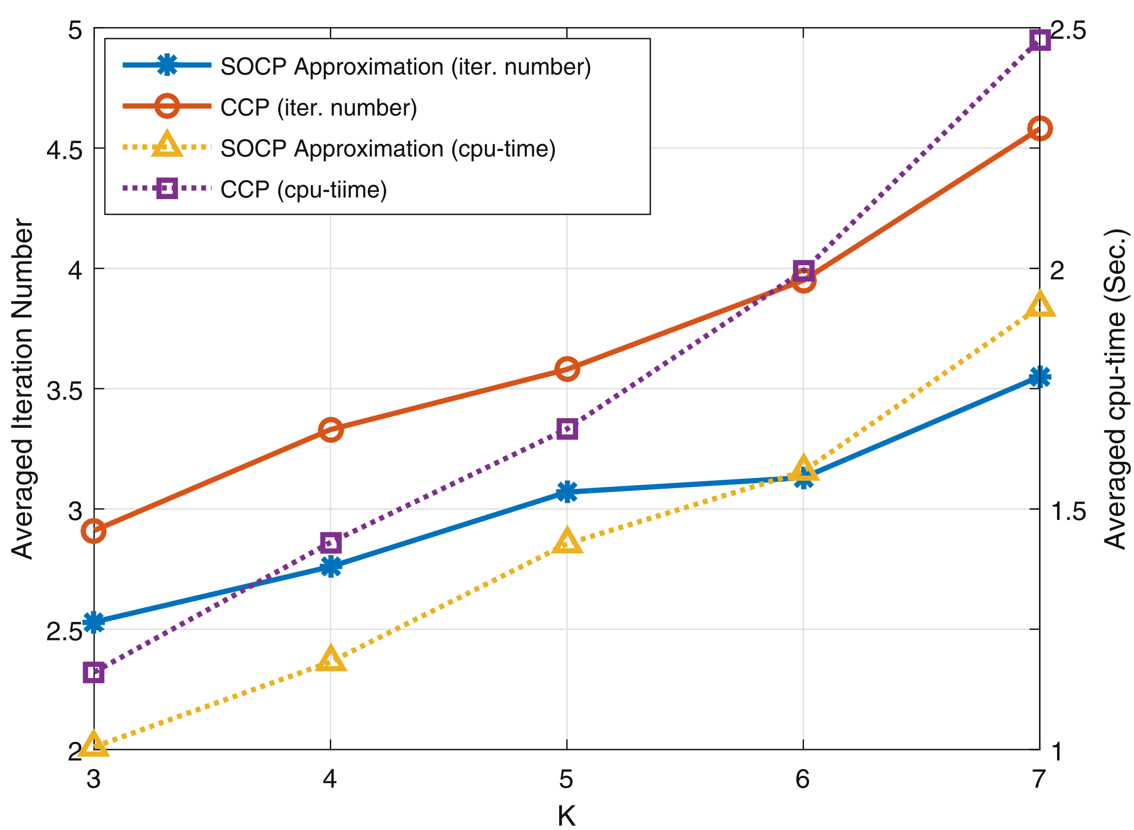

Fig. 4. The average iteration number and the average cpu time versus number of antennas K, with $\zeta = 10$, $\eta = 1$, and $\gamma = 1$ dB.

classes of optimal values is marginal, which means that the approximation is very closed to the globally optimal value for the NOMA downlink beamforming problem.

Example 3. Let us compare the performance of our proposed method (Algorithm 1 herein) with that of the consensus-ADMM [16], when both the methods are applied to solve separable QCQP problem (5). In the consensus-ADMM, the iteration steps consists of those in Eq. (4) of [16], and are specified by applying the results in Section 3.1 and 3.2 therein.

Suppose that there are two users in an NOMA downlink system, with the distances between the BS and the users being $(d_1, d_2) = (6, 5)$ m, and with the angles of departure to the users being $(\theta_1, \theta_2) = (40^\circ, 30^\circ)$. The channel vector defined in (1) includes parameters $\zeta = 10$, $\beta_m = 1/d_m$ (i.e. $\eta = 1$), $m = 1, 2$, and a random complex Gaussian vector $\boldsymbol{u}_m$. In beamforming problem (5), the noise variance of each user is set to 1 (=0 dB), the per-antenna power upper bound $P_k = 15$ ($\approx$= 11.76 dB) for antenna k, and the threshold values of SINR are equal to 1 dB. Every data point in the figure is the average over 100 random channel realizations.

Fig. 3 examines how the optimal total transmission power and the cpu running time are affected by number of antennas K. As can been observed, our method (termed as "SOCP Approximation" in the figure) has lower transmission power, while the consensus-ADMM (termed as "C-ADMM" in the figure) is faster. In other words, there is trade-off between the transmission power and the running time. In fact, the solution by our method is close to an optimal solution while the consensus-ADMM outputs a suboptimal solution. On the other hand, the consensus-ADMM does not have to call the time-consuming optimization tool CVX [17] and is faster. Another observation is that the transmission power decreases as K grows. This is because when K is larger, the optimization search space becomes bigger so that minimization problem (5) has the less optimal value. We remark that the behavior of the plots is similar when more users or antennas are considered.

Example 4. We consider a two-user NOMA downlink system, and assume that the simulation settings are the same as those in Example 3. Fig. 4 demonstrates the performance and complexity of our proposed method (called "SOCP Approximation" in the figure) and the CCP (called "CCP" in the figure), in terms of iteration numbers and cpu running time. Each data point in the figure is the average value over 100 random channel realizations (see (1)). It is observed that the average iteration number using our method is less than the CCP, as well as the average cpu-time is shorter by applying our method than the CCP. This coincides with the previous claim that the SOCP problem (problem (13)) in every iteration of our method (Algorithm 1) is simpler than the SOCP problem (problem (A.5)) in each iteration of the CCP (see Appendix A.1). Also, as can be seen, both the average iteration number and cup-time increase when number of antennas K grows. This is reasonable since when K grows, the number of constraints and the dimension of optimization variables in either SOCP problem (13) or (A.5) become bigger, and thus sizes of the SOCP problems are larger, which requires heavier computational cost.

Lastly, we report that applying the two methods to solve separable QCQP problem (5) yields the same optimal value, in the sense that the two optimal values $v_1^\star$ and $v_2^\star$ (obtained by our method and the CCP, respectively) fulfill $|v_1^\star - v_2^\star| \leq 10^{-3}$ always. Therefore, it is of no practical interest to compare the performance of the two methods in terms of the optimal values (i.e., the total transmission power).

5. Conclusion

In this paper, the MISO NOMA downlink beamforming optimization problem with per-antenna power constraints has been studied. An SOCP based approximate algorithm has been proposed for the downlink beamforming problem. In words, a sequence of SOCPs are solved with the property that the feasible set of SOCP in the lth iteration is always included in that of SOCP in the $(l+1)$th iteration. This gives the sequence of SOCPs with non-increasing optimal values, which means that the approximation converges to a locally optimal value. In particular, we have proved that the SDP relaxation for the NOMA downlink beamforming problem with two users is tight if one of the NOMA SINR inequality constraints is strict at the optimality. The performance of the proposed beamforming design has been demonstrated by simulations, when comparing it with the SDP relaxation, the consensus-ADMM, and the CCP.

Declaration of Competing Interest

Authors declare that they have no conflict of interest.

Acknowledgments

The work of Yongwei Huang was supported in part by the National Natural Science Foundation of China under Grant 11871168.

Appendix A

A1. Applying the CCP to solve problem (5)

In this appendix, we wish to derive the SOCP problem, which is solved in each iterative step of the CCP.

Let

$$g_m(\boldsymbol{w}_n) = |\boldsymbol{h}_m^H \boldsymbol{w}_n|^2 = \boldsymbol{w}_n^H \boldsymbol{h}_m \boldsymbol{h}_m^H \boldsymbol{w}_n,$$

for $n \leq m \leq M, 1 \leq n \leq M$, which are equal to

$$\begin{bmatrix} \Re \boldsymbol{w}_n \\ \Im \boldsymbol{w}_n \end{bmatrix}^H \begin{bmatrix} \Re(\boldsymbol{h}_m \boldsymbol{h}_m^H) & -\Im(\boldsymbol{h}_m \boldsymbol{h}_m^H) \\ \Im(\boldsymbol{h}_m \boldsymbol{h}_m^H) & \Re(\boldsymbol{h}_m \boldsymbol{h}_m^H) \end{bmatrix} \begin{bmatrix} \Re \boldsymbol{w}_n \\ \Im \boldsymbol{w}_n \end{bmatrix}.$$

Since function $g_m(\boldsymbol{w}_n)$ are convex, hence we have at $\boldsymbol{w}_n^l$ (where l denotes the index of iteration)

$$g_m(\boldsymbol{w}_n) \geq g_m(\boldsymbol{w}_n^l) + \nabla g_m(\boldsymbol{w}_n^l)^T \begin{bmatrix} \Re(\boldsymbol{w}_n - \boldsymbol{w}_n^l) \\ \Im(\boldsymbol{w}_n - \boldsymbol{w}_n^l) \end{bmatrix}. \quad (A.1)$$

Here $\nabla g_m(\boldsymbol{w}_n^l)$ are equal to

$$2 \begin{bmatrix} \Re(\boldsymbol{h}_m \boldsymbol{h}_m^H) & -\Im(\boldsymbol{h}_m \boldsymbol{h}_m^H) \\ \Im(\boldsymbol{h}_m \boldsymbol{h}_m^H) & \Re(\boldsymbol{h}_m \boldsymbol{h}_m^H) \end{bmatrix} \begin{bmatrix} \Re \boldsymbol{w}_n^l \\ \Im \boldsymbol{w}_n^l \end{bmatrix}.$$

Thereby, the right hand side of (A.1) are linear with respective to $\boldsymbol{w}_n$, and are employed to approximate $g_m(\boldsymbol{w}_n)$. In other words, SINR constraints (5c) are approximated by the following convex representations:

$$g_m(\boldsymbol{w}_n^l) + \nabla g_m(\boldsymbol{w}_n^l)^T \begin{bmatrix} \Re(\boldsymbol{w}_n - \boldsymbol{w}_n^l) \\ \Im(\boldsymbol{w}_n - \boldsymbol{w}_n^l) \end{bmatrix} \geq \gamma_n \sum_{i=n+1}^{M} |\boldsymbol{h}_m^H \boldsymbol{w}_i|^2 + \gamma_n \sigma_m^2. \quad (A.2)$$

After tedious calculation, convex constraints (A.2) are reexpressed into:

$$2\Re(\boldsymbol{w}_n^H \boldsymbol{h}_m \boldsymbol{h}_m^H \boldsymbol{w}_n^l) - |\boldsymbol{h}_m^H \boldsymbol{w}_n^l|^2 \geq \gamma_n \sum_{i=n+1}^{M} |\boldsymbol{h}_m^H \boldsymbol{w}_i|^2 + \gamma_n \sigma_m^2,$$

which are equivalent to the following SOC constraints:

$$(\Re(\boldsymbol{w}_n^H \boldsymbol{h}_m \boldsymbol{h}_m^H \boldsymbol{w}_n^l) + 1)/\sqrt{2} \geq \sqrt{\gamma_n \sum_{i=n+1}^{M} |\boldsymbol{h}_m^H \boldsymbol{w}_i|^2 + (\Re(\boldsymbol{w}_n^H \boldsymbol{h}_m \boldsymbol{h}_m^H \boldsymbol{w}_n^l) - 1)^2/2 + \gamma_n \sigma_m^2 + |\boldsymbol{h}_m^H \boldsymbol{w}_n^l|^2} \tag{A.3}$$

for $n \leq m \leq M, 1 \leq n \leq M$. Evidently, SOC constraints (A.3) are variations of the following standard SOC

$$\left\{ \begin{bmatrix} t \\ \boldsymbol{x} \end{bmatrix} \in \mathbb{R} \times \mathbb{R}^{2(M-n)+2} \mid t \geq \|\boldsymbol{x}\| \right\}. \tag{A.4}$$

Therefore, problem (13) with (13c) replaced with (A.3) is the SOCP problem to be solved in every iteration of the CCP (cf. [14, Algorithm 1.1]). In other words, the following SOCP problem is solved in each iteration of the CCP:

$$\underset{\{\boldsymbol{w}_m\}}{\text{minimize}} \sum_{m=1}^{M} \boldsymbol{w}_m^H \boldsymbol{w}_m \tag{A.5a}$$

$$\text{subject to } (5b), (A.3) \text{ satisfied.} \tag{A.5b}$$

A2. Applying the FPP-SCA algorithm to solve problem (5)

In the FPP-SCA algorithm, problem (5) with slack variables is solved, namely, the following problem is solved:

$$\underset{\{\boldsymbol{w}_m\},\{r_k\},\{s_{mn}\}}{\text{minimize}} \sum_{m=1}^{M} \boldsymbol{w}_m^H \boldsymbol{w}_m + \lambda(\|\boldsymbol{r}\|_1 + \|\boldsymbol{s}\|_1) \tag{A.6a}$$

$$\text{subject to } \left(\sum_{m=1}^{M} |\boldsymbol{e}_k^H \boldsymbol{w}_m|^2 \right) \leq P_k + r_k, \; k = 1, \ldots, K, \tag{A.6b}$$

$$s_{mn} \geq -|\boldsymbol{h}_m^H \boldsymbol{w}_n|^2 + \gamma_n \sum_{i=n+1}^{M} |\boldsymbol{h}_m^H \boldsymbol{w}_i|^2 + \gamma_n \sigma_m^2, \tag{A.6c}$$

$$n \leq m \leq M, \; 1 \leq n \leq M,$$

$$r_k \geq 0, \; s_{mn} \geq 0. \tag{A.6d}$$

Then, linearize the concave term $-|\boldsymbol{h}_m^H \boldsymbol{w}_n|^2$ in (A.6c) at the point $\boldsymbol{w}_n^l$, and obtain the approximate SOC constraints for (A.6c):

$$s_{mn} + 2\Re(\boldsymbol{w}_n^H \boldsymbol{h}_m \boldsymbol{h}_m^H \boldsymbol{w}_n^l) \geq \gamma_n \sum_{i=n+1}^{M} |\boldsymbol{h}_m^H \boldsymbol{w}_i|^2 + \gamma_n \sigma_m^2 + |\boldsymbol{h}_m^H \boldsymbol{w}_n^l|^2, \tag{A.7}$$

which is equivalent to the following standard SOC constraint:

$$(\Re(\boldsymbol{w}_n^H \boldsymbol{h}_m \boldsymbol{h}_m^H \boldsymbol{w}_n^l) + \frac{s_{mn}}{2} + 1)/\sqrt{2} \geq \sqrt{\gamma_n \sum_{i=n+1}^{M} |\boldsymbol{h}_m^H \boldsymbol{w}_i|^2 + \left(\Re\left(\boldsymbol{w}_n^H \boldsymbol{h}_m \boldsymbol{h}_m^H \boldsymbol{w}_n^l\right) + \frac{s_{mn}}{2} - 1\right)^2 \Big/ 2 + \gamma_n \sigma_m^2 + |\boldsymbol{h}_m^H \boldsymbol{w}_n^l|^2}. \tag{A.8}$$

Clearly, (A.8) is a variation of the following SOC:

$$\left\{ \begin{bmatrix} t \\ \boldsymbol{x} \end{bmatrix} \in \mathbb{R} \times \mathbb{R}^{2(M-n)+2} \mid t \geq \|\boldsymbol{x}\| \right\}. \tag{A.9}$$

And then, the following SOCP is solved in each iterative step in the FPP-SCA algorithm:

$$\underset{\{\boldsymbol{w}_m\},\{r_k\},\{s_{mn}\}}{\text{minimize}} \sum_{m=1}^{M} \boldsymbol{w}_m^H \boldsymbol{w}_m + \lambda(\|\boldsymbol{r}\|_1 + \|\boldsymbol{s}\|_1) \tag{A.10a}$$

$$\text{subject to } (A.6b), (A.8), (A.6d) \text{ satisfied.} \tag{A.10b}$$

It is noteworthy that when the slack variables $\boldsymbol{r} = \boldsymbol{0}$ and $\boldsymbol{s} = \boldsymbol{0}$, SOCP (A.10) is same as SOCP (A.5), which is solved in each iterative step of the CCP (therefore, the FPP-SCA algorithm can be viewed as an extension of the CCP).

References

[1] Z. Ding, Y. Liu, J. Choi, Q. Sun, M. Elkashlan, C.-L. I., H.V. Poor, Application of non-orthogonal multiple access in LTE and 5G network, IEEE Commun. Mag. 55 (2) (2017) 185–191.

[2] Z. Ding, X. Lei, G.K. Karagiannidis, R. Schober, J. Yuan, V.K. Bhargava, A survey on non-orthogonal multiple access for 5Gnetworks: research challenges and future trends, IEEE J. Sel. Areas Commun. 35 (10) (2017) 2181–2194.

[3] W. Yu, T. Lan, Transmitter optimization for the multi-antenna downlink with per-antenna power constraints, IEEE Trans. Signal Process. 55 (6) (2007) 2646–2660.

[4] J. Choi, Minimum power multicast beamforming with superposition coding for multiresolution broadcast and application to NOMA systems, IEEE Trans. Commun. 63 (3) (2015) 791–800.

[5] M.F. Hanif, Z. Ding, T. Ratnarajah, G.K. Karagiannidis, A minorization-maximization method for optimizing sum rate in the downlink of non-orthogonal multiple access systems, IEEE Trans. Signal Process. 64 (1) (2016) 76–88.

[6] Z. Chen, Z. Ding, X. Dai, G.K. Karagiannidis, On the application of quasi-degradation to MISO-NOMA downlink, IEEE Trans. Signal Process. 64 (23) (2016) 6174–6189.

[7] X. Shi, J.S. Thompson, R. Liu, M. Safari, P. Cao, Efficient optimization algorithms for multi-user beamforming with superposition coding, IEEE Trans. Commun. 66 (12) (2018) 5902–5915.

[8] M. Bengtsson, B. Ottersten, Optimal and suboptimal transmit beamforming, in: L.C. Godara (Ed.), Handbook of Antennas in Wireless Communications, CRC Press, Boca Raton, FL, 2001. Ch. 18.

[9] Y. Huang, D.P. Palomar, Rank-constrained separable semidefinite programming with applications to optimal beamforming, IEEE Trans. Signal Process. 58 (2) (2010) 664–678.

[10] J. Choi, On generalized downlink beamforming with NOMA, J. Commun. Netw. 19 (4) (2017) 319–328.

[11] Y. Liu, H. Xing, C. Pan, A. Nallanathan, M. Elkashlan, L. Hanzo, Multiple antenna assisted non-orthogonal multiple access, IEEE Wirel. Commun. 25 (2) (2018) 17–23.

[12] S. Kim, M. Kojima, Second order cone programming relaxation of nonconvex quadratic optimization problems, Optim. Methods Softw. 15 (3-4) (2001) 201–224.

[13] A. Ben-Tal, A. Nemirovski, Lectures on Modern Convex Optimization, Class Notes, Georgia Institute of Technology, 2020. Online: https://www2.isye.gatech.edu/~nemirovs/LMCOLN2020WithSol.pdf.

[14] T. Lipp, S. Boyd, Variations and extension of the convex-concave procedure, Optim. Eng. 17 (2016) 263–287.

[15] O. Mehanna, K. Huang, B. Gopalakrishnan, A. Konar, N.D. Sidiropoulos, Feasible point pursuit and successive approximation of non-convex QCQPs, IEEE Signal Process. Lett. 22 (7) (2015) 804–808.

[16] K. Huang, N.D. Sidiropoulos, Consensus-ADMM for general quadratically constrained quadratic programming, IEEE Trans. Signal Process. 64 (20) (2016) 5297–5310.

[17] CVX research, inc., 2011, CVX: Matlab software for disciplined convex programming, version 2.0. http://cvxr.com/cvx.