Deng *et al. EURASIP Journal on Wireless Communications and Networking* (2019) 2019:185
https://doi.org/10.1186/s13638-019-1501-6

EURASIP Journal on Wireless Communications and Networking

RESEARCH Open Access

Precoding design of NOMA-enabled D2D communication system with low latency

Ping Deng[1,2*], Wei Wu[1], Xinmeng Shen[1], Pei Li[1] and Baoyun Wang[3]

Abstract

This letter investigated a device-to-device (D2D) communication underlaying cellular system, where one wireless device (WD) can directly transmit information to two inner users under non-orthogonal multiple access (NOMA) protocol with low latency. Considering the user fairness and reliability, both users can adopt successive interference cancellation (SIC) to decode information by the given decoding order. Subject to predefined quality-of-service (QoS) requirement and the constraint of WD's transmit power, the maximal rate of latter decoded information and the corresponding precoding vectors are obtained by applying the technique of semidefinite relaxation (SDR) and the Langrangian duality method. Then, the suboptimal scheme based on singular value decomposition (SVD) is also proposed with a lower computational complexity. Numerical simulation shows that with the design of proper precoding vectors D2D communication system assisted by NOMA has a better performance than orthogonal multiple access (OMA).

Keywords: Non-orthogonal multiple access (NOMA), Low latency, Semidefinite relaxation, Langrangian duality, Singular value decomposition (SVD)

1 Introduction

Non-orthogonal multiple access (NOMA) is one of the promising multiple access to realize the challenging requirements of 5G [1, 2], such as massive connectivity, high data rate, and low latency. It has proved to be a viable solution for future dense networks and Internet of Things (IoT) devices. Unlike conventional orthogonal multiple access, NOMA uses the power domain to serve multiple users at different power levels at the same time, code, and frequency [3], in which superposition coding and successive interference cancellation (SIC) are employed [4]. Many various NOMA designs combined with multiple-input multiple-output (MIMO) [5], cooperative relaying [6] and millimeter-wave communications [7], have appeared in recent researches. In [8], the random opportunistic beamforming, which is a signal processing technique used in various wireless systems for directional communications [9], is first proposed for the MIMO NOMA systems, and the transmitter generated multiple beams and superposed multiple users within each beam. In [10], a beamforming design based on zero-forcing and user pairing scheme are proposed for the downlink multi-user NOMA system, assuming the perfect channel state information (CSI) is available at the transmitter. The integration of NOMA and multi-user beamforming thus has the potential to capture the benefits of both NOMA and beamforming.

Device-to-device (D2D) communication makes it possible for users in proximity to communicate with each other directly rather than relying on base stations (BSs) [11], and thus, it is an available way for reliable and low-latency communication. In [12] and [14], mode selection in underlay D2D networks is studied, while [13] investigates an efficient way of reusing the downlink resources for cellular and D2D mode communication. A step further from D2D pairs, [15] studies D2D groups that use NOMA as their transmission technique to serve multiple D2D receivers. Zhao et al. [16] consider the setting of an uplink single-cell cellular network communications. In order to further improve the outage performance of the NOMA-weak user in a user pair and reduce cooperative

*Correspondence: dengp@njupt.edu.cn
[1]College of Communication and Information, Nanjing University of Posts and Telecommunications, Xinmofan Road 66, Nanjing 210003, China
[2]College of Automation and College Of Artificial Intelligence, Nanjing University of Posts and Telecommunications, Xinmofan Road 66, Nanjing 210003, China
Full list of author information is available at the end of the article

delay, [17] focuses on full-duplex D2D-aided cooperative NOMA.

2 Method

This paper considers the D2D communication underlaying cellular system in a multiple cellular networks, which takes advantages of NOMA and D2D systems to increase the available throughput of the wireless networks. The superposed signals are sent from the multi-antenna wireless device (WD) to two single-antenna users under NOMA protocol. In particular, taking the user fairness into account, we study the case in which both users can adopt SIC to decode information by the given decoding order. To guarantee predefined quality-of-service (QoS) requirement and the constraint of WD's transmit power, the maximal rate of latter decoded information and the corresponding precoding vectors are obtained by applying the technique of semidefinite relaxation and the Langrangian duality method. The suboptimal solution based on single value decomposition (SVD) is proposed with lower computational complexity. Then, simulations are also provided to verify the performance of the proposed NOMA-enabled D2D schemes.

The rest of this letter is organized as follows. In Section 3, we introduce the system model and formulate the problem . In Section 4, we derive the optimal solution to this optimization problem. In Section 5, we propose the suboptimal solution based on SVD. In Section 6, we show simulation results that justify the performance of the proposed approaches. Our conclusions are included in Section 7.

Notations: Scalars are denoted by lowercase letters, vectors are denoted by boldface lowercase letters, and matrices are denoted by boldface uppercase letters. For a square matrix $\mathbf{A}$, $tr(\mathbf{A})$, $rank(\mathbf{A})$, and $\mathbf{A}^H$ denote its trace, rank, and conjugate, respectively. $\mathbf{A} \succeq 0$ and $\mathbf{A} \preceq 0$ represent that $\mathbf{A}$ is a positive semidefinite matrix and a negative semidefinite matrix, respectively. $\|\mathbf{x}\|$ denotes the Euclidean norm of a complex vector $\mathbf{x}$. $E[\cdot]$ denotes the statistical expectation. $[\cdot]^+$ means $\max(0, \cdot)$. $\mathbb{P}(\cdot)$ defines an outage probability event. The distribution of a circularly symmetric complex Gaussian (CSCG) random vector with mean vector $\mathbf{x}$ and covariance matrix Σ is denoted by $CN(0, \Sigma)$, and $\sim$ stands for "distributed as". $\mathbb{C}^{x\times y}$ denotes the space of $x \times y$ complex matrices.

3 System model and problem formulation

3.1 System model

In this paper, one D2D communication underlaying cellular system in a multiple cellular networks is considered, as illustrated in Fig. 1. One BS can serve a set of cellular users by WDs. Different cellular networks are allocated with orthogonal resource blocks, such as time, frequency, code, and space, in order to eliminate the inter interference between different cellular networks. Actually, WD is also regarded as a special user with relaying function and has a higher priority to decode its own message compared to s_1 and s_2, which are transmitted to two ordinary users. So WD has direct links with two ordinary users, respectively, while no direct link between the BS and each ordinary user is assumed due to significant path loss [18]. We focus on a single-cell downlink transmission scenario, where a WD can receive the signals from the BS and then transmit the superposed signals using the well-known amplify-and-forward (AF) protocol [19] to corresponding users under NOMA protocol. The WD is equipped N antennas,

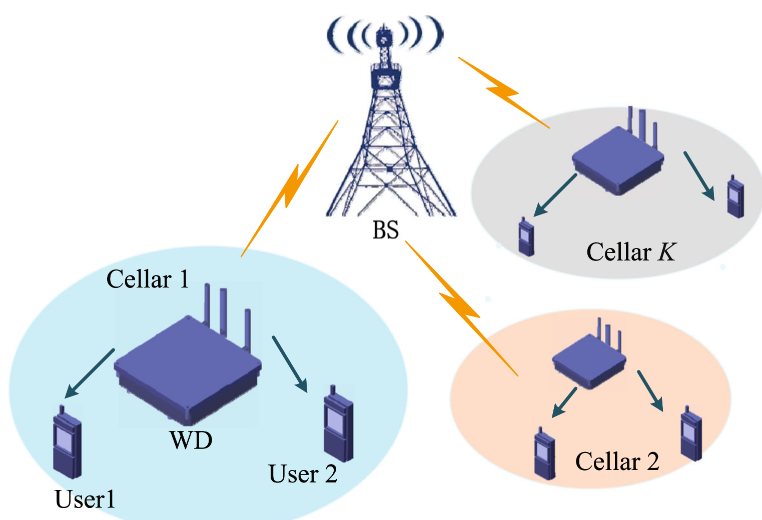

Fig. 1 System model. A D2D communication networks including one BS and K underlaying systems, in which a WD can transmit information to two users under NOMA protocol

and each user has a single antenna for the facility cost. Considering the complexity of the system, only two users are served at the same resource block. Note that it is of practical significance to choose two users to perform NOMA since NOMA systems are strongly interference-limited [20]. It is often more appropriate to group two users together to perform NOMA with user pairing [21] to realize reliable and low latency communication. It is assumed that the users' channel state information is perfectly available at the WD. We denote two types of information to both users by s_1 and s_2, respectively. It is assumed that s_1 and s_2 are independent and identically distributed (i.i.d.) circularly symmetric complex Gaussian (CSCG) random signals with unit average power [22], i.e., $E[\,|s_1|^2] = E[\,|s_2|^2] = 1$. Therefore, the complex baseband transmitted signal of WD can be expressed as:

$$\mathbf{x} = \mathbf{w}_1 s_1 + \mathbf{w}_2 s_2 \tag{1}$$

where the $\mathbf{w}_1, \mathbf{w}_2 \in \mathbb{C}^{N\times 1}$ are the precoding vectors of s_1 and s_2, respectively [23]. And the observations at two users are given by:

$$y_1 = \mathbf{h}_1^H \mathbf{x} + n_1 = \mathbf{h}_1^H(\mathbf{w}_1 s_1 + \mathbf{w}_2 s_2) + n_1 \tag{2}$$

$$y_2 = \mathbf{h}_2^H \mathbf{x} + n_2 = \mathbf{h}_2^H(\mathbf{w}_1 s_1 + \mathbf{w}_2 s_2) + n_2 \tag{3}$$

where $\mathbf{h_1}, \mathbf{h_2} \in \mathbb{C}^{N\times 1}$ denote the complex Gaussian channel vector of user 1 and user 2, respectively, which are independent and identically distributed (*i.i.d*) fading channels. n_1 and n_2 are additive Gaussian noise (AGN), satisfying $n_1, n_2 \sim CN(0, \sigma^2)$.

3.2 Problem formulation

First, by taking the fairness of users and the decoding order into consideration, it is assumed that both users can use SIC to decode s_2 first, then subtract it from the observation before s_1 is decoded. The signal-to-interference-plus-noise ratio (SINR) of two users to decode s_1 and s_2 are respectively given by:

$$\mathrm{SINR}_{i,1} = \frac{|\mathbf{h}_i^H \mathbf{w}_1|^2}{\sigma^2}, i = 1, 2 \tag{4}$$

$$\mathrm{SINR}_{i,2} = \frac{|\mathbf{h}_i^H \mathbf{w}_2|^2}{|\mathbf{h}_i^H \mathbf{w}_1|^2 + \sigma^2}, i = 1, 2 \tag{5}$$

Then, according to the decoding order, s_2 has a higher priority to be decoded. Actually, the cognitive radio concept is used here [24]. So the achievable rate of s_2 is dependent on the minimal SINR of s_2 which is decoded by each user, and its achievable rate can be expressed as:

$$R_2 = \min[\,\log_2(1 + \mathrm{SINR}_{1,2}), \log_2(1 + \mathrm{SINR}_{2,2})] \tag{6}$$

In fact, s_1, which is intend for user 1, is inevitably decoded by user 2, and the achievable rate of s_1 is considered to subtract the interception by user 2. So the achievable rate of s_1 is written by [25]:

$$R_1 = \left[\log_2\left(1 + \mathrm{SINR}_{1,1}\right) - \log_2\left(1 + \mathrm{SINR}_{2,1}\right)\right]^+ \tag{7}$$

It can be verified that there is an effective rate of s_1 only if the channel condition of user 1 is no worse than that of use 2.

Next, we will discuss the transmit power of WD. Suppose that the energy for receiving and amplifying the information in WD is supplied by extra power, and the transmit power P of WD is almost used to forward two types of information. So the transmit power must be satisfied [26]:

$$E\left[\mathbf{x}^H \mathbf{x}\right] = \|\mathbf{w}_1\|^2 + \|\mathbf{w}_2\|^2 \le P \tag{8}$$

At last, in this paper, we aim to maximize the achievable rate of s_1 subject to the predefined QoS of user 2 and the given transmit power of the WD. The optimization problem can be formulated as:

$$\max_{\mathbf{w}_1, \mathbf{w}_2} R_1 \tag{9a}$$

$$s.t.\;\; R_2 \ge \gamma_M \tag{9b}$$

$$\|\mathbf{w}_1\|^2 + \|\mathbf{w}_2\|^2 \le P \tag{9c}$$

where $\gamma_M = 2^{2R_M - 1}$, and R_M is the target rate of s_2 to satisfy the corresponding requirement of QoS.

Note that due to the non-convex nature of (9a) and (9b), problem (9) is undoubtedly non-convex in its current form. In the following subsection, we will find the optimal solution based on the analysis and transformation of problem (9).

4 The optimal solution

In order to solve the above non-convex problem, we consider the nontrivial case of the problem (9), in which the objective function is positive and can be rewritten as:

$$\begin{aligned} &\log_2(1 + \mathrm{SINR}_{1,1}) - \log_2(1 + \mathrm{SINR}_{2,1}) \\ &= \log_2 \frac{1 + \mathrm{SINR}_{1,1}}{1 + \mathrm{SINR}_{2,1}} = \log_2 \frac{\sigma^2 + |\mathbf{h}_1^H \mathbf{w}_1|^2}{\sigma^2 + |\mathbf{h}_2^H \mathbf{w}_1|^2} \end{aligned} \tag{10}$$

It is obvious that with the same constraints, $\log_2 \frac{\sigma^2 + |\mathbf{h}_1^H \mathbf{w}_1|^2}{\sigma^2 + |\mathbf{h}_2^H \mathbf{w}_1|^2}$ and $\frac{\sigma^2 + |\mathbf{h}_1^H \mathbf{w}_1|^2}{\sigma^2 + |\mathbf{h}_2^H \mathbf{w}_1|^2}$ have the same optimal solution.

Meanwhile, (9b) is divided into two constraints about the SINR of s_2. So the problem (9) has the same optimal solution with the following problem:

$$\max_{\mathbf{w}_1,\mathbf{w}_2} \frac{\sigma^2 + \left|\mathbf{h}_1^H\mathbf{w}_1\right|^2}{\sigma^2 + \left|\mathbf{h}_2^H\mathbf{w}_1\right|^2} \tag{11a}$$

$$s.t.\ \frac{\left|\mathbf{h}_1^H\mathbf{w}_2\right|^2}{\left|\mathbf{h}_1^H\mathbf{w}_1\right|^2+\sigma^2} \geq \gamma_M \tag{11b}$$

$$\frac{\left|\mathbf{h}_2^H\mathbf{w}_2\right|^2}{\left|\mathbf{h}_2^H\mathbf{w}_1\right|^2+\sigma^2} \geq \gamma_M \tag{11c}$$

$$\|\mathbf{w}_1\|^2 + \|\mathbf{w}_2\|^2 \leq P \tag{11d}$$

The semidefinite relaxation (SDR) [27] is applied to obtain the optimal solution of (9). Define $\mathbf{W}_1 = \mathbf{w}_1\mathbf{w}_1^H, \mathbf{W}_2 = \mathbf{w}_2\mathbf{w}_2^H, \mathbf{H}_1 = \mathbf{h}_1\mathbf{h}_1^H, \mathbf{H}_2 = \mathbf{h}_2\mathbf{h}_2^H$, and ignore the constraint of $rank\,(\mathbf{W}_1) = rank\,(\mathbf{W}_2) = 1$, we can obtain:

$$\max_{\mathbf{W}_1,\mathbf{W}_2} \frac{\sigma^2 + tr\,(\mathbf{H}_1\mathbf{W}_1)}{\sigma^2 + tr\,(\mathbf{H}_2\mathbf{W}_1)} \tag{12a}$$

$$s.t.\ \frac{tr\,(\mathbf{H}_1\mathbf{W}_2)}{tr\,(\mathbf{H}_1\mathbf{W}_1)+\sigma^2} \geq \gamma_M \tag{12b}$$

$$\frac{tr\,(\mathbf{H}_2\mathbf{W}_2)}{tr\,(\mathbf{H}_2\mathbf{W}_1)+\sigma^2} \geq \gamma_M \tag{12c}$$

$$tr\,(\mathbf{W}_2) + tr\,(\mathbf{W}_1) \leq P \tag{12d}$$

The problem (12) we proposed is a fractional programming obviously. The Dinkelbach method is widely adopted in solving the fractional programming. Thus, we use this method to transform and solve the problem. With a continuous auxiliary variable t [28], we reformulate the objective function in (12a) as:

$$F(t) = \max_{\{\mathbf{W}_1,\mathbf{W}_2\}\in\Psi} \sigma^2 + tr(\mathbf{H}_1\mathbf{W}_1) - t[\sigma^2 + tr(\mathbf{H}_2\mathbf{W}_1)] \tag{13}$$

where $\left\{\Psi \middle| \frac{tr(\mathbf{H}_1\mathbf{W}_2)}{tr(\mathbf{H}_1\mathbf{W}_1)+\sigma^2} \geq \gamma_M, \frac{tr(\mathbf{H}_2\mathbf{W}_2)}{tr(\mathbf{H}_2\mathbf{W}_1)+\sigma^2} \geq \gamma_M, tr(\mathbf{W}_1) + tr(\mathbf{W}_2) \leq P\right\}$ is the feasible set in (12). We have the following lemma, whose proof can refer to [29]:

Lemma 1 *$F(t)$ is a strictly decreasing and continuous function, and it has a unique zero solution, which is denoted as t^*. Then, the optimal value of the objective function in (12a) is t^*.*

From Lemma 1, we can see that the optimal solution can be obtained by solving (12) if we know t^* in advance. Though t^* is unknown at first, Lemma 1 tells us that the Dinkelbach method by round search [28] can be used to find the root of $F(t)$ efficiently. In each iteration, the non-negative variable t will update its value and finally reach the optimal denoting as t^*.

Therefore, in the following, we will optimize (12) for a given t at first. For convenience, we rewrite (13) into the following form:

$$F(t) = \max_{\mathbf{W}_1,\mathbf{W}_2} \sigma^2 + tr\,(\mathbf{H}_1\mathbf{W}_1) - t\left[\sigma^2 + tr\,(\mathbf{H}_2\mathbf{W}_1)\right] \tag{14a}$$

$$s.t.\ tr\,(\mathbf{H}_1\mathbf{W}_2) \geq \gamma_M\left[tr\,(\mathbf{H}_1\mathbf{W}_1)+\sigma^2\right] \tag{14b}$$

$$tr\,(\mathbf{H}_2\mathbf{W}_2) \geq \gamma_M\left[tr\,(\mathbf{H}_2\mathbf{W}_1)+\sigma^2\right] \tag{14c}$$

$$tr\,(\mathbf{W}_1) + tr\,(\mathbf{W}_2) \leq P \tag{14d}$$

Meanwhile, problem (14) is a convex semidefinite problem (SDP) and can be efficiently solved by convex optimization solvers, e.g., CVX [30].

Proposition 1 *The optimal solution to problem (14) satisfies $rank(\mathbf{W}_1^*) = rank(\mathbf{W}_2^*) = 1$.*

Proof Obviously, the problem (14) is a separate SDP with three generalized constraints. According to [31], the optimal solution $(\mathbf{W}_1^*, \mathbf{W}_2^*)$ to (14) always satisfies $rank^2(\mathbf{W}_1^*) + rank^2(\mathbf{W}_2^*) \leq 3$. Here, we consider the non-trivial case where $\mathbf{W}_1^* \neq 0$ and $\mathbf{W}_2^* \neq 0$, then $rank(\mathbf{W}_1^*) = rank(\mathbf{W}_2^*) = 1$ can be obtained. Proposition 1 is proved, and this implies that the SDR here is tight. □

Let α_1, α_2, and α_3 denote the dual variables associated with constraints (14b), (14c), and (14d), respectively. The dual problem of (14) is expanded as:

$$d_{\min} = \min_{\alpha_1,\alpha_2,\alpha_3} \alpha_3 P - (\alpha_1+\alpha_2)\gamma_M\sigma^2 + (1-t)\sigma^2 \tag{15a}$$

$$s.t.\ \mathbf{A} \leq 0, \mathbf{B} \leq 0, \tag{15b}$$

$$\alpha_1 \geq 0, \alpha_2 \geq 0, \alpha_3 \geq 0 \tag{15c}$$

where

$$\mathbf{A} = -\alpha_3\mathbf{I} + (1-\alpha_1\gamma_M)\mathbf{H}_1 - (t+\alpha_2\gamma_M)\mathbf{H}_2, \tag{16}$$

$$\mathbf{B} = -\alpha_3\mathbf{I} + \alpha_1\mathbf{H}_1 + \alpha_2\mathbf{H}_2, \tag{17}$$

and $d_{\min}$ is the value of problem (15).

Proposition 2 *The optimal dual solution α_3^* to problem (15) satisfies $\alpha_3^* > 0$.*

Proof We show that the optimal solution $\alpha_3^* > 0$ by contradiction. Assume that $\alpha_3^* = 0$, since the Lagrangian dual variables are all non-negative. The matrix $\mathbf{B}^*$ is negative semidefinite, i.e., $\alpha_1^*\mathbf{H}_1 + \alpha_2^*\mathbf{H}_2 \preceq 0$. Moreover, (14) is convex and satisfies Slater's condition, and the duality gap between (14) and (15) is zero. Thus $F(t) = (1-t)\sigma^2$. According to Lemma 1, the optimal solution to problem (14) is the same with the problem (12) when $F(t^*) = 0$. So $t^* = 1$, and the maximal rate of R_1 is $\log_2 t^* = 0$. It is not reasonable, and then $\alpha_3^* > 0$ must be true. Proposition 2 is proved. □

With the optimal dual solution $\left(\alpha_1^*, \alpha_2^*, \alpha_3^*\right)$ and optimal value $d_{\min}^*$ obtained by solving problem (15), we can derive $\mathbf{A}^*$ and $\mathbf{B}^*$, respectively, by substituting $\left(\alpha_1^*, \alpha_2^*, \alpha_3^*\right)$ into (16) and (17). Moreover, the complementary slackness condition of (15b) yields to $\mathbf{A}^*\mathbf{W}_1^* = 0, \mathbf{B}^*\mathbf{W}_2^* = 0$. Since $rank(\mathbf{W}_1^*) = 1$ and $rank(\mathbf{W}_2^*) = 1$, we have $rank(\mathbf{A}^*) = N-1$ and $rank(\mathbf{B}^*) = N-1$. Let $\mathbf{u}_1$ and $\mathbf{u}_2$ be the basis of the null space of $\mathbf{A}^*$ and $\mathbf{B}^*$, respectively, and define $\hat{\mathbf{W}}_1 = \mathbf{u}_1\mathbf{u}_1^H, \hat{\mathbf{W}}_2 = \mathbf{u}_2\mathbf{u}_2^H$, then we have:

$$\begin{cases} (1-t)\sigma^2 + \tau_1^2 tr[(\mathbf{H}_1 - t\mathbf{H}_2)\hat{\mathbf{W}}_1] = d_{\min}^* \\ \tau_2^2 tr(\hat{\mathbf{W}}_2) + \tau_1^2 tr(\hat{\mathbf{W}}_1) = P \end{cases} \tag{18}$$

where τ_1 and τ_2 are the power allocation coefficients for transmitting information s_1 and s_2, respectively. And

$$\begin{cases} \tau_1 = \sqrt{\dfrac{d_{\min}^* - (1-t)\sigma^2}{tr\left[(\mathbf{H}_1 - t\mathbf{H}_2)\,\hat{\mathbf{W}}_1\right]}}, \\ \tau_2 = \sqrt{P - \tau_1^2}. \end{cases} \tag{19}$$

Thus, the optimal precoding vectors are $\mathbf{w}_1^* = \tau_1\mathbf{u}_1, \mathbf{w}_2^* = \tau_2\mathbf{u}_2$ with given t.

Note that $2N$ complex variables are to be optimized for problem (12), while only three real variables for problem (15). So problem (15) has the lower computational complexity than problem (12). Actually, the complexity reduction is significant as the number of antennas at WD increases. Detailed steps of proposed optimal algorithm are summarized as Algorithm 1.

Algorithm 1 The optimal solution to problem (11)

1: Initialize t satisfying $F(t) \geq 0$ and tolerance ε
2: **while** $(|F(t)| > \varepsilon)$ **do**
3: Solve problem (15) to obtain $\left(\alpha_1^*, \alpha_2^*, \alpha_3^*\right)$
4: Calculate $\mathbf{A}^*$ and $\mathbf{B}^*$ according to (16) and (17), respectively.
5: Calculate $\mathbf{w}_1^*$, and $\mathbf{w}_2^*$ according to (19)
6: $t \leftarrow \frac{\sigma^2 + |\mathbf{h}_1^H\mathbf{w}_1^*|^2}{\sigma^2 + |\mathbf{h}_2^H\mathbf{w}_1^*|^2}$
7: **end while**
8: **return** $\mathbf{w}_1^*$ and $\mathbf{w}_2^*$

5 The SVD-based suboptimal solution

As described in the previous section, we can derive the optimal solution to problem (11) by using fractional programming and solving dual problem. But the round search for finding optimal t^* reduces the feasibility of the optimal solution to a certain extent in practice. In this section, we propose a suboptimal solution based on SVD to further reduce the computational complexity.

When $N \geq 2$, the SVD-based precoding scheme can be used to eliminate the interference caused by s_1 at the WD by restricting precoding vector $\mathbf{w}_1$ to satisfy $\mathbf{h}_2^H\mathbf{w}_1 = 0$ [32], which simplifies the precoding vector design. With SVD-based precoding vector, user 2 cannot decode the information s_1. It implies that the precoding vector $\mathbf{w}_1$ must lie in the null space of $\mathbf{h}_2$. Let the SVD of $\mathbf{h}_2$ be expressed as $\mathbf{h}_2^H = \mathbf{u}\Lambda\mathbf{v}^H = \mathbf{u}\Lambda[\mathbf{v}_1\mathbf{v}_0]^H$, where $\mathbf{u} \in \mathbb{C}^{1\times1}, \mathbf{v} \in \mathbb{C}^{N\times N}$ are orthogonal left and right singular vectors of $\mathbf{h}_2$, respectively, and $\Lambda \in \mathbb{C}^{1\times N}$ contains one positive singular value of $\mathbf{h}_2$. $\mathbf{v}_0 \in \mathbb{C}^{N\times(N-1)}$, which satisfies $\mathbf{v}_0^H\mathbf{v}_0 = \mathbf{I}_{N-1}$, is the last $N-1$ columns of $\mathbf{v}$ and forms an orthogonal basis for the null space of $\mathbf{h}_2^H$. The SVD-based precoding vector $\mathbf{w}_1$ can be expressed as $\mathbf{w}_1 = \mathbf{v}_0\tilde{\mathbf{w}}_1$, where $\tilde{\mathbf{w}}_1$ denotes the new vector to be designed, and the corresponding precoding vector of s_2 to be designed is $\tilde{\mathbf{w}}_2$. It is obvious to observe that in order for the SVD-based solution to be feasible, we must have $N \geq 2$. Problem (11) is consequently formulated as:

$$\max_{\tilde{\mathbf{w}}_1, \tilde{\mathbf{w}}_2} 1 + \frac{1}{\sigma^2}|\mathbf{h}_1^H\mathbf{v}_0\tilde{\mathbf{w}}_1|^2 \tag{20a}$$

$$s.t.\ \ |\mathbf{h}_1^H\tilde{\mathbf{w}}_2|^2 \geq \gamma_M\left(|\mathbf{h}_1^H\mathbf{v}_0\tilde{\mathbf{w}}_1|^2 + \sigma^2\right) \tag{20b}$$

$$|\mathbf{h}_2^H\tilde{\mathbf{w}}_2|^2 \geq \gamma_M\sigma^2 \tag{20c}$$

$$||\tilde{\mathbf{w}}_2||^2 + ||\tilde{\mathbf{w}}_1||^2 \leq P \tag{20d}$$

Define $\tilde{\mathbf{w}}_1 = \tilde{\mathbf{w}}_1\tilde{\mathbf{w}}_1^H$, $\tilde{\mathbf{w}}_2 = \tilde{\mathbf{w}}_2\tilde{\mathbf{w}}_2^H$, $\tilde{\mathbf{H}}_1 = \mathbf{v}_0^H\mathbf{h}_1\mathbf{h}_1^H\mathbf{v}_0$, we have the SVD-based SDP:

$$\max_{\tilde{\mathbf{w}}_1, \tilde{\mathbf{w}}_2} 1 + \frac{1}{\sigma^2}tr(\tilde{\mathbf{H}}_1\tilde{\mathbf{w}}_1) \tag{21a}$$

$$s.t.\ \ tr(\mathbf{H}_1\tilde{\mathbf{w}}_2) \geq \gamma_M[tr(\tilde{\mathbf{H}}_1\tilde{\mathbf{w}}_1) + \sigma^2] \tag{21b}$$

$$tr(\mathbf{H}_2\tilde{\mathbf{w}}_2) \geq \gamma_M\sigma^2 \tag{21c}$$

$$tr(\tilde{\mathbf{w}}_1) + tr(\tilde{\mathbf{w}}_2) \leq P \tag{21d}$$

Obviously, the achieved optimal solution also satisfies the rank-one constraint. Let $\tilde{\alpha}_1$, $\tilde{\alpha}_2$, and $\tilde{\alpha}_3$ denote dual variables, and its dual problem is given by:

$$d^* = \min_{\tilde{\alpha}_1, \tilde{\alpha}_2, \tilde{\alpha}_3} 1 + \tilde{\alpha}_3 P - (\tilde{\alpha}_1 + \tilde{\alpha}_2)\gamma_M\sigma^2 \tag{22a}$$

$$s.t.\ \ \tilde{\mathbf{A}} \preceq 0, \tilde{\mathbf{B}} \preceq 0, \tilde{\alpha}_1 \geq 0, \tilde{\alpha}_2 \geq 0, \tilde{\alpha}_3 > 0 \tag{22b}$$

where

$$\tilde{\mathbf{A}} = -\tilde{\alpha}_3\mathbf{I} + \frac{1}{\sigma^2}\tilde{\mathbf{H}}_1 - \tilde{\alpha}_1\tilde{\mathbf{H}}_1\gamma_M \tag{23}$$

$$\tilde{\mathbf{B}} = -\tilde{\alpha}_3\mathbf{I} + \tilde{\alpha}_1\mathbf{H}_1 + \tilde{\alpha}_2\mathbf{H}_2 \tag{24}$$

Different from problem (12), problem (20) is convex and the dual gap between (21) and (22) is also zero. In the same way as the previous section, we can also solve the problem (20) by its Lagrangian dual problem (22) for complexity reduction. With the SVD-based solution $(\tilde{\alpha}_1^*, \tilde{\alpha}_2^*, \tilde{\alpha}_3^*)$ achieved by problem (22), we can derive $\tilde{\mathbf{A}}^*$ and $\tilde{\mathbf{B}}^*$ according to (23) and (24). Let $\tilde{\mathbf{u}}_1$ and $\tilde{\mathbf{u}}_2$ be the basis of the null space of $\tilde{\mathbf{A}}^*$ and $\tilde{\mathbf{B}}^*$, respectively, and define $\hat{\mathbf{W}}_1 = \tilde{\mathbf{u}}_1\tilde{\mathbf{u}}_1^H, \hat{\mathbf{W}}_2 = \tilde{\mathbf{u}}_2\tilde{\mathbf{u}}_2^H$. Similar to (18), we have the precoding vectors based on SVD as $\tilde{\mathbf{w}}_1^* = \tilde{\tau}_1\tilde{\mathbf{u}}_1, \tilde{\mathbf{w}}_2^* = \tilde{\tau}_2\tilde{\mathbf{u}}_2$, where

$$\begin{cases} \tilde{\tau}_1 = \sqrt{\dfrac{(d^* - 1)\sigma^2}{tr\left(\mathbf{H}_1\hat{\mathbf{W}}_1\right)}}, \\ \tilde{\tau}_2 = \sqrt{P - \tilde{\tau}_1^2} \end{cases} \tag{25}$$

And the detailed steps of proposed SVD-based suboptimal algorithm are presented as Algorithm 2. Compared with Algorithm 1, the proposed SVD-based suboptimal scheme in Algorithm 2 further reduces the computational complexity without the Dinkelbach method.

Algorithm 2 The SVD-based suboptimal solution to problem (11)

1: Set $\mathbf{v}_0 = null(\mathbf{h}_2)$, where '$null(\cdot)$' is a MATLAB function which computes the orthonormal basis for the null space of a matrix using SVD.
2: Set $\tilde{\mathbf{H}}_1 = \mathbf{v}_0^H\mathbf{h}_1\mathbf{h}_1^H\mathbf{v}_0$.
3: Solve problem (22) to obtain $\left(\tilde{\alpha}_1^*, \tilde{\alpha}_2^*, \tilde{\alpha}_3^*\right)$
4: Calculate $\tilde{\mathbf{A}}^*$ and $\tilde{\mathbf{B}}^*$ according to (23) and (24), respectively.
5: Calculate $\tilde{\mathbf{w}}_1^*$, and $\tilde{\mathbf{w}}_2^*$ according to (25)
6: **return** $\tilde{\mathbf{w}}_1^*$ and $\tilde{\mathbf{w}}_2^*$

6 Simulation results and discussions

In this section, we numerically evaluate the performance of the proposed optimal and suboptimal schemes. The simulation parameters are listed in Table 1. It is assumed that the channels from WD to two ordinary users are deep fading channels and the information signal attenuations are 55 dB and 60 dB, respectively, corresponding to an identical distance of 15 m and 20 m. We set the parameter in the Dinkelbach method $\varepsilon = 10^{-4}$ and the power of noise $\sigma^2 = -50$ dBm. The channels $\mathbf{h}_1$ and $\mathbf{h}_2$ are assumed to be quasi-static flat Rayleigh fading, and each element of them follows an independent complex Gaussian distribution $CN(0, 1)$. All the simulation results are averaged over 1000 channel realizations. The optimal scheme and SVD-based scheme in this section, respectively, mean the optimal precoding vector scheme and the SVD-based precoding vector scheme of WD.

Table 1 Simulation parameters

Parameter	Value
Attenuation of $\mathbf{h}_1$	55 dB
Attenuation of $\mathbf{h}_2$	60 dB
Power of noise (σ^2)	− 50 dBm
Parameter in Dinkelbach (ε)	10^{-4}
Number of antennas at WD (N)	4, 6
Transmit power of WD (P)	20, 25 dBm
Target rate R_M	4, 6 bps/Hz
Number of channel realizations	1000

Figure 2 shows the maximal rate performance of different schemes versus the transmit power P under the number of antennas at WD $N = 4$ and $R_M = 4$ bps/Hz. Without loss of generality, time division multiple access (TDMA) is used to a representative of orthogonal multiple access (OMA), in which WD only serves single user in one time slot with joint power allocation among two time slots for two users. It can be observed that the proposed optimal and suboptimal scheme outperform traditional TDMA in terms of maximal rate of R_1, and this performance advantage is more obvious in the high transmit power region, though in TDMA scheme the s_1 is not interrupted by s_2. And it is noted that the proposed SVD-based suboptimal scheme only has a slight performance loss compared to the optimal scheme.

Figure 3 compares the maximal rate of R_1 versus the number of antennas at WD for different schemes under the transmit power of WD $P = 25$ dBm, when the target rate of s_2, i.e., R_M, is 4 bps/Hz and 6 bps/Hz, respectively. It is obviously noted that the maximal rate of R_1 is enhanced as the number of antennas grows. However, the growth trend gradually becomes slow. Besides, the gap between proposed optimal scheme and SVD-based schemes in terms of the maximal rate of R_1 is reducing with the increasing R_M.

In Fig. 4, the maximal rate region of R_1 versus R_M is characterized for different schemes with the number of antennas at WD $N = 4$, when the transmit power of WD is 25 dBm and 20 dBm, respectively. It is also noted that the optimal scheme achieves better rate regions than the SVD-based scheme. Furthermore, the higher target rate of s_2 requires, the smaller gap between the optimal and suboptimal scheme is. Meanwhile, the impact of transmission power at WD on the achieved rate regions for different

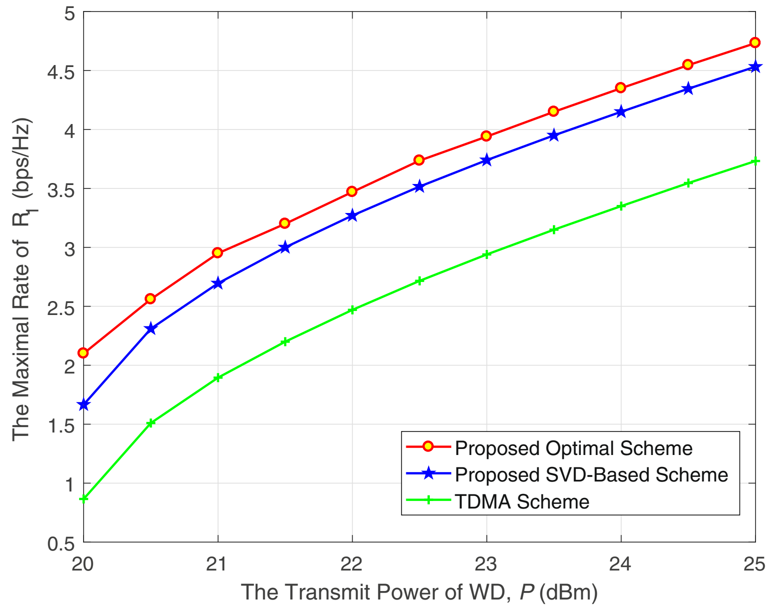

Fig. 2 Simulation 1. The maximal rate of R_1 versus the transmit power at WD for different schemes with $N = 4$ and $R_M = 4$ bps / Hz

schemes is also shown in Fig. 4. We can find that under certain transmit power of WD, when R_M is large, the rate of R_1 may be zero. The reason is that all power should be allocated to precoding vector $\mathbf{w}_2$ to first satisfy the target rate demand of s_2. So the precoding vector $\mathbf{w}_1$ has little effect on the system performance no matter it is designed in optimal solution or suboptimal solution. And the rate of R_1 becomes zero almost at the same value of target rate R_M for the optimal and SVD-based schemes.

Finally, Fig. 5 presents the outage performance of R_1 when R_M varies from 2 to 11 bps/Hz with the transmit power of WD $P = 25$ dBm and the number of antennas at WD $N = 4$. Given the transmit power of WD P, the number of antennas N, and the target rate R_M, the outage probability is $p_{\text{out}}(P, R_M, N) \triangleq \mathbb{P}(R_1 = 0)$. Especially, we set $P = 25$ dBm and $N = 4$. It is observed that the proposed optimal and suboptimal solutions achieve a similar performance to each other, and our proposed two

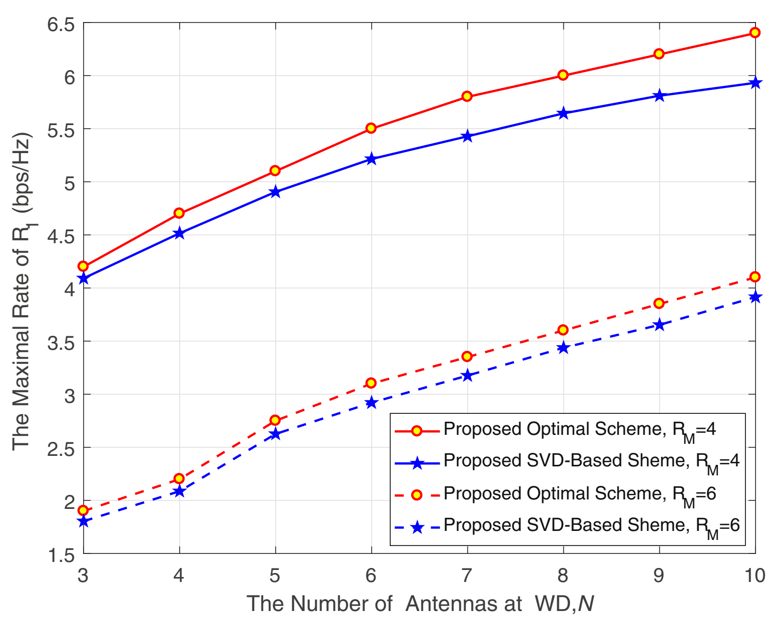

Fig. 3 Simulation 2. The maximal rate of R_1 versus the number of antennas at WD for different schemes with $P = 25$ dBm

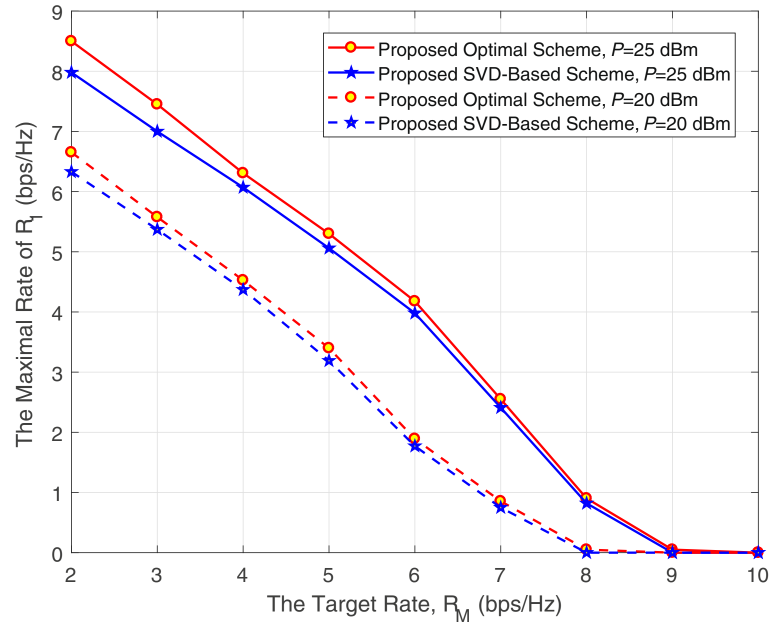

Fig. 4 Simulation 3. The maximal rate of R_1 versus R_M for different schemes with $N = 4$

schemes significantly decrease the outage probability of rate of R_1 compared with the TDMA scheme.

7 Conclusion

In this paper, an optimization problem of precoding vectors for two-user D2D communication underlaying system enabled by NOMA is investigated. Given the target rate R_M and the transmit power of WD P, the maximal rate of R_1 and corresponding optimal precoding vectors have been obtained. Then, the suboptimal solution based on SVD is proposed for complexity reduction. Finally, simulation results are provided to show the proposed optimal

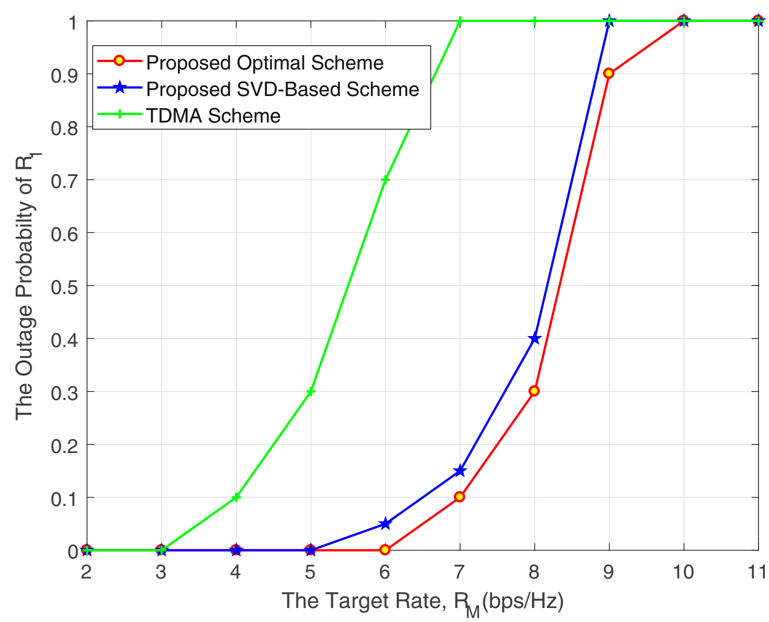

Fig. 5 Simulation 4. The outage probability of R_1 versus R_M for different schemes with N=4 and $P = 25$ dBm

and suboptimal precoding algorithms can outperform OMA scheme, such as TDMA.

Abbreviations
AF: Amplify-and-forward; AGN: Additive Gaussian noise; BS: Base station; CSI: Channel state information; CVX: Matlab software for disciplined convex programming; D2D: Device-to-device; IoT: Internet of Things; MIMO: Multiple-input multiple-output; NOMA: Non-orthogonal multiple access; OMA: Orthogonal multiple access; QoS: Quality of service; SDP: Semidefinite problem; SDP: Semidefinite relaxation; SIC: Successive interference cancellation; SINR: Signal-to-interference-plus-noise-ratio; SVD: Singular value decomposition; TDMA: Time division multiple access; WD: Wireless device

Acknowledgements
Not applicable.

Authors' contributions
PD is the main author of the current paper. PD contributed to the development of the ideas, design of the study, theory, result analysis, and article writing. WW contributed to the development of the ideas, design of the study, theory, and article writing. PD and PL conceived and designed the experiments. XS and PL performed the experiments. BW undertook revision works of the paper. All authors read and approved the final manuscript.

Funding
This paper was supported in part by the National Natural Science Foundation of China under Grant 61271232, the Natural Science Foundation of Jiangsu Province of China under BK20180757, the Project of Educational Commission of Jiangsu Province of China under 18KJB510028, the Introducing Talent Research Start-Up Fund of Nanjing University of Posts and Telecommunications under NY218100, and the Postgraduate Research & Practice Innovation Program of Jiangsu Province under Grant CXZZ13_0487.

Availability of data and materials
Data sharing is not applicable to this article as no datasets were generated or analyzed during the current study.

Competing interests
The authors declare that they have no competing interests.

Author details
[1]College of Communication and Information, Nanjing University of Posts and Telecommunications, Xinmofan Road 66, Nanjing 210003, China. [2]College of Automation and College Of Artificial Intelligence, Nanjing University of Posts and Telecommunications, Xinmofan Road 66, Nanjing 210003, China. [3]College of Overseas Education, Nanjing University of Posts and Telecommunications, Xinmofan Road 66, Nanjing 210003, China.

Received: 12 April 2019 Accepted: 24 June 2019
Published online: 19 July 2019

References

1. Z. Ding, et al., Application of non-orthogonal multiple access in LTE and 5G networks. IEEE Commun. Mag. **55**(2), 185-191 (2017)
2. R. Hu, Y. Qian, An energy efficient and spectrum efficient wireless heterogeneous network framework for 5G systems. IEEE Commun. Mag. **52**(5), 94–101 (2014)
3. Y. Saito, Y. Kishiyama, A. Benjebbour, T. Nakamura, A.LK. Higuchi, in *Proc. IEEE 77th Vehicular Technology Conference (VTC Spring)*. Non-orthogonal multiple access (NOMA) for cellular future radio access. (IEEE, Dresden, 2013), 1–5
4. K. Higuchi, A. Benjebbour, Non-orthogonal multiple access (NOMA) with successive interference cancellation. IEICE Trans. Commum. **E98-B**(3), 403–414 (2015)
5. Z. Ding, F. Adachi, H Poor.The application of MIMO to non-orthogonal multiple access. IEEE Trans. Wirel. Commum. Lett. **15**(1), 537–552 (2016)
6. D. Wan, M. Wen, F. Ji, Y. Liu, Y. Huang, Cooperative NOMA systems with partial channel state information over Nakagami- m fading channels. IEEE Trans. Commun. **66**(3), 947–958 (2018)
7. B. Wang, L. Dai, M. Xiao, *Millimeter Wave NOMA. Encyclopedia of Wireless Networks*. (Springer, Cham, 2018)
8. A. Benjebbour, Y. Saito, Y. Kishiyama, A. Li, A. Harada, T. Nakamura, in *Proc. 2013 International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS)*. Concept and practical considerations of non-orthogonal multiple access (NOMA) for future radio access. (IEEE, Naha, 2013), 770–774
9. W. Wu, B. Wang, Robust secrecy beamforming for wireless information and power transfer in multiuser MISO communication system. EURASIP J Wireless Com Network. **2015**(1), 161 (2015)
10. B. Kim, S. Lim, H. Kim, S. Suh, J. Kwun, S. Choi, C. Lee, S. Lee, D. Hong, in *Proc. MILCOM 2013-2013 IEEE Military Communications Conference*. Non-orthogonal multiple access in a downlink multiuser beamforming system (IEEE, San Diego, 2013), pp. 1278-1283
11. H. ElSawy, E. Hossain, M. Alouini, Analytical modeling of mode selection and power control for underlay D2D communication in cellular networks. IEEE Trans. Commun. **62**(11), 4147–4161 (2014)
12. K. Doppler, C. Yu, C. Ribeiro, P. Janis, in *Proc. 2010 IEEE Wireless Communication and Networking Conference*. Mode selection for device-to-device communication underlaying an LTE-advanced network. (IEEE, Sydney, 2010), 1–6
13. K. Akkarajitsakul, P. Phunchongharn, E. Hossain, V.K. Bhargava, in *Proc 2012 IEEE International Conference on Communication Systems (ICCS)*. Mode selection for energy-efficient D2D communications in LTE-advanced networks: A coalitional game approach (IEEE, Singapore, 2012), pp. 488–492
14. D. Zhu, J. Wang, A. Swindlehurst, C. Zhao, Downlink resource reuse for device-to-device communications underlaying cellular networks. IEEE Signal Process. Lett. **21**(5), 531–534 (2014)
15. J. Zhao, Y. Liu, K. Chai, Y. Chen, M. Elkashlan, J. Alonso-Zarate, in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*. NOMA-based D2D communications: towards 5G. (IEEE, Washington, 2016), 1–6
16. J. Zhao, Y. Liu, K. Chai, Y. Chen, M. Elkashlan, Joint subchannel and power allocation for NOMA enhanced D2D communications. IEEE Trans. Commun. **65**(11), 5081–5094 (2017)
17. Z. Zhang, Z. Ma, M. Xiao, Z. Ding, P. Fan, Full-duplex device-to-device-aided cooperative nonorthogonal multiple access. IEEE Trans. Veh. Technol. **66**(5), 4467–4471 (2017)
18. C. Li, H. Yang, F. Sun, J. Cioffi, L. Yang, Multiuser overhearing for cooperative two-way multiantenna relays. IEEE Trans. Veh. Technol. **65**(5), 3796–3802 (2016)
19. C. Li, S. Zhang, P. Liu, F. Sun, J. Cioffi, L. Yang, Overhearing protocol design exploiting inter-cell interference in cooperative green networks. IEEE Trans. Veh. Technol. **65**(1), 441–446 (2016)
20. Z. Ding, X. Lei, G. Karagiannidis, R. Schober, J. Yuan, V. Bhargava, A survey on non-orthogonal multiple access for 5G networks: research challenges and future trends. IEEE J. Select. Areas Commun. **35**(10), 2181–2195 (2017)
21. Z. Ding, P. Fan, H. Poor, Impact of user pairing on 5G non-orthogonal multiple access downlink transmissions. IEEE Trans. Veh. Technol. **65**(8), 6010–6023 (2016)
22. C. Li, P. Liu, C. Zou, F. Sun, J. Cioffi, L. Yang, Spectral-efficient cellular communications with coexistent one- and two-hop transmissions. IEEE Trans. Veh. Technol. **65**(8), 6010–6023 (2016)
23. C. Li, H. Yang, F. Sun, J. Cioffi, L. Yang, Adaptive overhearing in two-way multi-antenna relay channels. IEEE Signal Process. Lett. **23**(1), 117–120 (2016)
24. F. Zhou, Y. Wu, R. Hu, Y. Wang, K.-K. Wong, Energy-efficient NOMA heterogeneous cloud radio access networks. IEEE Netw. **32**(2), 152–160 (2017)
25. W. Wu, F. Zhou, P.ei. Li, P. Deng, B. Wang, V.CM. Leung, in *Proc IEEE Inter. Conf. Commun. (ICC)*. Energy-efficient secure NOMA-enabled mobile edge computing networks (IEEE, Shanghai, 2019)
26. P. Gover, A. Sahai, in *Proc. IEEE International Symposium on Information Theory*. Shannon meets Tesla: wireless information and power transfer. (IEEE, Austin, 2010), 2363–2367
27. Z. Luo, W. Ma, A.M. So, Y. Ye, S. Zhang, Semidefinite relaxation of quadratic optimization problems. IEEE Signal Process. Mag. **27**(3), 20–34 (2010)
28. H. Zhang, Y. Huang, S. Li, L. Yang, Energy-efficient precoder design for MIMO wiretap channels. IEEE Commun. Lett. **18**(9), 1559–1562 (2014)
29. J. Xu, L. Qiu, Energy efficiency optimization for MIMO broadcast channels. IEEE Trans. Wirel. Commun. **12**(2), 690–701 (2013)
30. M. Grant, S. Boyd, CVX: Matlab software for disciplined convex programming, version 2.0 beta (2013). http://cvxr.com/cvx

31. Y. Huang, D. Palomar, Rank-constrained separable semidefinite programming with applications to optimal beamforming. IEEE Trans. Signal Proc. **58**(2), 664–678 (2010)
32. C. Li, F. Sun, J. Cioffi, L. Yang, Energy efficient MIMO relay transmissions via joint power allocations. IEEE Trans. Circ. Syst. **61**(7), 531–535 (2014)

Publisher's Note