Received June 9, 2020, accepted June 28, 2020, date of publication July 6, 2020, date of current version November 19, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3007526

Interference Channels With Full-Duplex Amplify-and-Forward Receiver Cooperations

DAN WANG[1], (Graduate Student Member, IEEE), JIANHAO HUANG[2], (Graduate Student Member, IEEE), CHUAN HUANG[2], (Member, IEEE), AND GUOWEI SHI[3]

[1]National Key Laboratory of Science and Technology on Communications, University of Electronic Science and Technology of China, Chengdu 611731, China
[2]School of Science and Engineering, The Chinese University of Hong Kong, Shenzhen 518172, China
[3]China Academy of Telecommunication Research, Beijing 102209, China

Corresponding author: Chuan Huang (huangchuan@cuhk.edu.cn)

This work was supported in part by the National Key R&D Program of China with grant No. 2018YFB1800800, by the Key Area R&D Program of Guangdong Province with grant No. 2018B030338001, by Shenzhen Outstanding Talents Training Fund, and by Guangdong Research Project No. 2017ZT07X152.

ABSTRACT This paper considers a two-transmitter and two-receiver interference channel (IC), where each transmitter sends a message to its desired receiver. In particular, the full-duplex (FD) amplify-and-forward (AF) protocol is adopted to build up the receiver cooperations: each receiver can receive the signals from the two transmitters and two receivers, and after self-interference (SI) cancellation and message decoding, it forwards them to its counterpart. With the above scheme, the equivalent channel model is analyzed, and the statistics of the accumulated residual interference and noise (ARIN), generated by the imperfect SI cancellation and the AF scheme at the receivers, is calculated by mathematical induction. Then, the achievable rate regions of both the single-user and joint decoding schemes are derived. It is proved that one-side cooperation, i.e., only one of the two receivers forwards its counterpart's signal, is optimal to achieve the best system performance. Next, to characterize the obtained rate regions, the rate maximization problems are formulated and approximately solved by a sequential parametric convex approximation (SPCA) method. Simulation results show that the proposed scheme can improve the achievable rate compared to the conventional non-cooperative scheme in several typical scenarios.

INDEX TERMS Interference channel (IC), full-duplex (FD), amplify-and-forward (AF), receiver cooperations.

I. INTRODUCTION

For the upcoming 5G wireless communications, a large number of base stations (BSs) are densely deployed in one area, and the distances between the mobile terminals and BSs can be significantly reduced, which is promising to achieve higher energy efficiency and larger throughput [1]. However, too close proximity of these small cells generates more complicated inter-cell interference than the conventional cellular systems. For instance, two mobile terminals, which are close to each other and located at the edges of two cells, respectively, will interfere with each other when they share the same frequency band, and this scenario is modeled as an interference channel (IC) [2], [3].

For a two-user Gaussian IC, the best known achievable result [4] was obtained by superposition coding [5], where the transmitters independently encode their messages as superpositions of two sub-messages: public message, which is decoded at both the receivers, and private message, which is only decoded at the intended receiver and is simply treated as interference at the undesired receiver. For a general IC, [6] derived new outer bounds and characterized the capacity region of the IC to within one bit by using the same scheme in [4]. Afterward, the outer bounds were further developed in [7]–[9] and the sum rates of the IC were characterized in the weak and strong interference regimes, respectively. In particular, when the cross interference of the IC is very weak [7], the optimal scheme to achieve the channel capacity is the single-user decoding [8], in which each receiver decodes its desired message and treats the interference from the other transmitter simply as noise. When the cross interference is very strong [7], joint decoding scheme, which treats the network as a multiple access channel at each receiver and decodes both the two messages, is applied at each receiver to decode both the source messages to achieve the capacity of the IC [9]–[13].

In the aforementioned IC setup, these receivers cannot communicate with each other and receiver cooperations were

The associate editor coordinating the review of this manuscript and approving it for publication was Cunhua Pan.

not considered. However, via out-of-band or in-band links, receivers may own the capability to talk to each other and exchange certain amount of source information [14]–[17], and the receiver cooperations can be built up. Due to the different duplex modes at the receivers, receiver cooperations can be divided into two categories: out-of-band conferencing and in-band full-duplex (FD) communications. For the out-of-band receiver cooperations, the receivers communicate with each other via the conferencing links, which are orthogonal to the other transmitter-receiver links. Receiver cooperations with the decode-and-forward (DF) protocol was studied in [14], where the overall transmission period was divided into the transmit phase for the transmitters and the cooperative phase for the receivers, and it was showed that the two-user IC with receiver cooperations outperforms the time division multiple access system. The authors in [15] focused on an improved compress-and-forward (CF) protocol and studied the interference mitigate ability through limited receiver cooperations by using the same method in [6], and it was revealed that extra degree-of-freedom and power gains were obtained by introducing receiver cooperations.

On the other hand, for receiver cooperations with in-band FD communications, the receivers can simultaneously transmit and receive in the same frequency band. In contrast with the out-of-band case, it economizes the frequency resources while causing more complicated self-interference (SI) at the receivers [16], [17]. The authors in [16] investigated the outer bounds on the capacity region for the two-user Gaussian IC with in-band receiver cooperations, and both the CF and DF schemes were applied to derive the achievable rates for both the asynchronous and synchronous cases. Under a similar setup, the authors in [17] employed a coding scheme based on superposition [4] and derived some new upper bounds on the sum rate for the case of symmetric in-band receiver cooperations. All the above related works assumed ideal SI cancellation at each FD receiver. However, the practical SI cancellation is always imperfect, and the impact of the residual SI on the achievable rates of the considered channel needs to be carefully investigated.

In this paper, we consider a two-transmitter and two-receiver IC with receiver cooperations. As studied in the previous work [18], the cooperative links between the two receivers were established by in-band FD communications: The transmitter-receiver communications and the receiver cooperations work in the same frequency band; and the SI cancellation at the receivers is imperfect, and the residual SI is modeled as additive noise. We adopt the low-complexity amplify-and-forward (AF) protocol at each receiver, where each receiver amplifies the received signals from the two transmitters and forwards them to the counterpart receiver. Under the above setup, each receiver will receive the same messages at two consecutive time slots, the messages at the former time slot directly from the transmitters while the latter messages forwarded by the counterpart receiver. To improve the achievable rates for the considered system, a new decoding strategy is proposed, which supports each receiver to decode messages by combining the received signals at two consecutive time slots, and thus potential power gains can be achieved [19]. Meanwhile, the residual SI and the additive noise at each receiver are forwarded to its counterpart and cannot be canceled under the AF scheme, such that they will be accumulated over time. Compared with the transmitter cooperations [20], the potential power gains for the receiver cooperations completely come from the cooperations between the receivers, without introducing the beamforming technique. The main contributions of this paper are summarized as follows:

1) First, we propose an in-band FD receiver cooperations framework with the AF scheme for a two-user Gaussian IC and analyze the equivalent channel model. In particular, each receiver receives the same messages at two consecutive time slots; after SI cancellation, a new decoding strategy is proposed at each receiver to decode its desired message by combining the received signals at two consecutive time slots; then, each receiver removes the decoded message and forwards the remainders to its counterpart by using the AF scheme. We reveal that the considered system is equivalent to a two-tap IC [21], with the additive noise being the accumulated residual interference and noise (ARIN).
2) Then, the statistics of the ARIN is analyzed. The ARIN evolving over time is shown to form a Markov chain. As time goes to infinity, the ARIN will converge to a stationary state by properly designing the transmit power of the two transmitters and the two receivers. Under the obtained stationary state, the power of ARIN is computed.
3) Finally, under the stationary state, the achievable rates for the two-user Gaussian IC are derived based on both the single-user and joint decoding schemes, respectively. We prove that one-side cooperation, i.e., only one of the two receivers forwards its counterpart's signal, is optimal to achieve the best system performance. To characterize the achievable rate regions, the rate maximization problem is formulated and approximately solved by a sequential parametric convex approximation (SPCA) method: First, change the design variables, and the constraint functions are equivalently transformed into the concave-convex form; then, the lower bound of the new constraints are obtained by the concave-convex procedure (CCCP); finally, the original optimization problem is solved by successively optimizing a sequence of convex approximations of the original problems.

The remainder of this paper is organized as follows. Section II introduces the system model. Section III analyzes the statistics of the ARIN, and derives the achievable rate regions for single-user and joint decoding schemes, and then proposes an efficient algorithm to characterize the

achievable rate regions. Section IV presents the simulation results. Section V concludes this paper.

Notations: Bold-face upper-case letters, e.g., $\boldsymbol{X}$, denote matrices, and bold-face lower-case letters, e.g., $\boldsymbol{x}$, denote vectors. In addition, $\boldsymbol{X}^T$ and $\boldsymbol{X}^H$ denote the transpose and conjugate transpose of matrix $\boldsymbol{X}$. $\log(x)$, $|x|$ and x^* denote the base-2 logarithm, the 2-norm and the conjugate of x, respectively. $\mathbb{E}(X)$ represents the mathematical expectation of a random variable X.

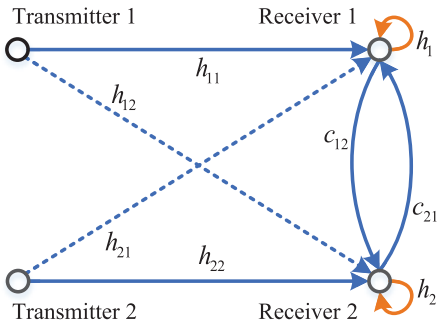

FIGURE 1. Gaussian IC with two full-duplex receivers.

II. SYSTEM MODEL

In this paper, a Gaussian IC with two transmitters and two receivers is considered, as shown in Fig. 1, where the two transmitters send messages to the receivers, and the two receivers can cooperatively decode the desired messages. In particular, receiver cooperations are established between the two receivers by in-band FD communications: Each receiver can simultaneously transmit (i.e., forward the received signals to its counterpart by the AF scheme) and receive in the same frequency band, and each receiver is capable to receive the signals from the two transmitters, its counterpart receiver, and itself. The signal directly received from itself is treated as SI, which needs to be cancelled before further processions. The cooperative links between the two receivers are with the same frequency band as the others of the IC. In this paper, the transmissions for the source messages of N blocks are considered, and they are encoded individually into codewords $x_k(1), x_k(2), \cdots, x_k(N)$. Due to the one-block delay produced by the AF operation at receivers, the transmissions of N blocks cost $N+1$ time slots. As N goes to infinity, the considered system will reach a stationary state, and the asymptotic performance will be analyzed in this paper.

Under the above setup, the transmissions and receptions for the considered Gaussian IC are proposed and shown in Fig. 2, where the counterpart of receiver j, $j \in \{1, 2\}$, is denoted as receiver $\bar{j}, \bar{j} \in \{2, 1\}$. At the first time slot, transmitter j sends signal $t_j(1) = \mu_j x_j(1)$ to the two receivers with μ_j being the transmission parameter at transmitter j, and each receiver receives the signals $t_1(1)$ and $t_2(1)$ from the two transmitters, as well as the additive noise $n_j(1)$. At the second time slot, receiver j forwards the received signal $r_j(1) = h_{1j}t_1(1) + h_{2j}t_2(1) + n_j(1)$ obtained at the first time slot to its counterpart via the AF scheme, where $h_{kj}, k = 1, 2$, is the channel coefficient from transmitter k to receiver j as shown in Fig. 1. Meanwhile, receiver j receives the signals $s_j(2) = \omega_j r_j(1)$ and $s_{\bar{j}}(2) = \omega_{\bar{j}} r_{\bar{j}}(1)$ from itself and its counterpart with ω_j being the transmission parameter at receiver j, as well as the transmit signals $t_1(2)$ and $t_2(2)$ from the two transmitters. For the received signal $r_j(2) = h_{1j}t_1(2) + h_{2j}t_2(2) + h_j s_j(2) + c_{\bar{j}j} s_{\bar{j}}(2) + n_j(2)$ (The channel coefficients h_j and $c_{\bar{j}j}$ will be defined later in (5) and (7)), $s_j(2)$ is treated as SI, with significantly large power compared with the other parts of the received signals. As a rule of thumb, due to the imperfect SI suppression, the residual SI denoted as $\hat{e}_j^I(2)$ is modeled as additive noise and forwarded to the counterpart receiver during the further transmissions. After the SI cancellation, a heuristic thought motivates us to decode and completely cancel the signals $x_1(1)$ and $x_2(1)$ before directly forwarding the received signal.[1] Then, the residual parts of the received signals consisting of $t_1(2)$, $t_2(2)$ and the residual SI and noise $\hat{e}_j(2) = \hat{e}_j^I(2) + n_j(2) + c_{\bar{j}j}\omega_{\bar{j}} n_{\bar{j}}(1)$ will be forwarded to the counterpart receiver by using the AF scheme,[2] while $\hat{e}_j(2)$ is remained in the signals at the rest of transmission blocks and generates the accumulated residual interference and noise (ARIN). At the first two time slots, both $x_1(1)$ and $x_2(1)$ are received twice at each receiver and potential power gains can be achieved by properly combining the two copies of each signal [19]. Compared with the DF scheme (i.e., decode the desired messages and forward a noiseless version of these messages) [22], the adopted AF-based scheme is simpler to be implemented.

More generally, the signal transmissions and receptions at the i-th time slot are introduced in the sequel.

A. TRANSMISSIONS AT THE TRANSMITTERS

At the i-th time slot, transmitter j sends signal $t_j(i)$ to the two receivers, and the transmit signal $t_j(i)$ is given as

$$t_j(i) = \mu_j x_j(i), \quad j = 1, 2, \tag{1}$$

where μ_j is the transmission parameter at transmitter j, and signal $x_j(i)$ is with unit power, i.e., $\mathbb{E}\left(|x_j(i)|^2\right) = 1$. Besides, we define the transmission parameter region as

$$\mathcal{P} = \left\{(\mu_1, \mu_2) : 0 \le |\mu_1|^2 \le P_{T_1}, 0 \le |\mu_2|^2 \le P_{T_2}\right\}, \tag{2}$$

where P_{T_1} and P_{T_2} are the maximum power values for transmitters 1 and 2, respectively.

[1] Note that signals $x_1(1)$ and $x_2(1)$ are decoded at the second time slot, and then are completely canceled. It is worth pointing out that this elimination process for signals $x_1(1)$ and $x_2(1)$ is not SI cancellation.

[2] For the sake of simplicity, only the transmitted signals at receivers for the joint decoding are shown in Fig. 2. To obtain the transmitted signals at receivers for the single-user decoding given in (15), we only need to add the term of $c_{\bar{j}j}\omega_{\bar{j}} h_{\bar{j}\bar{j}} t_{\bar{j}}(i-2)$ to the last row of Fig. 2.

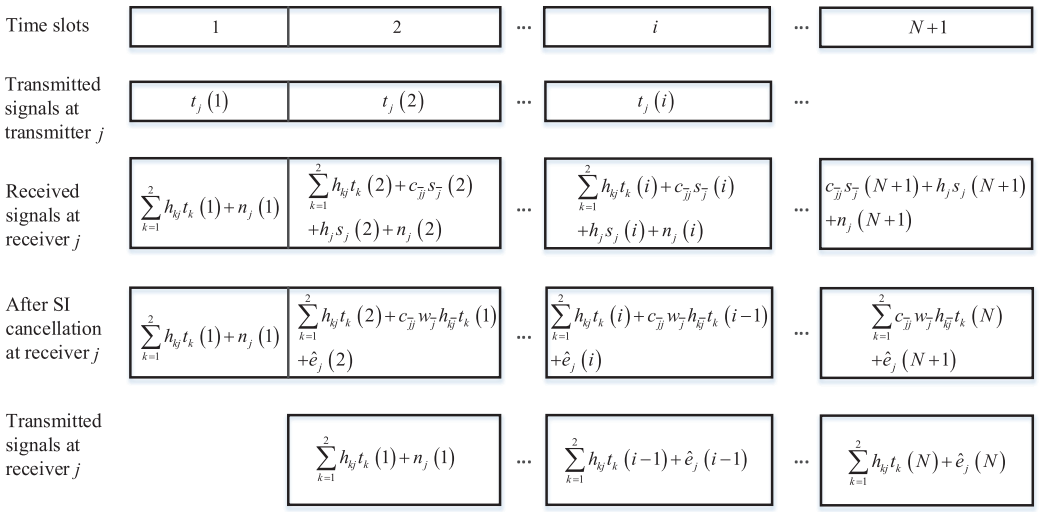

FIGURE 2. Signal transmissions and receptions over $N+1$ time slots.

B. TRANSMISSIONS AT THE RECEIVERS

At the i-th time slot, receiver j sends signal $y_j(i-1)$ after SI cancellation and other processes to its counterpart receiver $\bar{j}$, where $y_j(i-1)$ contains the transmit signals from the two transmitters and the ARIN at the previous time slot (Forward single-user and joint decoding schemes are adopted in this paper, which support each receiver to decode signals in sequence. For the two decoding schemes introduced in Section III-A, $y_j(i)$ are given in (15) and (19)), respectively). Thus, the transmit signal $s_j(i)$ at receiver j is given as

$$s_j(i) = \omega_j y_j(i-1), \quad j = 1, 2, \tag{3}$$

where ω_j is the transmission parameter at receiver j to be designed in this paper. The transmit signal $s_j(i)$ satisfies a power constraint, i.e.,

$$\mathbb{E}\left(|s_j(i)|^2\right) = |\omega_j|^2 \mathbb{E}\left(|y_j(i-1)|^2\right) \le P_{R_j}, \tag{4}$$

where P_{R_j} is the power budget at receiver j.

C. SI CANCELLATION AT RECEIVERS

At the i-th time slot, the received SI signal at receiver j from its local transmitter is defined as

$$e_j^I(i) = h_j s_j(i), \tag{5}$$

where h_j denotes the SI channel coefficient for receiver j. Without loss of generality, the SI signal $e_j^I(i)$ in (5) can be partly canceled [23]. Then, residual part of $e_j^I(i)$ is denoted as $\hat{e}_j^I(i)$. Following a similar SI cancellation process in [24], the residual SI $\hat{e}_j^I(i)$ is modeled as independent and identically distributed (i.i.d.) complex symmetric circularly Gaussian (CSCG) noise with mean zero and variance $\hat{P}_j$ in this paper, i.e.,

$$\hat{e}_j^I(i) \sim \mathcal{CN}(0, \hat{P}_j), \tag{6}$$

where $\hat{P}_j$ is the power of $\hat{e}_j^I(i)$. In particular, when $\omega_j = 0$, $\hat{P}_j$ is set to be zero.[3]

D. RECEPTIONS AT THE RECEIVERS

At the i-th time slot, receiver j receives signals from the two transmitters, the counterpart receiver $\bar{j}$, and itself. After SI cancellation, the received signals at receiver j becomes

$$d_j(i) = h_{1j}t_1(i) + h_{2j}t_2(i) + c_{\bar{j}j}s_{\bar{j}}(i) + \hat{e}_j^I(i) + n_j(i), \tag{7}$$

where $h_{kj}, k = 1, 2$, denotes the channel coefficient from transmitter k to receiver j, $c_{\bar{j}j}$ denotes the cooperative channel coefficient from the receiver $\bar{j}$ to receiver j, $\hat{e}_j^I(i)$ is the residual SI signal defined in (5), and $n_j(i)$ is the i.i.d. CSCG noise with mean zero and unit variance. Then, we define the ARIN at receiver j as

$$\hat{e}_j(i) = e_j^N(i) + \hat{e}_j^I(i) + n_j(i), \tag{8}$$

with $e_j^N(i) = c_{\bar{j}j}\omega_{\bar{j}}\hat{e}_{\bar{j}}(i-1)$ being the forwarded ARIN from the counterpart receiver $\bar{j}$. In (8), it is observed that the ARIN at the i-th time slot are determined by the ARIN at the $(i-1)$-th time slot, and thus the ARIN forms a Markov chain [25]. As $i \to \infty$, the ARIN will go to a stationary state by properly designing the transmission parameters of the two transmitters and two receivers.

It is worthy pointing out that forward single-user and joint decoding schemes are adopted in this paper, which support each receiver to decode signals in sequence. We combine the received signals at two continuous time slots, i.e., $d_j(i)$ and $d_j(i-1)$, to decode $x_k(i-1)$ at the i-th time slot, which will be introduced detailedly in Section III-A. For the joint decoding, each receiver decodes both of the two messages from the two transmitters. Thus, the forwarded signal at the i-th time

[3] If $\omega_j = 0$, receiver j does not forward any received signal and does not suffer from any SI. Thus, the power of the residual SI is zero.

at receiver j only includes the transmit signals from the two transmitters and the ARIN at the previous time slot, which is given in (19). By substituting (1), (3), (8) and (19) into (7), $d_j(i)$ defined in (7) under the joint decoding is rewritten as

$$d_j(i) = h_{1j}\mu_1 x_1(i) + h_{2j}\mu_2 x_2(i) + c_{\bar{j}\bar{j}}\omega_{\bar{j}} h_{1\bar{j}}\mu_1 x_1(i-1) + c_{\bar{j}\bar{j}}\omega_{\bar{j}} h_{2\bar{j}}\mu_2 x_2(i-1) + \hat{e}_j(i). \quad (9)$$

However, single-user decoding supports each receiver only to decode its desired message, and the undesired message will be forwarded to the counterpart receiver. Thus, the forwarded signal at receiver j at the i-th time slot also contains $x_{\bar{j}}(i-1)$, which is give in (15). Similarly, By substituting (1), (3), (8) and (15) into (7), $d_j(i)$ defined in (7) under the single-user decoding is rewritten as

$$\begin{aligned} d_j(i) = {} & h_{1j}\mu_1 x_1(i) + h_{2j}\mu_2 x_2(i) + c_{\bar{j}\bar{j}}\omega_{\bar{j}} h_{1\bar{j}}\mu_1 x_1(i-1) \\ & + c_{\bar{j}\bar{j}}\omega_{\bar{j}} h_{2\bar{j}}\mu_2 x_2(i-1) + \hat{e}_j(i) \\ & + c_{\bar{j}\bar{j}}\omega_{\bar{j}} c_{j\bar{j}}\omega_j h_{j\bar{j}}\mu_j x_j(i-2), \end{aligned} \quad (10)$$

where $x_j(i-2)$ has been decoded at receiver j at the $(i-1)$-th time slot. When being forwarded back to receiver j again at the i-th time slot, it as known signal can be completely canceled. Thus, the term of $x_j(i-2)$ can be removed from $d_j(i)$ in (10), and then (10) becomes (9).

For the sake of consistency, we rewrite (9) and (10) in an universal form, i.e.,

$$d_j(i) = h_{1j}\mu_1 x_1(i) + h_{2j}\mu_2 x_2(i) + c_{\bar{j}\bar{j}}\omega_{\bar{j}} h_{1\bar{j}}\mu_1 x_1(i-1) + c_{\bar{j}\bar{j}}\omega_{\bar{j}} h_{2\bar{j}}\mu_2 x_2(i-1) + \hat{e}_j(i). \quad (11)$$

In (11), $x_1(i)$ and $x_2(i)$ are directly from transmitters 1 and 2, respectively; $x_1(i-1)$ and $x_2(i-1)$ are forwarded by the counterpart receiver $\bar{j}$; $\hat{e}_j(i)$ is the ARIN at receiver j.

Remark 1: From (11), the equivalent channel model is actually a two-tap IC [21]. Due to the one-block delay produced by the AF operation at the receivers, signal $x_k(i-1), k = 1, 2$, at the $(i-1)$-th time slot interferes with signal $x_k(i)$ at the i-th time slot. If receiver j directly forwards signal $d_j(i)$ to its counterpart, signal $x_k(i-1)$ will be accumulated over time such that it causes large interference to the useful signals. To solve this problem, $x_k(i-1)$ will be removed before decoding the message in codeword $x_k(i)$.

III. ACHIEVABLE RATE REGIONS

In this section, a decoding scheme based on the forward decoding is first proposed. Then, the statistics of the ARIN is analyzed. Based on the single-user and joint decoding schemes, the corresponding achievable rate regions for the considered IC with receiver cooperations are characterized.

A. FORWARD DECODING

From the two-tap channel model in (11), it is easy to see that jointly decoding these N-block messages based on the received signals at all time slots is optimal for the considered IC with receiver cooperations. However, with this optimal decoding scheme, each receiver is required to buffer all the received signals and the decoding complexity increases exponentially with N increasing, which makes the decoding too complicated to be implemented. Therefore, a suboptimal decoding scheme based on the forward decoding is proposed: At each time, only one pair of source messages $x_1(i)$ and $x_2(i)$ are decoded, based on the received signals $d_j(i)$ and $d_j(i+1), j = 1, 2$, at two consecutive time slots; and each receiver only needs to buffer the received signals $d_j(i)$ and $d_j(i+1), j = 1, 2$. For the backward decoding, where each receiver decodes the source message from the last time slot to the first time slot, it is also required to buffer all the received signals, and thus we will not consider it in this paper.

The above mentioned forward decoding scheme supports the receivers to decode the source messages in sequence. At the $(i+1)$-th time slot, $x_k(i-1), k = 1, 2$, has been successfully decoded, and thus it can be completely canceled from $d_j(i)$.

Based on the above analysis, single-user and joint decoding schemes [26] are considered, and the corresponding achievable rates are characterized for the considered IC with receiver cooperations.

1) Single-user decoding: Each receiver decodes its desired message and treats the interference from other transmitter as noise [27]. Based on the channel model in (11), the equivalent channel input and output relationship of the single-user decoding scheme is given as

$$\boldsymbol{d}_j(i) = \boldsymbol{h}_{jj} x_j(i) + \boldsymbol{z}_j^S(i), \quad (12)$$

where $\boldsymbol{d}_j(i) = \left[d_j(i+1), d_j(i)\right]^T$, $\boldsymbol{h}_{jj} = \left[b_{jj}, a_{jj}\right]^T$, with $a_{jj} = h_{jj}\mu_j$ and $b_{jj} = c_{\bar{j}\bar{j}}\omega_{\bar{j}} h_{\bar{j}\bar{j}}\mu_j$. At the $(i+1)$-th time slot, signals $x_1(i-1)$ and $x_2(i-1)$ are completely canceled, while signals $x_1(i+1)$ and $x_2(i+1)$, as well as the undesired signal $x_{\bar{j}}(i)$ are treated as interference. Hence, the additive noise in (12) is given as $\boldsymbol{z}_j^S(i) = \left[z_j^S(i+1), z_{\bar{j}}^S(i)\right]^T$, with

$$z_j^S(i+1) = a_{1j}x_1(i+1) + a_{2j}x_2(i+1) + b_{\bar{j}\bar{j}}x_{\bar{j}}(i) + \hat{e}_j(i+1), \quad (13)$$

$$z_{\bar{j}}^S(i) = a_{\bar{j}\bar{j}}x_{\bar{j}}(i) + \hat{e}_{\bar{j}}(i), \quad (14)$$

where $a_{kj} = h_{kj}\mu_k$ and $b_{kj} = c_{\bar{j}\bar{j}}\omega_{\bar{j}} h_{k\bar{j}}\mu_k$, $k = 1, 2$. For the single-user decoding, all the remaining signals in (13) will be forwarded at the next time slot. Thus, the forwarded signal $y_j(i)$ defined in (3) for the single-user decoding is rewritten as

$$y_j(i) = a_{1j}x_1(i) + a_{2j}x_2(i) + \hat{e}_j(i) + b_{\bar{j}\bar{j}}x_{\bar{j}}(i-1). \quad (15)$$

By substituting (1), (3), and (15) into (7), detailed expression for $d(i)$ given in (10) is derived. Note that $x_{\bar{j}}(i-2)$ in (10) has been decoded at receiver $\bar{j}$ at the $(i-1)$-th time slot. When being forwarded back to receiver j again at the i-th time slot, $x_j(i-2)$ as known signal can be completely canceled at receiver $\bar{j}$. Thus, by canceling the term of $x_j(i-2)$ at receiver j, (10) is rewritten as (11).

2) Joint decoding: Each receiver decodes both the two messages from the two transmitters [10]–[13]. The equivalent channel input and output relationship of the joint decoding scheme is given as

$$\boldsymbol{d}_j(i) = \boldsymbol{h}_{1j}x_1(i) + \boldsymbol{h}_{2j}x_2(i) + \boldsymbol{z}_j^J(i), \tag{16}$$

where the additive noise is given as $\boldsymbol{z}_j^J(i) = \left[z_j^J(i+1), z_j^J(i)\right]^T$, with

$$z_j^J(i+1) = a_{1j}x_1(i+1) + a_{2j}x_2(i+1) + \hat{e}_j(i+1), \tag{17}$$

$$z_j^J(i) = \hat{e}_j(i), \tag{18}$$

where $\boldsymbol{h}_{kj}$ is defined as $\boldsymbol{h}_{kj} = \left[b_{kj}, a_{kj}\right]^T$, and $\boldsymbol{d}_j$, a_{kj}, and b_{kj}, $k = 1, 2$, are defined the same as the previous decoding scheme. Similarly, all the remaining signals in (17) will be forwarded at the next time slot. The forwarded signal $y_j(i)$ defined in (3) for the joint decoding is rewritten as

$$y_j(i) = a_{1j}x_1(i) + a_{2j}x_2(i) + \hat{e}_j(i). \tag{19}$$

By substituting (1), (3), and (19) into (7), detailed expression for $d(i)$ given in (9) (the same as (11)) is also derived.

B. STATISTICS OF ARIN

After the decoding of message $x_k(i-1)$, the transmit signals $y_j(i)$ at receivers for the single-user decoding and joint decoding are obtained in (15) and (19), respectively. To compute the power of $y_j(i)$, the statistics of ARIN $\hat{e}_j(i)$ at the stationary state, i.e., N and $i \to \infty$, is first analyzed in this subsection.

Proposition 1: The ARIN process given in (8) is a Markov chain. Thus, the general expression of the ARIN $\hat{e}_j(i)$ is calculated by mathematical induction as (20), shown at the bottom of this page, where $\hat{e}_j^I(i) = 0, n_j(i) = 0, j = 1, 2$, for $i \le 0$, $\lfloor x \rfloor$ rounds x down to an integer, and a is defined as

$$a = c_{\bar{j}j}\omega_{\bar{j}}c_{j\bar{j}}\omega_j. \tag{21}$$

Proposition 2: The power of ARIN at the i-th time slot with N and $i \to \infty$ is finite, if and only if $|a| < 1$.

Proof: In (20), the terms in $(\cdot)$ are the CSCG noises with finite and non-zero power and are mentioned in the previous section. Then, we only focus on a. If $|a| \ge 1$, the power of ARIN will go to infinity, which contradicts with the constraint in (4). When $|a| < 1$, the power of ARIN $\hat{e}_j(i)$ converges to a finite constant. ■

Next, the statistics of the ARIN is computed with $|a| < 1$ as both N and i go to infinity.

Proposition 3: As both N and i go to infinity, the correlation function of ARIN is computed as

$$\mathbb{E}(\hat{e}_j(i)\hat{e}_j^*(i+\delta)) \to A_j(\delta) = \begin{cases} \frac{a^{-\delta}}{1-|a|^2}\left(1+\hat{P}_j+c_{jj}^2\omega_j^2(\hat{P}_{\bar{j}}+1)\right), & \delta \text{ is even, and } \delta \le 0; \\ \frac{(a^*)^{\delta}}{1-|a|^2}\left(1+\hat{P}_j+c_{jj}^2\omega_j^2(\hat{P}_{\bar{j}}+1)\right), & \delta \text{ is even, and } \delta \ge 0; \\ 0, & \delta \text{ is odd}; \end{cases} \tag{22}$$

where a is defined in (21), and $\hat{P}_j$ and $\hat{P}_{\bar{j}}$ are the variance of $\hat{e}_j^I$ and $\hat{e}_{\bar{j}}^I$, respectively.

Proof: Please see Appendix A. ■

Corollary 1: By setting $\delta = 0$ in Proposition 3, when $|a| < 1$ and both N and i go to infinity, the power of ARIN is given as

$$\mathbb{E}(|\hat{e}_j(i)|^2) \to A_j(0) = \frac{1 + \hat{P}_j + |c_{\bar{j}j}\omega_{\bar{j}}|^2(\hat{P}_{\bar{j}}+1)}{1-|a|^2}. \tag{23}$$

Then, the power of $y_j(i)$ in (15) and (19) can be easily computed with the result of (23). The power constraints at receiver j defined in (4) for the single-user decoding and joint decoding are respectively rewritten as

$$|\omega_j|^2\left(|a_{1j}|^2 + |a_{2j}|^2 + |b_{\bar{j}j}|^2 + A_j\right) \le P_{Rj}, \tag{24}$$

$$|\omega_j|^2\left(|a_{1j}|^2 + |a_{2j}|^2 + A_j\right) \le P_{Rj}, \tag{25}$$

where we shorten $A_j(0)$ defined in (23) as A_j. Then, by substituting (23) into (24) and (25), respectively, we define the corresponding transmission parameter regions for the two receivers as

$$\mathcal{W}^S = \{(\omega_1, \omega_2) : |a| < 1, (24),\ j = 1, 2\}, \tag{26}$$

$$\mathcal{W}^J = \{(\omega_1, \omega_2) : |a| < 1, (25),\ j = 1, 2\}, \tag{27}$$

where $\mathcal{W}^S$ and $\mathcal{W}^J$ are the transmission parameter regions for the single-user decoding and joint decoding schemes, respectively.

C. ACHIEVABLE RATE REGIONS

In this subsection, the covariance matrices of the additive noise $\boldsymbol{z}_j^S$ and $\boldsymbol{z}_j^J$ given in (12) and (16) are respectively computed for both the single-user and joint decoding schemes. Then, the corresponding achievable rate regions are characterized for the considered IC with receiver cooperations.

1) For the single-user decoding scheme, the covariance matrix $\boldsymbol{Q}_j$ of the additive noise $\boldsymbol{z}_j^S$ is computed as

$$\boldsymbol{Q}_j = \boldsymbol{Q}_j^S = \mathbb{E}\left(\boldsymbol{z}_j^S \cdot \left(\boldsymbol{z}_j^S\right)^H\right)$$

$$\hat{e}_j(i) = \sum_{n=0}^{\lfloor i/2 \rfloor} a^n\left(\hat{e}_j^I(i-2n) + c_{\bar{j}j}\omega_{\bar{j}}\hat{e}_{\bar{j}}^I(i-2n-1) + c_{\bar{j}j}\omega_{\bar{j}}n_{\bar{j}}(i-2n-1) + n_j(i-2n)\right). \tag{20}$$

$$= \begin{bmatrix} A_j + \left|b_{\bar{j}j}\right|^2 + \left|a_{1j}\right|^2 + \left|a_{2j}\right|^2 & b_{\bar{j}j}a_{\bar{j}j}^* \\ b_{\bar{j}j}^* a_{\bar{j}j} & A_j + \left|a_{\bar{j}j}\right|^2 \end{bmatrix}, \quad (28)$$

where $\boldsymbol{z}_j^S$ is defined in (13)-(14), and A_j is given in (24). Then, based on the equivalent channel model in (12), we define the rate region $\mathcal{C}(\mathcal{P}, \mathcal{W})$, subject to the power constraints at both the transmitters and receivers. The rate region of a two-user Gaussian IC is defined as the closure of the set of rate pairs (R_1, R_2) for which both the receivers can decode their own messages. By the result of the two-user Gaussian IC [28], the rate region for the single-user decoding scheme is given by (29). To characterize $\mathcal{C}(\mathcal{P}, \mathcal{W})$, we fix R_1 to maximize R_2. Then, the optimal problem, subject to the rate constraints in (29), as shown at the bottom of this page, and the transmit power constraints in (2) and (26), is formulated as

$$\max_{\{\mu_1,\mu_2,\omega_1,\omega_2\}} \quad R_2 \quad (30)$$

$$\text{s.t. } R_1 \le \log\left(1 + \boldsymbol{h}_{11}^H \boldsymbol{Q}_1^{-1} \boldsymbol{h}_{11}\right) \quad (31)$$

$$R_2 \le \log\left(1 + \boldsymbol{h}_{22}^H \boldsymbol{Q}_2^{-1} \boldsymbol{h}_{22}\right) \quad (32)$$

$$(\mu_1, \mu_2) \in \mathcal{P}, \quad (\omega_1, \omega_2) \in \mathcal{W}^S. \quad (33)$$

2) For the joint decoding scheme, the covariance matrix of $\boldsymbol{z}_j^J$ is computed as

$$\boldsymbol{Q}_j = \boldsymbol{Q}_j^J = \mathbb{E}\left(\boldsymbol{z}_j^J \cdot \left(\boldsymbol{z}_j^J\right)^H\right) \quad (34)$$

$$= \begin{bmatrix} A_j + \left|a_{1j}\right|^2 + \left|a_{2j}\right|^2 & 0 \\ 0 & A_j \end{bmatrix},$$

where $\boldsymbol{z}_j^J$ is given in (17)-(18), and A_j is given in (24). Since the joint decoding scheme supports both the receivers to decode the two messages from the two transmitters, this IC acts the same as two multiple-access channels (MACs), and the corresponding rate region is the intersection of the rate regions of the two MACs [10], i.e., (35), as shown at the bottom of this page. Similarly, by fixing R_1 and maximizing R_2, the corresponding optimization problem, subject to the rate constraints in (35) and the transmit power constraints in (2) and (26), is formulated as

$$\max_{\{\mu_1,\mu_2,\omega_1,\omega_2\}} \quad R_2 \quad (36)$$

$$\text{s.t. } R_1 \le \log\left(1 + \boldsymbol{h}_{1j}^H \boldsymbol{Q}_j^{-1} \boldsymbol{h}_{1j}\right) \quad (37)$$

$$R_2 \le \log\left(1 + \boldsymbol{h}_{2j}^H \boldsymbol{Q}_j^{-1} \boldsymbol{h}_{2j}\right) \quad (38)$$

$$R_1 + R_2 \le \log\left(1 + \sum_{k=1}^{2} \boldsymbol{h}_{kj}^H \boldsymbol{Q}_j^{-1} \boldsymbol{h}_{kj}\right) \quad (39)$$

$$(\mu_1, \mu_2) \in \mathcal{P}, (\omega_1, \omega_2) \in \mathcal{W}^J, \quad j = 1, 2. \quad (40)$$

By substituting (28) and (34) into (31)-(33) and (37)-(40), respectively, it is observed that the complex parameters h_{kj}, $c_{\bar{j}j}$, μ_j, and ω_j, are either with the form of $|x|^2$ or $x + x^*$, ($x \in \{h_{kj}, c_{\bar{j}j}, \mu_j, \omega_j\}$). Therefore, the phases of these parameters will not affect the optimal values of problems (30) and (36). For the sake of simplicity, h_{kj}, $c_{\bar{j}j}$, μ_j, and ω_j, $k = 1, 2, j = 1, 2$, in problems (30) and (36) are treated as real numbers in the sequel.

D. ONE-SIDE COOPERATION

In this subsection, we prove that one-side cooperation is optimal to achieve the best system performance for the considered IC with receiver cooperations.

Proposition 4: For both the single-user and joint decoding schemes, the optimal value of problems (30) and (36) is obtained when either $\omega_1 = 0$ or $\omega_2 = 0$.

Proof: Please see Appendix B. ■

Remark 2: From the proof of proposition 4, it is observed that: When the cooperative channel gains, i.e., c_{12} and c_{12}, are relatively strong, only the receiver, whose received signal has better signal-to-interference-plus-noise ratio (SINR), forwards its counterpart's signal, while the other receiver does not forward anything. The results can also be derived by the well-known date-processing inequality [29] from the perspective of the information theory.

From Proposition 4, it is observed that when the best system performance is achieved, there exists $\omega_1 = 0$ or $\omega_2 = 0$. Hence, we have $a = 0$, and the considered condition $a < 1$ is always satisfied. For the sake of simplicity, $a = 0$ is

$$\mathcal{C}(\mathcal{P}, \mathcal{W}) \triangleq \bigcup_{\substack{(\mu_1,\mu_2)\in\mathcal{P} \\ (\omega_1,\omega_2)\in\mathcal{W}}} \left\{ (R_1, R_2) \left| \begin{array}{l} R_1 \le \log\left(1 + \boldsymbol{h}_{11}^H \boldsymbol{Q}_1^{-1} \boldsymbol{h}_{11}\right) \\ R_2 \le \log\left(1 + \boldsymbol{h}_{22}^H \boldsymbol{Q}_2^{-1} \boldsymbol{h}_{22}\right) \end{array} \right. \right\}. \quad (29)$$

$$\mathcal{C}(\mathcal{P}, \mathcal{W}) \triangleq \bigcup_{\substack{(\mu_1,\mu_2)\in\mathcal{P} \\ (\omega_1,\omega_2)\in\mathcal{W}}} \left\{ (R_1, R_2) \left| \begin{array}{l} R_1 \le \log\left(1 + \boldsymbol{h}_{1j}^H \boldsymbol{Q}_j^{-1} \boldsymbol{h}_{1j}\right) \\ R_2 \le \log\left(1 + \boldsymbol{h}_{2j}^H \boldsymbol{Q}_j^{-1} \boldsymbol{h}_{2j}\right) \\ R_1 + R_2 \le \log\left(1 + \sum_{k=1}^{2} \boldsymbol{h}_{kj}^H \boldsymbol{Q}_j^{-1} \boldsymbol{h}_{kj}\right) \end{array} \right., j = 1, 2 \right\}. \quad (35)$$

considered in the sequel. Then, the power of ARIN A_j given in (23) is simplified as

$$A_j = 1 + \hat{P}_j + c_{\bar{j}\bar{j}}^2 \omega_{\bar{j}}^2 (\hat{P}_{\bar{j}} + 1). \tag{41}$$

Also, by setting $a = 0$, the transmission parameter regions of the receivers in (26) and (27) for the single-user and joint decoding schemes can be rewritten in an universal form, i.e.,

$$\mathcal{W} = \mathcal{W}^S = \mathcal{W}^J = \{(\omega_1, \omega_2) : (25), (41),\ j = 1, 2\} \tag{42}$$

Since all the constraints (31)-(33) and (37)-(40) are non-convex, in the next subsection, we propose a suboptimal algorithm to compute the local optimal solutions for problems (30) and (36).

E. ALGORITHM

We adopt a SPCA method [30] to solve problems (30) and (36). The basic idea of this method is that each non-convex constraint in problems (30) and (36) is replaced by its lower bound, and then the convex approximations of original problems (30) and (36) are successively optimized. The main steps are summarized as follows:

1) Change the design variables in problems (30) and (36), and then problems (30) and (36) are equivalently transformed into new problems with concave-convex constraints;
2) Obtain the lower bounds for these concave-convex constraints, and then the original constraints are approximated by new convex constraints.

(I) Changing variables: Since the composition of exponential and linear functions is convex [31], we utilize the exponential functions to change the design variables μ_1, μ_2, ω_1, and ω_2, i.e.,

$$\mu_1 = e^{\frac{1}{2}\alpha_1}, \mu_2 = e^{\frac{1}{2}\alpha_2}, \omega_1 = e^{\frac{1}{2}\beta_1}, \omega_2 = e^{\frac{1}{2}\beta_2}, \tag{43}$$

where we consider $\alpha_j = \beta_j = -\infty$ for the case of $u_j = \omega_j = 0, j = 1, 2$. By substituting the new design variables into power constraints in (2) and (42), the new transmission parameter regions defined in (2) and (42) are transformed as

$$\Theta_T = \left\{(\alpha_1, \alpha_2) : \eta \le \alpha_j \le \ln P_{T_j}, j = 1, 2\right\}, \tag{44}$$

$$\Theta_R = \left\{(\beta_1, \beta_2) : \eta \le \beta_j \le \frac{1}{2} \ln \frac{P_{R_j}}{M_j}, j = 1, 2\right\}, \tag{45}$$

where $M_j = h_{1j}^2 e^{\alpha_1} + h_{2j}^2 e^{\alpha_2} + 1 + c_{\bar{j}\bar{j}}^2 e^{\beta_{\bar{j}}} (\hat{P}_{\bar{j}} + 1)$, and $\eta \to -\infty$. To simplify the analysis, we set $\eta = -2 \times 10^6$ in the sequel. From (44) and (45), it is observed that Θ_T and Θ_R are compact convex sets.

Then, we denote $\boldsymbol{\nu} = [\alpha_1, \alpha_2, \beta_1, \beta_2, R_2]$ as the new design variable vector, and then problems (30) and (36) are transformed into new problems with concave-convex constraints over $\boldsymbol{\nu}$.

1) For the single-user decoding scheme, problem (30) is transformed into

$$\max_{\boldsymbol{\nu}} \ R_2 \tag{46}$$

$$\text{s.t.} \ f_{x_1}(\boldsymbol{\nu}) + f_{e_1}(\boldsymbol{\nu}) \ge 0 \tag{47}$$

$$f_{x_2}(\boldsymbol{\nu}) + f_{e_2}(\boldsymbol{\nu}) \ge 0 \tag{48}$$

$$(\alpha_1, \alpha_2) \in \Theta_T, \ \ (\beta_1, \beta_2) \in \Theta_R, \tag{49}$$

where $f_{x_j}(\boldsymbol{\nu})$ and $f_{e_j}(\boldsymbol{\nu})$ are given in (50)-(51), with $\gamma_{j\bar{j}} = c_{j\bar{j}}^2$, and $g_{kj} = h_{kj}^2, k = 1, 2$.

2) For the joint decoding scheme, problem (36) is transformed into

$$\max_{\boldsymbol{\nu}} \ R_2 \tag{52}$$

$$\text{s.t.} \ f_{x_1^j}(\boldsymbol{\nu}) + f_{e_1^j}(\boldsymbol{\nu}) \ge 0 \tag{53}$$

$$f_{x_2^j}(\boldsymbol{\nu}) + f_{e_2^j}(\boldsymbol{\nu}) \ge 0 \tag{54}$$

$$f_{x_s^j}(\boldsymbol{\nu}) + f_{e_s^j}(\boldsymbol{\nu}) \ge 0 \tag{55}$$

$$(\alpha_1, \alpha_2) \in \Theta_T, (\beta_1, \beta_2) \in \Theta_R, \quad j = 1, 2, \tag{56}$$

where $f_{x_k^j}(\boldsymbol{\nu}), f_{x_s^j}(\boldsymbol{\nu}), f_{e_k^j}(\boldsymbol{\nu})$ and $f_{e_s^j}(\boldsymbol{\nu})$ are given in (57)-(60), as shown at the bottom of this page.

$$f_{x_j}(\boldsymbol{\nu}) = 2^{R_j} \gamma_{j\bar{j}} g_{\bar{j}\bar{j}} g_{\bar{j}j} e^{2\alpha_{\bar{j}} + \beta_{\bar{j}}} + \gamma_{j\bar{j}} g_{j\bar{j}} e^{\alpha_j + \beta_j} \left(g_{\bar{j}j} e^{\alpha_{\bar{j}}} + \gamma_{\bar{j}j} e^{\beta_{\bar{j}}} (\hat{P}_{\bar{j}} + 1) + \hat{P}_j + 1\right)$$
$$+ \left(g_{1j} e^{\alpha_1} + g_{2j} e^{\alpha_2} + \gamma_{j\bar{j}} e^{\beta_{\bar{j}}} (\hat{P}_{\bar{j}} + 1) + \hat{P}_j + 1\right) \left(\gamma_{j\bar{j}} g_{j\bar{j}} e^{\alpha_j + \beta_j} + g_{1j} e^{\alpha_1} + g_{2j} e^{\alpha_2} + \gamma_{\bar{j}j} e^{\beta_{\bar{j}}} (\hat{P}_{\bar{j}} + 1) + \hat{P}_j + 1\right), \tag{50}$$

$$f_{e_j}(\boldsymbol{\nu}) = -\gamma_{j\bar{j}} h_{j\bar{j}}^* h_{\bar{j}\bar{j}}^* h_{\bar{j}j} e^{\beta_{\bar{j}} + \alpha_j + \alpha_{\bar{j}}} - \left(\gamma_{j\bar{j}} h_{j\bar{j}}^* h_{\bar{j}\bar{j}}^* h_{\bar{j}j} e^{\beta_{\bar{j}} + \alpha_j + \alpha_{\bar{j}}}\right)^* - \gamma_{j\bar{j}} g_{j\bar{j}} g_{\bar{j}j} e^{\beta_{\bar{j}} + 2\alpha_{\bar{j}}}$$
$$- 2^{R_j} \left(\gamma_{j\bar{j}} e^{\beta_{\bar{j}}} (\hat{P}_{\bar{j}} + 1) + \hat{P}_j + 1 + g_{2j} e^{\alpha_2} + g_{1j} e^{\alpha_1} + \gamma_{j\bar{j}} g_{j\bar{j}} e^{\alpha_{\bar{j}} + \beta_{\bar{j}}}\right) \left(1 + \hat{P}_j + \gamma_{j\bar{j}} e^{\beta_{\bar{j}}} (\hat{P}_{\bar{j}} + 1) + g_{\bar{j}j} e^{\alpha_{\bar{j}}}\right). \tag{51}$$

$$f_{x_k^j}(\boldsymbol{\nu}) = \left(\gamma_{j\bar{j}} e^{\beta_{\bar{j}}} (\hat{P}_{\bar{j}} + 1) + \hat{P}_j + 1 + g_{1j} e^{\alpha_1} + g_{2j} e^{\alpha_2}\right) \left(g_{kj} e^{\alpha_k} + \gamma_{j\bar{j}} e^{\beta_{\bar{j}}} (\hat{P}_{\bar{j}} + 1) + \hat{P}_j + 1\right)$$
$$+ \left(\gamma_{j\bar{j}} g_{k\bar{j}} e^{\beta_{\bar{j}} + \alpha_k} + \gamma_{j\bar{j}} e^{\beta_{\bar{j}}} (\hat{P}_{\bar{j}} + 1) + \hat{P}_j + 1\right), \tag{57}$$

$$f_{x_s^j}(\boldsymbol{\nu}) = \left(\gamma_{j\bar{j}} e^{\beta_{\bar{j}}} (\hat{P}_{\bar{j}} + 1) + \hat{P}_j + 1 + g_{1j} e^{\alpha_1} + g_{2j} e^{\alpha_2}\right)^2 \left(\gamma_{j\bar{j}} g_{1\bar{j}} e^{\beta_{\bar{j}} + \alpha_1} + \gamma_{j\bar{j}} g_{2\bar{j}} e^{\beta_{\bar{j}} + \alpha_2} + \gamma_{j\bar{j}} e^{\beta_{\bar{j}}} (\hat{P}_{\bar{j}} + 1) + \hat{P}_j + 1\right), \tag{58}$$

$$f_{e_k^j}(\boldsymbol{\nu}) = -2^{R_j} \left(\gamma_{j\bar{j}} e^{\beta_{\bar{j}}} (\hat{P}_{\bar{j}} + 1) + \hat{P}_j + 1\right) \left(\gamma_{j\bar{j}} e^{\beta_{\bar{j}}} (\hat{P}_{\bar{j}} + 1) + \hat{P}_j + 1 + g_{1j} e^{\alpha_1} + g_{2j} e^{\alpha_2}\right), \tag{59}$$

$$f_{e_s^j}(\boldsymbol{\nu}) = -2^{R_1 + R_2} \left(\gamma_{j\bar{j}} e^{\beta_{\bar{j}}} (\hat{P}_{\bar{j}} + 1) + \hat{P}_j + 1\right) \left(\gamma_{j\bar{j}} e^{\beta_{\bar{j}}} (\hat{P}_{\bar{j}} + 1) + \hat{P}_j + 1 + g_{1j} e^{\alpha_1} + g_{2j} e^{\alpha_2}\right). \tag{60}$$

It is obvious that $f_{x_j}(\boldsymbol{v}), f_{x_k^j}(\boldsymbol{v})$ and $f_{x_s^j}(\boldsymbol{v})$ are strictly convex functions on $\boldsymbol{v}$, while $f_{e_j}(\boldsymbol{v}), f_{e_k^j}(\boldsymbol{v})$ and $f_{e_s^j}(\boldsymbol{v})$ are strictly concave. Thus, all the new constraints (47)-(48) and (53)-(55), are the sum of convex and concave functions, and then problems (46) and (52) are also non-convex.

(II) Constraints approximation: In order to transform the concave-convex constraints (47)-(48) and (53)-(55) into convex forms, a CCCP is adopted to replace the concave-convex functions by their lower bounds [32]. More precisely, we denote the concave-convex functions as an universal form, i.e.,

$$g(\boldsymbol{v}) = f_x(\boldsymbol{v}) + f_e(\boldsymbol{v}), \tag{61}$$

where $f_x(\boldsymbol{v})$ is a convex function, and $f_e(\boldsymbol{v})$ is concave on $\boldsymbol{v}$. By utilizing the first-order Taylor approximation, the lower bound of $g(\boldsymbol{v})$, denoted as $G(\boldsymbol{v}, \boldsymbol{y})$, is given as

$$G(\boldsymbol{v}, \boldsymbol{y}) = f_x(\boldsymbol{y}) + \nabla f_x(\boldsymbol{y})(\boldsymbol{v} - \boldsymbol{y})^T + f_e(\boldsymbol{v}), \tag{62}$$

where $\nabla f_x(\boldsymbol{y})$ denotes the gradient of $f_x(\boldsymbol{y})$ at the point $\boldsymbol{y}$.

By using the same method in (61)-(62), the constraints (47)-(48) and (53)-(55) can also be transformed into convex forms, and then problems (46) and (52) can be approximated to the following convex problems (63) and (67), respectively.

1) For the single-user decoding scheme, problem (46) is solved by successively optimizing the approximate version, i.e.,

$$(\mathbf{P}_s^l): \max_{\boldsymbol{v}} \ R_2 \tag{63}$$

$$\text{s.t.} \ f_{x_1}(\boldsymbol{y}^l) + \nabla f_{x_1}(\boldsymbol{y}^l)(\boldsymbol{v} - \boldsymbol{y}^l)^T + f_{e_1}(\boldsymbol{v}) \geq 0 \tag{64}$$

$$f_{x_2}(\boldsymbol{y}^l) + \nabla f_{x_2}(\boldsymbol{y}^l)(\boldsymbol{v} - \boldsymbol{y}^l)^T + f_{e_2}(\boldsymbol{v}) \geq 0 \tag{65}$$

$$(\alpha_1, \alpha_2) \in \Theta_S, \ \ (\beta_1, \beta_2) \in \Theta_R, \tag{66}$$

where $(\mathbf{P}_s^l)$ denotes the l-th ($l \geq 1$) problem for the single-user decoding scheme, and the fix vector $\boldsymbol{y}^l$ in (64)-(65) is the solution of the problem $(\mathbf{P}_s^{l-1})$.

2) For the joint decoding scheme, we solve problem (52) by a sequence of convex problems as the following form:

$$(\mathbf{P}_j^l): \max_{\boldsymbol{v}} \ R_2 \tag{67}$$

$$\text{s.t.} \ f_{x_1^j}(\boldsymbol{y}^l) + \nabla f_{x_1^j}(\boldsymbol{y}^l)(\boldsymbol{v} - \boldsymbol{y}^l)^T + f_{e_1^j}(\boldsymbol{v}) \geq 0 \tag{68}$$

$$f_{x_2^j}(\boldsymbol{y}^l) + \nabla f_{x_2^j}(\boldsymbol{v}^l)(\boldsymbol{v} - \boldsymbol{y}^l)^T + f_{e_2^j}(\boldsymbol{v}) \geq 0 \tag{69}$$

$$f_{x_s^j}(\boldsymbol{y}^l) + \nabla f_{x_s^j}(\boldsymbol{y}^l)(\boldsymbol{v} - \boldsymbol{y}^l)^T + f_{e_s^j}(\boldsymbol{v}) \geq 0 \tag{70}$$

$$(\alpha_1, \alpha_2) \in \Theta_S, \ \ (\beta_1, \beta_2) \in \Theta_R, \tag{71}$$

where $\mathbf{P}_j^l$ denotes the l-th problem for the joint decoding scheme, and $\boldsymbol{y}^l$ in (68)-(70) depends on the solution of the problem $(\mathbf{P}_j^{l-1})$.

It is observed that problems (63) and (67) are convex over $\boldsymbol{v} = [\alpha_1, \alpha_2, \beta_1, \beta_2, R_2]$, which can be efficiently solved by some optimization tools, e.g., CVX [33].

The SPCA algorithm solves problems (46) and (52) by a sequence of problems (63) and (67), respectively. At the l-th iteration of this algorithm, we update vector $\boldsymbol{y}^l$ by using the optimal point obtained in the previous iteration, i.e., $\boldsymbol{y}^l = \boldsymbol{v}^{l-1} (l \geq 1)$. When the algorithm converges, the solutions for the original problems (30) and (36) are obtained. We summarize the SPCA algorithm in Table 1.

In [30], it has been proved that when the approximation function of $g(\boldsymbol{v})$, i.e., $G(\boldsymbol{v}, \boldsymbol{y})$, satisfies three conditions, i.e., $g(\boldsymbol{v}) \geq G(\boldsymbol{v}, \boldsymbol{y})$, $g(\boldsymbol{v}^{l-1}) = G(\boldsymbol{v}^{l-1}, \boldsymbol{y})$, and $\nabla g(\boldsymbol{v}^{l-1}) = \nabla G(\boldsymbol{v}^{l-1}, \boldsymbol{y})$, at the l-th iteration, the generated sequence $\{\boldsymbol{v}^l\}$ by using the SPCA algorithm will converge to a stationary point which satisfies the Karush-Kuhn-Tucker conditions of the original problems. In this paper, it is easy to verify that all the constraints in problems (63) and (67) satisfy the above three conditions. Thus, we can obtain a local optimal solution for the original problems (30) and (36) by the proposed algorithm.

TABLE 1. Algorithm 1: The SPCA algorithm for solving problems (30) and (36).

Input: $R_1, \epsilon, P_j, P_{Rj}, h_{kj}, c_{j\bar{j}}, c_{\bar{j}j}$.
Output: R_2^*.
1: Change the design variables by (43).
2: **Initialize**
3: Set $l = 0$ and $\epsilon = 0.001$;
4: Choose a starting point $\boldsymbol{v}^0 = [\alpha_1^0, \alpha_2^0, \beta_1^0, \beta_2^0, R_2^0]$ for problems (46) and (52), respectively;
5: **repeat**
6: Set $l = l + 1$;
7: Set $\boldsymbol{y}^l = \boldsymbol{v}^{l-1}$;
8: Structure the r-th problems $(\mathbf{P}_s^l)$ and $(\mathbf{P}_j^l)$ by using the CCCP, as shown in (63) and (67)
9: Compute the solution $\boldsymbol{v}^l = [\alpha_1^l, \alpha_2^l, \beta_1^l, \beta_2^l, R_2^l]$ of the convex problems $(\mathbf{P}_s^l)$ and $(\mathbf{P}_j^l)$ at the l-th step, respectively;
10: **until** $||\boldsymbol{v}^l - \boldsymbol{v}^{l-1}|| < \epsilon$.

IV. SIMULATION RESULTS

Simulation results are demonstrated to verify our analysis about the achievable rate regions for the single-user and joint decoding schemes. As a rule of thumb, we mainly consider two typical scenarios: 1) the cross interference is relatively strong [34], i.e., $h_{12} > h_{11}$ and $h_{21} > h_{22}$; and 2) the cross interference is relatively weak [35], i.e., $h_{12} \leq h_{11}$ and $h_{21} \leq h_{22}$. For the sake of simplicity, we focus on the symmetric case, i.e., $h_{11} = h_{22}$, $h_{12} = h_{21}$, and $c_{12} = c_{21}$. More precisely, three channel conditions are set up. Scenario $\mathrm{A_I}$: under the strong cross interference scenario, set $h_{11} = h_{22} = 0.1$ and $h_{12} = h_{21} = 0.3$; Scenario $\mathrm{A_{II}}$: under the strong cross interference scenario, set $h_{11} = h_{22} = 0.1$ and $h_{12} = h_{21} = 0.2$; Scenario B: under the weak cross interference scenario, set $h_{11} = h_{22} = 0.3$ and $h_{12} = h_{21} = 0.1$. Moreover, the maximum power budgets at the two transmitters and two receivers are set as $P_{T_1} = P_{T_2} = P_{R_1} = P_{R_2} = 20$ dB.

First, the achievable rate region of the considered IC with an ideal SI cancellation are compared with the upper bounds (6)-(10) of [17] by placing the parameters of $\mathrm{A_I}$ with $c_{12} = c_{21} = 1$. In Fig. 3, it is observed that the maximum sum rate can be achieved within around $0.7 - 0.8$ bit of the upper bound for both the single-user and joint decoding schemes,

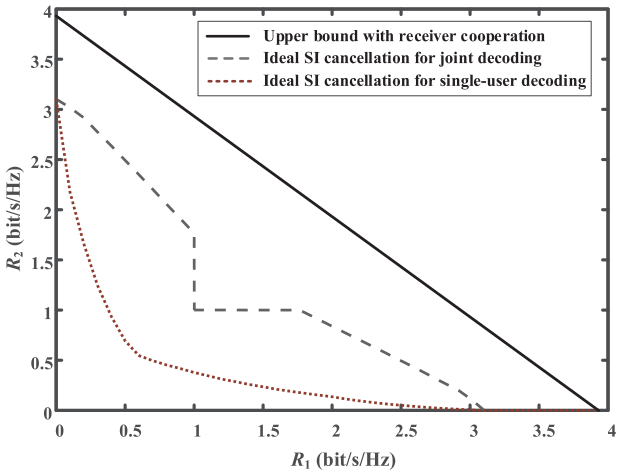

FIGURE 3. Achievable rate regions for ideal SI cancellation cases: $c_{12} = c_{21} = 1$, $h_{11} = h_{22} = 0.1$, and $h_{12} = h_{21} = 0.3$.

respectively. Under the symmetric channel condition, R_2 is improved by the receiver cooperations when $0 \leq R_1 < 1$. When $R_1 = R_2 = 1$, the considered IC with receiver cooperations degenerates into the non-cooperative systems, since the two transmit rates R_1 and R_2 cannot be improved simultaneously by one-side cooperation. Besides, it is worthy pointing out that R_1 can also be improved when $0 \leq R_2 < 1$, due to the symmetric channel.

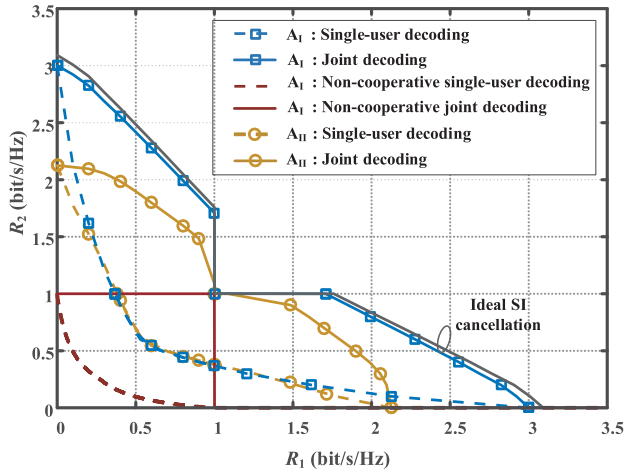

FIGURE 4. Achievable rate regions in strong interference scenarios ($h_{12} = h_{21} \geq h_{11} = h_{22}$): $c_{12} = c_{21} = 1$, A_I: $h_{11} = h_{22} = 0.1$, $h_{12} = h_{21} = 0.3$, A_{II}: $h_{11} = h_{22} = 0.1$, $h_{12} = h_{21} = 0.2$, and $\hat{P} = -10$dB.

Fig. 4 describes the achievable rate regions of the considered IC for the single-user and joint decoding schemes under different channel conditions. Compared with scenarios A_I and A_{II} in Fig. 4, it is obversed that giving higher cross interference, receiver cooperations performs better. On the other hand, the non-cooperative systems [19] with disabled receiver cooperations, are compared with our proposed system. Fig. 4 shows that the proposed receiver cooperations outperforms the non-cooperative system and the joint decoding scheme outperforms the single-user decoding scheme, when the cross interference is relatively strong. Moreover, Fig. 4 also reveals that under the symmetric channel condition, the proposed receiver cooperations are disabled when $R_1 = R_2$. That is because the one-cooperation does not support both R_1 and R_2 to be improved under the symmetric channel condition. To show the optimal transmission parameters of the two receivers for different transmission rates, we plot the parameters ω_1 and ω_2 as functions of R_1 and R_2 for the single-user and joint decoding schemes under Scenario A_I by the enumeration method. Fig. 5 and Fig. 6 show that when we fix one of the two transmission rates, the other one always reaches its maximum at $\omega_1 = 0$ or $\omega_2 = 0$.

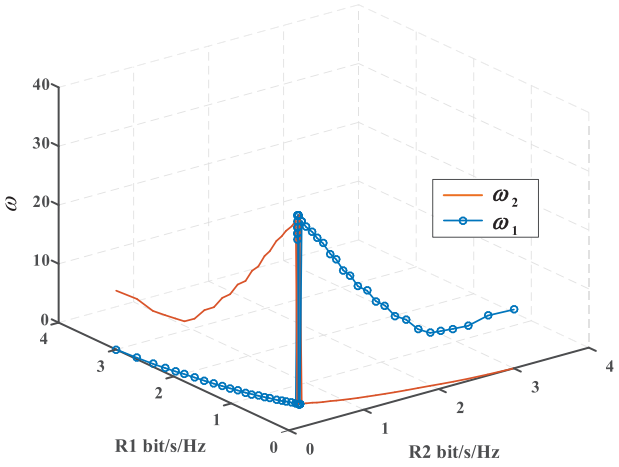

FIGURE 5. Optimal transmission parameters of the receivers for different transmission rates under Scenario A_I with single-user decoding.

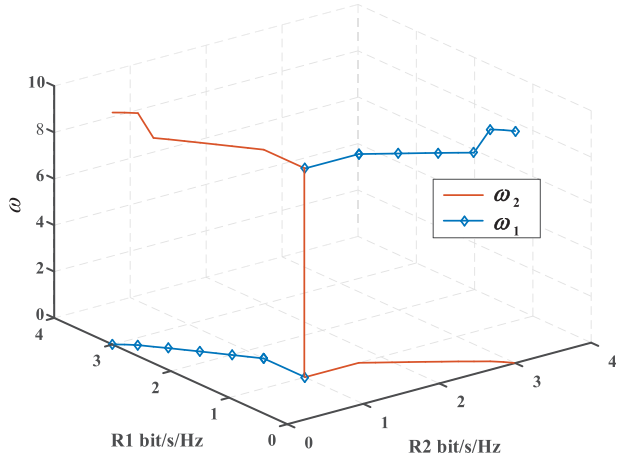

FIGURE 6. Optimal transmission parameters of the receivers for different transmission rates under Scenario A_I with joint decoding.

Fig. 7 describes a relatively weak cross interference scenorio. For the joint decoding, the rate region of this case is equivalent to that of scenorio A_I in Fig 4. This is because

the receivers always forward the signals from the channels of 0.3 rather than channels of 0.1 to their counterpart receivers in both strong and weak interference scenorios. However, for single-user decoding scheme, receiver cooperations performs as non-cooperative with weak cross interference. This is because the message is directly decoded at its desired receiver, while the counterpart receiver does not forward anything under the weak cross channel condition. Furthermore, the results in Fig. 7 also show that the single-user decoding scheme dominates the joint decoding scheme, which is coincident with the results in [7] for the weak interference regime.

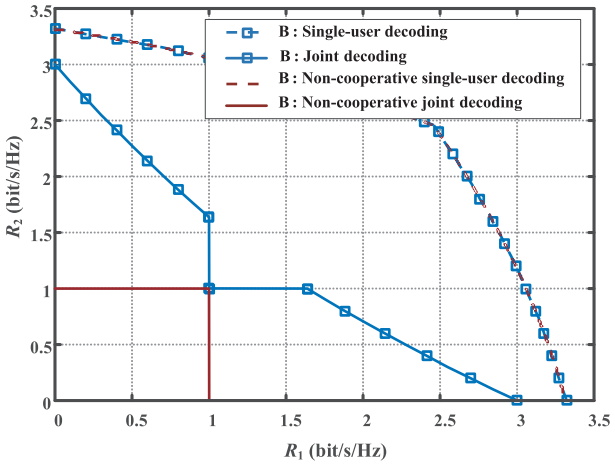

FIGURE 7. Achievable rate regions in weak interference scenario ($h_{12} = h_{21} \leq h_{11} = h_{22}$): $c_{12} = c_{21} = 1$, B: $h_{11} = h_{22} = 0.3$, $h_{12} = h_{21} = 0.1$, and $P = -10$dB.

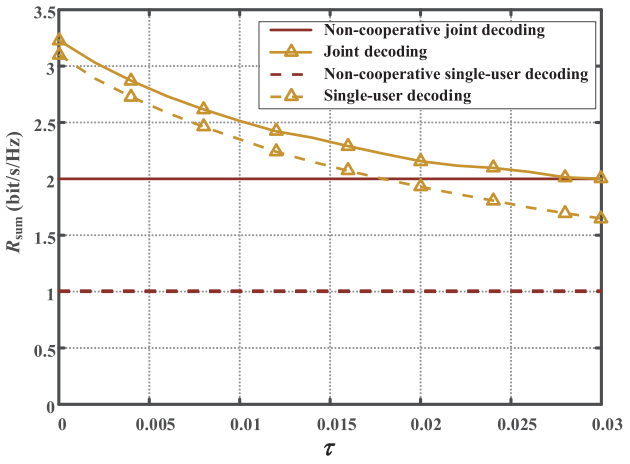

FIGURE 8. The effect of SI cancellation on the maximum sum rate R_{sum}.

Then, we define the ratio $\tau = \frac{\hat{P}}{P_R}$ to study the effects of the SI cancellation on the maximum sum rate R_{sum}. Under scenario A_I, Fig. 8 shows that with τ increasing, the R_{sum} decreases since the power of the residual SI weakens the system performance. Besides, we notice that when the SI cancellation is perfect, i.e., $\tau = 0$, the sum rate has maximum values around 3.2 bits/s/Hz and 3.3 bits/s/Hz, respectively, for the single-user and joint decoding schemes. As τ increases, R_{sum} decreases gradually. In particular, when τ approaches around 0.028, the considered system for the joint decoding scheme cannot bring any improvements.

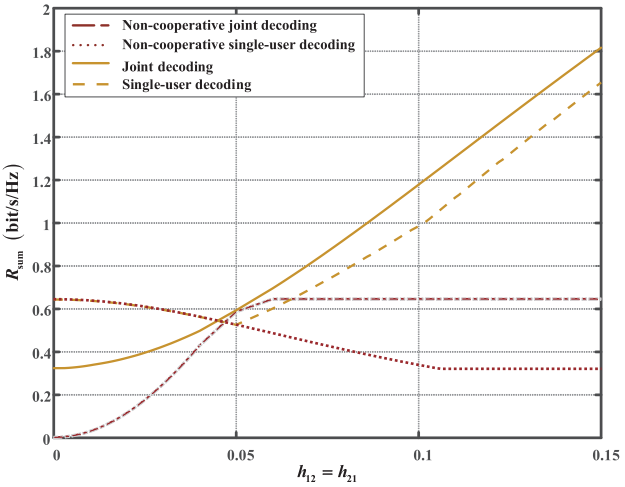

FIGURE 9. The effect of the cross-channel coefficients $h_{12} = h_{21}$ on the maximum sum rate R_{sum}.

Next, the effects of the cross-channels coefficients on the maximum sum rate are examined. The cooperative channel coefficients are set as $c_{12} = c_{21} = 1$. We fix the straight-channel coefficients as $h_{11} = h_{22} = 0.05$, and then change the cross channel coefficients from 0 to 0.15. From Fig. 9, it is observed that when the cross links increase, the sum rate improvement induced by AF scheme will be enhanced. However, when the cross-channel coefficients are smaller than the straight-channel coefficients, the receiver cooperations are equivalent to the non-cooperative case for single-user decoding scheme, but perform better for the joint decoding scheme. Note that when $h_{12} = h_{21} = 0$, the sum rate in non-cooperative joint decoding decreases to 0. This is because the transmit rate of each transmitter in the non-cooperative joint decoding needs to satisfy the constraint: $R_k = \min\left(\log(1 + \frac{h_{k1}^2 \mu_k^2}{1}), \log(1 + \frac{h_{k2}^2 \mu_k^2}{1})\right), k = 1, 2.$ When $h_{12} = h_{21} = 0$, there do exist $R_1 = R_2 = 0$, such that the sum rate for non-cooperative joint decoding is 0.

Finally, the effect of cooperative channels coefficients c_{12} and c_{21} on the maximum sum rate R_{sum} are studied under scenario A_I. Here, we set $c_{12} = c_{21}$. Fig. 10 shows that as the cooperative channel coefficient increases, the gain introduced by receiver cooperations increases monotonously and converge to a constant 3.4 bits/s/Hz. For the single-user decoding scheme, the maximum sum rate R_{sum} increases by nearly 1 bits/s/Hz as c_{12} and c_{21} increase from 0 to 0.3. For the joint decoding scheme, with c_{12} and c_{21} increasing from 0.3 to 0.8, the maximum sum rate increases nearly 1.1 bits/s/Hz. Note that when c_{12} and c_{21} change from 0 to 0.3, the joint decoding scheme performs as the non-cooperative

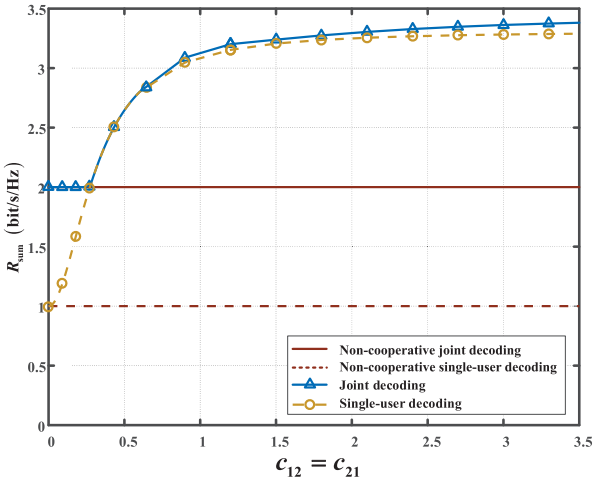

FIGURE 10. The effect of the cooperative channel coefficients $c_{12} = c_{21}$ on the maximum sum rate R_{sum}.

system. That is because under a relatively weak cooperative channel condition, the power gain introduced by the receiver cooperations extremely decays due to the cooperative channel fading, such that the receiver cooperations perform as the non-cooperative system.

V. CONCLUSION

In this paper, a receiver cooperations scheme with in-band FD communications was proposed in a two-transmitter two-receiver Gaussian IC. With the AF scheme, corresponding achievable rates regions for the single-user and joint decoding schemes were characterized, respectively. We proved that one-side cooperation, i.e., only one of the two receivers forwards its counterpart's signal, is optimal to achieve the best system performance. A SPCA algorithm was proposed to effectively solve the optimization problems with properly designing the transmission parameters at the two transmitters and two receivers, and then the achievable rate regions for the two decoding schemes are characterized. Moreover, some simulation results showed that the receiver cooperations can achieve a higher transmission rate in several typical scenarios.

APPENDIX A
THE PROOF OF PROPOSITION 3

To prove Proposition 3, we first compute the general term of the correlation coefficient by Proposition 2.

Since $\hat{e}_j^I(i), \hat{e}_{\bar{j}}^I(i), n_j(i), n_{\bar{j}}(i)$ are independent for different receivers and time slots, we can immediately obtain that $\mathbb{E}\left[\hat{e}_j(i)\hat{e}_j^*(i+\delta)\right] = 0$ with δ being odd.

When δ is even, the correlation coefficient of the AIRN for receiver j is computed as

$$\begin{aligned}&\mathbb{E}\left[\hat{e}_j(i)\hat{e}_j^*(i+\delta)\right]\\ &= \mathbb{E}\sum_{n=0}^{\lfloor i/2\rfloor}\sum_{m=0}^{\lfloor (i+\delta)/2\rfloor}\mathbb{I}_{\{\delta=2m-2n\}}a^n(a^*)^m\\ &\quad\cdot\begin{pmatrix}\hat{e}_j^I(i-2n)(\hat{e}_j^I)^*(i+\delta-2m)\\ +|c_{\bar{j}j}\omega_{\bar{j}}|^2\hat{e}_{\bar{j}}^I(i-2n-1)(\hat{e}_{\bar{j}}^I)^*(i+\delta-2m-1)\\ +|c_{\bar{j}j}\omega_{\bar{j}}|^2 n_{\bar{j}}(i-2n-1)n_{\bar{j}}^*(i+\delta-2m-1)\\ +n_j(i-2n)n_j^*(i+\delta-2m)\end{pmatrix},\end{aligned} \tag{72}$$

with $\mathbb{I}_{\{\delta=2m-2n\}}$ being an indicator function.

In (20), $\hat{e}_j^I, \hat{e}_{\bar{j}}^I, n_j$ and $n_{\bar{j}}$ are modeled as Gaussian noise, and thus (72) is rewritten as (73), as shown at the bottom of this page, with $i \to \infty$, where signals n_j and $n_{\bar{j}}$ are with unit

$$\mathbb{E}(\hat{e}_j(i)\hat{e}_j^*(i+\delta)) \to A_j(\delta) = \begin{cases}\dfrac{a^{-\delta}}{1-|a|^2}\left(1+\hat{P}_j+c_{\bar{j}j}^2\omega_{\bar{j}}^2(\hat{P}_{\bar{j}}+1)\right), & \delta \text{ is even, and } \delta \le 0;\\ \dfrac{(a^*)^{\delta}}{1-|a|^2}\left(1+\hat{P}_j+c_{\bar{j}j}^2\omega_{\bar{j}}^2(\hat{P}_{\bar{j}}+1)\right), & \delta \text{ is even, and } \delta \ge 0.\end{cases} \tag{73}$$

$$\Upsilon_{sum}^* = \max_{\{\mu_1,\mu_2,\omega_1,\omega_2\}\in\mathcal{S}_1}\Upsilon_{sum1} \tag{78}$$

$$= \max_{\{\mu_1,\mu_2,\omega_1,\omega_2\}\in\mathcal{S}_1}\sum_{k=1}^{2}\boldsymbol{h}_{k1}^H\boldsymbol{Q}_1^{-1}\boldsymbol{h}_{k1} \tag{79}$$

$$= \max_{\{\mu_1,\mu_2,\omega_1,\omega_2\}\in\mathcal{S}_1}\left(\frac{b_{11}^2+b_{21}^2}{A_1+a_{11}^2+a_{21}^2}+\frac{a_{11}^2+a_{21}^2}{A_1}\right) \tag{80}$$

$$= \max_{\{\mu_1,\mu_2,\omega_1,\omega_2\}\in\mathcal{S}_1}\frac{c_{21}^2\omega_2^2(h_{12}^2\mu_1^2+h_{22}^2\mu_2^2)}{\frac{1+\hat{P}_1+c_{21}^2\omega_2^2(\hat{P}_2+1)}{1-|a|^2}+h_{11}^2\mu_1^2+h_{21}^2\mu_2^2}+\frac{h_{11}^2\mu_1^2+h_{21}^2\mu_2^2}{\frac{1+\hat{P}_1+c_{21}^2\omega_2^2(\hat{P}_2+1)}{1-|a|^2}} \tag{81}$$

$$= \max_{\{\mu_1,\mu_2,\omega_1=0,\omega_2\}\in\mathcal{S}_1}\left(\frac{c_{21}^2\omega_2^2(h_{12}^2\mu_1^2+h_{22}^2\mu_2^2)}{c_{21}^2\omega_2^2(\hat{P}_2+1)+1+h_{11}^2\mu_1^2+h_{21}^2\mu_2^2}+\frac{h_{11}^2\mu_1^2+h_{21}^2\mu_2^2}{c_{21}^2\omega_2^2(\hat{P}_2+1)+1}\right) \tag{82}$$

$$= \Upsilon_{sum1}(\mu_1^*, \mu_2^*, \omega_1=0, \omega_2^*), \tag{83}$$

variance, and $\hat{P}_j$ and $\hat{P}_{\bar{j}}$ are the power of signals $\hat{e}_j^I$ and $\hat{e}_{\bar{j}}^I$, respectively.

With the above analysis, Proposition 3 is proved.

APPENDIX B
THE PROOF OF PROPOSITION 4

To prove Proposition 4, we need to prove that for a fixed R_1, the optimal values of problems (30) and (36) are always obtained when $\omega_1\omega_2 = 0$. In this section, we take (36) as example, and (30) can be proved in the same way.

For a fixed R_1 satisfying (37), the maximum R_2 denoted as R_2^* in (36) is computed by

$$R_2^* = \max_{\{\mu_1,\mu_2,\omega_1,\omega_2\}} \min(f_1, f_2), \tag{74}$$

where $(\mu_1, \mu_2) \in \mathcal{P}$, $(\omega_1, \omega_2) \in \mathcal{W}$, $f_1 = \log(1 + \Upsilon_2)$, and $f_2 = \log(1 + \Upsilon_{sum}) - R_1$, with $\Upsilon_{sum} = \min(\Upsilon_{sum1}, \Upsilon_{sum2})$, $\Upsilon_2 = \min(\Upsilon_{21}, \Upsilon_{22})$, $\Upsilon_{sumj} = \sum_{k=1}^{2} \boldsymbol{h}_{kj}^H \boldsymbol{Q}_j^{-1} \boldsymbol{h}_{kj}$, and $\Upsilon_{kj} = \boldsymbol{h}_{kj}^H \boldsymbol{Q}_j^{-1} \boldsymbol{h}_{kj}, j = 1, 2, k = 1, 2$.

Next, we will prove that (74) is equivalent to

$$R_2^* = \max_{\{\mu_1,\mu_2,\omega_1\omega_2=0\}} \min(f_1, f_2). \tag{75}$$

By comparing with f_1 and f_2 defined in (74), we will prove that for the following two cases, i.e., $f_1 > f_2$, and $f_1 \le f_2$, (75) is always true. Note that for a given fixed R_1, one of the two cases is always true for all $(\mu_1, \mu_2) \in \mathcal{P}$, $(\omega_1, \omega_2) \in \mathcal{W}^J$.

1) Case I: When $f_2(\mu_1, \mu_2, \omega_1, \omega_2) < f_1(\mu_1, \mu_2, \omega_1, \omega_2)$, (74) is equivalent to

$$R_2^* = \max_{\{\mu_1,\mu_2,\omega_1,\omega_2\}} f_2(\mu_1, \mu_2, \omega_1, \omega_2), \tag{76}$$

In order to prove that (75) is always true for case I, we need to prove that (76) is obtained with $\omega_1\omega_2 = 0$. Since f_2 in (76) is monotonically increasing as Υ_{sum} increases, we only need to maximize Υ_{sum}. It is observed that the maximal Υ_{sum} (denoted as Υ_{sum}^*) satisfies

$$\Upsilon_{sum}^* = \max_{\{\mu_1,\mu_2,\omega_1,\omega_2\}} \min(\Upsilon_{sum1}, \Upsilon_{sum2}), \tag{77}$$

By comparing with the two terms Υ_{sum1} and Υ_{sum2} in (77), we consider the following three conditions, i.e., $\Upsilon_{sum1} < \Upsilon_{sum2}$, $\Upsilon_{sum1} > \Upsilon_{sum2}$, and $\Upsilon_{sum1} = \Upsilon_{sum2}$.

a) If $\Upsilon_{sum1} < \Upsilon_{sum2}$, we define $\mathcal{S}_1 = \{(\mu_1, \mu_2, \omega_1, \omega_2) | \Upsilon_{sum1} < \Upsilon_{sum2}\}$. (77) is equivalent to (78)-(83), where (78) is obtained by the condition of $\Upsilon_{sum1} < \Upsilon_{sum2}$, (79) is obtain due to $\Upsilon_{sumj} = \sum_{k=1}^{2} h_{kj}^H Q_j^{-1} h_{kj}$, (80) is obtained by substituting $\boldsymbol{h}_{k1}$ and $\boldsymbol{Q}_1$ into (79), (81) is obtained by substituting A_1, a_{k1} and b_{k1} into (80). It can be observed that the terms in (81) increase with $\omega_1 = 0$, under the condition of $c_{21}\omega_2 c_{12}\omega_1 < 1$. We set $\hat{P}_1 = 0$ with $\omega_1 = 0$, since there is no residual SI at receiver 1 as discussed in Section II-D. Then, we derive (82) with $\omega_1 = 0$. Next, we prove that $\Upsilon_{sum1} < \Upsilon_{sum2}$ is true with $\omega_1 = 0$ in (84)-(88), where (84) is obtained by setting $\omega_1 = 0$. Note that there is always a constraint $h_{11}^2\mu_1^2 + h_{21}^2\mu_2^2 < h_{12}^2\mu_1^2 + h_{22}^2\mu_2^2$ for all $\mu_1, \mu_2 \in \mathcal{S}_1$ by sloving the inequality $\Upsilon_{sum1} < \Upsilon_{sum2}$. Due to the space limitation, we omit the detailed calculation. We chose a small $\hat{P}_2$, which lets $(h_{11}^2\mu_1^2 + h_{21}^2\mu_2^2)(1 + \hat{P}_2) < h_{12}^2\mu_1^2 + h_{22}^2\mu_2^2$ hold. With the term of $h_{11}^2\mu_1^2 + h_{21}^2\mu_2^2$ increasing, there is no gap between $h_{11}^2\mu_1^2 + h_{21}^2\mu_2^2$ and $h_{12}^2\mu_1^2 + h_{22}^2\mu_2^2$, such that the receiver cooperative system degenerate into the non-cooperative system as discussed in case I-c. Thus, (85) is obtained by magnifying the two terms in (84), and then by using the fact of $1 + \hat{P}_2 \neq 0$. (86) holds by calculating (85). (87) holds due to $1 + c_{21}^2\omega_2^2(\hat{P}_2 + 1) \neq 0$. (88) holds by substituting $\omega_1 = 0$ into Υ_{sum2}. From (84)-(88), it can be concluded that $\Upsilon_{sum1} < \Upsilon_{sum2}$ is true when $\omega_1 = 0$. Thus, (82) holds, and (83) holds with $\mu_1^*, \mu_2^*, \omega_1 = 0$, and ω_2^*.

From (78)-(83), Υ_{sum}^* is achieved when $\omega_1 = 0$. Thus, we derive that R_2^* is achieved when $\omega_1 = 0$, and (76) holds with $\omega_1\omega_2 = 0$ for $\Upsilon_{sum1} < \Upsilon_{sum2}$.

We have proved that Υ_{sum}^* given in (78) is obtained when $\omega_1 = 0$. The power gain induced by receiver cooperations is related to the cooperative link c_{21} and the transmit parameters ω_2, as shown in (82). However, with weak cooperative link c_{21}, we have to adjust ω_2 to gaurantee the power gain. If the maximal ω_2 cannot gaurantee the power gain induced by receiver cooperations, the considered IC degenerates into non-cooperative system, i.e., $\omega_1 = \omega_2 = 0$.

The results can also be obtained in the following cases.

b) If $\Upsilon_{sum1} > \Upsilon_{sum2}$, we define $\mathcal{S}_2 = \{(\mu_1, \mu_2, \omega_1, \omega_2) | \Upsilon_{sum1} > \Upsilon_{sum2}\}$. By using the same method in (78)-(83), it is easy to prove that

$$\begin{aligned}\Upsilon_{sum}^* &= \max_{\{\mu_1,\mu_2,\omega_1,\omega_2\}\in\mathcal{S}_2} \Upsilon_{sum2} \\ &= \Upsilon_{sum2}(\mu_1^*, \mu_2^*, \omega_1^*, \omega_2 = 0).\end{aligned} \tag{89}$$

From (89), Υ_{sum}^* is achieved when $\omega_2 = 0$. Thus, we derive that R_2^* is achieved when $\omega_2 = 0$, and (76) holds with $\omega_1\omega_2 = 0$ for $\Upsilon_{sum1} > \Upsilon_{sum2}$.

c) If $\Upsilon_{sum1} = \Upsilon_{sum2}$, we define $\mathcal{S}_3 = \{(\mu_1, \mu_2, \omega_1, \omega_2) | \Upsilon_{sum1} = \Upsilon_{sum2}\}$. Then, Υ_{sum}^* can be obtain as (90)-(95), where (90) is obtained due to the condition of $\Upsilon_{sum1} = \Upsilon_{sum2}$, (91) is obtained by the definition of Υ_{sum1} as the same as (81), (92) holds by magnifying the first term in (91), and the equality is achieved with $\omega_1 = \omega_2 = 0$. In order to solve the function of $\Upsilon_{sum1} = \Upsilon_{sum2}$, all terms in Υ_{sum1} and Υ_{sum2} are expanded and compared, and then it can be found

that there is always one condition of $h_{11}^2\mu_1^2 + h_{21}^2\mu_2^2 = h_{12}^2\mu_1^2 + h_{22}^2\mu_2^2$ at least. Due to the space limitation, we omit the detailed calculation. Thus, (93) holds. (94) holds due to $a^2 \geq 0$, and the equality is achieved with $\omega_1\omega_2 = 0$ (which is same as $a^2 = 0$). (95) holds due to $1 + \hat{P}_1 + c_{21}^2\omega_2^2(\hat{P}_2 + 1) > 0$. From (90)-(95), it is obvious that Υ_{sum}^* is achieved with $\omega_1 = \omega_2 = 0$. Thus, we derive that R_2^* is achieved with $\omega_1 = \omega_2 = 0$, and (76) holds with $\omega_1\omega_2 = 0$ for $\Upsilon_{sum1} = \Upsilon_{sum2}$. Same result can also be obtained by maximize Υ_{sum2}. It is worth pointing out that the considered IC with receiver cooperations performs as the non-cooperative system for the condition of $\Upsilon_{sum1} = \Upsilon_{sum2}$.

With the above analysis, we have proved that (76) is always true with $\omega_1\omega_2 = 0$ for the three different conditions, and (75) holds for case I.

2) Case II: When $f_2(\mu_1, \mu_2, \omega_1, \omega_2) \geq f_1(\mu_1, \mu_2, \omega_1, \omega_2)$, (74) is equivalent to

$$R_2^* = \max_{\{\mu_1,\mu_2,\omega_1,\omega_2\}} f_1(\mu_1, \mu_2, \omega_1, \omega_2). \tag{96}$$

Next, in order to prove (75) is always true for case II, we need to prove that (96) always holds with $\omega_1\omega_2 = 0$. By using the same method in case I, we only need to maximize Υ_2 in (96) to obtain R_2^*. It is observed that the maximal Υ_2 (denoted as Υ_2^*) satisfies

$$\Upsilon_2^* = \max_{\{\mu_1,\mu_2,\omega_1,\omega_2\}} \min\left(\Upsilon_{21}, \Upsilon_{22}\right), \tag{97}$$

By comparing with the two terms Υ_{21} and Υ_{22} in (97), we consider the following three conditions, i.e., $\Upsilon_{21} < \Upsilon_{22}$, $\Upsilon_{21} > \Upsilon_{22}$, and $\Upsilon_{21} = \Upsilon_{22}$. By using the same analysis in (78)-(83), we prove that (96) is always true with $\omega_1\omega_2 = 0$, and then (75) holds for case II.

Under the two different cases, we have proved that (75) is always true for case I-case III. For the joint decoding scheme, it has been proved that under a fixed R_1, R_2^* is achieved with $\omega_j\omega_{\bar{j}} = 0$. Moreover, if both ω_1 and ω_2 are equal to zero, the considered IC with receiver cooperations degenerates into the conventional IC without cooperation. When we fix R_2, the same conclusion can also be obtained for the maximum R_1. The same conclusions can also be proved for the single-user decoding by using the same method.

With the above analysis, Proposition 4 is proved.

ACKNOWLEDGMENT

A conference version of this article was presented at the IEEE GLOBECOM 2018 [18].

$$\Upsilon_{sum1} = \frac{c_{21}^2\omega_2^2(h_{12}^2\mu_1^2 + h_{22}^2\mu_2^2)}{1 + c_{21}^2\omega_2^2(\hat{P}_2 + 1) + h_{11}^2\mu_1^2 + h_{21}^2\mu_2^2} + \frac{h_{11}^2\mu_1^2 + h_{21}^2\mu_2^2}{1 + c_{21}^2\omega_2^2(\hat{P}_2 + 1)} \tag{84}$$

$$< \frac{c_{21}^2\omega_2^2(h_{12}^2\mu_1^2 + h_{22}^2\mu_2^2)(1 + \hat{P}_2)}{(1 + c_{21}^2\omega_2^2(\hat{P}_2 + 1))(1 + \hat{P}_2)} + \frac{h_{12}^2\mu_1^2 + h_{22}^2\mu_2^2}{(1 + c_{21}^2\omega_2^2(\hat{P}_2 + 1))(1 + \hat{P}_2)} \tag{85}$$

$$= \frac{(1 + c_{21}^2\omega_2^2(1 + \hat{P}_2))(h_{12}^2\mu_1^2 + h_{22}^2\mu_2^2)}{(1 + c_{21}^2\omega_2^2(\hat{P}_2 + 1))(1 + \hat{P}_2)} \tag{86}$$

$$= \frac{h_{12}^2\mu_1^2 + h_{22}^2\mu_2^2}{1 + \hat{P}_2}, \tag{87}$$

$$= \Upsilon_{sum2} \tag{88}$$

$$\Upsilon_{sum}^* = \max_{\{\mu_1,\mu_2,\omega_1,\omega_2\}\in\mathcal{S}_3} \Upsilon_{sum1} \tag{90}$$

$$= \max_{\{\mu_1,\mu_2,\omega_1,\omega_2\}\in\mathcal{S}_3} \frac{c_{21}^2\omega_2^2(h_{12}^2\mu_1^2 + h_{22}^2\mu_2^2)(1 - a^2)}{1 + \hat{P}_1 + c_{21}^2\omega_2^2(\hat{P}_2 + 1) + (1 - a^2)(h_{11}^2\mu_1^2 + h_{21}^2\mu_2^2)} + \frac{(h_{11}^2\mu_1^2 + h_{21}^2\mu_2^2)(1 - a^2)}{1 + \hat{P}_1 + c_{21}^2\omega_2^2(\hat{P}_2 + 1)} \tag{91}$$

$$\leq \max_{\{\mu_1,\mu_2,\omega_1,\omega_2\}\in\mathcal{S}_3} \frac{(h_{12}^2\mu_1^2 + h_{22}^2\mu_2^2)(\hat{P}_1 + c_{21}^2\omega_2^2(\hat{P}_2 + 1))}{1 + \hat{P}_1 + c_{21}^2\omega_2^2(\hat{P}_2 + 1)} + \frac{(h_{11}^2\mu_1^2 + h_{21}^2\mu_2^2)(1 - a^2)}{1 + \hat{P}_1 + c_{21}^2\omega_2^2(\hat{P}_2 + 1)} \tag{92}$$

$$= \max_{\{\mu_1,\mu_2,\omega_1,\omega_2\}\in\mathcal{S}_3} \frac{(h_{11}^2\mu_1^2 + h_{21}^2\mu_2^2)(\hat{P}_1 + c_{21}^2\omega_2^2(\hat{P}_2 + 1))}{1 + \hat{P}_1 + c_{21}^2\omega_2^2(\hat{P}_2 + 1)} + \frac{(h_{11}^2\mu_1^2 + h_{21}^2\mu_2^2)(1 - a^2)}{1 + \hat{P}_1 + c_{21}^2\omega_2^2(\hat{P}_2 + 1)} \tag{93}$$

$$\leq \max_{\{\mu_1,\mu_2,\omega_1,\omega_2\}\in\mathcal{S}_3} \frac{(h_{11}^2\mu_1^2 + h_{21}^2\mu_2^2)(1 + \hat{P}_1 + c_{21}^2\omega_2^2(\hat{P}_2 + 1))}{1 + \hat{P}_1 + c_{21}^2\omega_2^2(\hat{P}_2 + 1)} \tag{94}$$

$$= \max_{\{\mu_1,\mu_2,\omega_1,\omega_2\}\in\mathcal{S}_3} h_{11}^2\mu_1^2 + h_{21}^2\mu_2^2. \tag{95}$$

REFERENCES

[1] X. Ge, S. Tu, G. Mao, and C. X. Wang, "5g ultra-dense cellular networks," *IEEE Trans. Wireless Commun.*, vol. 23, no. 1, pp. 72–79, Feb. 2016.

[2] C. E. Shannon, "Two-way communication channels," in *Proc. 4th Berkeley Symp. Math., Statist., Probab.*, vol. 1, pp. 611–644, Oct. 1961.

[3] A. B. Carleial, "Interference channels," *IEEE Trans. Inf. Theory*, vol. IT-24, no. 1, pp. 60–70, Jan. 1978.

[4] T. Han and K. Kobayashi, "A new achievable rate region for the interference channel," *IEEE Trans. Inf. Theory*, vol. 27, no. 1, pp. 49–60, Jan. 1981.

[5] T. Cover, "An achievable rate region for the broadcast channel," *IEEE Trans. Inf. Theory*, vol. IT-21, no. 4, pp. 399–404, Jul. 1975.

[6] R. H. Etkin, D. N. C. Tse, and H. Wang, "Gaussian interference channel capacity to within one bit," *IEEE Trans. Inf. Theory*, vol. 54, no. 12, pp. 5534–5562, Dec. 2008.

[7] A. S. Motahari and A. K. Khandani, "Capacity bounds for the Gaussian interference channel," *IEEE Trans. Inf. Theory*, vol. 55, no. 2, pp. 620–643, Feb. 2009.

[8] V. S. Annapureddy and V. V. Veeravalli, "Gaussian interference networks: Sum capacity in the low-interference regime and new outer bounds on the capacity region," *IEEE Trans. Inf. Theory*, vol. 55, no. 7, pp. 3032–3050, Jul. 2009.

[9] X. Shang, G. Kramer, and B. Chen, "A new outer bound and the noisy-interference sum–rate capacity for Gaussian interference channels," *IEEE Trans. Inf. Theory*, vol. 55, no. 2, pp. 689–699, Feb. 2009.

[10] T. Cover and C. Leung, "An achievable rate region for the multiple-access channel with feedback," *IEEE Trans. Inf. Theory*, vol. IT-27, no. 3, pp. 292–298, May 1981.

[11] F. Baccelli, A. El Gamal, and D. N. C. Tse, "Interference networks with Point-to-Point codes," *IEEE Trans. Inf. Theory*, vol. 57, no. 5, pp. 2582–2596, May 2011.

[12] J. Blomer and N. Jindal, "Transmission capacity of wireless ad hoc networks: Successive interference cancellation vs. Joint detection," in *Proc. IEEE Int. Conf. Commun.*, Dresden, Germany, Jun. 2009, pp. 1–5.

[13] J. Lee, D. Toumpakaris, and W. Yu, "Interference mitigation via joint detection," *IEEE J. Sel. Areas Commun.*, vol. 29, no. 6, pp. 1172–1184, Jun. 2011.

[14] B. W. Khoueiry and M. R. Soleymani, "Destination cooperation in interference channels," in *Proc. IEEE Int. Conf. Consum. Electron. (ICCE)*, Las Vegas, NV, USA, Jan. 2012, pp. 211–212.

[15] I.-H. Wang and D. N. C. Tse, "Interference mitigation through limited receiver cooperation," *IEEE Trans. Inf. Theory*, vol. 57, no. 5, pp. 2913–2940, May 2011.

[16] A. Host-Madsen, "Capacity bounds for cooperative diversity," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1522–1544, Apr. 2006.

[17] V. M. Prabhakaran and P. Viswanath, "Interference channels with destination cooperation," *IEEE Trans. Inf. Theory*, vol. 57, no. 1, pp. 187–209, Jan. 2011.

[18] D. Wang, J. Huang, and C. Huang, "Full-duplex Amplify-and-Forward receiver cooperations for interference channels," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Abu Dhabi, United Arab Emirates, Dec. 2018, pp. 206–212.

[19] D. Tse and P. Viswanath, *Fundamentals Wireless Communication*. Cambridge, U.K.: Cambridge Univ. Press, 2005.

[20] J. Huang, D. Wang, and C. Huang, "Interference channel with full-duplex amplify-and-forward transmitter cooperations," *IEEE Access*, vol. 7, pp. 118994–119008, 2019.

[21] S. Galli, "A simple two-tap statistical model for the power line channel," in *Proc. ISPLC*, Rio de Janeiro, Brazil, Mar. 2010, pp. 242–248.

[22] H. Xiao, Z. Zhang, and A. T. Chronopoulos, "Performance analysis of multi-source multi-destination cooperative vehicular networks with the hybrid Decode-Amplify-Forward cooperative relaying protocol," *IEEE Trans. Intell. Transp. Syst.*, vol. 19, no. 9, pp. 3081–3086, Sep. 2018.

[23] X. Quan, Y. Liu, S. Shao, C. Huang, and Y. Tang, "Impacts of phase noise on digital self-interference cancellation in full-duplex communications," *IEEE Trans. Signal Process.*, vol. 65, no. 7, pp. 1881–1893, Apr. 2017.

[24] H. Cui, M. Ma, L. Song, and B. Jiao, "Relay selection for two-way full duplex relay networks with Amplify-and-Forward protocol," *IEEE Trans. Wireless Commun.*, vol. 13, no. 7, pp. 3768–3777, Jul. 2014.

[25] E. B. Dynkin, *Markov Processes*. Berlin, Germany: Springer, 1965.

[26] C. Huang, M. Zeng, and S. Cui, "Source power allocation and relaying design for two-hop interference networks with relay conferencing," *IEEE Trans. Wireless Commun.*, vol. 12, no. 7, pp. 3213–3225, Jul. 2013.

[27] M. Griot, A. I. Vila Casado, W.-Y. Weng, H. Chan, J. Wang, and R. D. Wesel, "Nonlinear trellis codes for binary-input binary-output multiple-access channels with single-user decoding," *IEEE Trans. Commun.*, vol. 60, no. 2, pp. 364–374, Feb. 2012.

[28] I. Sason, "On achievable rate regions for the Gaussian interference channel," *IEEE Trans. Inf. Theory*, vol. 50, no. 6, pp. 1345–1356, Jun. 2004.

[29] T. M. Cover and J. A. Thomas, *Elements Information Theory*. New York, NY, USA: Wiley, 1991.

[30] A. Beck, A. Ben-Tal, and L. Tetruashvili, "A sequential parametric convex approximation method with applications to nonconvex truss topology design problems," *J. Global Optim.*, vol. 47, no. 1, pp. 29–51, May 2010.

[31] S. P. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.

[32] M. Razaviyayn, M. Hong, and Z.-Q. Luo, "A unified convergence analysis of block successive minimization methods for nonsmooth optimization," *SIAM J. Optim.*, vol. 23, no. 2, pp. 1126–1153, Jan. 2013.

[33] A. El Gamal and Y.-H. Kim, "Lecture notes on network information theory," 2010, *arXiv:1001.3404*. [Online]. Available: http://arxiv.org/abs/1001.3404

[34] H. Cui, V. C. M. Leung, D. Pan, and H. Wang, "Delay analysis for distributed opportunistic cooperative communication under strong interference channel," in *Proc. IEEE 83rd Veh. Technol. Conf. (VTC Spring)*, Nanjing, China, May 2016, pp. 1–5.

[35] D. Zahavi and R. Dabora, "On cooperation and interference in the weak interference regime," *IEEE Trans. Inf. Theory*, vol. 63, no. 6, pp. 3894–3922, Jun. 2017.

DAN WANG (Graduate Student Member, IEEE) received the B.S. degree in electrical engineering from the Chongqing University of Posts and Telecommunications, Chongqing, China, in 2017. She is currently pursuing the Ph.D. degree with the University of Electronic Science and Technology of China, Chengdu, China. Her current research interests include game theory in communication systems and full-duplex communications. She was a TPC Member of GLOBECOM 2019. She is currently a Reviewer of the IEEE WIRELESS COMMUNICATION LETTERS.

JIANHAO HUANG (Graduate Student Member, IEEE) received the B.S. degree in electrical engineering from Harbin Engineering University, Harbin, China, in 2017. He is currently a Research Assistant with the School of Science and Engineering, The Chinese University of Hong Kong, Shenzhen, China. His current research interests include compress sensing in communication systems and full-duplex cooperation communications. He was a TPC Member of GLOBECOM 2019. He is also a Reviewer of the IEEE WIRELESS COMMUNICATION LETTERS.

CHUAN HUANG (Member, IEEE) received the B.S. degree in mathematics and the M.S. degree in communications engineering from the University of Electronic Science and Technology of China and the Ph.D. degree in electrical engineering from Texas A&M University, College Station, TX, USA, in 2012. From August 2012 to December 2013, he was a Postdoctoral Research Fellow and promoted as an Assistant Research Professor, from December 2013 to July 2014, at Arizona State University, Tempe, AZ, USA. He was also a Visiting Scholar with the National University of Singapore and a Research Associate with Princeton University. His current research interests include wireless communications and signal processing. He is currently serving as an Editor for the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, IEEE ACCESS, and the IEEE WIRELESS COMMUNICATIONS LETTERS.

GUOWEI SHI received the M.S. and Ph.D. degrees in electrical engineering from Northwestern Polytechnical University, China, in 1998 and 2001, respectively. He is currently a Senior Research Staff with the China Academy of Telecommunication Research, Beijing, China. He has involved in network optimization, network architecture, mesh networking, and M2M communications. His research interests include communication theory, networking, and information theory.