Capacity Bounds for a Cognitive MIMO Gaussian Z-Interference Channel

— Source link ⎘

Yong Peng, Dinesh Rajan

Institutions: Southern Methodist University

Topics: Channel capacity, MIMO, Node (circuits), Single antenna interference cancellation and Communication channel

Related papers:

- Achievable Rate of Gaussian Cognitive Z-Interference Channel with Partial Side Information

- Gaussian Z-interference channel with a relay link: Achievable rate region and asymptotic sum capacity

- On the Maximum Achievable Sum-Rate With Successive Decoding in Interference Channels

- MIMO Truncated Shannon Bound for system level capacity evaluation of wireless networks

- Achievable Rate Region for Gaussian MIMO MAC With Partial CSI

Capacity Bounds for a Cognitive MIMO Gaussian Z-Interference Channel

Yong Peng, *Student Member, IEEE*, and Dinesh Rajan, *Senior Member, IEEE*

Abstract—In this paper, we compute an achievable rate of a multiple-input–multiple-output (MIMO) Gaussian Z-interference channel (ZIC), as shown in Fig. 1, when transmit node C has imperfect cognitive knowledge of the signal sent by transmit node A. First, we compute the capacity of this channel, assuming noncausal but noisy knowledge at node C of node A's signal. We then compute the achievable rate for a causal cognitive strategy: This achievable rate is derived using a two-phase transmission scheme, in which node C uses a combination of a linear MMSE (LMMSE) estimator and a dirty-paper code, and node D employs a combination of an LMMSE estimator and a partial interference canceler. The achievable rate is studied in two different cases: 1) Node C operates in full-duplex mode, and 2) node C operates in half-duplex mode. To quantify the performance of the proposed strategy, we compute simple lower and upper bounds on the capacity of this channel. Similar to an interference channel, the achievable rate of the cognitive ZIC nonmonotonically varies with the interference. Specifically, the achievable rate first decreases with the channel gain between nodes A and D and then begins to increase beyond a certain threshold. The difference in the achievable rate between full- and half-duplex transmissions is also numerically evaluated.

Index Terms—Achievable rate, dirty-paper coding (DPC), interference mitigation, random-coding error exponent.

I. INTRODUCTION

CONSIDER a Z-channel, as shown in Fig. 1, in which the primary transmitter (node A) communicates with its intended receiver (node B). There is also a secondary transmitter (node C) that wishes to communicate with its receiver (node D) on the same frequency as the primary nodes. Each node has multiple antennas. Furthermore, node C can cognitively monitor the interfering signal from node A, hence reducing its effect at node D. We call nodes C and D the cognitive transmitter and receiver, respectively. We focus on the case when nodes C and D are relatively closer to node A than node B. Such a scenario might occur, for instance, when node A is a cellular base station and nodes C and D are two nearby nodes, whereas node B is at the cell edge. Since we assume that node B is much farther away from the other nodes, we do not explicitly consider the interference that node C causes at node B. Since
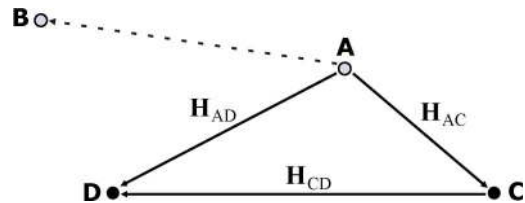
Fig. 1. System model.

node D receives a combination of both the intended signal from node C and the interfering signal from node A and node B is interference free, this system model is sometimes also referred to as a Z-interference channel (ZIC) [1]–[4], a "Z" channel [5], [6], or a one-sided interference channel. Since we also assumed that node A's signal is cognitively known at node C, we call the network in Fig. 1 a Gaussian cognitive ZIC.

Cognitive radio has been identified as a promising method for its efficient utilization of spectrum [7]–[10]. Research on cognitive spectrum sharing has broadly been classified into underlay, interweave, and overlay schemes [11]. In underlay schemes, the cognitive user selects a strategy (spread signal over very large bandwidths as in ultrawideband systems) to ensure that the additional interference to the primary user is minimized. In interweave systems, the cognitive user adaptively detects and utilizes the time-varying spectral *holes* due to inactivity of the primary user. In overlay schemes, simultaneous transmissions of primary and cognitive users are typically allowed using variations of dirty-paper coding (DPC) [12], [13]. This paper fits in the latter category and proposes a noisy version of DPC.

A simple model for a cognitive radio channel is constructed using two transmitter–receiver pairs, with one pair being the primary and the other being cognitive. Capacity bounds of cognitive radio networks have intensively been studied for the single-antenna case [11], [14]–[17]. An achievable region and an outer bound for a multiple-input–multiple-output (MIMO) Gaussian cognitive channel is derived in [18]. A new transmission scheme for the Gaussian MIMO cognitive radio channel is proposed in [19], where the channel gains are no longer deterministic and are only imperfectly known at the transmitter.

For the ZIC, the achievable region and sum capacity without cognition are studied in [2] and [3], where the two receive nodes can cooperate through a rate-limited relay link. In [6], the capacity region of the Gaussian ZIC is investigated when the crossover channel gain is small (less than the direct channel gain). With cognition, the capacity bounds of compound nonergodic fading channels with side information at the cognitive transmitter for finite alphabet channels is derived in [20]. The

capacity region of a discrete memoryless cognitive ZIC with side information of the interference at the cognitive transmitter is evaluated in [21].

In all these prior works [2], [3], [6], [11], [14]–[21] on cognitive radio channels, the authors assume that the cognitive transmitter has *perfect* and *noncausal* side information of the interference. The cognitive radio channel with *causal* side information is considered in [22], where the achievable rates of a standard two-user single-input–single-output (SISO) interference channel with cognition are evaluated for both noncausal and causal cases. The authors have also assumed that the cognitive transmitter can obtain *perfect* side information from the primary transmitter.

In this paper, we compute the achievable rates of the MIMO Gaussian cognitive ZIC.[1] Our study differs from the previous works in two key aspects.

1) We investigate transmission schemes of the MIMO Gaussian cognitive ZIC, where only *partial* and *causal* side information of the interference is available at the cognitive transmitter.

2) Unlike [22], the primary users do not cooperate with the cognitive users in any fashion, and the cognitive users can only estimate the interference from the primary transmitter based on the observed signal. Furthermore, we assume that the primary receiver is far apart such that the interference created by the cognitive transmitter can be neglected.

Specifically, let node A communicate with its receiver, i.e., node B, at rate R using the transmit signal with covariance matrix $\mathbf{Q}$. Let the covariance matrix of the transmit signal of node C be equal to $\mathbf{P}$. A simple lower bound $R_{\mathrm{CD,lb}}$ on the rate that nodes C and D can communicate is

$$R_{\mathrm{CD,lb}} = \log\left(\frac{\left|\mathbf{I}_{n_R} + \mathbf{H}_{\mathrm{AD}}\mathbf{Q}\mathbf{H}_{\mathrm{AD}}^{\dagger} + \mathbf{H}_{\mathrm{CD}}\mathbf{P}\mathbf{H}_{\mathrm{CD}}^{\dagger}\right|}{\left|\mathbf{I}_{n_R} + \mathbf{H}_{\mathrm{AD}}\mathbf{Q}\mathbf{H}_{\mathrm{AD}}^{\dagger}\right|}\right) \quad (1)$$

which is achieved by treating the signal from node A as noise at node D. Similarly, a trivial upper bound on this rate is obtained (if either node C or D has perfect noncausal knowledge of node A's signal) as

$$R_{\mathrm{CD,ub}} = \log\left|\mathbf{I}_{n_R} + \mathbf{H}_{\mathrm{CD}}\mathbf{P}\mathbf{H}_{\mathrm{CD}}^{\dagger}\right|. \quad (2)$$

We first compute the capacity of this channel when the signal from node A is noncausally observed at both nodes C and D, with certain Gaussian observation noise. We show that the interference at node D can be written as the linear combination of the following: 1) a linear MMSE (LMMSE) estimate of node A's signal based on the observations of node A's signal at nodes C and D and 2) a residual noise. We prove that by using DPC at node C and interference cancellation at node D, in conjunction with these LMMSE estimators, the interference in part 1 can be eliminated completely. Furthermore, this scheme achieves the channel capacity extended from [24] for the Gaussian partial

side-information case. For simplicity, we refer to this scheme as noisy DPC.

We then apply the noisy DPC derived in Section II to the MIMO Gaussian cognitive ZIC, where the interference from node A can only be causally observed at nodes C and D. We compute achievable rates and outer bounds on the capacity of this channel in two different cases: 1) Node C operates in full-duplex mode, and 2) node C operates in half-duplex mode.

The achievable rate is derived by applying a combination of a causal noisy DPC at the transmitter and interference cancellation at the receiver. We use a two-phase transmission strategy. In phase 1, node C obtains causal knowledge of the interference from node A. Then, node C attempts to decode node A's signal based on the portion of the codeword it observed. In phase 2, node C is in transmit mode only. Node D attempts to decode node A's signal based on the codeword it observed after phase 2. The probability of decoding errors is bounded using Gallager's random-coding exponent [25]. These decoding errors result in a noisy estimate of the interference at nodes C and D. We use this noisy estimate to compute an achievable rate R_{CD} by applying the results from the noncausal case. We evaluate the increase of R_{CD} by using the proposed causal transmission schemes in Section III. The numerical results also allow us to quantify various components of the gain (noisy DPC/interference cancellation) in various channel conditions.

Our main contributions are summarized in the list that follows.

1) We derive the capacity of a MIMO Gaussian channel with additive Gaussian interference when the interference is noncausal but imperfectly known at the transmitter and the receiver.

2) We propose novel transmission schemes for the cognitive Gaussian ZIC that significantly increase the cognitive user's achievable rate using partial and causal knowledge of the interference.

The rest of this paper is organized as follows: We evaluate the capacity of a MIMO Gaussian channel with noncausal noisy side information in Section II. We then apply the results to the cognitive Gaussian ZIC and quantify the increase of achievable rates in Section III. Numerical examples are provided in Section IV, and Section V concludes this paper.

II. CAPACITY OF A MULTIPLE-INPUT–MULTIPLE-OUTPUT GAUSSIAN CHANNEL WITH NOISY SIDE INFORMATION

In [12], Costa considered a single-user/single-antenna Gaussian channel with independent additive interference at the receiver and assumed that the interference is perfectly known at the transmitter. He proposed a new coding strategy, which is commonly referred to as DPC, that achieves rates as if the interfering source is not present. Following Costa's work, several extensions of DPC have been studied, e.g., colored Gaussian noise [26], arbitrary distributions of interference [13], deterministic sequences [27], and capacity of the Gaussian broadcast channel [28]–[30]. The case when the interference is perfectly known to the encoder and a noisy version is known to the decoder is considered in [31], mainly focusing on discrete memoryless channels. The only result in [31] for a Gaussian

---

[1]An earlier version [23] of this paper discussed the SISO full-duplex transmission case.
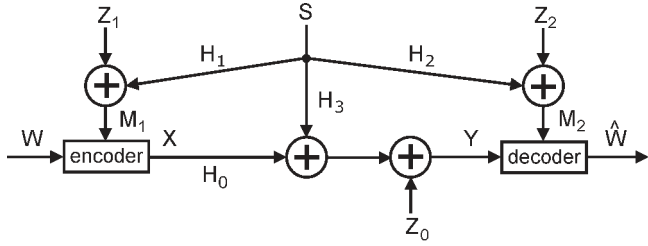
Fig. 2. Gaussian channel with noisy side information of the interference at both the encoder and the decoder.

channel reveals no additional gain due to the presence of the noisy estimate at the decoder, since perfect knowledge is available at the encoder, and DPC can be used. In contrast, in this section, we study the case when only noisy knowledge of the interference is available at both the transmitter and the receiver with multiple antennas.

A. Channel Model

Consider a MIMO Gaussian channel with additive interference, as depicted in Fig. 2. Let there be n_T transmit antennas, n_R receive antennas, and n_I antennas at the interfering source. The complex received signal matrix for n uses of the channel, i.e., $\mathbf{Y} \in \mathbb{C}^{n_R \times n}$, is corrupted by both the Gaussian noise $\mathbf{Z}_0 \in \mathbb{C}^{n_R \times n}$ and an independent interference signal matrix $\mathbf{S} \in \mathbb{C}^{n_I \times n}$ and is given by

$$\mathbf{Y} = \mathbf{H}_0 \mathbf{X} + \mathbf{H}_3 \mathbf{S} + \mathbf{Z}_0 \qquad (3)$$

where $\mathbf{H}_0 \in \mathbb{C}^{n_R \times n_T}$ and $\mathbf{H}_3 \in \mathbb{C}^{n_R \times n_I}$ are the channel matrices from the transmitter to the receiver and from the interference to the receiver, respectively, and $\mathbf{X} \in \mathbb{C}^{n_T \times n}$ is the transmitting signal matrix. Side information $\mathbf{M}_1 = \mathbf{H}_1 \mathbf{S} + \mathbf{Z}_1$ and $\mathbf{M}_2 = \mathbf{H}_2 \mathbf{S} + \mathbf{Z}_2$ of the interference is assumed to be non-causally known at the transmitter and the receiver, respectively, where $\mathbf{H}_1 \in \mathbb{C}^{n_T \times n_I}$ and $\mathbf{H}_2 \in \mathbb{C}^{n_R \times n_I}$ represent, respectively, the transmitter–interference and receiver–interference channels, and $\mathbf{Z}_1 \in \mathbb{C}^{n_T \times n}$ and $\mathbf{Z}_2 \in \mathbb{C}^{n_R \times n}$ are the respective Gaussian noises corrupting $\mathbf{S}$ at the transmitter and the receiver. We assume that the channel matrices $\mathbf{H}_i$, $i = 0, 1, 2,$ and 3 are all deterministic and are known to both the transmitter and the receiver. Since we assume noncausal noisy side information of $\mathbf{S}$ at the transmitter and the receiver, we do not restrict ourselves to the case that $\mathbf{S}$ is somehow known to the receiver through the same channel $\mathbf{H}_3$ as (3); hence, we distinguish $\mathbf{H}_2$ and $\mathbf{H}_3$.[2] Furthermore, we assume that each element of the complex Gaussian noises $\mathbf{Z}_i$, $i = 1, 2,$ and 3 is zero-mean unit-variance independent identically distributed (i.i.d.), with independent equal-variance real and imaginary parts.

The transmitter sends an index $W \in \{1, 2, \ldots, K\}$ to the receiver in n uses of the channel at rate $R = (1/n) \log K$ nats per transmission. Based on the index W and the partial side information $\mathbf{M}_1$, the transmitter picks one codeword from an (e^{nR}, n) codebook to transmit. For each of the n transmissions,

the transmit signal vector $\mathbf{x} \in \mathbb{C}^{n_T \times 1}$ and the interfering signal vector $\mathbf{s} \in \mathbb{C}^{n_I \times 1}$ are both independently generated with the following distributions: $\mathbf{x} \sim \mathcal{N}(\mathbf{0}, \mathbf{P})$ and $\mathbf{s} \sim \mathcal{N}(\mathbf{0}, \mathbf{Q})$, where the covariance matrix $\mathbf{P}$ of the transmit signal vector is determined based on $\mathbf{H}_0$ to achieve MIMO channel capacity [32]. The transmit signal has an average power constraint P_T, i.e.,

$$\frac{1}{n} \mathbb{E} \left[\mathrm{tr}(\mathbf{X}^\dagger \mathbf{X}) \right] = \frac{1}{n} \mathrm{tr}(\mathbf{P} \otimes \mathbf{I}_n) = \mathrm{tr}(\mathbf{P}) \le P_T \qquad (4)$$

where $\mathbf{I}_n$ is the identity matrix of size n, and $\otimes$ represents the Kronecker product. Similarly, the average power of the interference signal $\mathbf{S}$ is constrained by P_I, which can be expressed as

$$\frac{1}{n} \mathbb{E} \left[\mathrm{tr}(\mathbf{S}^\dagger \mathbf{S}) \right] = \frac{1}{n} \mathrm{tr}(\mathbf{Q} \otimes \mathbf{I}_n) = \mathrm{tr}(\mathbf{Q}) \le P_I. \qquad (5)$$

Letting $\hat{W}$ be the estimate of W at the receiver, an error occurs if $\hat{W} \ne W$.

B. Channel Capacity and Achievability

Theorem 1: Considering a channel with the received signal given in (3), let independent partial side information $\mathbf{M}_1 = \mathbf{H}_1 \mathbf{S} + \mathbf{Z}_1$ and $\mathbf{M}_2 = \mathbf{H}_2 \mathbf{S} + \mathbf{Z}_2$ of the interference $\mathbf{S}$ be available at the transmitter and the receiver, respectively. The covariance matrices of the signals at the transmitter and the interfering source for each of the channel uses are, respectively, $\mathbf{P}$ and $\mathbf{Q}$, and each element of the noise matrices $\mathbf{Z}_i$, $i = 1, 2,$ and 3 are zero-mean unit-variance i.i.d. Gaussian. The capacity of this channel is given by

$$C = \log \left(\frac{\left| \mathbf{I}_{n_R} + \mathbf{H}_3 \mathbf{Q} \Phi \mathbf{H}_3^\dagger + \mathbf{H}_0 \mathbf{P} \mathbf{H}_0^\dagger \right|}{\left| \mathbf{I}_{n_R} + \mathbf{H}_3 \mathbf{Q} \Phi \mathbf{H}_3^\dagger \right|} \right) \qquad (6)$$

where $\Phi = (\mathbf{I}_{n_I} + (\mathbf{H}_1^\dagger \mathbf{H}_1 + \mathbf{H}_2^\dagger \mathbf{H}_2)\mathbf{Q})^{-1}$.

1) Capacity Outer Bound: Let $\mathbf{y} \in \mathbb{C}^{n_R \times 1}$, $\mathbf{m}_1 \in \mathbb{C}^{n_T \times 1}$, and $\mathbf{m}_2 \in \mathbb{C}^{n_R \times 1}$ be the received signal vector and the transmitter and receiver side partial information of the interference signal vectors, respectively, for one use of the channel. Since both $\mathbf{x} \sim \mathcal{N}(\mathbf{0}, \mathbf{P})$ and $\mathbf{s} \sim \mathcal{N}(\mathbf{0}, \mathbf{Q})$ are independently generated for each channel use and all elements in $\mathbf{Z}_i$'s are zero-mean unit-variance i.i.d. Gaussian, we can express the distributions of the vectors as $\mathbf{y} \sim \mathcal{N}(\mathbf{0}, \mathbf{N}_y)$, where $\mathbf{N}_y = \mathbf{I}_{n_R} + \mathbf{H}_0 \mathbf{P} \mathbf{H}_0^\dagger + \mathbf{H}_3 \mathbf{Q} \mathbf{H}_3^\dagger$, $\mathbf{m}_1 \sim \mathcal{N}(\mathbf{0}, \mathbf{N}_1)$, where $\mathbf{N}_1 = \mathbf{I}_{n_T} + \mathbf{H}_1 \mathbf{Q} \mathbf{H}_1^\dagger$, and $\mathbf{m}_2 \sim \mathcal{N}(\mathbf{0}, \mathbf{N}_2)$, where $\mathbf{N}_2 = \mathbf{I}_{n_R} + \mathbf{H}_2 \mathbf{Q} \mathbf{H}_2^\dagger$. It is clear that the channel capacity cannot exceed $\max_{p(\mathbf{x}|\mathbf{m}_1, \mathbf{m}_2)} I(\mathbf{x}; \mathbf{y}|\mathbf{m}_1, \mathbf{m}_2)$, which is the capacity when $\mathbf{M}_1$ and $\mathbf{M}_2$ are known at both the transmitter and the receiver. Thus, a capacity bound is given by

$$I(\mathbf{x}; \mathbf{y}|\mathbf{m}_1, \mathbf{m}_2)$$
$$= I(\mathbf{x}; \mathbf{y}, \mathbf{m}_1, \mathbf{m}_2) - I(X; \mathbf{m}_1, \mathbf{m}_2)$$
$$\le I(\mathbf{x}; \mathbf{y}, \mathbf{m}_1, \mathbf{m}_2) \qquad (7)$$
$$= H(\mathbf{x}) + H(\mathbf{y}, \mathbf{m}_1, \mathbf{m}_2) - H(\mathbf{x}, \mathbf{y}, \mathbf{m}_1, \mathbf{m}_2)$$
$$= \log(2\pi e)^{2(n_T + n_R)} |\mathbf{P}|$$

---

[2]In Section III, we set $\mathbf{H}_2 = \mathbf{H}_3$, since the receiver uses its received signal to estimate the interference.

$$\times \begin{vmatrix} \mathbf{N}_y & \mathbf{H}_3\mathbf{PH}_1^\dagger & \mathbf{H}_3\mathbf{PH}_2^\dagger \\ \mathbf{H}_1\mathbf{PH}_3^\dagger & \mathbf{N}_1 & \mathbf{H}_1\mathbf{PH}_2^\dagger \\ \mathbf{H}_2\mathbf{PH}_3^\dagger & \mathbf{H}_2\mathbf{PH}_1^\dagger & \mathbf{N}_2 \end{vmatrix}$$

$$- \log(2\pi e)^{2(n_T + n_R)}$$

$$\times \begin{vmatrix} \mathbf{P} & \mathbf{PH}_0^\dagger & \mathbf{0} & \mathbf{0} \\ \mathbf{H}_0\mathbf{P} & \mathbf{N}_y & \mathbf{H}_3\mathbf{PH}_1^\dagger & \mathbf{H}_3\mathbf{PH}_2^\dagger \\ \mathbf{0} & \mathbf{H}_1\mathbf{PH}_3^\dagger & \mathbf{N}_1 & \mathbf{H}_1\mathbf{PH}_2^\dagger \\ \mathbf{0} & \mathbf{H}_2\mathbf{PH}_3^\dagger & \mathbf{H}_2\mathbf{PH}_1^\dagger & \mathbf{N}_2 \end{vmatrix} \quad (8)$$

$$= \log\left(\frac{\left|\mathbf{I}_{n_R} + \mathbf{H}_3\mathbf{Q}\Phi\mathbf{H}_3^\dagger + \mathbf{H}_0\mathbf{PH}_0^\dagger\right|}{\left|\mathbf{I}_{n_R} + \mathbf{H}_3\mathbf{Q}\Phi\mathbf{H}_3^\dagger\right|}\right). \quad (9)$$

Note that the inequality in (7) is actually a strict equality since $I(\mathbf{x}; \mathbf{m}_1, \mathbf{m}_2) = 0$. The detailed derivation from (7)–(9) is given in the Appendix.

2) Proof of Achievability: We now show that the capacity bound in (9) is achievable. The proof is based on a direct application of DPC by treating the LMMSE estimation of the interference at the receiver as side information; hence, we only outline the key steps. Using LMMSE, the interference at the receiver, i.e., $\mathbf{H}_3\mathbf{S}$, can be expressed as

$$\mathbf{H}_3\mathbf{S} = \mathbf{H}_3\mathbf{A}_1\mathbf{M}_1 + \mathbf{H}_3\mathbf{A}_2\mathbf{M}_2 + \hat{\mathbf{Z}} \quad (10)$$

where $\mathbf{A}_1 = (\mathbf{I}_{n_I} + \mathbf{Q}(\mathbf{H}_1^\dagger\mathbf{H}_1 + \mathbf{H}_2^\dagger\mathbf{H}_2))^{-1}\mathbf{Q}\mathbf{H}_1^\dagger$ and $\mathbf{A}_2 = (\mathbf{I}_{n_I} + \mathbf{Q}(\mathbf{H}_1^\dagger\mathbf{H}_1 + \mathbf{H}_2^\dagger\mathbf{H}_2))^{-1}\mathbf{Q}\mathbf{H}_2^\dagger$ are the LMMSE coefficient matrices, and $\hat{\mathbf{Z}}$ is the residual estimation error.

Since $\mathbf{S}$, $\mathbf{M}_1$, and $\mathbf{M}_2$ are all Gaussian matrices, the estimation noise matrix $\hat{\mathbf{Z}}$ is also Gaussian and is orthogonal to $\mathbf{M}_1$ and $\mathbf{M}_2$ due to LMMSE. This channel can be treated as a Gaussian channel with perfect side information $\mathbf{H}_3\mathbf{A}_1\mathbf{M}_1$ at the transmitter and perfect knowledge of part of the interference $\mathbf{H}_3\mathbf{A}_2\mathbf{M}_2$ at the receiver. Therefore, the transmitter can apply DPC to remove $\mathbf{H}_3\mathbf{A}_1\mathbf{M}_1$ from the received signal. The receiver can directly subtract $\mathbf{H}_3\mathbf{A}_2\mathbf{M}_2$ from the received signal. Thus, the remaining noise of the channel is $\hat{\mathbf{Z}} + \mathbf{Z}_0$. Letting $\hat{\mathbf{z}}$ be the estimation noise vector at any time instant, after some algebra, the covariance $\hat{\mathbf{N}}$ of $\hat{\mathbf{z}}$ is found to be

$$\hat{\mathbf{N}} = \mathbf{H}_3\mathbf{Q}\Phi\mathbf{H}_3^\dagger. \quad (11)$$

Since $\hat{\mathbf{Z}}$ and $\mathbf{Z}_0$ are independent of each other, the covariance matrix of the total noise is $\mathbf{I}_{n_R} + \hat{\mathbf{N}}$. The achievable rate is then given by $\log(|\mathbf{I}_{n_R} + \hat{\mathbf{N}} + \mathbf{H}_0\mathbf{PH}_0^\dagger|/|\mathbf{I}_{n_R} + \hat{\mathbf{N}}|)$, which is exactly the channel capacity expression given in (9). For easy reference, we call this capacity-achieving scheme the *noisy DPC*.

C. Special Case: SISO

We now briefly discuss the single-antenna case and compare it with Costa's standard DPC model. To simplify the comparison with [12], we normalize the channel gains instead of the Gaussian channel noises to 1.
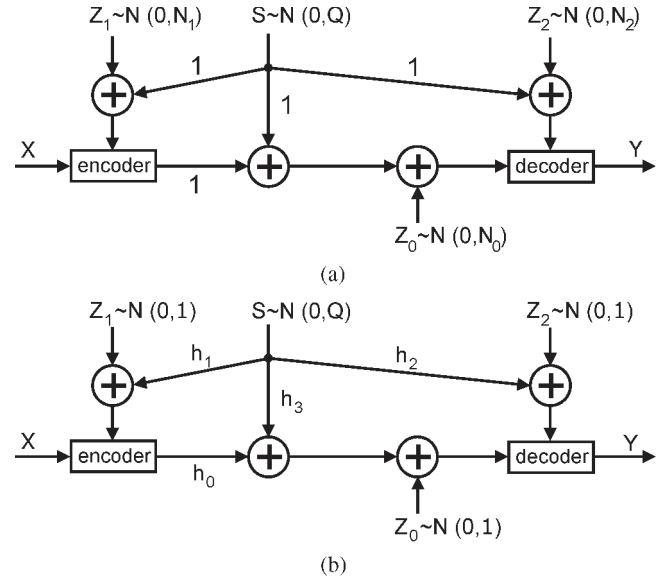

Fig. 3. SISO Gaussian channel with noisy side information at both the encoder and the decoder.

Consider a SISO Gaussian channel with partial side information, as shown in Fig. 3(a), where $Z_0 \sim \mathcal{N}(0, N_0)$, $Z_1 \sim \mathcal{N}(0, N_1)$, and $Z_2 \sim \mathcal{N}(0, N_2)$ are the noise variable added to the channel, the side information at the transmitter, and the side information at the receiver, respectively. By normalizing the variance Z_0, Z_1, and Z_2 to 1, the equivalent channel gains between the transmitter and the receiver and between the interfering source and the receiver are $h_0 = h_3 = 1/\sqrt{N_0}$. The observation channels between the transmitter and the interfering source and between the receiver and the interfering source have channel gains $h_1 = 1/\sqrt{N_1}$ and $h_2 = 1/\sqrt{N_2}$, respectively. The channel model after normalizing the noise variances is depicted in Fig. 3(b).

Therefore, we can express the capacity of the SISO Gaussian channel with partial side information in terms of N_0, N_1, and N_2 by substituting $\mathbf{H}_0$ and $\mathbf{H}_3$ with $1/\sqrt{N_0}$ in (6) and by substituting $\mathbf{H}_1$ and $\mathbf{H}_2$ with $1/\sqrt{N_1}$ and $1/\sqrt{N_2}$, respectively. The resultant capacity is given by $I(X; Y|M_1, M_2) = C(P/(\mu Q + N_0))$, where $0 \le \mu = 1/(1 + (Q/N_1) + (Q/N_2)) \le 1$. Letting $N_2 = \infty$ and $N_1 = \infty$, we can obtain the channel capacities when the observation of the interference is available either at the transmitter or the receiver, and those capacities are, respectively

$$I(XY|M_1) = C\left(\frac{P}{QN_1/(Q+N_1) + N_0}\right) \quad (12)$$

$$I(XY|M_2) = C\left(\frac{P}{QN_2/(Q+N_2) + N_0}\right). \quad (13)$$

Note that when $N_1 = 0$, the channel model further reduces to Costa's DPC channel model [12]. Furthermore, by setting $N_1 = N_2$ in (12) and (13), we can see that the observation of $\mathbf{S}$ made at the transmitter and the receiver are equivalent in the achievable rate, as long as the corrupting Gaussian noises have the same statistics.

III. COGNITIVE MULTIPLE-INPUT–MULTIPLE-OUTPUT GAUSSIAN Z-INTERFERENCE CHANNEL

We present our main results in this section. First, we introduce the cognitive system model of interest. Then, we apply noisy DPC to the causal case and provide practical transmission schemes that can increase the achievable rate of the Gaussian cognitive ZIC.

A. System Model

The considered cognitive ZIC model is shown in Fig. 1. We make similar assumptions as in Section II that all the channel gains are deterministic and that the noises are all i.i.d. Gaussian with covariance matrices normalized to identity. In Fig. 1, nodes A and B are the primary transmitter and receiver, respectively. Node A transmits an index $V \in \{1, 2, \dots, L\}$ using n_I antennas to the receiver in n uses of the channel at rate $R = (1/n) \log L$ nats per transmission. The codeword node A transmits is from an (e^{nR}, n) random Gaussian codebook, i.e., each element of the codeword is randomly generated according to Gaussian i.i.d. across time and with covariance matrix $\mathbf{Q} \in \mathbb{C}^{n_I \times n_I}$ across different antennas. The covariance matrix $\mathbf{Q}$ is determined by the channel matrix $\mathbf{H}_{\mathrm{AB}}$ between nodes A and B, which can achieve the MIMO Gaussian channel capacity. Nodes C and D are the cognitive transmitter and receiver, with n_T and n_R antennas, respectively. We assume that nodes C and D are relatively closer to node A than node B. The channel matrices between nodes C and D, nodes A and C, and nodes A and D are $\mathbf{H}_{\mathrm{CD}} \in \mathbb{C}^{n_R \times n_T}$, $\mathbf{H}_{\mathrm{AC}} \in \mathbb{C}^{n_T \times n_I}$, and $\mathbf{H}_{\mathrm{AD}} \in \mathbb{C}^{n_R \times n_I}$, respectively. Similarly, the covariance matrix $\mathbf{P} \in \mathbb{C}^{n_T \times n_T}$ of node C's transmit symbols across different antennas is determined by $\mathbf{H}_{\mathrm{CD}}$. Letting P_A and P_C be the average power constraints of nodes A and C, respectively, which are defined similarly as in (4) and (5).

It is clear that if noncausal noisy side information is available at the transmitter/receiver, using the noisy DPC scheme proposed in Section II can significantly increase the achievable rate. However, the noncausal assumption is not realistic for interference channels in practice. In Section III-B, we show that, in the causal case, we can also apply noisy DPC to obtain higher achievable rates under certain channel conditions.

B. Achievable Rates

Theorem 2: Consider a Gaussian ZIC, as shown in Fig. 1. If node C can operate in full-duplex mode, then it can communicate with node D at the rate given by

$$R_{\mathrm{CD}} = \max(R_{\mathrm{CD},1}, R_{\mathrm{CD},2}, R_{\mathrm{CD},3}, R_{\mathrm{CD},4}) \qquad (14)$$

where $R_{\mathrm{CD},1}$, $R_{\mathrm{CD},2}$, $R_{\mathrm{CD},3}$, and $R_{\mathrm{CD},4}$ are given in (24), (28), (32), and (33), respectively. If node C is constrained to operate in half-duplex mode, then it can communicate with node D at the rate given by

$$R'_{\mathrm{CD}} = \max\left(R'_{\mathrm{CD},1}, R'_{\mathrm{CD},2}, R_{\mathrm{CD},3}, R_{\mathrm{CD},4}\right) \qquad (15)$$

where $R'_{\mathrm{CD},1}$ and $R'_{\mathrm{CD},2}$ are given in (25) and (29), respectively.

Proof: We construct four different transmission schemes and compute the maximum achievable rate across all schemes. These four schemes are as follows:

1) noisy DPC at node C with partial interference cancellation at node D;
2) only noisy DPC at node C;
3) only partial interference cancellation at node D;
4) baseline transmission without interference mitigation at nodes C and D.

Schemes 3 and 4 have been well studied in the literature. Our main contribution lies in schemes 1 and 2, which are based on an extension of the noncausal noisy DPC proposed in Section II to the causal case. We first qualitatively explain the proposed causal noisy DPC scheme. As noted in Section III-A, node A chooses its codeword from an (e^{nR}, n) random Gaussian codebook under certain channel conditions. Node C can decode the message node A sent by using the first m symbols ($0 < m < n, m \in \mathbb{N}$). The upper bound on the decoding error probability can be calculated by using the theory of error exponents [25]. Node C can then use the decoded signal to estimate the remaining $n - m$ symbols that node A transmits and treat it as the noisy side information. Therefore, we can apply the results in Section II to compute the increased achievable rate of the cognitive ZIC.

We now elaborate the derivations of the achievable rates under each of the four schemes for both full- and half-duplex transmissions.

Scheme 1 (Noisy DPC at Node C With Partial Interference Cancellation at Node D): We first evaluate the full-duplex achievable rate of scheme 1 and then extend the result to the half-duplex case.

Full-duplex: The proposed scheme has two phases: Phase 1 lasts for m symbol periods, and phase 2 lasts for $n - m$ symbol periods. The covariance matrices of the signal transmitted by node C at each time instant across all antennas are $\mathbf{P}_1 = (n/m)\alpha\mathbf{P}$ and $\mathbf{P}_2 = (n/n - m)(1 - \alpha)\mathbf{P}$ in phases 1 and 2, respectively, where $0 \le \alpha \le 1$ is the power-allocation factor.

The total achievable rate between nodes A and C in phase 1 is given by

$$R_{\mathrm{AC}}(m) = m \log \left| \mathbf{I}_{n_T} + \mathbf{H}_{\mathrm{AC}} \mathbf{Q} \mathbf{H}_{\mathrm{AC}}^{\dagger} \right|. \qquad (16)$$

For values of m such that $nR \le R_{\mathrm{AC}}(m)$, node C can decode node A's signal using the symbols received in phase 1 with an error probability bounded by $P_{e,C} \le \exp(-mE_C(m))$, where $E_C(m)$ is Gallager's random-coding error exponent [25]. The error exponent $E_C(m)$ can be computed as a special case of [33] for the deterministic MIMO Gaussian channel and is given by

$$E_C(m) = \max_{\substack{0 \le \rho \le 1 \\ r \ge 1, 0 < |\mathbf{I}_I - r\mathbf{Q}| \le 1}} (1 + \rho)\left(rP_A + \log|\mathbf{I}_{n_I} - r\mathbf{Q}|\right)$$

$$+ \rho \log \left| \mathbf{I}_{n_R} + \frac{\mathbf{H}_{\mathrm{AC}}(\mathbf{Q}^{-1} - r\mathbf{I}_{n_I})^{-1}\mathbf{H}_{\mathrm{AC}}^{\dagger}}{1 + \rho} \right| - \frac{\rho n}{m} R \qquad (17)$$

where ρ and r are free parameters to be optimized [25].

Thus, the estimate of node A's codeword at node C, i.e., $\tilde{\mathbf{M}}_1$, is a combination of the true codeword matrix $\mathbf{S}$ and an estimation noise matrix $\tilde{\mathbf{Z}}_1$, i.e., $\tilde{\mathbf{M}}_1 = \mathbf{S} + \tilde{\mathbf{Z}}_1$. Let $\tilde{\mathbf{z}}_1$ be the estimation noise vector at one time instant across different antennas. The covariance matrix $\tilde{\mathbf{N}}_1$ of $\tilde{\mathbf{z}}_1$ can be computed as

$$\tilde{\mathbf{N}}_1 = P_{e,C}\mathbb{E}\left[(\mathbf{s}-\tilde{\mathbf{s}})(\mathbf{s}-\tilde{\mathbf{s}})^\dagger\right] = 2P_{e,C}\mathbf{Q} \qquad (18)$$

where $\mathbf{s}$ is the random Gaussian vector that node A transmits at one time instant, and $\tilde{\mathbf{s}}$ is its false estimate at node C. Node C can then apply noisy DPC by treating $\tilde{\mathbf{M}}_1$ as the noisy side information to reduce the interference of the remaining $n - m$ transmissions.

The total achievable rate $R_{\mathrm{AD}}(m,\alpha)$ between nodes A and D during phases 1 and 2 is given by

$$R_{\mathrm{AD}}(m,\alpha) = m\,\log\left(\frac{\left|\mathbf{I}_{n_R}+\mathbf{H}_{\mathrm{AD}}\mathbf{Q}\mathbf{H}_{\mathrm{AD}}^\dagger+\mathbf{H}_{\mathrm{CD}}\mathbf{P}_1\mathbf{H}_{\mathrm{CD}}^\dagger\right|}{\left|\mathbf{I}_{n_R}+\mathbf{H}_{\mathrm{CD}}\mathbf{P}_2\mathbf{H}_{\mathrm{CD}}^\dagger\right|}\right)$$
$$+ (n-m)\,\log\left(\frac{\left|\mathbf{I}_{n_R}+\mathbf{H}_{\mathrm{AD}}\mathbf{Q}\mathbf{H}_{\mathrm{AD}}^\dagger+\mathbf{H}_{\mathrm{CD}}\mathbf{P}_2\mathbf{H}_{\mathrm{CD}}^\dagger\right|}{\left|\mathbf{I}_{n_R}+\mathbf{H}_{\mathrm{CD}}\mathbf{P}_2\mathbf{H}_{\mathrm{CD}}^\dagger\right|}\right). \qquad (19)$$

For values of m such that $nR \le R_{\mathrm{AD}}(m,\alpha)$, node D can decode node A's message with error probability $P_{e,D} \le \exp(-nE_D(m,\alpha))$ at the end of phase 2, where $E_D(m,\alpha)$ is the random-coding error exponent for node D to decode node A's signal. The computation of $E_D(m,\alpha)$ is more involved since node D experiences time-varying interference as a result of node C using different powers in phases 1 and 2. Applying [34] and assuming deterministic channel matrices, $E_D(m,\alpha)$ is given in (20), shown at the bottom of the page, where $\Xi = \mathbf{H}_{\mathrm{AD}}(\mathbf{Q}^{-1} - r\mathbf{I}_{n_I}^{-1})^{-1}\mathbf{H}_{\mathrm{AD}}^\dagger$. Similar to the case of node C, we can express node D's estimation of node A's codeword as $\tilde{\mathbf{M}}_2 = \mathbf{S} + \tilde{\mathbf{Z}}_2$, where $\tilde{\mathbf{Z}}_2$ is the estimation noise matrix. Letting $\tilde{\mathbf{z}}_2$ be the estimation noise vector at one time instant across different antennas, the covariance matrix of $\tilde{\mathbf{z}}_2$ is $\tilde{\mathbf{N}}_2 = 2P_{e,D}\mathbf{Q}$. Node D can thus treat $\tilde{\mathbf{M}}_2$ as its noisy side information and apply noisy DPC to partially cancel out the interference.

We now apply the result in Theorem 1 to calculate the achievable rate for scheme 1. Recall that we assumed that the channel noise covariance matrices are all normalized to identity when calculating the channel capacity for the noncausal case. To utilize the result in (6), we also normalize the estimation noise covariance matrices of the causal case. Let $\mathbf{U}_1$ be a scaled version of the unitary matrix that can normalize $\tilde{\mathbf{N}}_1$ to $\mathbf{I}_{n_T}$, i.e.,

$\mathbf{U}_1\tilde{\mathbf{N}}_1\mathbf{U}_1^\dagger = \mathbf{I}_{n_T}$. Equivalently, we can express the noisy side information at the transmitter as $\tilde{\mathbf{M}}_1 = \mathbf{U}_1\mathbf{S} + \mathbf{U}_1\tilde{\mathbf{Z}}_1$, where $\mathbf{U}_1$ can be interpreted as the channel matrix between nodes A and C. Similarly, at node D, let $\mathbf{U}_2$ be a scaled version of the unitary matrix that can normalize $\tilde{\mathbf{N}}_2$ to $\mathbf{I}_{n_R}$; we can express the noisy side information at the receiver as $\tilde{\mathbf{M}}_2 = \mathbf{U}_2\mathbf{S} + \mathbf{U}_2\tilde{\mathbf{Z}}_2$. Thus, we can calculate the achievable rate for the causal case by simply replacing the channel matrices $\mathbf{H}_1$, $\mathbf{H}_2$, and $\mathbf{H}_3$ in (6) with the corresponding equivalent channel matrices $\mathbf{U}_1$, $\mathbf{U}_2$, and $\mathbf{H}_{\mathrm{AD}}$, respectively. Since $\mathbf{U}_1\tilde{\mathbf{N}}_1\mathbf{U}_1^\dagger = \mathbf{I}_{n_T}$, we have

$$\mathbf{U}_1^\dagger\mathbf{U}_1 = \tilde{\mathbf{N}}_1^{-1} = \frac{1}{2P_{e,C}}\mathbf{Q}^{-1}. \qquad (21)$$

Since $\mathbf{U}_2\tilde{\mathbf{N}}_2\mathbf{U}_2^\dagger = \mathbf{I}_{n_R}$, we have

$$\mathbf{U}_2^\dagger\mathbf{U}_2 = \tilde{\mathbf{N}}_2^{-1} = \frac{1}{2P_{e,D}}\mathbf{Q}^{-1}. \qquad (22)$$

After obtaining the noisy side information of node A in phase 1, node C can apply DPC, and node D can apply partial interference cancellation in phase 2. Substituting $\mathbf{H}_1^\dagger\mathbf{H}_1$ and $\mathbf{H}_2^\dagger\mathbf{H}_2$ in (6) with the values of $\mathbf{U}_1^\dagger\mathbf{U}_1$ and $\mathbf{U}_2^\dagger\mathbf{U}_2$ computed in (21) and (22), respectively, and replacing $\mathbf{H}_3$ by $\mathbf{H}_{\mathrm{AD}}$, the achievable rate $R_{\mathrm{CD},1b}(m,\alpha)$ in phase 2 is given by

$$R_{\mathrm{CD},1b}(m,\alpha)$$
$$= \log\left(\frac{\left|\mathbf{I}_{n_R} + \mu_{\mathrm{TR}}(m,\alpha)\mathbf{H}_{\mathrm{AD}}\mathbf{Q}\mathbf{H}_{\mathrm{AD}}^\dagger + \mathbf{H}_{\mathrm{CD}}\mathbf{P}_2\mathbf{H}_{\mathrm{CD}}^\dagger\right|}{\left|\mathbf{I}_{n_R} + \mu_{\mathrm{TR}}(m,\alpha)\mathbf{H}_{\mathrm{AD}}\mathbf{Q}\mathbf{H}_{\mathrm{AD}}^\dagger\right|}\right) \qquad (23)$$

where

$$\mu_{\mathrm{TR}}(m,\alpha) = \frac{1}{1+0.5\exp\left(mE_C(m)\right)+0.5\exp\left(nE_D(m,\alpha)\right)}.$$

Since node D can go back to cancel node A's interference at the end of phase 2, the achievable rate $R_{\mathrm{CD},1a}(m,\alpha)$ can be calculated using similar method as in calculating (23) and is given by

$$R_{\mathrm{CD},1a}(m,\alpha)$$
$$= \log\left(\frac{\left|\mathbf{I}_{n_R} + \mu_R(m,\alpha)\mathbf{H}_{\mathrm{AD}}\mathbf{Q}\mathbf{H}_{\mathrm{AD}}^\dagger + \mathbf{H}_{\mathrm{CD}}\mathbf{P}_1\mathbf{H}_{\mathrm{CD}}^\dagger\right|}{\left|\mathbf{I}_{n_R} + \mu_R(m,\alpha)\mathbf{H}_{\mathrm{AD}}\mathbf{Q}\mathbf{H}_{\mathrm{AD}}^\dagger\right|}\right)$$

where

$$\mu_R(m,\alpha) = \frac{1}{1 + 0.5\exp\left(nE_D(m,\alpha)\right)}.$$

$$E_D(m,\alpha) = \max_{\substack{0\le\rho\le1 \\ r\ge1,\,0<|\mathbf{I}_I-r\mathbf{Q}|\le1}} (1+\rho)\left(rP_A + \log|\mathbf{I}_{n_I} - r\mathbf{Q}|\right)$$
$$+\rho\,\log\left(\frac{m}{n}\frac{\left|(1+\rho)\left(\mathbf{I}_{n_R}+\mathbf{H}_{\mathrm{CD}}\mathbf{P}_1\mathbf{H}_{\mathrm{CD}}^\dagger\right)+\Xi\right|}{\left|(1+\rho)\left(\mathbf{I}_{n_R}+\mathbf{H}_{\mathrm{CD}}\mathbf{P}_1\mathbf{H}_{\mathrm{CD}}^\dagger\right)\right|} + \frac{n-m}{n}\frac{\left|(1+\rho)\left(\mathbf{I}_{n_R}+\mathbf{H}_{\mathrm{CD}}\mathbf{P}_2\mathbf{H}_{\mathrm{CD}}^\dagger\right)+\Xi\right|}{\left|(1+\rho)\left(\mathbf{I}_{n_R}+\mathbf{H}_{\mathrm{CD}}\mathbf{P}_2\mathbf{H}_{\mathrm{CD}}^\dagger\right)\right|}\right) - \rho R \qquad (20)$$

Thus, the average achievable rate $R_{\mathrm{CD},1}$ using scheme 1 with full-duplex transmission is given by

$$R_{\mathrm{CD},1} = \max_{\substack{0 < m < n \\ 0 \leq \alpha \leq 1}} \frac{m}{n} R'_{\mathrm{CD},1a}(m,\alpha) + \frac{n-m}{n} R'_{\mathrm{CD},1b}(m,\alpha)$$

$$\text{s.t.} \quad nR \leq R_{\mathrm{AC}}(m) \text{ and } nR \leq R_{\mathrm{AD}}(m,\alpha). \quad (24)$$

If there are no values of m, $0 < m < n$, such that $nR \leq R_{\mathrm{AC}}(m)$ and $nR \leq R_{\mathrm{AD}}(m,\alpha)$, we set $R_{\mathrm{CD},1} = R_{\mathrm{CD,lb}}$.

Half-duplex: In half-duplex transmission, node C has to be silent in phase 1 to obtain the noisy side information from node A. Hence, by setting $R_{\mathrm{CD},1a}(m,\alpha) = 1$ and $\alpha = 0$ in (24), the achievable rate $R'_{\mathrm{CD},1}$ using scheme 1 with half-duplex transmission is given by

$$R'_{\mathrm{CD},1} = \max_{\substack{0 < m < n \\ \text{s.t. } nR \leq R_{\mathrm{AC}}(m) \text{ and } nR \leq R_{\mathrm{AD}}(m,0)}} \frac{n-m}{n}$$

$$\times \log \left(\frac{\left| \mathbf{I}_{n_R} + \mu_{\mathrm{TR}}(m,0)\mathbf{H}_{\mathrm{AD}}\mathbf{Q}\mathbf{H}_{\mathrm{AD}}^{\dagger} + \frac{n}{n-m}\mathbf{H}_{\mathrm{CD}}\mathbf{P}\mathbf{H}_{\mathrm{CD}}^{\dagger} \right|}{\left| \mathbf{I}_{n_R} + \mu_{\mathrm{TR}}(m,0)\mathbf{H}_{\mathrm{AD}}\mathbf{Q}\mathbf{H}_{\mathrm{AD}}^{\dagger} \right|} \right). \quad (25)$$

Note that there is no constraint that node C must use codewords of length $n - m$ since node A uses codewords of length n. Node C can code over multiple codewords of A to achieve its desired probability of error.

Scheme 2 (Only Noisy DPC at Node C): Scheme 2 is similar to scheme 1, except that node D does not apply partial interference cancellation. This scheme is potentially superior in case the channel quality between nodes A and D is not as good as the channel quality between nodes A and C.

Full-duplex: The calculation of the achievable rate of scheme 2 is similar to scheme 1, except that $E_D(m,\alpha)$ does not exist since node D does not decode node A's signal. In phase 1, the achievable rate $R_{\mathrm{CD},2a}(m)$ is given by

$$R_{\mathrm{CD},2a}(m) = \log \left(\frac{\left| \mathbf{I}_{n_R} + \mathbf{H}_{\mathrm{AD}}\mathbf{Q}\mathbf{H}_{\mathrm{AD}}^{\dagger} + \mathbf{H}_{\mathrm{CD}}\mathbf{P}_1\mathbf{H}_{\mathrm{CD}}^{\dagger} \right|}{\left| \mathbf{I}_{n_R} + \mathbf{H}_{\mathrm{AD}}\mathbf{Q}\mathbf{H}_{\mathrm{AD}}^{\dagger} \right|} \right). \quad (26)$$

In phase 2, node C uses noisy DPC to reduce the interference, and the achievable rate $R'_{\mathrm{CD},2b}(m)$ is given by

$$R_{\mathrm{CD},2b}(m)$$
$$= \log \left(\frac{\left| \mathbf{I}_{n_R} + \mu_T(m)\mathbf{H}_{\mathrm{AD}}\mathbf{Q}\mathbf{H}_{\mathrm{AD}}^{\dagger} + \mathbf{H}_{\mathrm{CD}}\mathbf{P}_2\mathbf{H}_{\mathrm{CD}}^{\dagger} \right|}{\left| \mathbf{I}_{n_R} + \mu_T(m)\mathbf{H}_{\mathrm{AD}}\mathbf{Q}\mathbf{H}_{\mathrm{AD}}^{\dagger} \right|} \right) \quad (27)$$

for values of m such that $nR \geq R_{\mathrm{AC}}(m)$, where

$$\mu_T(m) = \frac{1}{1 + 0.5 \exp\left(m E_C(m)\right)}.$$

Thus, the average achievable rate $R_{\mathrm{CD},2}$ using scheme 2 with full-duplex transmission is

$$R_{\mathrm{CD},2} = \max_{\substack{0 < m < n \\ \text{s.t. } nR \leq R_{\mathrm{AC}}(m)}} \frac{m}{n} R_{\mathrm{CD},2a}(m) + \frac{n-m}{n} R_{\mathrm{CD},2b}(m). \quad (28)$$

Half-duplex: Since node C has to be silent in phase 1 to obtain the noisy side information, we set $R_{\mathrm{CD},2a}(m) = 0$ in (28). The average achievable rate $R'_{\mathrm{CD},2}$ using scheme 2 with half-duplex transmission is then given by

$$R'_{\mathrm{CD},2} = \max_{\substack{0 < m < n, \\ \text{s.t. } nR \leq R_{\mathrm{AC}}(m)}} \frac{n-m}{n}$$

$$\times \log \left(\frac{\left| \mathbf{I}_{n_R} + \mu_T(m)\mathbf{H}_{\mathrm{AD}}\mathbf{Q}\mathbf{H}_{\mathrm{AD}}^{\dagger} + \frac{n}{n-m}\mathbf{H}_{\mathrm{CD}}\mathbf{P}\mathbf{H}_{\mathrm{CD}}^{\dagger} \right|}{\left| \mathbf{I}_{n_R} + \mu_T(m)\mathbf{H}_{\mathrm{AD}}\mathbf{Q}\mathbf{H}_{\mathrm{AD}}^{\dagger} \right|} \right). \quad (29)$$

We set $R_{\mathrm{CD},2} = R'_{\mathrm{CD},2} = R_{\mathrm{CD,lb}}$ if no values of m satisfy $nR \leq R_{\mathrm{AC}}(m)$.

Scheme 3 (Only Partial Interference Cancellation at Node D): Scheme 3 is similar to scheme 1, except that node C does not apply noisy DPC. This scheme is potentially superior if the channel quality between nodes A and C is worse than the channel quality between nodes A and D. This scheme is also similar to a multiple-access channel (formed by nodes A, C, and D), except that node D decodes node A's signal only to increase the achievable rate between nodes C and D. Furthermore, since node C does not employ any decoding, the achievable rate is the same, even if node C can only operate in half-duplex mode. Following similar derivations as shown in scheme 1, the achievable rate $R_{\mathrm{CD},3a}(m,\alpha)$ in phase 1 using scheme 3 is given by

$$R_{\mathrm{CD},3a}(m,\alpha)$$
$$= \log \left(\frac{\left| \mathbf{I}_{n_R} + \mu_R(m,\alpha)\mathbf{H}_{\mathrm{AD}}\mathbf{Q}\mathbf{H}_{\mathrm{AD}}^{\dagger} + \mathbf{H}_{\mathrm{CD}}\mathbf{P}_1\mathbf{H}_{\mathrm{CD}}^{\dagger} \right|}{\left| \mathbf{I}_{n_R} + \mu_R(m,\alpha)\mathbf{H}_{\mathrm{AD}}\mathbf{Q}\mathbf{H}_{\mathrm{AD}}^{\dagger} \right|} \right) \quad (30)$$

and the achievable rate $R_{\mathrm{CD},3b}(m,\alpha)$ in phase 2 is given by

$$R_{\mathrm{CD},3b}(m,\alpha)$$
$$= \log \left(\frac{\left| \mathbf{I}_{n_R} + \mu_R(m,\alpha)\mathbf{H}_{\mathrm{AD}}\mathbf{Q}\mathbf{H}_{\mathrm{AD}}^{\dagger} + \mathbf{H}_{\mathrm{CD}}\mathbf{P}_2\mathbf{H}_{\mathrm{CD}}^{\dagger} \right|}{\left| \mathbf{I}_{n_R} + \mu_R(m,\alpha)\mathbf{H}_{\mathrm{AD}}\mathbf{Q}\mathbf{H}_{\mathrm{AD}}^{\dagger} \right|} \right). \quad (31)$$

The average achievable rate $R_{\mathrm{CD},3}$ using scheme 3 for both full- and half-duplex transmissions is then given by

$$R_{\mathrm{CD},3} = \max_{\substack{0 < m < n \\ 0 \leq \alpha \leq 1}} \frac{m}{n} R_{\mathrm{CD},3a}(m,\alpha) + \frac{n-m}{n} R_{\mathrm{CD},3b}(m,\alpha)$$

$$\text{s.t.} \quad nR \leq R_{\mathrm{AD}}(m,\alpha). \quad (32)$$

We set $R_{\mathrm{CD},2} = R'_{\mathrm{CD},2} = R_{\mathrm{CD,lb}}$ if no values of m satisfy $nR < R_{\mathrm{AD}}(m,\alpha)$.

Scheme 4 (Baseline Transmission Without Interference Mitigation at Node C or D): Since neither node C nor D does interference mitigation, the achievable rate $R_{\mathrm{CD},4}$ of scheme 4 is

$$R_{\mathrm{CD},4} = R_{\mathrm{CD,lb}}. \quad (33)$$

∎

Remark 1: A closed-form expression of the random-coding error exponent for a MIMO Gaussian block-fading channel is given in [33]. However, since no channel knowledge is assumed at the transmitter in [33], the total power P_T is equally allocated to all transmit antennas with covariance matrix $\mathbf{Q} = (P_T/n_T)\mathbf{I}_{n_T}$. Thus, instead of optimizing r, which is shown inside both logarithm and matrix inversion operation, the authors made the approximation of optimizing $\beta = n_T - rP_T$ for analytical tractability [33, eqs. (7) and (8)]. Since we assumed perfect channel-state information and optimal power allocation, we cannot make similar approximations to derive closed-form expressions of the error exponents. Instead, we numerically calculate the optimal error exponent and analyze the resultant achievable rate. The closed-form expression of the error exponent for the SISO case is provided in [25, eqs. (7.4.33) and (7.4.35)], using similar approximations as in [33]. Since the result in [25] can easily be applied to the single-antenna situation of our case, we omit the expression for simplicity and refer the interested reader to the original literature.

Remark 2: The random-coding error exponent for the MIMO Gaussian deterministic channel case is also discussed in [32], where the parameter r is set to 0. However, as stated in [25], this choice of r is not optimal, since the codewords with small power would dominate the bound on error probability with $r = 0$ in the calculation, these codewords are not included in the ensemble of codes. Furthermore, the Gaussian codewords we choose are optimal in maximizing the error exponent only if the rate approaches the channel capacity [33]. Otherwise, a distribution concentrated on a "thin spherical shell" yields better results [25], [32]. However, we use the Gaussian codewords due to mathematical tractability.

Remark 3: The covariance matrices $\tilde{\mathbf{N}}_1$ and $\tilde{\mathbf{N}}_2$ are, in general, not Gaussian. However, it is known that, for a given noise variance, the capacity of an additive noise channel is minimized if the noise is Gaussian. Thus, by modeling the decoding error as Gaussian, we derive an achievable rate using noisy DPC, which serves as the lower bound on the capacity.

IV. NUMERICAL EXAMPLES

In this section, we provide numerical examples to illustrate the increase in the achievable rate using noisy DPC. We consider both SISO and MIMO cases. In the SISO case, let h_{AB}, h_{CD}, h_{AC}, and h_{AD} denote the complex channel gains between nodes A and B, C and D, A and C, and A and D, respectively. We set $|h_{AB}| = 0.25$, $|h_{CD}| = 1$, and the average power constraints $P_A = P_C = 100$. In the MIMO case, let each node have two antennas, and $P_A = P_C = 10$. Let $\mathbf{H}_{AB}$ be the channel matrix between the primary transmitter and receiver: nodes A and B. The real and imaginary parts of all elements in $\mathbf{H}_{AB}$, $\mathbf{H}_{CD}$, $\mathbf{H}_{AC}$, and $\mathbf{H}_{AD}$ are independently generated according to $\mathcal{N}(0, 1)$. We multiply the factors λ_{AB}, λ_{AC}, and λ_{AD} to $\mathbf{H}_{AB}$, $\mathbf{H}_{AC}$, and $\mathbf{H}_{AD}$, respectively, and change their values to vary the channel qualities. We also set $\lambda_{AB} = 0.25$, $\lambda_{CD} = 1$, and $n = 50$, unless otherwise specified.

Fig. 4(a) and (b) shows the variation of R_{CD} and R'_{CD} with respect to the channel gains between nodes A and D in the SISO and MIMO cases, respectively. In the SISO case, as we
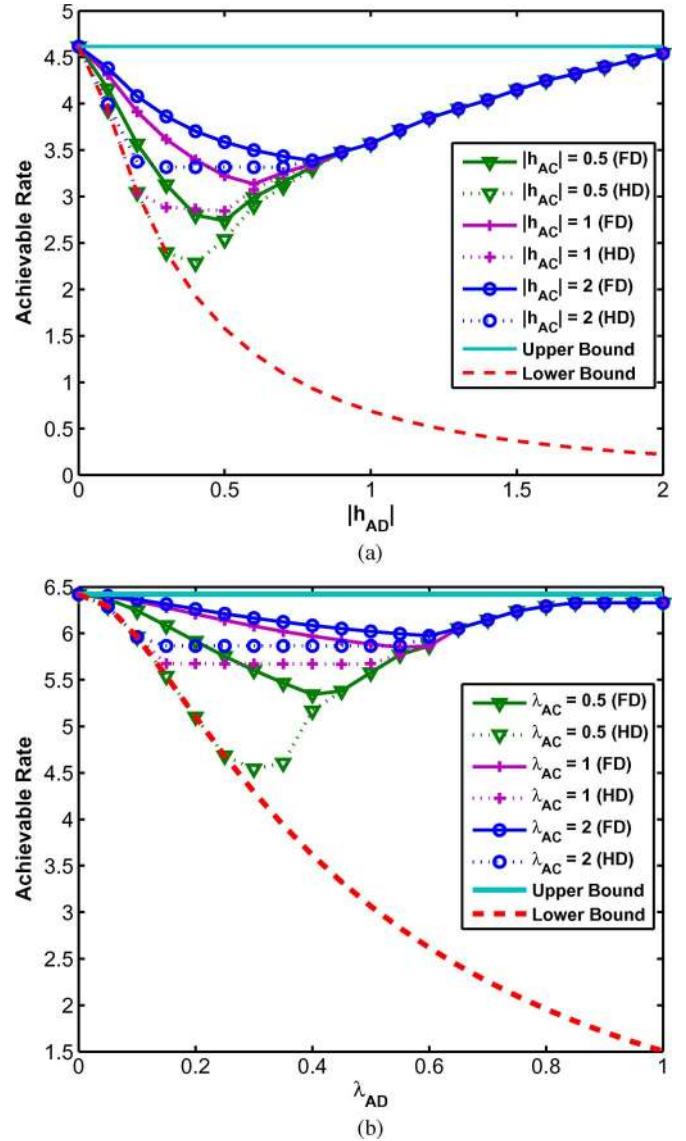

Fig. 4. Variation of achievable rates with (a) $|h_{AD}|$ for different values of $|h_{AC}|$ in the SISO case and (b) λ_{AD} for different values of λ_{AC} in the MIMO case (FD: full-duplex; HD: half-duplex).

can see from Fig. 4(a), the achievable rates with full-duplex transmission are greater than the lower bound when $|h_{AD}| > 0$, since all values of $|h_{AC}|$ are greater than $|h_{AB}|$; hence, scheme 2 (noisy DPC at node C) can be applied to increase the achievable rate in phase 2 (interference cancellation at node D is not possible for small values of $|h_{AD}|$). However, for half-duplex transmission, since node C has to be silent in phase 1 to apply noisy DPC, the fractional penalty on the rate loss cannot always be compensated by the gains due to better decoding. Thus, only when $|h_{AC}|$ is sufficiently large ($|h_{AC}| = 2$), using scheme 2 alone under half-duplex assumption increases the achievable rate. Notice also the nonmonotonic variation of R_{CD} and R'_{CD} with $|h_{AD}|$, which can be explained as follows. The interference of the channel increases with $|h_{AD}|$; when $|h_{AD}|$ is small, the transmitter side interference cancellation (scheme 2) cannot offset the increased interference. Although R_{CD} and R'_{CD} can be greater than $R_{CD,lb}$, they still decrease as $|h_{AD}|$
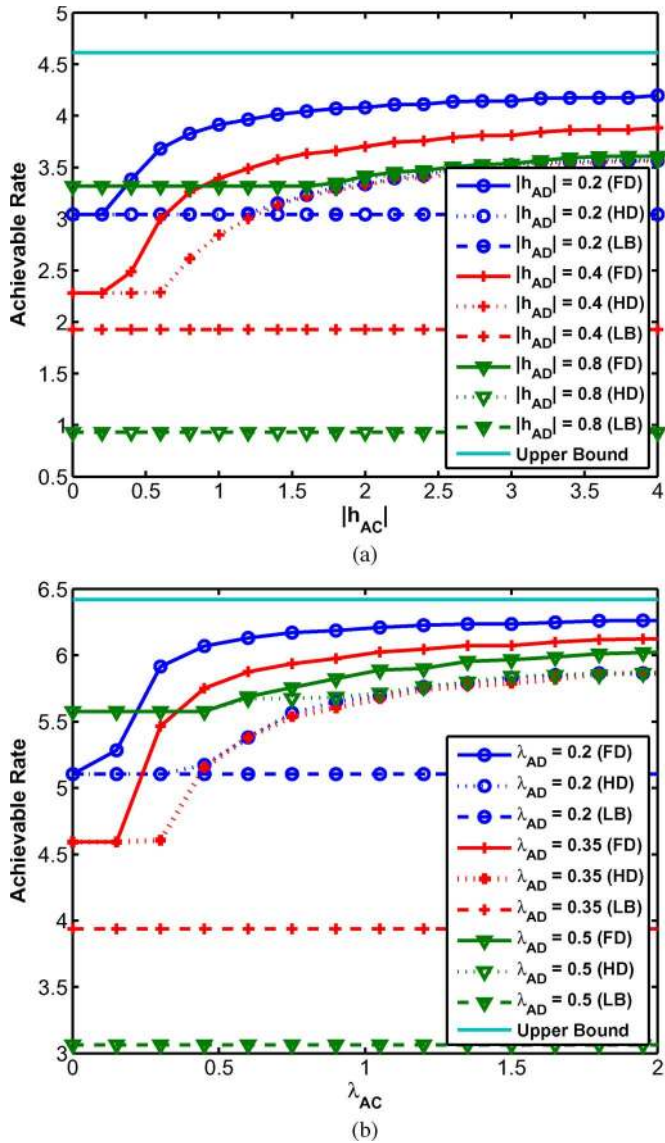
Fig. 5. Variation of achievable rates with (a) $|h_{AC}|$ for different values of $|h_{AD}|$ in the SISO case and (b) λ_{AC} for different values of λ_{AD} in the MIMO case (FD: full-duplex; HD: half-duplex; LB: lower bound).
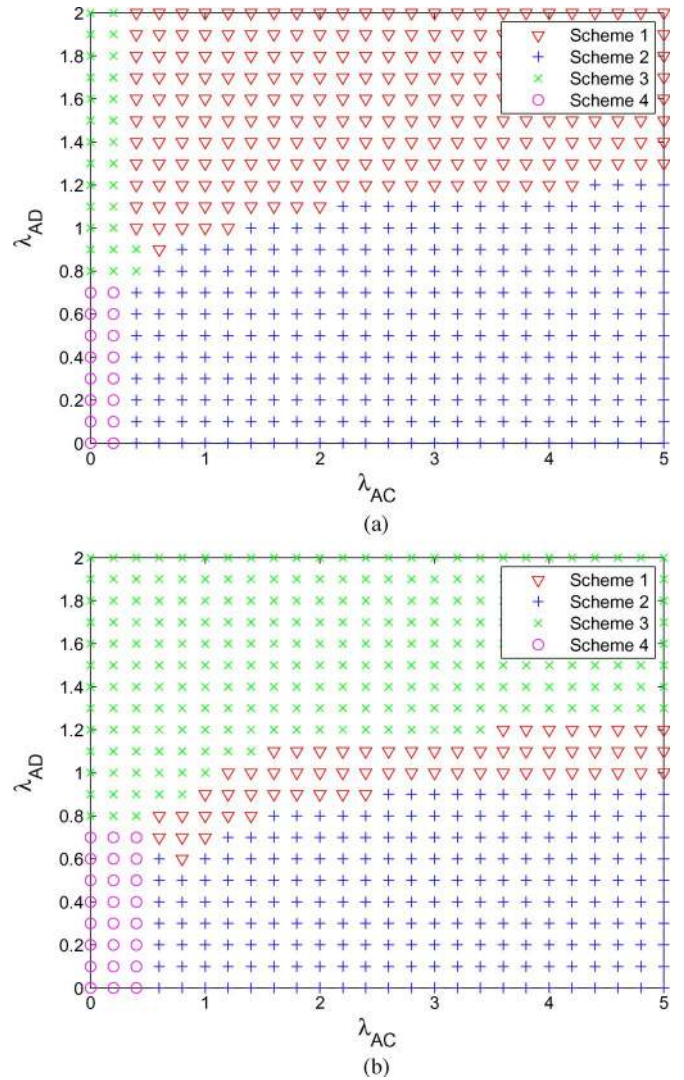

Fig. 6. Scheme that achieves the highest rate with different λ_{AC} and λ_{AD} for (a) full-duplex transmission and (b) half-duplex transmission.

increases. When $|h_{AD}|$ is large, both noisy DPC and receiver side partial interference cancellation (scheme 1) can be applied. Thus, R_{CD} and R'_{CD} start to increase after some thresholds when scheme 1 can be applied. Furthermore, when $|h_{AD}|$ is sufficiently large, node D can decode and subtract node A's codeword by treating node C's signal as interference. Therefore, R_{CD} and R'_{CD} will eventually approach $R_{CD,ub}$. Similar trends are shown for the MIMO case in Fig. 4(b).

Fig. 5(a) and (b) shows the variation of R_{CD} and R'_{CD} with respect to the channel gains between nodes A and C in the SISO and MIMO cases, respectively. For the SISO case, first consider when $|h_{AD}| = 0.2$. Since $|h_{AD}| < |h_{AB}|$, node D cannot reliably decode node A's signal. Now, for small values of $|h_{AC}|$, node C cannot also decode node A's signal, and the achievable rate is then $R_{CD,lb}$. As $|h_{AC}|$ increases, noisy DPC (scheme 2) can be applied, as shown in the figure, and R_{CD} immediately increases when $|h_{AC}| > |h_{AB}|$. However, R'_{CD} increases only

when $|h_{AC}|$ is sufficiently large due to a similar reason we explained in illustrating Fig. 4(a). When $|h_{AD}| = 0.4$, since the interference increases, the achievable rates decrease. However, since node D can partially decode node A's signal (scheme 3), R_{CD} and R'_{CD} are both greater than the lower bound, even when $|h_{AC}|$ is very small. When $|h_{AD}| = 0.8$, node D can cancel out more interference to achieve transmission rates higher than the case when $|h_{AD}| = 0.2$ by using scheme 3 alone. Note, however, that a larger $|h_{AD}|$ value causes more interference at node D, which is reflected in the decrease of $R_{CD,lb}$. Thus, for a given $|h_{AC}|$, the achievable rate can be lower or higher, depending on the value of $|h_{AD}|$. Similar behavior of R_{CD} and R'_{CD} with respect to λ_{AC} is observed for the MIMO case, as shown in Fig. 5(b).

In Fig. 6(a) and (b), the scheme that achieves the highest transmission rate with respect to different λ_{AC} and λ_{AD} are plotted for full- and half-duplex MIMO cases, respectively. In the full-duplex case, it is clear that, when both λ_{AC} and λ_{AD} are small, no scheme can achieve rates larger than $R_{CD,lb}$; hence, scheme 4 is optimal. Scheme 3 achieves the highest rate
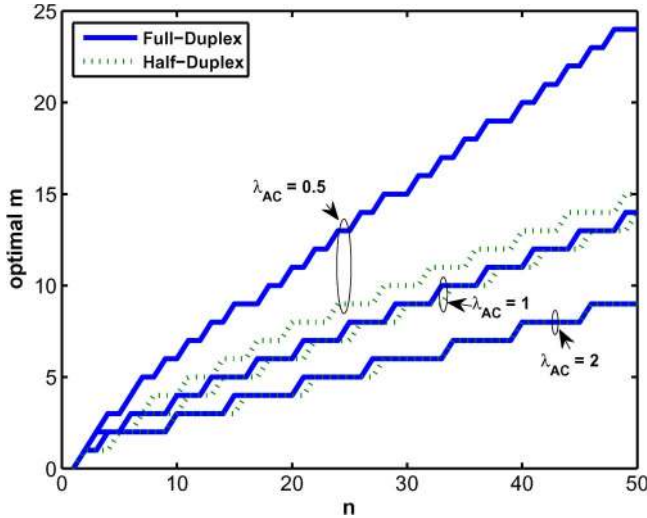
Fig. 7. Optimal m with respect to n for different λ_{AC}.

when λ_{AD} is large, since node D can do partial interference cancellation. When λ_{AC} is large, node C can apply noisy DPC to increase the achievable rate; thus, scheme 2 is the best. When both λ_{AC} and λ_{AD} are large, both nodes C and D can decode node A's signal and do interference cancellation; in this case, scheme 1 achieves the highest transmission rate. Note that there is no clear-cut threshold of λ_{AC} for scheme 1 to be better than scheme 2, since when λ_{AC} is large but λ_{AD} is not large enough, there are fewer values of m and α (such that $R_{\mathrm{AD}}(m, \alpha) \geq nR$) for scheme 1 to achieve a larger rate than the lower bound. However, once λ_{AC} is large enough (such that $R_{\mathrm{AC}}(m) \geq nR$), the achievable rate of scheme 2 can be optimized over all possible power allocations between phases 1 and 2. A similar argument can be used to explain the choices between schemes 1 and 3. In the half-duplex case, when λ_{AD} is small, the value of λ_{AC} for scheme 2 to be better than scheme 4 is larger than the full-duplex case. Since node C cannot transmit in phase 1 in scheme 2, the rate increase in phase 2 by applying noisy DPC cannot compensate for the rate loss in phase 1 if λ_{AC} is large enough. Furthermore, unlike the full-duplex case, scheme 3 is the best, even if both λ_{AC} and λ_{AD} are large, because when λ_{AD} is large, node D can decode node A's signal with a small error probability and cancel out its effect. When scheme 1 is used in the half-duplex case, there is a rate loss in phase 1. Thus, scheme 3 achieves the transmission rate that is large enough to be greater than scheme 1 when λ_{AD} is large. However, when λ_{AC} is large while λ_{AD} is not large enough, noisy DPC can provide extra gain to the achievable rate; thus, scheme 1 is the best within this region of channel gains.

Fig. 7 shows the optimal choice of m with respect to the codeword length n for different values of λ_{AC}. As we can see, the optimal m almost linearly increases with n. However, the optimal m decreases with λ_{AC}, since as the channel quality between nodes A and C increases, less time is needed in phase 1 for node C to decode node A's signal. Furthermore, the optimal m for the half-duplex transmission is smaller than the full-duplex transmission, since in the half-duplex case, there is less time node C can transmit as m increases.

V. CONCLUSION

In this paper, we first considered a MIMO Gaussian channel where noisy side information of the additive interference is known to both the transmitter and the receiver. We calculated the capacity outer bound and proved its achievability. We then applied this result to a practical cognitive ZIC. We constructed new coding schemes using noisy DPC and quantified the resultant achievable rates. We also presented numerical examples that show the significant increase in the achievable rate compared with prior results.

As shown in this paper, the achievable rate is highly dependent on the estimation error. A better lower bound may be achieved by modeling the estimation error differently. Use of non-Gaussian codebooks and its impact on the performance of the cognitive channel should be studied in future work. Furthermore, it would be interesting to exploit the potential applications of noisy DPC on other interference networks.

APPENDIX
DERIVATION FROM (7)–(9)

Let

$$
\mathbf{D}_1 = \begin{bmatrix} \mathbf{N}_y & \mathbf{H}_3\mathbf{P}\mathbf{H}_1^\dagger & \mathbf{H}_3\mathbf{P}\mathbf{H}_2^\dagger \\ \mathbf{H}_1\mathbf{P}\mathbf{H}_3^\dagger & \mathbf{N}_1 & \mathbf{H}_1\mathbf{P}\mathbf{H}_2^\dagger \\ \mathbf{H}_2\mathbf{P}\mathbf{H}_3^\dagger & \mathbf{H}_2\mathbf{P}\mathbf{H}_1^\dagger & \mathbf{N}_2 \end{bmatrix}. \quad (34)
$$

We can then write

$$
H(\mathbf{x}) + H(\mathbf{y}, \mathbf{m}_1, \mathbf{m}_2) = \log(2\pi e)^{2(n_T + n_R)} \left(|\mathbf{P}||\mathbf{D}_1|\right). \quad (35)
$$

Applying the rule for calculating the determination of the block matrix, we have

$$
\begin{aligned}
&H(\mathbf{x}, \mathbf{y}, \mathbf{m}_1, \mathbf{m}_2) \\
&= \log(2\pi e)^{2(n_T + n_R)} \\
&\quad \times \left(|\mathbf{P}| \left|\mathbf{D}_1 - \begin{bmatrix} \mathbf{H}_0\mathbf{P} \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix} \mathbf{P}^{-1} \begin{bmatrix} \mathbf{P}\mathbf{H}_0 & \mathbf{0} & \mathbf{0} \end{bmatrix}\right|\right) \\
&= \log(2\pi e)^{2(n_T + n_R)} \left(|\mathbf{P}||\mathbf{D}_2|\right) \quad (36)
\end{aligned}
$$

where

$$
\mathbf{D}_2 = \begin{bmatrix} \mathbf{I}_{n_R} + \mathbf{H}_3\mathbf{Q}\mathbf{H}_3^\dagger & \mathbf{H}_3\mathbf{P}\mathbf{H}_1^\dagger & \mathbf{H}_3\mathbf{P}\mathbf{H}_2^\dagger \\ \mathbf{H}_1\mathbf{P}\mathbf{H}_3^\dagger & \mathbf{I}_{n_T} + \mathbf{H}_1\mathbf{Q}\mathbf{H}_1^\dagger & \mathbf{H}_1\mathbf{P}\mathbf{H}_2^\dagger \\ \mathbf{H}_2\mathbf{P}\mathbf{H}_3^\dagger & \mathbf{H}_2\mathbf{P}\mathbf{H}_1^\dagger & \mathbf{I}_{n_T} + \mathbf{H}_2\mathbf{Q}\mathbf{H}_2^\dagger \end{bmatrix}. \quad (37)
$$

Thus, we can express the channel capacity as $I(\mathbf{x}; \mathbf{y}|\mathbf{m}_1, \mathbf{m}_2) = H(\mathbf{x}) + H(\mathbf{y}, \mathbf{m}_1, \mathbf{m}_2) - H(\mathbf{x}, \mathbf{y}, \mathbf{m}_1, \mathbf{m}_2) = \log(|\mathbf{D}_1|/|\mathbf{D}_2|)$. Again, applying the determination of the block matrix, we have

$$
I(\mathbf{xy}|\mathbf{m}_1, \mathbf{m}_2) = \log\left(\frac{\left|\mathbf{I}_{n_R} + \mathbf{D}_3 + \mathbf{H}_0\mathbf{P}\mathbf{H}_0^\dagger\right|}{|\mathbf{I}_{n_R} + \mathbf{D}_3|}\right) \quad (38)
$$

in which

$$
\mathbf{D}_3 = \mathbf{H}_3\mathbf{Q}\mathbf{H}_2^\dagger - \begin{bmatrix} \mathbf{H}_3\mathbf{Q}\mathbf{H}_1^\dagger & \mathbf{H}_3\mathbf{Q}\mathbf{H}_2^\dagger \end{bmatrix} \mathbf{D}_4^{-1} \begin{bmatrix} \mathbf{H}_1\mathbf{Q}\mathbf{H}_3^\dagger \\ \mathbf{H}_2\mathbf{Q}\mathbf{H}_3^\dagger \end{bmatrix} \quad (39)
$$

where

$$\mathbf{D}_4 = \begin{bmatrix} \mathbf{I}_{n_T} + \mathbf{H}_1 \mathbf{Q} \mathbf{H}_1^\dagger & \Theta \\ \Theta^\dagger & \mathbf{I}_{n_T} + \mathbf{H}_2 \mathbf{Q} \mathbf{H}_2^\dagger \end{bmatrix} \quad (40)$$

where $\Theta = \mathbf{H}_1 \mathbf{P} \mathbf{H}_2^\dagger$. We now evaluate $\mathbf{D}_4^{-1}$. Applying block matrix inversion, we have

$$\mathbf{D}_4^{-1} = \begin{bmatrix} \mathbf{D}_6^{-1} & -\mathbf{D}_6^{-1} \Theta \mathbf{D}_5^{-1} \\ -\mathbf{D}_5^{-1} \Theta^\dagger \mathbf{D}_6^{-1} & \mathbf{D}_5^{-1} \left(\mathbf{I}_{n_R} + \Theta^\dagger \mathbf{D}_6^{-1} \Theta \mathbf{D}_5^{-1} \right) \end{bmatrix} \quad (41)$$

where $\mathbf{D}_5 = \mathbf{I}_{n_R} + \mathbf{H}_2 \mathbf{Q} \mathbf{H}_2^\dagger$, and

$$\mathbf{D}_6 = \mathbf{I}_{n_T} + \mathbf{H}_1 \left[\mathbf{Q} - \mathbf{Q} \mathbf{H}_2 \left(\mathbf{I}_{n_R} + \mathbf{H}_2 \mathbf{Q} \mathbf{H}_2^\dagger \right)^{-1} \mathbf{H}_2 \mathbf{Q} \right] \mathbf{H}_1^\dagger. \quad (42)$$

Applying the matrix inversion lemma reversely to (42), $\mathbf{D}_6$ simplifies to

$$\mathbf{D}_6 = \mathbf{I}_{n_T} + \mathbf{H}_1 \mathbf{Q} \left(\mathbf{I}_{n_R} + \mathbf{H}_2^\dagger \mathbf{H}_2 \mathbf{Q} \right) \mathbf{H}_1^\dagger. \quad (43)$$

Now, we substitute $\mathbf{D}_4^{-1}$ in (39) with the expression given by (41), which yields

$$\mathbf{D}_3 = \mathbf{H}_3 \mathbf{Q} \Big[\mathbf{I}_{n_I} - \mathbf{D}_7 \mathbf{H}_1^\dagger \mathbf{D}_6^{-1} \mathbf{H}_1 \mathbf{Q} \mathbf{D}_7$$

$$- \mathbf{H}_2^\dagger \left(\mathbf{I}_{n_I} + \mathbf{H}_2 \mathbf{Q} \mathbf{H}_2^\dagger \right)^{-1} \mathbf{H}_2 \mathbf{Q} \Big] \mathbf{H}_3^\dagger \quad (44)$$

where

$$\mathbf{D}_7 = \mathbf{I}_{n_I} - \mathbf{H}_2^\dagger \left(\mathbf{I}_{n_R} + \mathbf{H}_2 \mathbf{Q} \mathbf{H}_2^\dagger \right)^{-1} \mathbf{H}_2 \mathbf{Q}. \quad (45)$$

Applying the matrix inversion lemma reversely to (45), $\mathbf{D}_7$ simplifies to

$$\mathbf{D}_7 = \left(\mathbf{I}_{n_I} + \mathbf{H}_2^\dagger \mathbf{H}_2 \mathbf{Q} \right)^{-1}. \quad (46)$$

Substituting $\mathbf{D}_7$ in (44) with the expression given by (46) and again applying the matrix inversion lemma reversely, $\mathbf{D}_3$ simplifies to

$$\mathbf{D}_3 = \left(\mathbf{I}_{n_I} + \left(\mathbf{H}_1^\dagger \mathbf{H}_1 + \mathbf{H}_2^\dagger \mathbf{H}_2 \right) \mathbf{Q} \right)^{-1}. \quad (47)$$

Substituting $\mathbf{D}_3$ in (38) with the expression given by (47) yields the capacity outer bound in (9).

REFERENCES

[1] M. Costa, "On the Gaussian interference channel," *IEEE Trans. Inf. Theory*, vol. IT-31, no. 5, pp. 607–615, Sep. 1985.

[2] L. Zhou and W. Yu, "Gaussian Z-interference channel with a relay link: Achievable rate region and asymptotic sum capacity," in *Proc. Int. Symp. Inf. Theory Appl.*, Auckland, New Zealand, Dec. 2008.

[3] L. Zhou and W. Yu, Gaussian Z-Interference Channel With a Relay Link: Type II Channel and Sum Capacity Bound, preprint. [Online]. Available: http://ita.ucsd.edu/workshop/09/files/paper/paper 422.pdf

[4] N. Liu and A. Goldsmith, Capacity Regions and Bounds for a Class of Z-Interference Channels, Dec. 2008. [Online]. Available: http://arxiv.org/abs/0808.0876

[5] H. Chong, M. Motani, and H. Garg, Capacity theorems for the 'Z channel,'" *IEEE Trans. Inf. Theory*, vol. 53, no. 4, pp. 1348–1365, Apr. 2007.

[6] N. Liu and S. Ulukus, "On the capacity region of the Gaussian Z-channel," in *Proc. IEEE GlobeCom*, Nov./Dec. 2004, vol. 1, pp. 415–419.

[7] F. C. Commission, "Spectrum policy task force," Rep. ET Docket, Nov. 2002.

[8] J. Mitola and G. Maguire, "Cognitive radio: Making software radios more personal," *IEEE Pers. Commun.*, vol. 6, no. 4, pp. 13–18, Aug. 1999.

[9] G. Staple and K. Werbach, "The end of spectrum scarcity," *IEEE Spectr.*, vol. 41, no. 3, pp. 48–52, Mar. 2004.

[10] S. Haykin, "Cognitive radio: Brain-empowered wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 2, pp. 201–220, Feb. 2005.

[11] S. Srinivasa and S. Jafar, "The throughput potential of cognitive radio: A theoretical perspective," *IEEE Commun. Mag.*, vol. 45, no. 5, pp. 73–79, 2007.

[12] M. Costa, "Writing on dirty paper," *IEEE Trans. Inf. Theory*, vol. 29, no. 3, pp. 439–441, May 1983.

[13] A. Cohen and A. Lapidoth, "Generalized writing on dirty paper," in *Proc. IEEE Int. Symp. Inf. Theory*, Jun./Jul. 2002, p. 227.

[14] A. Jovicic and P. Viswanath, "Cognitive radio: An information-theoretic perspective," in *Proc. Int. Symp. Inf. Theory*, Jul. 2006, pp. 2413–2417.

[15] F. Awan and M. Hanif, "A unified view of information-theoretic aspects of cognitive radio," in *Proc. 5th Int. Conf. Inf. Technol.*, Apr. 2008, pp. 327–331.

[16] S. Srinivasa, S. Jafar, and N. Jindal, "On the capacity of the cognitive tracking channel," in *Proc. Int. Symp. Inf. Theory*, Jul. 2006, pp. 2077–2081.

[17] S. Jafar and S. Srinivasa, "Capacity limits of cognitive radio with distributed and dynamic spectral activity," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 3, pp. 529–537, Apr. 2007.

[18] S. Sridharan and S. Vishwanath, "On the capacity of a class of MIMO cognitive radios," *IEEE J. Sel. Topics Signal Process.*, vol. 2, no. 1, pp. 103–117, Feb. 2008.

[19] C. Vaze and M. Varanasi, "Dirty paper coding for the MIMO cognitive radio channel with imperfect CSIT," in *Proc. Int. Symp. Inf. Theory*, Jun./Jul. 2009, pp. 2532–2536.

[20] P. Mitran, N. Devroye, and V. Tarokh, "On compound channels with side information at the transmitter," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1745–1755, Apr. 2006.

[21] N. Liu, I. Maric, A. Goldsmith, and S. Shamai, "Bounds and capacity results for the cognitive Z-interference channel," in *Proc. Int. Symp. Inf. Theory*, Jun./Jul. 2009, pp. 2422–2426.

[22] N. Devroye, P. Mitran, and V. Tarokh, "Achievable rates in cognitive radio channels," *IEEE Trans. Inf. Theory*, vol. 52, no. 2, pp. 1813–1827, May 2006.

[23] Y. Peng and D. Rajan, "Achievable rates of Gaussian cognitive Z-interference channel with partial side information," in *Proc. IEEE GlobeCom*, Dec. 2009.

[24] T. Cover and M. Chiang, "Duality between channel capacity and rate distortion with two-sided state information," *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1629–1638, Jun. 2002.

[25] R. Gallager, *Information Theory and Reliable Communication*. Hoboken, NJ: Wiley, 1968.

[26] W. Yu, A. Sutivong, D. Julian, T. M. Cover, and M. Chiang, "Writing on colored paper," in *Proc. IEEE Int. Symp. Inf. Theory*, Jun. 2001, p. 302.

[27] U. Erez, S. Shamai, and R. Zamir, "Capacity and lattice strategies for cancelling known interference," *IEEE Trans. Inf. Theory*, vol. 51, no. 11, pp. 3820–3833, Nov. 2005.

[28] G. Caire and S. Shamai, "On achievable rates in a multi-antenna Gaussian broadcast channel," in *Proc. IEEE Int. Symp. Inf. Theory*, Jun. 2001, p. 147.

[29] W. Yu and J. M. Cioffi, "Trellis precoding for the broadcast channel," in *Proc. IEEE GlobeCom*, Nov. 2001, vol. 2, pp. 1344–1348.

[30] H. Weingarten, Y. Steinberg, and S. Shamai, "The capacity region of the Gaussian multiple-input multiple-output broadcast channel," *IEEE Trans. Inf. Theory*, vol. 52, no. 9, pp. 3929–3964, Sep. 2006.

[31] M. Mazzotti and Z. Xiong, "Effects of noisy side information at the decoder in dirty-paper and dirty-tape coding," in *Proc. Inf. Theory Workshop*, Oct. 2006, pp. 390–394.

[32] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *Eur. Trans. Telecommun.*, vol. 10, no. 6, pp. 585–595, Nov./Dec. 1999.

[33] M. Sarkar, H. Shin, and M. Win, "Random coding exponent for MIMO channels," in *Proc. IEEE Veh. Technol. Conf.*, May 2008, pp. 559–563.

[34] H. Shin and M. Win, Gallager's exponent for MIMO channels: A reliability-rate trade-off, Jul. 2006. [Online]. Available: http://arxiv.org/abs/cs/0607095

Yong Peng (S'09) received the B.Eng. degree in communications engineering from the University of Electronic Science and Technology of China, Chengdu, China, in 2003 and the M.S. degree in electrical engineering from New Jersey Institute of Technology, Newark, in 2005. He is currently working toward the Ph.D. degree in electrical engineering with Southern Methodist University (SMU), Dallas, TX.

Since 2006, he has been a Research Assistant with the Wireless Networking Laboratory, SMU. His research interests include cognitive radio, multiple-antenna systems, and multiuser information theory.

Dinesh Rajan (S'99–M'02–SM'07) received the B.Tech. degree in electrical engineering from the Indian Institute of Technology, Madras, India, in 1997 and the M.S. and Ph.D. degrees in electrical and computer engineering from Rice University, Houston, TX, in 1999 and 2002, respectively.

Since 2002, he has been with the Department of Electrical Engineering, Southern Methodist University, Dallas, TX, where he is currently an Associate Professor. His current research interests include communications theory, wireless networks, information theory, and computational imaging.

Dr. Rajan was a recipient of the National Science Foundation Faculty Early Career Development Award in 2006.