

# Sum Capacity of MIMO Interference Channels in the Low Interference Regime

V. Sreerkanth Annapureddy, *Student Member, IEEE*, and Venugopal V. Veeravalli, *Fellow, IEEE*

**Abstract**—Using Gaussian inputs and treating interference as noise at the receivers has recently been shown to be sum capacity achieving for the two-user single-input single-output (SISO) Gaussian interference channel in a low interference regime, where the interference levels are below certain thresholds. In this paper, such a low interference regime is characterized for multiple-input multiple-output (MIMO) Gaussian interference channels. Conditions are provided on the direct and cross channel gain matrices under which using Gaussian inputs and treating interference as noise at the receivers is sum capacity achieving. For the special cases of the symmetric multiple-input single-output (MISO) and single-input multiple-output (SIMO) Gaussian interference channels, more explicit expressions for the low interference regime are derived. In particular, the threshold on the interference levels that characterize low interference regime is related to the input SNR and the angle between the direct and cross channel gain vectors. It is shown that the low interference regime can be quite significant for MIMO interference channels, with the low interference threshold being at least as large as the sine of the angle between the direct and cross channel gain vectors for the MISO and SIMO cases.

**Index Terms**—Capacity, genie, interference channel, multiple-input multiple-output (MIMO), outer bound, treating interference as noise.

## I. INTRODUCTION

**B**REAKING the interference barrier is an important step in achieving higher throughput in wireless networks. Towards this end, there has been a renewed interest in information-theoretic studies of interference networks in recent years. A canonical problem for such an information-theoretic analysis is the two-user single-input single-output (SISO) Gaussian interference channel. This problem was first studied more than thirty years ago, and the capacity region was determined in the

strong (and very strong) interference regimes, where the interference-to-noise ratio (INR) is larger than the signal-to-noise ratio (SNR) at each receiver [1]–[3].

Establishing the capacity region in the more commonly encountered weak interference regime, where the INR is smaller than the SNR, is still mostly an open problem. Nevertheless, the best known inner bound on the capacity region proposed by Han and Kobayashi [2] was recently shown to be within one bit of the capacity region [4]. In the Han-Kobayashi scheme, the users split their messages into private and common messages, and each user jointly decodes its own message and the common message of the interfering user. This is in general a sophisticated scheme, requiring multi-user encoders and decoders and coordination between the users. While such techniques are promising and being implemented in advanced systems, the traditional way to combat the interference is to treat interference as noise when the interference is weak, and to orthogonalize the users in time or frequency when the interference is moderate. Interestingly, treating interference as noise with Gaussian inputs was shown to be capacity achieving in a low (very weak) interference regime [5]–[7], which is the counterpart of the very strong interference regime [1]. Our goal in this paper is to characterize such a low interference regime for multiple-input multiple-output (MIMO) Gaussian interference channels.

The channel of interest is described in Fig. 1. There has been some previous work on studying inner and outer bounds on the capacity region of this channel [8]–[13]. In particular, Teletar and Tse [10] showed that an appropriate extension of the SISO Han-Kobayashi inner bound is within a finite number (equal to the number of receive antennas) of bits of the capacity region. In [13], Shang *et al.* studied the low interference regime of the parallel interference channel. The low interference regime of the general MIMO interference channel was analyzed by Shang *et al.* in [11], [12], where they showed that using Gaussian inputs and treating interference as noise is sum capacity achieving if the channel matrices satisfy a low interference regime condition for all input covariance matrices satisfying certain average transmit power constraints. For the special case of multiple-input single-output (MISO) interference channel, this result is not useful because the required low interference regime condition can be satisfied only if the input covariance matrices are unit-rank, but the result in [11] requires the condition to be satisfied for all input covariance matrices satisfying the transmit power constraints. Bernd *et al.* studied the low interference regime of MISO interference channel under the assumption that single-mode beamforming is applied at the transmitters [14].

For the general MIMO interference channel, we show that it is sufficient for the low interference regime condition to be

Manuscript received September 10, 2009; revised October 09, 2010; accepted December 31, 2010. Date of current version April 20, 2011. This work was supported in part by the NSF awards CNS-081670 and CCF-0904619, through the University of Illinois, and in part by grants from Intel and Texas Instruments. This paper was presented in part at the Asilomar Conference on Signals, Systems, and Computers, Pacific Grove, CA, November 2008, and the Information Theory and Applications (ITA) Workshop, University of California, San Diego, January 2009.

The authors are with the Coordinated Science Laboratory, Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign, Urbana, IL 61801 USA (e-mail: sreerkanthav@gmail.com, vvv@illinois.edu).

Communicated by S. Vishwanath, Associate Editor for the special issue on "Interference Networks".

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TIT.2011.2119250

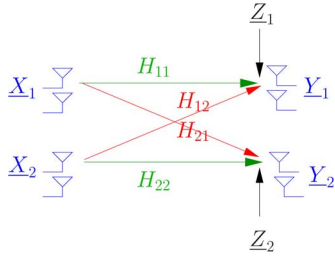


Fig. 1. Two-user MIMO Gaussian interference channel.

satisfied only for the optimal input covariance matrices (those that maximize the achievable sum rate assuming Gaussian inputs and treating interference as noise) if they are full rank. For the MISO interference channel, this result cannot be exploited to establish a low interference regime because the optimal input covariance matrices are unit-rank, i.e., transmit beamforming maximizes the achievable sum-rate for the MISO interference channel [15]. We analyze the special case of symmetric MISO interference channel explicitly and derive a simple expression that characterizes the low interference regime. In particular, we relate the threshold on the interference levels, that characterize low interference regime, to the input SNR and the angle between the direct and cross channel gain vectors. We show that the same low interference regime holds for the dual single-input multiple-output (SIMO) interference channel also.

We establish that the low interference regime can be quite significant for MIMO interference channels, with the low interference threshold being at least as large as the sine of the angle between the direct and cross channel gain vectors for the MISO and SIMO cases.

#### A. Notation and Organization

We use underlined letters to denote vectors, and superscripts to denote sequences in time. For example, we use  $X$  to denote a scalar,  $\underline{X}$  to denote a vector, and  $X^n$  and  $\underline{X}^n$  to denote sequences of length  $n$  of scalars and vectors, respectively. We use  $\text{Cov}(X)$  to denote the variance of a random variable  $X$ , and  $\text{Cov}(X|Y)$  denote the minimum mean square error in estimating the random variable  $X$  from the random variable  $Y$ , with similar notation for random vectors. We use  $\mathcal{N}(\mu, \sigma^2)$  to denote the Gaussian distribution with mean  $\mu$  and variance  $\sigma^2$ , and  $\mathcal{N}(\underline{\mu}, \underline{\Sigma})$  to denote the Gaussian vector distribution with mean  $\underline{\mu}$  and covariance matrix  $\underline{\Sigma}$ . We use  $h(\cdot)$  to denote the differential entropy of a continuous random variable or vector and  $I(\cdot; \cdot)$  to denote the mutual information. For real symmetric matrices  $A$  and  $B$ , we use  $A \succeq B$  to denote that  $A - B$  is positive semidefinite.

The rest of the paper is organized as follows. In Section II, we introduce the model for MIMO interference channel that we study. In Section III, we present the results on the general MIMO interference channel. In Sections IV and V, we characterize the low interference regime for MISO and SIMO interference channels, respectively. In Section VI, we discuss some properties of the low interference regime. In Section VII, we provide some concluding remarks.

## II. PROBLEM STATEMENT AND SUMMARY OF CONTRIBUTIONS

Consider the two-user MIMO Gaussian interference channel (Fig. 1)

$$\begin{aligned} \underline{Y}_1 &= H_{11}\underline{X}_1 + H_{12}\underline{X}_2 + \underline{Z}_1 \\ \underline{Y}_2 &= H_{21}\underline{X}_1 + H_{22}\underline{X}_2 + \underline{Z}_2 \end{aligned} \quad (1)$$

with inputs  $\underline{X}_1, \underline{X}_2$  and the corresponding outputs  $\underline{Y}_1, \underline{Y}_2$ . The channel gain matrices  $\{H_{ij}\}$  are arbitrary but fixed and real. The receiver noise terms  $\underline{Z}_1 \sim \mathcal{N}(0, I)$  and  $\underline{Z}_2 \sim \mathcal{N}(0, I)$ , are assumed to be independent and identically distributed (i.i.d.) in time. The average transmit power constraints on users 1 and 2 are  $P_1$  and  $P_2$ , respectively. Using  $\underline{X}_{it}$  to denote the channel input of the user  $i$  at time  $t$ , the average transmit power constraints can be written as

$$\frac{1}{n} \sum_{t=1}^n \mathbb{E} [\underline{X}_{it}\underline{X}_{it}^T] \in Q_i$$

where

$$Q_i = \{Q_i : Q_i \succeq 0, \text{tr}(Q_i) \leq P_i\}. \quad (2)$$

We are interested in characterizing the low interference regime where using Gaussian inputs and treating interference as noise at the receivers achieves the sum capacity.

#### A. Achievable Sum-Rate

Let  $R_{\text{TIN}}^{\text{IC}}(Q_1, Q_2)$  denote the sum-rate achieved by using Gaussian inputs with input covariance matrix  $Q_i$  at transmitter  $i$  and treating interference as noise at the receivers

$$R_{\text{TIN}}^{\text{IC}}(Q_1, Q_2) = \sum_{i=1}^2 I(\underline{X}_{iG}; \underline{Y}_{iG})$$

where the subscript  $G$  on the inputs and outputs is used to indicate that Gaussian inputs are used. The input covariance matrices used should be obvious from the context throughout the paper.

The following result holds since any achievable sum rate is a lower bound on the sum capacity.

*Lemma 1:* The sum capacity  $C_{\text{sum}}^{\text{IC}}$  of the MIMO interference channel (1) is lower bounded by

$$C_{\text{sum}}^{\text{IC}} \geq R_{\text{TIN}}^{\text{IC}}(Q_1^*, Q_2^*) \quad (3)$$

where  $(Q_1^*, Q_2^*)$  are optimal covariance matrices

$$(Q_1^*, Q_2^*) \in \arg \max_{Q_i \in Q_i} R_{\text{TIN}}^{\text{IC}}(Q_1, Q_2). \quad (4)$$

In the rest of the paper, we investigate when the inequality in Lemma 1 is achieved with equality, i.e., when treating interference as noise achieves the sum capacity. The optimization over the input covariance matrices is the new feature that makes the problem more difficult than its SISO counterpart, since little is known about the structure of the optimal covariance matrices  $(Q_1^*, Q_2^*)$  for general MIMO interference channels. For MISO interference channels, it was shown in [15] that the optimal

covariance matrices are unit rank; this fact is exploited in Section IV to derive explicit conditions for the low interference regime.

### B. Genie-Aided Outer Bound

The outer bound is based on a genie giving side information to the receivers. Let  $\underline{S}_i$  denote the side information given to the receiver  $i$

$$\begin{aligned}\underline{S}_1 &= H_{21}\underline{X}_1 + \underline{W}_1 \\ \underline{S}_2 &= H_{12}\underline{X}_2 + \underline{W}_2\end{aligned}\quad (5)$$

Let  $R_{\text{TIN}}^{\text{GA-IC}}(Q_1, Q_2, \Psi)$  denote the sum rate achievable by treating interference as noise for the genie-aided interference channel using Gaussian inputs with input covariance matrix  $Q_i$  at transmitter  $i$

$$R_{\text{TIN}}^{\text{GA-IC}}(Q_1, Q_2, \Psi) = \sum_{i=1}^2 I(\underline{X}_{iG}; \underline{Y}_{iG}, \underline{S}_{iG}).$$

where  $\Psi$  is used to denote the genie parameters. Let  $C_{\text{sum}}^{\text{GA-IC}}(\Psi)$  denote the sum capacity of the genie-aided interference channel. The steps required in the outer bound are as follows:

$$\begin{aligned}C_{\text{sum}}^{\text{IC}} &\stackrel{(a)}{\leq} C_{\text{sum}}^{\text{GA-IC}}(\Psi) \\ &\stackrel{(b)}{=} \max_{Q_i \in \tilde{Q}_i} R_{\text{TIN}}^{\text{GA-IC}}(Q_1, Q_2, \Psi) \\ &\stackrel{(c)}{=} R_{\text{TIN}}^{\text{GA-IC}}(Q_1^*, Q_2^*, \Psi) \\ &\stackrel{(d)}{=} R_{\text{TIN}}^{\text{IC}}(Q_1^*, Q_2^*) \\ &\stackrel{(e)}{\leq} C_{\text{sum}}^{\text{IC}}.\end{aligned}\quad (6)$$

Step (a) is obvious since providing the receivers with extra information can only improve the sum capacity. Step (e) follows from Lemma 1. Steps (b), (c), and (d) are the nontrivial steps:

- **Useful genie:** The genie is chosen to satisfy step (b), i.e., treating interference as noise, and using Gaussian inputs, achieve the sum capacity of the genie-aided interference channel. We say that such a genie is *useful*.
- **Step (c):** The genie is chosen so that the optimal covariance matrices  $(Q_1^*, Q_2^*)$  maximizing the achievable sum-rate of the original interference channel (4), also maximize the achievable sum-rate of the genie-aided interference channel.
- **Smart genie:** The genie is chosen to satisfy step (d), i.e., it does not give away any extra information when the coding strategy is fixed to using Gaussian inputs with covariance matrices  $(Q_1^*, Q_2^*)$  and treating interference as noise. We say that a genie is *smart* with respect to (w.r.t.) the covariance matrices  $(Q_1, Q_2)$  if  $R_{\text{TIN}}^{\text{GA-IC}}(Q_1, Q_2, \Psi) = R_{\text{TIN}}^{\text{IC}}(Q_1, Q_2)$ . So, step (d) says that the genie is chosen to be smart w.r.t.  $(Q_1^*, Q_2^*)$ .

These steps were first introduced in [5]–[7] for the SISO interference channel, where a genie is constructed satisfying all the steps in (6) when the channel parameters satisfy the following condition for low interference regime:

$$\left| \frac{h_{12}}{h_{22}} (1 + h_{21}^2 P_1) \right| + \left| \frac{h_{21}}{h_{11}} (1 + h_{12}^2 P_2) \right| \leq 1$$

In this paper, we intend to generalize this outer bound to MIMO interference channels. The key step, new to the MIMO interference channels, is the optimization step (c), which is implicitly taken care of in the case of SISO interference channel. We now summarize the known results on this problem [11], [12]:

- **Covariance constraint:** In [12], Shang *et al.* studied the MIMO interference channel with the average transmit covariance constraint

$$\mathcal{Q}_i = \{Q_i : Q_i \preceq \tilde{Q}_i\} \quad (7)$$

instead of the average power constraint given in (2). With the outer bound techniques used in [5]–[7], [11], [12] and this paper, when the genie is useful, the sum-rate  $R_{\text{TIN}}^{\text{GA-IC}}(Q_1, Q_2, \Psi)$  is concave in  $(Q_1, Q_2)$ , and satisfies the following property:

$$\begin{aligned}R_{\text{TIN}}^{\text{GA-IC}}(Q_1, Q_2, \Psi) &\leq R_{\text{TIN}}^{\text{GA-IC}}(\tilde{Q}_1, \tilde{Q}_2, \Psi) \\ \forall Q_1 \preceq \tilde{Q}_1, Q_2 \preceq \tilde{Q}_2.\end{aligned}\quad (8)$$

With covariance constraint (7), this property (8) automatically takes care of the optimization step (c), and thus the problem simplifies to finding conditions under which a genie exists that is useful and smart w.r.t.  $(\tilde{Q}_1, \tilde{Q}_2)$  (See Theorem 6 in [12]).

- **Power constraint:** In [11], Shang *et al.* applied the above insight from the covariance constraint problem to the power constraint problem leading to the following observation (See Theorem 1 of [11]). If for every  $(\tilde{Q}_1, \tilde{Q}_2)$  satisfying the power constraint  $\text{tr}(\tilde{Q}_i) \leq P_i$ , there exists a genie that is both useful and smart w.r.t.  $(\tilde{Q}_1, \tilde{Q}_2)$ , then using Gaussian inputs and treating interference as noise achieves the sum capacity.

However, the proof outline in (6) only requires the existence of a genie that is useful and smart w.r.t.  $(Q_1^*, Q_2^*)$ . Asking for the existence of such a genie for every  $(\tilde{Q}_1, \tilde{Q}_2)$  satisfying  $\text{tr}(\tilde{Q}_i) \leq P_i$  can be too restrictive in general. For example, in the case of MISO interference channel, a genie can be smart w.r.t.  $(\tilde{Q}_1, \tilde{Q}_2)$  only if they are unit rank. Hence, the condition in [11] cannot be satisfied by a MISO interference channel no matter how the channel parameters are chosen. In this paper, we strengthen Theorem 1 of [11] in some special cases. In particular, we make the following contributions:

- **MIMO interference channel:** In Section III, we show that if the optimal covariance matrices  $(Q_1^*, Q_2^*)$  are full rank, it is just sufficient to check for the existence of a genie that is useful and smart w.r.t.  $(Q_1^*, Q_2^*)$ . To better understand why this is so, consider the optimization problem

$$\max_{a \leq x \leq b} f(x).$$

Suppose  $f(x)$  is not concave, and has a local maximum point at  $x^* \in (a, b)$ . This implies  $f'(x^*) = 0$ . Now, if there exists a concave function  $g(x)$  such that  $g(x^*) = f(x^*)$ ,  $g(x) \geq f(x), \forall x \in [a, b]$ , then we can show that  $x^*$  achieves global maximum for both  $f(\cdot)$  and  $g(\cdot)$ . First, observe that  $x^* \in (a, b)$ , and

$$x^* = \arg \min_{a \leq x \leq b} (g - f)(x) \quad (9)$$

which implies that  $(g - f)'(x^*) = 0$ . This, combined with  $f'(x^*) = 0$ , implies that  $g'(x^*) = 0$ . Since  $g(\cdot)$  is concave, it follows that  $x^*$  achieves global maximum of  $g(\cdot)$ , and

$$f(x) \leq g(x) \leq g(x^*) = f(x^*), \forall a \leq x \leq b.$$

By replacing  $f$  and  $g$  with  $R_{\text{TIN}}^{\text{IC}}$  and  $R_{\text{TIN}}^{\text{GA-IC}}$ , respectively, we use the same idea to show that the step (c) in (6) is automatically satisfied if the steps (b) and (d) are satisfied and  $(Q_1^*, Q_2^*)$  are full rank. To understand why the full rank condition is necessary, observe that the argument in the above example fails if  $x^* = a$  or  $b$ , because then (9) does not necessary imply that  $g'(x^*) - f'(x^*) = 0$ .

- **Symmetric MISO interference channel:** For the MISO interference channel, the optimal covariance matrices  $(Q_1^*, Q_2^*)$  are unit rank. So, just the existence of a genie that is useful and smart w.r.t.  $(Q_1^*, Q_2^*)$  does not guarantee the validity of the optimization step (c). In Section IV, we consider the symmetric MISO interference channel and show that if the channel parameters satisfy an additional condition, beyond what is required for the existence of useful and smart genie, then step (c) holds. We also derive an explicit threshold condition for the low interference regime.
- **Symmetric SIMO interference channel:** In the case of the SIMO interference channel, we have only one transmit antenna, and the optimization step (c) does not even come into the play. Similar to the SISO counterpart, the problem is to only find the conditions under which a genie exists that is both useful and smart. In Section V, we simplify these conditions in the symmetric case, and show that they result in a threshold condition that is identical to that obtained in the case of the dual MISO interference channel.

### III. MIMO INTERFERENCE CHANNEL

Let  $\underline{S}_i$  denote the side information given to the receiver  $i$

$$\begin{aligned} \underline{S}_1 &= H_{21}X_1 + \underline{W}_1 \\ \underline{S}_2 &= H_{12}X_2 + \underline{W}_2 \end{aligned} \quad (10)$$

where

$$\begin{bmatrix} \underline{Z}_i \\ \underline{W}_i \end{bmatrix} \sim \mathcal{N}\left(\mathbf{0}, \begin{bmatrix} I & A_i \\ A_i^\top & \Sigma_i \end{bmatrix}\right).$$

Since the covariance matrix has to be positive semi-definite, the genie parameters have to satisfy

$$\begin{bmatrix} I & A_i \\ A_i^\top & \Sigma_i \end{bmatrix} \succeq 0. \quad (11)$$

*Lemma 2 (Theorem 7.7.6 in [16]):* If  $\Sigma \succ 0$ , then the following two statements are equivalent:

$$\begin{aligned} \begin{bmatrix} I & A \\ A^\top & \Sigma \end{bmatrix} &\succeq 0 \\ I - A\Sigma^{-1}A^\top &\succeq 0. \end{aligned}$$

Recall that  $\Psi$  denotes the genie parameters collectively, and  $R_{\text{TIN}}^{\text{GA-IC}}(Q_1, Q_2, \Psi)$  denotes the sum rate achievable by treating interference as noise for the genie-aided interference channel using Gaussian inputs with input covariance matrix  $Q_i$  at transmitter  $i$ . We now extend Lemmas 12 and 13 in [7] (the so called *usefulness* and *smartness* conditions) to the MIMO interference channel. Lemmas 3 and 4 present conditions under which steps (b) and (d) in (6) hold true, respectively.

*Lemma 3 (Useful Genie):* If the genie ( $\Psi$ ) satisfies the following conditions:

$$\begin{aligned} \Sigma_1 &\preceq I - A_2\Sigma_2^{-1}A_2^\top \\ \Sigma_2 &\preceq I - A_1\Sigma_1^{-1}A_1^\top \end{aligned} \quad (12)$$

then the sum capacity of the genie-aided channel is given by

$$C_{\text{sum}}^{\text{GA-IC}}(\Psi) = \max_{Q_i \in \mathcal{Q}_i} R_{\text{TIN}}^{\text{GA-IC}}(Q_1, Q_2, \Psi)$$

and  $R_{\text{TIN}}^{\text{GA-IC}}(Q_1, Q_2, \Psi)$  is concave in  $(Q_1, Q_2)$ . Furthermore, if the inequalities in (12) are met with equality, then

$$C_{\text{sum}}^{\text{GA-IC}}(\Psi) = \max_{Q_i \in \mathcal{Q}_i} \sum_{i=1}^2 h(\underline{Y}_{iG} | \underline{S}_{iG}) - h(\underline{W}_i).$$

*Proof:* See Appendix B. ■

*Remark 1:* When the cross channel matrices  $(H_{12}, H_{21})$  are not full-column rank, the conditions (12) can be relaxed as we do later in Lemma 5 for the SIMO interference channel.

*Lemma 4 (Smart Genie):* For any input covariance matrices  $(Q_1, Q_2)$

$$R_{\text{TIN}}^{\text{GA-IC}}(Q_1, Q_2, \Psi) = R_{\text{TIN}}^{\text{IC}}(Q_1, Q_2).$$

if and only if (iff)

$$\begin{aligned} (A_1^\top (H_{12}Q_2H_{12}^\top + I)^{-1}H_{11} - H_{21})Q_1 &= 0 \\ (A_2^\top (H_{21}Q_1H_{21}^\top + I)^{-1}H_{22} - H_{12})Q_2 &= 0. \end{aligned} \quad (13)$$

*Proof:* See Appendix C. ■

*Theorem 1:* The sum capacity of the MIMO interference channel (1) is achieved by using Gaussian inputs and treating interference as noise at the receivers, and is given by

$$C_{\text{sum}}^{\text{IC}} = R_{\text{TIN}}^{\text{IC}}(Q_1^*, Q_2^*)$$

if there exists a full rank local optimal solution  $(Q_1^*, Q_2^*)$  to the optimization problem

$$\max_{Q_i \in \mathcal{Q}_i} R_{\text{TIN}}^{\text{IC}}(Q_1, Q_2) \quad (14)$$

and there exist matrices  $A_1, A_2, \Sigma_1 \succ 0, \Sigma_2 \succ 0$  satisfying the usefulness and smartness conditions corresponding to  $(Q_1^*, Q_2^*)$ , i.e.,

$$\begin{aligned} \Sigma_1 &\preceq I - A_2\Sigma_2^{-1}A_2^\top \\ \Sigma_2 &\preceq I - A_1\Sigma_1^{-1}A_1^\top \\ (A_1^\top (H_{12}Q_2^*H_{12}^\top + I)^{-1}H_{11} - H_{21})Q_1^* &= 0 \\ (A_2^\top (H_{21}Q_1^*H_{21}^\top + I)^{-1}H_{22} - H_{12})Q_2^* &= 0. \end{aligned} \quad (15)$$

*Proof:* Since

$$\begin{aligned} I - A_2 \Sigma_2^{-1} A_2^\top &\succeq \Sigma_1 \succ 0 \\ I - A_1 \Sigma_1^{-1} A_1^\top &\succeq \Sigma_2 \succ 0 \end{aligned}$$

from Lemma 2, condition (11) is satisfied. The outline of the proof is given in (6). Here, we justify the steps (b), (c) and (d) in (6). Steps (b) and (d) follow from Lemmas 3 and 4, respectively, and it only remains to prove step (c). The Lagrangian associated with (14) is [17]

$$-R_{\text{TIN}}^{\text{IC}}(Q_1, Q_2) + \sum_{i=1}^2 \lambda_i (\text{tr}(Q_i) - P_i) - \text{tr}(M_i Q_i).$$

Since  $(Q_1^*, Q_2^*)$  is a local optimal solution to (14), there exist dual variables  $\lambda_i \geq 0$  and  $M_i \succeq 0$  satisfying the KKT conditions [17]

$$\begin{aligned} \nabla_{Q_i} R_{\text{TIN}}^{\text{IC}}(Q_1^*, Q_2^*) &= \lambda_i I - M_i \\ \lambda_i (\text{tr}(Q_i^*) - P_i) &= 0 \\ \text{tr}(M_i Q_i^*) &= 0. \end{aligned} \quad (16)$$

Now, let  $R_{\text{TIN}}^{\text{DIFF}}(Q_1, Q_2, \Psi)$  denote the extra sum rate achievable because of the genie

$$\begin{aligned} R_{\text{TIN}}^{\text{DIFF}}(Q_1, Q_2, \Psi) &= R_{\text{TIN}}^{\text{GA-IC}}(Q_1, Q_2, \Psi) - R_{\text{TIN}}^{\text{IC}}(Q_1, Q_2) \\ &= \sum_{i=1}^2 I(\underline{X}_{iG}; \underline{Y}_{iG} | \underline{S}_{iG}) \\ &\geq 0. \end{aligned}$$

Since the genie  $\Psi$  satisfies condition (13), from Lemma 4 we have

$$R_{\text{TIN}}^{\text{DIFF}}(Q_1^*, Q_2^*, \Psi) = 0.$$

Therefore,  $(Q_1^*, Q_2^*)$  is also a solution to the optimization problem

$$\min_{Q_i \succeq 0} R_{\text{TIN}}^{\text{DIFF}}(Q_1, Q_2, \Psi). \quad (17)$$

The Lagrangian associated with (17) is

$$R_{\text{TIN}}^{\text{DIFF}}(Q_1, Q_2, \Psi) - \sum_{i=1}^2 \text{tr}(M_i Q_i).$$

There exists dual variables  $N_i \succeq 0$  satisfying the KKT conditions

$$\begin{aligned} \nabla_{Q_i} R_{\text{TIN}}^{\text{DIFF}}(Q_1^*, Q_2^*) &= N_i \\ \text{tr}(N_i Q_i^*) &= 0. \end{aligned}$$

Since  $Q_i^*$  is full rank,  $\text{tr}(N_i Q_i^*) = 0$  implies  $N_i = 0$ , and hence

$$\nabla_{Q_i} R_{\text{TIN}}^{\text{DIFF}}(Q_1^*, Q_2^*) = 0. \quad (18)$$

Combining (16) and (18), we have

$$\begin{aligned} \nabla_{Q_i} R_{\text{TIN}}^{\text{GA-IC}}(Q_1^*, Q_2^*) &= \lambda_i I - M_i \\ \lambda_i (\text{tr}(Q_i^*) - P_i) &= 0 \\ \text{tr}(M_i Q_i^*) &= 0 \end{aligned} \quad (19)$$

which are nothing but the KKT conditions associated with the problem

$$\max_{Q_i \in \mathcal{Q}_i} R_{\text{TIN}}^{\text{GA-IC}}(Q_1, Q_2, \Psi). \quad (20)$$

From Lemma 3,  $R_{\text{TIN}}^{\text{GA-IC}}(Q_1, Q_2)$  is concave in  $Q_1$  and  $Q_2$ . Since the objective function is concave, the KKT conditions are not only necessary but also sufficient, and hence  $(Q_1^*, Q_2^*)$  is a globally optimal solution to the optimization problem (20). (See Section 5.5.3 in [17] for an explanation). This verifies step (c) of (6), and completes the proof of Theorem 1. ■

#### IV. MISO INTERFERENCE CHANNEL

Consider a symmetric two-user Gaussian MISO interference channel

$$\begin{aligned} Y_1 &= \underline{d}^\top \underline{X}_1 + h \underline{c}^\top \underline{X}_2 + Z_1 \\ Y_2 &= \underline{d}^\top \underline{X}_2 + h \underline{c}^\top \underline{X}_1 + Z_2 \end{aligned} \quad (21)$$

obtained by making the following substitutions in (1):

$$\begin{aligned} H_{11} &= H_{22} = \underline{d}^\top \\ H_{12} &= H_{21} = h \underline{c}^\top \\ P_1 &= P_2 = P \end{aligned}$$

where  $\underline{d}$  and  $\underline{c}$  are unit norm vectors, denoting the directions of direct link and cross link, respectively. Without any loss of generality, we can assume two transmit antennas and

$$\begin{aligned} \underline{c} &= [1 \ 0]^\top \\ \underline{d} &= [\cos \theta \ \sin \theta]^\top, \theta \in (0, \pi/2). \end{aligned}$$

by projecting  $\underline{X}_1$  and  $\underline{X}_2$  along appropriate basis vectors. See Appendix D for an explanation. The extreme case  $\theta = 0$  corresponds to the SISO interference channel, which was considered in [7]. The extreme case  $\theta = \pi/2$  corresponds to the no interference scenario, with the capacity region given by  $\{(R_1, R_2) : R_i \leq \frac{1}{2} \log(1 + P)\}$ .

##### A. Achievable Sum-Rate

The sum-rate achieved by using Gaussian inputs with input covariance matrix  $Q_i$  at transmitter  $i$  and treating interference as noise at the receivers is

$$\begin{aligned} R_{\text{TIN}}^{\text{IC}}(Q_1, Q_2) &= \sum_{i=1}^2 I(\underline{X}_{iG}; \underline{Y}_{iG}) \\ &= \frac{1}{2} \log \left( 1 + \frac{\underline{d}^\top Q_1 \underline{d}}{1 + h^2 \underline{c}^\top Q_2 \underline{c}} \right) \\ &\quad + \frac{1}{2} \log \left( 1 + \frac{\underline{d}^\top Q_2 \underline{d}}{1 + h^2 \underline{c}^\top Q_1 \underline{c}} \right). \end{aligned}$$

This problem of finding the optimal transmit beams for the MISO interference channel is studied in [15], [18] and [19]. In [15], it is shown that the optimal covariance matrices  $Q_1^*$  and

$Q_2^*$  are unit-rank. Even though the channel is symmetric across the users, since the sum-rate  $R_{\text{TIN}}^{\text{IC}}(Q_1, Q_2)$  is not necessarily concave in  $(Q_1, Q_2)$ , the optimal covariance matrices  $Q_1^*$  and  $Q_2^*$  need not be the same. In this Section, for the purpose of achievable sum-rate, we restrict ourselves to covariance matrices symmetric across the users, i.e.,

$$Q_1 = Q_2 = Q = P\underline{b}\underline{b}^\top$$

where  $\underline{b}$  is the unit norm transmit beamforming vector. With this assumption, the achievable sum-rate is given by

$$R_{\text{TIN}}^{\text{IC}}(Q_1, Q_2) = \log(1 + \text{SINR})$$

where SINR denotes the signal to interference and noise ratio

$$\text{SINR} = \frac{P(\underline{b}^\top \underline{d})^2}{1 + h^2 P(\underline{b}^\top \underline{c})^2} = \frac{\underline{b}^\top (P\underline{d}\underline{d}^\top) \underline{b}}{\underline{b}^\top (I + h^2 P\underline{c}\underline{c}^\top) \underline{b}}.$$

The optimal  $\underline{b}$  maximizing SINR is given by the generalized eigenvector of the matrix pair

$$(P\underline{d}\underline{d}^\top, I + h^2 P\underline{c}\underline{c}^\top)$$

corresponding to the largest generalized eigenvalue, given by

$$\underline{b} = \frac{(I + h^2 P\underline{c}\underline{c}^\top)^{-1} \underline{d}}{\|(I + h^2 P\underline{c}\underline{c}^\top)^{-1} \underline{d}\|}. \quad (22)$$

The corresponding SINR is given by

$$\begin{aligned} \text{SINR} &= \frac{P(\underline{b}^\top \underline{d})^2}{\underline{b}^\top (I + h^2 P\underline{c}\underline{c}^\top) \underline{b}} \\ &= P\underline{d}^\top (I + h^2 P\underline{c}\underline{c}^\top)^{-1} \underline{d} \\ &\stackrel{(a)}{=} P\underline{d}^\top \left( I - \frac{h^2 P}{1 + h^2 P} \underline{c}\underline{c}^\top \right) \underline{d} \\ &= P \left( 1 - \frac{h^2 P \cos^2 \theta}{1 + h^2 P} \right) \\ &= \frac{P \cos^2 \theta}{1 + h^2 P} + P \sin^2 \theta. \end{aligned} \quad (23)$$

where step (a) follows from the matrix inversion lemma

$$(A + UCV)^{-1} = A^{-1} - A^{-1}U(C^{-1} + VA^{-1}U)^{-1}VA^{-1}. \quad (24)$$

Therefore, a sum-rate of

$$\begin{aligned} R_{\text{TIN}}^{\text{IC}}(Q^*, Q^*) &= \log(1 + \text{SINR}) \\ &= \log \left( 1 + \frac{P \cos^2 \theta}{1 + h^2 P} + P \sin^2 \theta \right) \end{aligned}$$

is achievable.

We now characterize the low interference regime for the MISO interference channel, and provide explicit conditions on  $h$  and  $\theta$ , when  $R_{\text{TIN}}^{\text{IC}}(Q^*, Q^*)$  is equal to the sum capacity. We cannot use Theorem 1 to do this, because the optimal covariance matrices  $Q_1^*$  and  $Q_2^*$  are not full-rank, as required in the hypothesis of Theorem 1.

## B. Low Interference Regime

*Theorem 2:* The sum capacity of the MISO interference channel (21) is achieved by using Gaussian inputs, transmit beamforming along the vector  $\underline{b}$  given in (22), and treating interference as noise at the receivers, and is given by

$$C_{\text{sum}}^{\text{IC}} = \log \left( 1 + \frac{P \cos^2 \theta}{1 + h^2 P} + P \sin^2 \theta \right)$$

if the channel gain parameter  $h$  satisfies

$$|h| < h_0(\theta, P) \quad (25)$$

with  $h_0(\theta, P)$  being the positive solution to the implicit equation

$$h^2 - \sin^2 \theta = \left( \frac{\cos \theta}{1 + h^2 P} - h \right)_+^2. \quad (26)$$

(The notation  $x_+^2$  is used to denote  $(\max(0, x))^2$ .)

*Proof:* First, observe that the capacity region of the symmetric MISO interference channel (21) does not depend on the sign of the channel parameter  $h$ . By replacing  $\underline{X}_2$  and  $Y_2$  with  $-\underline{X}_2$  and  $-Y_2$ , respectively, we can convert the interference channel with negative  $h$  to the interference channel with positive  $h$ . Therefore, with out any loss of generality, we assume that  $h \geq 0$ .

The achievability part of Theorem 2 is established in Section IV-A

$$C_{\text{sum}}^{\text{IC}} \geq R_{\text{TIN}}^{\text{IC}}(Q^*, Q^*) = \log \left( 1 + \frac{P \cos^2 \theta}{1 + h^2 P} + P \sin^2 \theta \right).$$

We now prove the converse, i.e., that

$$C_{\text{sum}}^{\text{IC}} \leq R_{\text{TIN}}^{\text{IC}}(Q^*, Q^*).$$

The outline of the converse proof is given in (6). We justify the steps (b), (c) and (d) of (6) here. Specializing the genie (5) to the MISO interference channel, we get

$$\begin{aligned} S_1 &= h\underline{c}^\top \underline{X}_1 + W_1 \\ S_2 &= h\underline{c}^\top \underline{X}_2 + W_2. \end{aligned} \quad (27)$$

We restrict the genie to be symmetric across the users, with the genie parameters  $\Psi = \{\Sigma, a\}$  chosen to satisfy the usefulness condition (12) of Lemma 3 and the smartness condition (13) of Lemma 4. From (13), we get

$$\begin{aligned} \left( \frac{a}{1 + h^2 P |\underline{b}^\top \underline{c}|^2} \underline{d}^\top - h\underline{c}^\top \right) Q^* &= 0 \\ \left( \frac{a\underline{d}^\top}{1 + h^2 P |\underline{b}^\top \underline{c}|^2} - h\underline{c}^\top \right) \underline{b} &= 0 \end{aligned}$$

and thus

$$a = \frac{h\underline{c}^\top \underline{b}}{\underline{d}^\top \underline{b}} (1 + h^2 P |\underline{b}^\top \underline{c}|^2).$$

From (12), we get

$$\begin{aligned}\Sigma &\leq 1 - \frac{a^2}{\Sigma} \\ \Sigma^2 - \Sigma + a^2 &\leq 0\end{aligned}$$

We choose

$$\Sigma = 0.5 + 0.5\sqrt{1 - 4a^2}$$

so that (12) is satisfied with equality. Thus

$$\begin{aligned}a &= \mathbb{E}[W_i Z_i] = h \frac{\underline{b}^\top \underline{c}}{\underline{b}^\top \underline{d}} (1 + h^2 P \underline{b}^\top \underline{c}) \\ \Sigma &= \mathbb{E}[W_i^2] = 0.5 + 0.5\sqrt{1 - 4a^2}.\end{aligned}\quad (28)$$

In Appendix F, we show that  $h < h_0(\theta, P)$  implies that

$$a < 0.5. \quad (29)$$

Thus, a real  $\Sigma > 0$  exists, and the conditions of Lemmas 3 and 4 are met. The steps (b) and (c) of (6) follow from Lemmas 3 and 4, respectively. Thus, it only remains to verify that step (c) of (6) holds true, i.e.,

$$(Q^*, Q^*) = \arg \max_{Q_i \in \mathcal{Q}_i} R_{\text{TIN}}^{\text{GA-IC}}(Q_1, Q_2, \Psi).$$

Since the condition (12) of Lemma 3 is met with equality, we have

$$R_{\text{TIN}}^{\text{GA-IC}}(Q_1, Q_2, \Psi) = \sum_{i=1}^2 h(\underline{Y}_{iG} | \underline{S}_{iG}) - h(W_i).$$

Furthermore,  $R_{\text{TIN}}^{\text{GA-IC}}(Q_1, Q_2, \Psi)$  is concave in  $(Q_1, Q_2)$ . Since the channel and the genie are symmetric across the users  $R_{\text{TIN}}^{\text{GA-IC}}(Q_1, Q_2, \Psi)$  is also symmetric across the users, i.e.,

$$R_{\text{TIN}}^{\text{GA-IC}}(Q_1, Q_2, \Psi) = R_{\text{TIN}}^{\text{GA-IC}}(Q_2, Q_1, \Psi).$$

Therefore, we have

$$\begin{aligned}R_{\text{TIN}}^{\text{GA-IC}}(Q_1, Q_2, \Psi) & \\ &= \frac{1}{2} R_{\text{TIN}}^{\text{GA-IC}}(Q_1, Q_2, \Psi) + \frac{1}{2} R_{\text{TIN}}^{\text{GA-IC}}(Q_2, Q_1, \Psi) \\ &\leq R_{\text{TIN}}^{\text{GA-IC}}(Q, Q, \Psi)\end{aligned}$$

where

$$Q = \left( \frac{Q_1 + Q_2}{2} \right).$$

Since the power constraints are also symmetric across users

$$\mathcal{Q}_1 = \mathcal{Q}_2 = \mathcal{Q} = \{Q : Q \succeq 0, \text{tr}(Q) \leq P\}$$

$Q_1 \in \mathcal{Q}$  and  $Q_2 \in \mathcal{Q}$  implies  $Q \in \mathcal{Q}$ . Therefore, we have shown that having symmetric  $Q$  across the users maximizes  $R_{\text{TIN}}^{\text{GA-IC}}(Q_1, Q_2, \Psi)$ . Hence, it is sufficient to prove that

$$Q^* = \arg \max_{Q \in \mathcal{Q}} R_{\text{TIN}}^{\text{GA-IC}}(Q, Q, \Psi) \quad (30)$$

Now

$$\begin{aligned}R_{\text{TIN}}^{\text{GA-IC}}(Q, Q, \Psi) &= \sum_{i=1}^2 h(\underline{Y}_{iG} | \underline{S}_{iG}) - h(W_i) \\ &= \sum_{i=1}^2 h(\underline{Y}_{iG} - \mu \underline{S}_{iG} | \underline{S}_{iG}) - h(W_i) \\ &\stackrel{(a)}{\leq} \sum_{i=1}^2 h(\underline{Y}_{iG} - \mu \underline{S}_{iG}) - h(W_i) \\ &= \log(\text{tr}(MQ) + k^2) - \log(\Sigma)\end{aligned}\quad (31)$$

where step (a) follows from the fact that conditioning reduces entropy, and the matrix  $M$  and the constant  $k^2$  are given by

$$\begin{aligned}M &= (\underline{d} - \mu h \underline{c})(\underline{d} - \mu h \underline{c})^\top + h^2 \underline{c} \underline{c}^\top \\ k^2 &= \mathbb{E}[(Z_1 - \mu W_1)^2].\end{aligned}\quad (32)$$

The constant  $\mu$  is chosen such that the step (a) is tight when  $Q = Q^*$ , which is true if  $\mu \underline{S}_{iG}$  is the MMSE estimate of  $\underline{Y}_{iG}$  given  $\underline{S}_{iG}$  when  $Q = Q^*$ , i.e., if

$$\begin{aligned}\mu &= \frac{\mathbb{E}[\underline{S}_{iG} \underline{Y}_{iG}]}{\mathbb{E}[\underline{S}_{iG}^2]} \\ &= \frac{a + h P \underline{b}^\top \underline{c} \underline{b}^\top \underline{d}}{\Sigma + h^2 P (\underline{b}^\top \underline{c})^2}.\end{aligned}\quad (33)$$

In Appendix F, we show that  $h < h_0(\theta, P)$  implies

$$\mu h < \cos(\theta). \quad (34)$$

In Appendix G, we show that  $\mu h < \cos(\theta)$  implies

$$\text{tr}(MQ) \leq \text{tr}(MQ^*), \quad \forall Q \in \mathcal{Q}. \quad (35)$$

From (31) and (35), we have

$$R_{\text{TIN}}^{\text{GA-IC}}(Q, Q, \Psi) \leq \log(\text{tr}(MQ^*) + k^2) - \log(\Sigma).$$

Moreover, since all the steps in (31) and (35) are tight when  $Q = Q^*$ , we have

$$\begin{aligned}R_{\text{TIN}}^{\text{GA-IC}}(Q^*, Q^*, \Psi) &= \log(\text{tr}(MQ^*) + k^2) - \log(\Sigma) \\ R_{\text{TIN}}^{\text{GA-IC}}(Q, Q, \Psi) &\leq R_{\text{TIN}}^{\text{GA-IC}}(Q^*, Q^*, \Psi), \quad \forall Q \in \mathcal{Q}.\end{aligned}$$

This verifies the step (c) of (6) and completes the converse proof.  $\blacksquare$

In the following section we establish that results identical to those in Theorem 2 also hold true for the dual SIMO interference channel, although the proofs are quite different.

## V. SIMO INTERFERENCE CHANNEL

Consider the symmetric SIMO interference channel

$$\begin{aligned}\underline{Y}_1 &= \underline{d} X_1 + h \underline{c} X_2 + \underline{Z}_1 \\ \underline{Y}_2 &= \underline{d} X_2 + h \underline{c} X_1 + \underline{Z}_2\end{aligned}\quad (36)$$

obtained by making the following substitutions in (1)

$$\begin{aligned} H_{11} &= H_{22} = \underline{d} \\ H_{12} &= H_{21} = h\underline{c} \\ P_1 &= P_2 = P \end{aligned}$$

where  $\underline{d}$  and  $\underline{c}$  are unit norm vectors, denoting the directions of direct link and cross link, respectively. Similar to the MISO Interference channel, without any loss of generality, we assume two receive antennas and

$$\begin{aligned} \underline{c} &= [1 \ 0]^\top \\ \underline{d} &= [\cos \theta \ \sin \theta]^\top, \quad \theta \in (0, \pi/2). \end{aligned}$$

by projecting  $\underline{Y}_1$  and  $\underline{Y}_2$  along appropriate basis vectors. See Appendix D for an explanation.

### A. Achievable Sum-Rate

Assume that each receiver does performs beamforming along the vector  $\underline{b}$  given in (22). By treating interference as noise at the receivers, the following signal to interference and noise ratio (SINR) is achieved.

$$\begin{aligned} \text{SINR} &= \frac{P(\underline{b}^\top \underline{d})^2}{1 + h^2 P(\underline{b}^\top \underline{c})^2} \\ &= \frac{P \cos^2 \theta}{1 + h^2 P} + P \sin^2 \theta. \end{aligned}$$

Therefore, a sum-rate of

$$\begin{aligned} R_{\text{TIN}}^{\text{IC}}(P, P) &= \log(1 + \text{SINR}) \\ &= \log \left( 1 + \frac{P \cos^2 \theta}{1 + h^2 P} + P \sin^2 \theta \right) \end{aligned}$$

is achievable.

### B. Low Interference Regime

*Theorem 3:* The sum capacity of the SIMO interference channel (36) is achieved by using Gaussian inputs, receive beamforming along the vector  $\underline{b}$  given in (22), and treating interference as noise at the receiver, and is given by

$$C_{\text{sum}}^{\text{IC}} = \log \left( 1 + \frac{P \cos^2 \theta}{1 + h^2 P} + P \sin^2 \theta \right)$$

if  $|h| < h_0(\theta, P)$ , where  $h_0(\theta, P)$  is the positive solution to the implicit equation

$$h^2 - \sin^2 \theta = \left( \frac{\cos \theta}{1 + h^2 P} - h \right)_+^2. \quad (37)$$

where the notation  $x_+^2$  is used to denote  $(\max(0, x))^2$ .

*Proof:* First, observe that the capacity region of the symmetric SIMO interference channel (36) does not depend on the sign of the channel parameter  $h$ . By replacing  $X_2$  and  $\underline{Y}_2$  with  $-X_2$  and  $-\underline{Y}_2$ , respectively, we can convert the interference

channel with negative  $h$  to the interference channel with positive  $h$ . Therefore, without any loss of generality, we assume that  $h \geq 0$ .

The achievability part of Theorem 3 is established in Section V.A, i.e.,

$$C_{\text{sum}}^{\text{IC}} \geq R_{\text{TIN}}^{\text{IC}}(P, P) = \log \left( 1 + \frac{P \cos^2 \theta}{1 + h^2 P} + P \sin^2 \theta \right).$$

We now prove the converse, i.e., that

$$C_{\text{sum}}^{\text{IC}} \leq R_{\text{TIN}}^{\text{IC}}(P, P).$$

The outline of the converse proof is given in (6). Specializing the genie (5) to the SIMO interference channel, we get

$$\begin{aligned} \underline{S}_1 &= h\underline{c}X_1 + \underline{W}_1 \\ \underline{S}_2 &= h\underline{c}X_2 + \underline{W}_2. \end{aligned} \quad (38)$$

We restrict the genie to be symmetric across the users, with the genie parameters  $\Psi = \{\Sigma, A\}$  defined as be chosen to be

$$\begin{aligned} A &= \mathbb{E} \left[ \underline{Z}_i \underline{W}_i^\top \right] = k\underline{v}\underline{c}^\top \\ \Sigma &= \mathbb{E} \left[ \underline{W}_i \underline{W}_i^\top \right] = \eta I \end{aligned} \quad (39)$$

where  $k \geq 0$  and  $\underline{v}$  is unit norm vector. For the genie parameters to be valid, they have to satisfy (11), and from Lemma 2, (11) is true iff

$$\eta \geq k^2 \quad (40)$$

As mentioned in Remark 1, Lemma 3 can be improved for the SIMO interference channel since the cross channel matrices are not full column rank.

*Lemma 5 (Useful Genie for SIMO Channel):* If the genie ( $\Psi$ ) satisfy

$$(\underline{c}^\top \Sigma^{-1} \underline{c})^{-1} \leq (\underline{c}^\top (I - A \Sigma^{-1} A^\top)^{-1} \underline{c})^{-1} \quad (41)$$

then the sum capacity of the genie-aided channel is given by

$$C_{\text{sum}}^{\text{GA-IC}}(\Psi) = R_{\text{TIN}}^{\text{GA-IC}}(P, P, \Psi).$$

*Proof:* See Appendix H. ■

Substituting (39) in (41), we get

$$\eta \leq \frac{\eta - k^2}{\eta - k^2(1 - (\underline{c}^\top \underline{v})^2)}. \quad (42)$$

From Lemma 4, if the condition (13) is satisfied at  $Q_1 = Q_2 = P$ , i.e., if

$$A^\top (h^2 P \underline{c}\underline{c}^\top + I)^{-1} \underline{d} - h\underline{c} = 0$$

equivalently, if

$$h = k\underline{v}^\top (I + h^2 P \underline{c}\underline{c}^\top)^{-1} \underline{d} \quad (43)$$



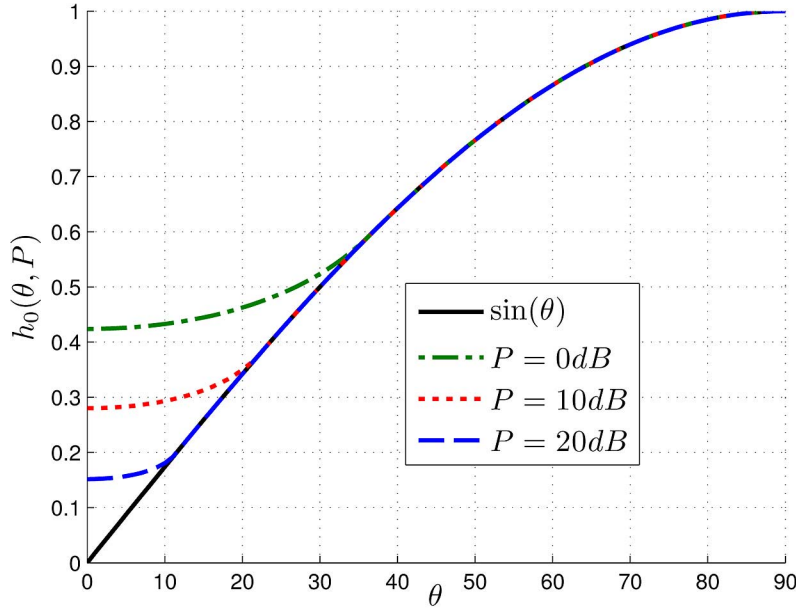


Fig. 2. Threshold on  $h$  characterizing the low interference regime of MISO and SIMO interference channels.

then

$$R_{\text{TIN}}^{\text{GA-IC}}(P, P, \Psi) = R_{\text{TIN}}^{\text{IC}}(P, P).$$

Therefore, if there exist  $\eta, k$  and  $\underline{v}$  satisfying (40), (42) and (43), then we have the required outer bound

$$\begin{aligned} \mathcal{C}_{\text{sum}}^{\text{IC}} &\leq \mathcal{C}_{\text{sum}}^{\text{GA-IC}} \\ &= R_{\text{TIN}}^{\text{GA-IC}}(P, P, \Psi) \\ &= R_{\text{TIN}}^{\text{IC}}(P, P). \end{aligned}$$

The existence of  $\eta, k$  and  $\underline{v}$  satisfying (40), (42) and (43) is proved in Appendix I, using the threshold condition of the theorem. ■

## VI. LOW INTERFERENCE REGIME

In Sections IV and V, we proved that using Gaussian inputs and treating interference as noise is sum capacity achieving for the symmetric MISO and SIMO interference channels if  $h \leq h_0(\theta, P)$  where  $h_0(\theta, P)$  is the positive solution to the equation

$$h_0^2 - \sin^2 \theta = \left( \frac{\cos \theta}{1 + h_0^2 P} - h_0 \right)_+^2. \quad (44)$$

Observe that, with  $\theta = 0$ , the threshold  $h_0(0, P)$  is the solution to the equation

$$h_0(1 + h_0^2 P) \leq 0.5$$

which is consistent with the threshold condition for the low interference regime of the SISO interference channel [7]. The threshold  $h_0(\theta, P)$  is plotted as a function of  $\theta$  for different values of  $P$  in Fig. 2. It can be observed that the threshold curve is always above the  $\sin(\theta)$  curve and approaches the  $\sin(\theta)$  curve as  $P$  becomes larger. These observations are summarized in the following Lemma.

*Lemma 6:* The threshold  $h_0(\theta, P)$  as defined in (44) satisfies the following properties.

1) Independent of the value of  $P$

$$h_0(\theta, P) \geq \sin \theta$$

2) For any fixed  $\theta$

$$\lim_{P \rightarrow \infty} h_0(\theta, P) = \sin \theta.$$

*Proof:* Any  $h_0 < \sin \theta$  cannot satisfy (44) since the RHS of (44) cannot be negative. Therefore, we have the first property. For any  $h_0 \geq \sin \theta$

$$\frac{\cos \theta}{1 + h_0^2 P} - h_0 \leq \frac{\cos \theta}{1 + P \sin^2 \theta} - \sin \theta$$

which is less than zero for sufficiently large  $P$ . Therefore, the RHS of (44) is zero for any  $h_0 \geq \sin \theta$  and hence  $h_0$  cannot be greater than  $\sin \theta$  for sufficiently large  $P$ . This proves the second property. ■

Theorems 2 and 3 along with the first property of the above Lemma lead to the following corollary.

*Corollary 1:* The sum capacity of the symmetric MISO/SIMO interference channel is achieved by using Gaussian inputs, transmit/receive beamforming, and treating interference as noise at the receivers, if  $h \leq \sin \theta$ .

## VII. CONCLUSION

In this paper, we derived sufficient conditions under which using Gaussian inputs and treating interference as noise at the receivers achieves the sum capacity of the two-user MIMO interference channel. The outer bound is based on a genie giving side information to the receivers. The genie is carefully chosen to satisfy the following properties.

- The genie should be useful, i.e., using Gaussian inputs and treating interference as noise at the receivers should achieve the sum capacity of the genie-aided channel.
- The genie should be smart at  $(Q_1^*, Q_2^*)$ , which are the optimal covariance matrices that maximize the achievable sum-rate of the interference channel. That is, when Gaussian inputs with covariance matrices  $(Q_1^*, Q_2^*)$  are used and interference is treated as noise at the receivers, the presence of the genie should not improve the achievable sum-rate.
- The optimal covariance matrices  $(Q_1^*, Q_2^*)$  for the original interference channel should also maximize the achievable sum-rate of the genie-aided channel.

The last property is the new feature of the MIMO interference channel relative to the SISO setting. One of the main contributions of this paper is to show that if  $(Q_1^*, Q_2^*)$  are full rank and if there exists a genie satisfying the first two properties, then the last property is automatically satisfied. If  $(Q_1^*, Q_2^*)$  are not full rank, then the genie might have to satisfy some additional constraints for the last property to be true.

The MISO interference channel is an example where the optimal covariance matrices  $(Q_1^*, Q_2^*)$  are not full rank, since transmit beamforming maximizes the achievable sum-rate. For the special case of the symmetric MISO interference channel, we derived the additional constraints and derived a threshold condition for the low interference regime. We also showed that the same threshold condition characterizes the low interference regime for the dual SIMO interference channel. The threshold condition is given by  $h \leq h_0(\theta, P)$ , where  $h_0(\theta, P)$  is larger than  $\sin \theta$ . This means that the low interference regime can be quite significant. For example, when  $\theta = 45^\circ$ , then using Gaussian inputs and treating interference as noise at the receivers achieves the sum capacity if INR ( $= h^2 P$ ) is 3 dB less than SNR ( $= P$ ).

## APPENDIX

### A. Some Useful Results

We now present a few results that will be useful in this paper.

*Lemma 7 (Lemma 1 in [7]):* Let  $\underline{X}$  be a random vector, and let  $\underline{Y}$  and  $\underline{S}$  be noisy observations of  $\underline{X}$

$$\begin{aligned}\underline{Y} &= A \underline{X} + \underline{Z} \\ \underline{S} &= C \underline{X} + \underline{W}\end{aligned}$$

where  $\underline{Z}$  and  $\underline{W}$  are correlated, zero-mean, Gaussian random vectors, and  $A$  and  $C$  are real valued matrices. Consider the random vector sequence  $\underline{X}^n = (\underline{X}_1, \dots, \underline{X}_n)$  with the covariance constraint  $\frac{1}{n} \sum_{j=1}^n \Sigma_{xj} \preceq \Sigma_x$ , where  $\Sigma_{xj}$  is the covariance matrix of  $\underline{X}_j$ . Furthermore, let  $\underline{Y}^n$  and  $\underline{S}^n$  be the corresponding observations when the noise vector sequences  $\underline{Z}^n$  and  $\underline{W}^n$  each have components that are i.i.d. in time. Then, we have

$$h(\underline{Y}^n | \underline{S}^n) \leq nh(\underline{Y}_G | \underline{S}_G)$$

where  $\underline{Y}_G$  and  $\underline{S}_G$  are  $\underline{Y}$  and  $\underline{S}$  when  $\underline{X} = \underline{X}_G \sim \mathcal{N}(\underline{0}, \Sigma_x)$ .

*Lemma 8 (Lemma 4 in [7]):* Let  $\underline{X}^n$  be a random vector sequence with an average covariance constraint, i.e.,  $\sum_{j=1}^n \Sigma_{xj} \preceq n \Sigma_x$ , and let  $\underline{Z}^n$  be an independent random vector sequence, with components that are i.i.d.  $\mathcal{N}(\underline{0}, \Sigma_z)$ . Then

$$h(\underline{X}^n) - h(\underline{X}^n + \underline{Z}^n) \leq nh(\underline{X}_G) - nh(\underline{X}_G + \underline{Z})$$

where  $\underline{X}_G \sim \mathcal{N}(\underline{0}, \Sigma_x)$ , and equality is achieved if  $\underline{X}^n = \underline{X}_G^n$ , where  $\underline{X}_G^n$  denotes the random sequence with components that are i.i.d.  $\mathcal{N}(\underline{0}, \Sigma_x)$ .

*Lemma 9 (Lemma 6 in [7]):* Let  $\underline{X}^n$  be a random vector sequence, and let  $\underline{Z}^n$  and  $\underline{W}^n$  be (possibly correlated) zero-mean Gaussian random vector sequences, independent of  $\underline{X}^n$  and i.i.d. in time. Then

$$h(\underline{X}^n + \underline{Z}^n | \underline{W}^n) = h(\underline{X}^n + \underline{V}^n)$$

where  $\underline{V}^n$  is i.i.d.  $\mathcal{N}(0, \text{Cov}(\underline{Z} | \underline{W}))$ .

*Lemma 10:* Let  $\underline{X}$  be a random vector with covariance matrix  $Q$  and  $\underline{Y}$  and  $\underline{S}$  be noisy observations of  $\underline{X}$

$$\begin{aligned}\underline{Y} &= A \underline{X} + \underline{Z} \\ \underline{S} &= B \underline{X} + \underline{W}\end{aligned}$$

where  $\underline{Z}$  and  $\underline{W}$  are correlated, zero-mean, Gaussian random vectors, and  $A$  and  $B$  are real-valued matrices. Then,  $h(\underline{Y}_G | \underline{S}_G)$  is concave in  $Q$ , where  $\underline{Y}_G$  and  $\underline{S}_G$  are  $\underline{Y}$  and  $\underline{S}$  when  $\underline{X} = \underline{X}_G = \mathcal{N}(\underline{0}, Q)$ .

*Proof:* Let  $T$  be a time sharing random variable taking value that takes values 1 and 2 with probabilities  $\lambda$  and  $1 - \lambda$ , respectively. Let  $\underline{X}_{1G}$ ,  $\underline{X}_{2G}$  and  $\tilde{\underline{X}}_G$  be independent Gaussian random vectors with covariance matrices  $Q_1$ ,  $Q_2$  and  $\lambda Q_1 + (1 - \lambda)Q_2$ , respectively and let  $\underline{Y}_{1G}$ ,  $\underline{Y}_{2G}$ ,  $\tilde{\underline{Y}}_G$ ,  $\underline{S}_{1G}$ ,  $\underline{S}_{2G}$  and  $\tilde{\underline{S}}_G$  denote the corresponding noisy observations. Now

$$\begin{aligned}\lambda h(\underline{Y}_{1G} | \underline{S}_{2G}) + (1 - \lambda)h(\underline{Y}_{2G} | \underline{S}_{2G}) \\ &= h(\underline{Y}_{TG} | \underline{S}_{TG}, T) \\ &\stackrel{(a)}{\leq} h(\underline{Y}_{TG} | \underline{S}_{TG}) \\ &\stackrel{(b)}{\leq} h(\tilde{\underline{Y}}_G | \tilde{\underline{S}}_G)\end{aligned}$$

where step (a) follows because conditioning reduces entropy and step (b) follows because Gaussian distribution maximizes the conditional differential entropy for a given covariance constraint [20, Lemma 1].  $\blacksquare$

*Lemma 11:* Let  $\underline{x}_1$  and  $\underline{x}_2$  be two vectors such that  $\underline{x}_1^\top \underline{x}_2 > 0$  and let  $\underline{x}$  be an eigenvector of the matrix

$$M = \underline{x}_1 \underline{x}_1^\top + \underline{x}_2 \underline{x}_2^\top$$

then  $\underline{x}$  corresponds to the largest eigenvalue iff  $(\underline{x}^\top \underline{x}_1)(\underline{x}^\top \underline{x}_2) > 0$ .

*Proof:*

$$\begin{aligned}M \underline{x} &= \lambda \underline{x} \\ \underline{x}_1 (\underline{x}_1^\top \underline{x}) + \underline{x}_2 (\underline{x}_2^\top \underline{x}) &= \lambda \underline{x}.\end{aligned}$$

Therefore, we have

$$\begin{aligned}\|\underline{x}_1\|^2(\underline{x}_1^\top \underline{x}) + (\underline{x}_1^\top \underline{x}_2)(\underline{x}_2^\top \underline{x}) &= \lambda(\underline{x}_1^\top \underline{x}) \\ \|\underline{x}_2\|^2(\underline{x}_2^\top \underline{x}) + (\underline{x}_1^\top \underline{x}_2)(\underline{x}_1^\top \underline{x}) &= \lambda(\underline{x}_2^\top \underline{x})\end{aligned}$$

i.e.,

$$\begin{aligned}(\underline{x}_1^\top \underline{x}_2)(\underline{x}_2^\top \underline{x}) &= (\lambda - \|\underline{x}_1\|^2)(\underline{x}_1^\top \underline{x}) \\ (\underline{x}_1^\top \underline{x}_2)(\underline{x}_1^\top \underline{x}) &= (\lambda - \|\underline{x}_2\|^2)(\underline{x}_2^\top \underline{x}).\end{aligned}$$

Therefore, we have

$$(\lambda - \|\underline{x}_1\|^2)(\lambda - \|\underline{x}_2\|^2) = (\underline{x}_1^\top \underline{x}_2)^2$$

and hence

$$\lambda = \frac{\|\underline{x}_1\|^2 + \|\underline{x}_2\|^2}{2} \pm \frac{\sqrt{(\|\underline{x}_1\|^2 - \|\underline{x}_2\|^2)^2 + 4(\underline{x}_1^\top \underline{x}_2)^2}}{2}.$$

Now, it can be seen that the larger eigenvalue  $\lambda_1$  satisfies

$$\lambda_1 > \max(\|\underline{x}_1\|^2, \|\underline{x}_2\|^2)$$

and hence, the corresponding eigenvector satisfies

$$(\underline{x}^\top \underline{x}_1)(\underline{x}^\top \underline{x}_2) > 0.$$

Similarly, it can be seen that the smaller eigenvalue  $\lambda_2$  satisfies

$$\lambda_2 < \min(\|\underline{x}_1\|^2, \|\underline{x}_2\|^2)$$

and hence, the corresponding eigenvector satisfies

$$(\underline{x}^\top \underline{x}_1)(\underline{x}^\top \underline{x}_2) < 0. \quad \blacksquare$$

### B. Proof of Lemma 3

Let  $\underline{X}_{iG} \sim \mathcal{N}(\underline{0}, Q_i)$  and  $\underline{Y}_G$  and  $\underline{S}_G$  be  $\underline{Y}$  and  $\underline{S}$  when  $\underline{X} = \underline{X}_{iG}$ , where

$$Q_i = \frac{1}{n} \sum_{t=1}^n \mathbb{E}[X_{it} X_{it}^\top] \in Q_i.$$

Starting with Fano's inequality, we have

$$\begin{aligned}n(C_{\text{sum}}^{\text{GA-IC}} - \epsilon_n) &\leq \sum_{i=1}^2 I(\underline{X}_i^n; \underline{Y}_i^n, \underline{S}_i^n) \\ &= \sum_{i=1}^2 I(\underline{X}_i^n; \underline{S}_i^n) + I(\underline{X}_i^n; \underline{Y}_i^n | \underline{S}_i^n) \\ &= \sum_{i=1}^2 h(\underline{S}_i^n) - h(\underline{S}_i^n | \underline{X}_i^n) \\ &\quad + h(\underline{Y}_i^n | \underline{S}_i^n) - h(\underline{Y}_i^n | \underline{S}_i^n, \underline{X}_i^n) \\ &\stackrel{(a)}{=} \sum_{i=1}^2 h(\underline{S}_i^n) - nh(\underline{W}_i) \\ &\quad + h(\underline{Y}_i^n | \underline{S}_i^n) - h(\underline{Y}_i^n | \underline{S}_i^n, \underline{X}_i^n) \\ &\stackrel{(b)}{\leq} \sum_{i=1}^2 h(\underline{S}_i^n) - nh(\underline{W}_i) \\ &\quad + nh(\underline{Y}_{iG} | \underline{S}_{iG}) - h(\underline{Y}_i^n | \underline{S}_i^n, \underline{X}_i^n)\end{aligned} \quad (45)$$

where step (a) follows from the fact that  $\underline{S}_i$  given  $\underline{X}_i$  is Gaussian noise and is independent of  $\underline{X}_i$  and step (b) follows from Lemma 7. Now consider

$$\begin{aligned}h(\underline{S}_1^n) - h(\underline{Y}_1^n | \underline{S}_1^n, \underline{X}_1^n) &= h(H_{21}\underline{X}_1^n + \underline{W}_1^n) - h(H_{21}\underline{X}_1^n + \underline{Z}_2^n | \underline{W}_2^n) \\ &\stackrel{(c)}{=} h(H_{21}\underline{X}_1^n + \underline{W}_1^n) - h(H_{21}\underline{X}_1^n + \underline{V}_1^n) \\ &\stackrel{(d)}{=} h(H_{21}\underline{X}_1^n + \underline{W}_1^n) - h(H_{21}\underline{X}_1^n + \underline{W}_1^n + \tilde{\underline{V}}_1^n) \\ &= -I(\tilde{\underline{V}}_1^n; H_{21}\underline{X}_1^n + \underline{W}_1^n + \tilde{\underline{V}}_1^n) \\ &\stackrel{(e)}{\leq} -nI(\tilde{\underline{V}}_1; H_{21}\underline{X}_{1G} + \underline{W}_1 + \tilde{\underline{V}}_1).\end{aligned} \quad (46)$$

where  $\underline{V}_1$  and  $\tilde{\underline{V}}_1$  are Gaussian random vectors independent of every other random vector with the following covariance matrices:

$$\begin{aligned}\underline{V}_1 &\sim \mathcal{N}(\underline{0}, I - A_2 \Sigma_2^{-1} A_2^\top) \\ \tilde{\underline{V}}_1 &\sim \mathcal{N}(\underline{0}, I - A_2 \Sigma_2^{-1} A_2^\top - \Sigma_1)\end{aligned}$$

Step (c) follows from Lemma 9, step (d) follows because  $\underline{W}_1 + \underline{V}_1$  has same distribution as  $\underline{V}_1$  and step (e) follows from Lemma 8. Similarly

$$\begin{aligned}h(\underline{S}_2^n) - h(\underline{Y}_2^n | \underline{S}_2^n, \underline{X}_2^n) &\leq -nI(\tilde{\underline{V}}_2; H_{12}\underline{X}_{2G} + \underline{W}_2 + \tilde{\underline{V}}_2).\end{aligned} \quad (47)$$

The conditions in (12) are required for the covariance matrices of  $\tilde{\underline{V}}_1$  and  $\tilde{\underline{V}}_2$  to be positive semi-definite. Substituting (46) and (47) in (45), we get

$$\begin{aligned}C_{\text{sum}}^{\text{GA-IC}} &\leq \sum_{i=1}^2 h(\underline{Y}_{iG} | \underline{S}_{iG}) - h(\underline{W}_i) \\ &\quad - I(\tilde{\underline{V}}_1; H_{21}\underline{X}_{1G} + \underline{W}_1 + \tilde{\underline{V}}_1) \\ &\quad - I(\tilde{\underline{V}}_2; H_{12}\underline{X}_{2G} + \underline{W}_2 + \tilde{\underline{V}}_2)\end{aligned}$$

Since all the inequalities in (45), (46) and (47) are met equality when  $\underline{X}_i^n = \underline{X}_{iG}^n$ , we also have

$$\begin{aligned}R_{\text{TIN}}^{\text{GA-IC}}(Q_1, Q_2) &= \sum_{i=1}^2 I(\underline{X}_{iG}; \underline{Y}_{iG}, \underline{S}_{iG}) \\ &= \sum_{i=1}^2 h(\underline{Y}_{iG} | \underline{S}_{iG}) - h(\underline{W}_i) \\ &\quad - I(\tilde{\underline{V}}_1; H_{21}\underline{X}_{1G} + \underline{W}_1 + \tilde{\underline{V}}_1) \\ &\quad - I(\tilde{\underline{V}}_2; H_{12}\underline{X}_{2G} + \underline{W}_2 + \tilde{\underline{V}}_2).\end{aligned} \quad (48)$$

Therefore, we proved that

$$C_{\text{sum}}^{\text{GA-IC}} \leq R_{\text{TIN}}^{\text{GA-IC}}(Q_1, Q_2), \text{ for some } Q_i \in Q_i$$

and hence

$$C_{\text{sum}}^{\text{GA-IC}} \leq \max_{Q_i \in Q_i} R_{\text{TIN}}^{\text{GA-IC}}(Q_1, Q_2).$$

We now prove that  $R_{\text{TIN}}^{\text{GA-IC}}(Q_1, Q_2)$  is concave in  $(Q_1, Q_2)$  using the expression (48). From Lemma 10, it immediately follows that  $h(\underline{Y}_{1G} | \underline{S}_{1G})$  and  $h(\underline{Y}_{2G} | \underline{S}_{2G})$  are concave in  $(Q_1, Q_2)$ . Also, observe that  $-I(\tilde{V}_2; H_{12}\underline{X}_{2G} + \underline{W}_2 + \tilde{V}_2)$  is concave in  $(Q_1, Q_2)$  since

$$\begin{aligned} & -I(\tilde{V}_2; H_{12}\underline{X}_{2G} + \underline{W}_2 + \tilde{V}_2) \\ &= -h(\tilde{V}_2) + h(\tilde{V}_2 | H_{12}\underline{X}_{2G} + \underline{W}_2 + \tilde{V}_2) \end{aligned}$$

where the first term does not depend on  $(Q_1, Q_2)$  and Lemma 10 implies the concavity of the second term in  $(Q_1, Q_2)$ . Similarly,  $-I(\tilde{V}_1; H_{21}\underline{X}_{1G} + \underline{W}_1 + \tilde{V}_1)$  is also concave in  $(Q_1, Q_2)$ .

### C. Proof of Lemma 4

$$\begin{aligned} R_{\text{TIN}}^{\text{GA-IC}}(Q_1, Q_2, \Psi) &= R_{\text{TIN}}^{\text{IC}}(Q_1, Q_2) \\ \Leftrightarrow I(\underline{X}_{iG}; \underline{Y}_{iG}, \underline{S}_{iG}) &= I(\underline{X}_{iG}; \underline{Y}_{iG}) \\ \Leftrightarrow I(\underline{X}_{iG}; \underline{S}_{iG} | \underline{Y}_{iG}) &= 0. \end{aligned}$$

Since  $\underline{X}_{1G}, \underline{Y}_{1G}$  and  $\underline{S}_{1G}$  are all Gaussian, from [7, Lemma 7],  $I(\underline{X}_{1G}; \underline{S}_{1G} | \underline{Y}_{1G}) = 0$  iff the MMSE estimate of  $\underline{S}_{1G}$  given  $(\underline{X}_{1G}, \underline{Y}_{1G})$  is equal to the MMSE estimate of  $\underline{S}_{1G}$  given  $\underline{Y}_{1G}$ . Using the property that the MMSE error is orthogonal to the observations  $\underline{X}_{1G}$  and  $\underline{Y}_{1G}$ , the above statement holds true iff there exist a matrix  $T$  such that

$$\begin{aligned} \mathbb{E}[(T\underline{Y}_{1G} - \underline{S}_{1G})\underline{X}_{1G}^\top] &= 0 \\ \mathbb{E}[(T\underline{Y}_{1G} - \underline{S}_{1G})\underline{Y}_{1G}^\top] &= 0. \end{aligned}$$

Since  $\underline{Y}_{1G} = H_{11}\underline{X}_{1G} + H_{12}\underline{X}_{2G} + \underline{Z}_1$ , the above equations are true iff

$$\begin{aligned} \mathbb{E}[(T\underline{Y}_{1G} - \underline{S}_{1G})\underline{X}_{1G}^\top] &= 0 \\ \mathbb{E}[(T\underline{Y}_{1G} - \underline{S}_{1G})(H_{12}\underline{X}_{2G} + \underline{Z}_1)^\top] &= 0 \end{aligned}$$

iff

$$\begin{aligned} (TH_{11} - H_{21})Q_1 &= 0 \\ T(H_{12}Q_2H_{12}^\top + I) - A_1^\top &= 0 \end{aligned}$$

Solving for  $T$  and substituting, we get

$$(A_1^\top(H_{12}Q_2H_{12}^\top + I)^{-1}H_{11} - H_{21})Q_1 = 0.$$

### D. Sufficiency of Two Antennas for the Two-User MISO and SIMO Interference Channels

We argue that, without any loss of generality, we can assume two transmit antennas and

$$\begin{aligned} \underline{c} &= [1 \ 0]^\top \\ \underline{d} &= [\cos \theta \ \sin \theta]^\top \end{aligned}$$

for the MISO interference channel (21). The argument for the sufficiency of two receive antennas for the SIMO interference channel follows in a similar fashion. Even though the argument is presented for the symmetric channel considered in this paper, the same applies for the asymmetric case as well.  $\blacksquare$

Consider the vector  $\underline{c}_\perp$  which is perpendicular to  $\underline{c}$  and is given by

$$\underline{c}_\perp = \frac{\underline{d} - \cos \theta \underline{c}}{\sin \theta}.$$

If  $\theta = 0$ , use any vector perpendicular to  $\underline{c}$  as  $\underline{c}_\perp$ . Now let  $U$  be an unitary matrix  $\underline{c}$  as its first column  $\underline{c}_\perp$  as its second column, and define  $\tilde{X}_i$  as

$$\tilde{X}_i = U^\top X_i$$

or equivalently  $X_i = U\tilde{X}_i$ .

Therefore, the MISO interference channel (21) can be re-written as

$$\begin{aligned} Y_1 &= \underline{d}^\top U\tilde{X}_1 + h\underline{c}^\top U\tilde{X}_2 + \underline{Z}_1 \\ &= \tilde{d}^\top \tilde{X}_1 + h\tilde{c}^\top \tilde{X}_2 + \underline{Z}_1 \end{aligned}$$

where

$$\begin{aligned} \tilde{c} &= [1 \ 0 \ 0 \ \dots \ 0]^\top \\ \tilde{d} &= [\cos \theta \ \sin \theta \ 0 \ \dots \ 0]^\top \end{aligned}$$

The interference channel outputs  $Y_1$  and  $Y_2$  do not depend on channel inputs  $\tilde{X}_{i3}$  to  $\tilde{X}_{iM}$ ,  $i = 1, 2$ , and hence, they can be dropped. Also, since  $U$  is the unitary matrix, the average sum power constraint remains the same.

### E. Alternative Expressions for MISO Genie Parameters

Let  $J$  denote the matrix  $I + h^2 P \underline{c} \underline{c}^\top$ . Therefore, from the matrix inversion Lemma, we have

$$\begin{aligned} J^{-1}\underline{d} &= \left( I - \frac{h^2 P}{1 + h^2 P \underline{c} \underline{c}^\top} \right) \underline{d} = \underline{d} - \frac{h^2 P \cos \theta}{1 + h^2 P} \underline{c} \\ \underline{c}^\top J^{-1}\underline{d} &= \cos \theta - \frac{h^2 P \cos \theta}{1 + h^2 P} = \frac{\cos \theta}{1 + h^2 P} \\ \underline{d}^\top J^{-1}\underline{d} &= 1 - \frac{h^2 P \cos^2 \theta}{1 + h^2 P} = \frac{1 + h^2 P \sin^2 \theta}{1 + h^2 P}. \end{aligned} \quad (49)$$

We will now provide alternative expressions for the parameters involved in MISO interference channel. These relations will be used in Appendices F and G.

*Claim 1:* The beamforming vector  $\underline{b}$  defined in (22) can be expressed as

$$\underline{b} = \frac{\underline{d}}{\|J^{-1}\underline{d}\|} - h^2 P \underline{b}^\top \underline{c} \underline{c}. \quad (50)$$

*Proof:* Using the notation  $J = I + h^2 P \underline{c} \underline{c}^\top$  in (22), we get

$$\underline{b} = \frac{J^{-1}\underline{d}}{\|J^{-1}\underline{d}\|}.$$

Therefore, we have

$$\begin{aligned} J\underline{b} &= \frac{\underline{d}}{\|J^{-1}\underline{d}\|} \\ \underline{b} + h^2 P \underline{b}^\top \underline{c} \underline{c} &= \frac{\underline{d}}{\|J^{-1}\underline{d}\|} \\ \underline{b} &= \frac{\underline{d}}{\|J^{-1}\underline{d}\|} - h^2 P \underline{b}^\top \underline{c} \underline{c}. \end{aligned}$$

*Claim 2:* The genie parameter  $a$  defined in (28) can be expressed as

$$\begin{aligned} a &= h \frac{\underline{b}^\top \underline{c}}{\underline{b}^\top \underline{d}} \underline{b}^\top J \underline{b} = h \frac{\underline{b}^\top \underline{c}}{\|\underline{J}^{-1} \underline{d}\|} = h \frac{\underline{c}^\top \underline{J}^{-1} \underline{d}}{\underline{d}^\top \underline{J}^{-2} \underline{d}} \\ &= \frac{h(1+h^2P)\cos\theta}{\cos^2\theta + (1+h^2P)^2\sin^2\theta}. \end{aligned} \quad (51)$$

*Proof:*

$$\begin{aligned} a &= h \frac{\underline{b}^\top \underline{c}}{\underline{b}^\top \underline{d}} \left(1 + h^2P |\underline{b}^\top \underline{c}|^2\right) \\ &= h \frac{\underline{b}^\top \underline{c}}{\underline{b}^\top \underline{d}} \underline{b}^\top J \underline{b} \\ &= h \frac{\underline{b}^\top \underline{c}}{\|\underline{J}^{-1} \underline{d}\|} \\ &= h \frac{\underline{c}^\top \underline{J}^{-1} \underline{d}}{\|\underline{J}^{-1} \underline{d}\|^2} \\ &= h \frac{\underline{c}^\top \underline{J}^{-1} \underline{d}}{\underline{d}^\top \underline{J}^{-2} \underline{d}} \end{aligned}$$

which, using (49), can be further simplified as

$$\begin{aligned} a &= h \frac{\frac{\cos\theta}{1+h^2P}}{\frac{\cos^2\theta}{(1+h^2P)^2} + \sin^2\theta} \\ &= \frac{h(1+h^2P)\cos\theta}{\cos^2\theta + (1+h^2P)^2\sin^2\theta}. \end{aligned}$$

*Claim 3:* The parameter  $\mu$  defined in (33) can be expressed as

$$\mu = \frac{1 - \Sigma + h^2P(\underline{b}^\top \underline{c})^2}{a} = \frac{\underline{b}^\top \underline{d}}{h\underline{b}^\top \underline{c}} - \frac{\Sigma}{a} \quad (52)$$

*Proof:* First, observe that

$$\begin{aligned} &(\Sigma + h^2P(\underline{b}^\top \underline{c})^2)(1 - \Sigma + h^2P(\underline{b}^\top \underline{c})^2) \\ &= \Sigma(1 - \Sigma) + h^2P(\underline{b}^\top \underline{c})^2 + (h^2P(\underline{b}^\top \underline{c})^2)^2 \\ &= a^2 + h^2P(\underline{b}^\top \underline{c})^2(1 + h^2P(\underline{b}^\top \underline{c})^2) \\ &= a^2 + hP\underline{b}^\top \underline{d} \underline{b}^\top \underline{c} \frac{h\underline{b}^\top \underline{c}}{\underline{b}^\top \underline{d}} (1 + h^2P(\underline{b}^\top \underline{c})^2) \\ &= a^2 + ahP\underline{b}^\top \underline{d} \underline{b}^\top \underline{c} \\ &= a(a + hP\underline{b}^\top \underline{d} \underline{b}^\top \underline{c}). \end{aligned}$$

Using the above equality in (33), we get

$$\begin{aligned} \mu &= \frac{a + hP\underline{b}^\top \underline{d} \underline{b}^\top \underline{c}}{\Sigma + h^2P(\underline{b}^\top \underline{c})^2} \\ &= \frac{1 - \Sigma + h^2P(\underline{b}^\top \underline{c})^2}{a} \\ &= \frac{\underline{b}^\top J \underline{b} - \Sigma}{a} \\ &= \frac{\underline{b}^\top \underline{d}}{h\underline{b}^\top \underline{c}} - \frac{\Sigma}{a}. \end{aligned}$$

where the last step uses (51).

*F. Proof of (29) and (34)*

*Lemma 12:* For any  $P, h$  and  $\theta \in (0, \frac{\pi}{2})$  such that

$$h < h_0(\theta, P)$$

where  $h_0(\theta, P)$ , as defined in (26), is the positive solution to the equation

$$h_0^2 - \sin^2\theta = \left(\frac{\cos\theta}{1+h_0^2P} - h_0\right)_+^2$$

the following conditions are satisfied:

$$\begin{aligned} a &< 0.5 \\ \mu h &< \cos\theta \end{aligned} \quad (53)$$

where  $a$  and  $\mu$  are as defined in (28) and (33), respectively.

*Proof:* First, observe that  $h < h_0(\theta, P)$  implies

$$\begin{aligned} h^2 - \sin^2\theta &< h_0^2 - \sin^2\theta \\ &= \left(\frac{\cos\theta}{1+h_0^2P} - h_0\right)_+^2 \\ &< \left(\frac{\cos\theta}{1+h^2P} - h\right)_+^2 \\ &< \left(\frac{\cos\theta}{1+h^2P} - h\right)^2 \end{aligned} \quad (54)$$

which is equivalent to

$$\begin{aligned} \frac{\cos^2\theta}{(1+h^2P)^2} - \frac{2h\cos\theta}{1+h^2P} + \sin^2\theta &> 0 \\ \cos^2\theta - 2h(1+h^2P)\cos\theta + (1+h^2P)\sin^2\theta &> 0 \\ \Rightarrow a = \frac{2h(1+h^2P)\cos\theta}{\cos^2\theta + (1+h^2P)\sin^2\theta} &< 0.5 \end{aligned}$$

where we used the expression for  $a$  from Claim 2 in Appendix E. Now it remains to prove that  $\mu h < \cos\theta$  is also satisfied. From Claim 3 in Appendix E, we have

$$\mu = \frac{\underline{b}^\top \underline{d}}{h\underline{b}^\top \underline{c}} - \frac{\Sigma}{a}$$

and thus

$$\begin{aligned} \mu h - \cos\theta &= \frac{\underline{b}^\top \underline{d}}{\underline{b}^\top \underline{c}} - \frac{h\Sigma}{a} - \cos\theta \\ &= \frac{\underline{d}^\top \underline{J}^{-1} \underline{d}}{\underline{c}^\top \underline{J}^{-1} \underline{d}} - \frac{h\Sigma}{a} - \cos\theta \\ &= \frac{1+h^2P\sin^2\theta}{\cos\theta} - \frac{h\Sigma}{a} - \cos\theta \\ &= \frac{(1+h^2P)\sin^2\theta}{\cos\theta} - \frac{h\Sigma}{a} \end{aligned}$$

where  $J$  is used to denote the matrix  $I + h^2P\underline{c}\underline{c}^\top$ . If  $h$  satisfies (54), then one of the following cases has to be true.

Case 1:  $h$  satisfies

$$h \leq \sin \theta.$$

Observe that

$$\begin{aligned} \frac{h\Sigma}{a} &= h \frac{0.5 + 0.5\sqrt{1-4a^2}}{a} \\ &= \frac{h}{2} \left( \frac{1}{a} + \sqrt{\frac{1}{a^2} - 4} \right). \end{aligned}$$

Using the expression

$$a = \frac{h(1+h^2P)\cos\theta}{\cos^2\theta + (1+h^2P)^2\sin^2\theta}$$

from Claim 2 in Appendix E, we have

$$\begin{aligned} \frac{1}{a^2} - 4 &= \left( \frac{\cos^2\theta + (1+h^2P)^2\sin^2\theta}{h(1+h^2P)\cos\theta} \right)^2 - 4 \\ &\stackrel{(b)}{\geq} \left( \frac{\cos^2\theta + (1+h^2P)^2\sin^2\theta}{h(1+h^2P)\cos\theta} \right)^2 \\ &\quad - 4 \left( \frac{\sin\theta(1+h^2P)\cos\theta}{h(1+h^2P)\cos\theta} \right)^2 \\ &= \left( \frac{\cos^2\theta - (1+h^2P)^2\sin^2\theta}{h(1+h^2P)\cos\theta} \right)^2 \end{aligned}$$

where step (b) follows because  $h < \sin \theta$ . Therefore, we have

$$\frac{h\Sigma}{a} = \frac{h}{2} \left( \frac{1}{a} + \sqrt{\frac{1}{a^2} - 4} \right) \geq \frac{(1+h^2P)\sin^2\theta}{\cos\theta}$$

and thus

$$\mu h - \cos\theta = \frac{(1+h^2P)\sin^2\theta}{\cos\theta} - \frac{h\Sigma}{a} \leq 0.$$

Case 2:  $h$  satisfies

$$\frac{\cos\theta}{1+h^2P} - h \geq 0. \quad (55)$$

Note that

$$\Sigma = 0.5 + 0.5\sqrt{1-a^2} > 0.5 > a.$$

Hence, we have

$$\begin{aligned} \mu h - \cos\theta &= \frac{(1+h^2P)\sin^2\theta}{\cos\theta} - \frac{h\Sigma}{a} \\ &< \frac{(1+h^2P)\sin^2\theta}{\cos\theta} - h \\ &\stackrel{(a)}{\leq} \frac{(1+h^2P)\sin^2\theta}{\cos\theta} - \frac{\sin^2\theta}{h} \\ &= \sin^2\theta \left( \frac{1+h^2P}{\cos\theta} - \frac{1}{h} \right) \\ &\stackrel{(b)}{\leq} 0 \end{aligned}$$

where steps (a) and (b) follow from (55). ■

G. Proof of (35)

In this Appendix, we show that if

$$\mu h < \cos\theta \quad (56)$$

then  $\text{tr}(MQ) \leq \text{tr}(MQ^*)$ ,  $\forall Q \in \mathcal{Q}$ , where  $\mu$  and  $M$  are defined in (33) and (32), respectively. For all  $Q \in \mathcal{Q}$ , we have

$$\begin{aligned} \text{tr}(MQ) &\leq \text{tr}(Q)\lambda_{\max}(M) \\ &\leq P\lambda_{\max}(M) \\ &\stackrel{(b)}{=} P\mathbf{b}^\top M\mathbf{b} \\ &= \text{tr}(MP\mathbf{b}\mathbf{b}^\top) \\ &= \text{tr}(MQ^*). \end{aligned} \quad (57)$$

Step (b) follows because  $\mathbf{b}$  is the eigenvector of the matrix  $M$  corresponding to the maximum eigenvalue  $\lambda_{\max}(M)$ . In proving that  $\mathbf{b}$  is an eigenvector of the matrix  $M$ , we need the following relation.

Claim 4:

$$\begin{aligned} (\mathbf{b}^\top \mathbf{d} - \mu h \mathbf{b}^\top \mathbf{c})\mu h - h^2 \mathbf{b}^\top \mathbf{c} \\ = h^2 P \mathbf{b}^\top \mathbf{c} (\mathbf{b}^\top \mathbf{d} - \mu h \mathbf{b}^\top \mathbf{c}) \|J^{-1} \mathbf{d}\|. \end{aligned} \quad (58)$$

*Proof:* Using the two different expressions for  $\mu$  in Claim 3 in Appendix E, we get

$$\left( \frac{\mathbf{b}^\top \mathbf{d}}{h \mathbf{b}^\top \mathbf{c}} - \mu \right) \left( \mu - \frac{h^2 P (\mathbf{b}^\top \mathbf{c})^2}{a} \right) = \frac{\Sigma(1-\Sigma)}{a^2} = 1.$$

Using (51), we get

$$\begin{aligned} \left( \frac{\mathbf{b}^\top \mathbf{d}}{h \mathbf{b}^\top \mathbf{c}} - \mu \right) (\mu - P \|J^{-1} \mathbf{d}\| h \mathbf{b}^\top \mathbf{c}) &= 1 \\ (\mathbf{b}^\top \mathbf{d} - \mu h \mathbf{b}^\top \mathbf{c})(\mu h - \|J^{-1} \mathbf{d}\| h^2 P \mathbf{b}^\top \mathbf{c}) &= h^2 \mathbf{b}^\top \mathbf{c}. \end{aligned}$$

We now prove that  $\mathbf{b}$  is an eigenvector of the matrix  $M$

$$\begin{aligned} M\mathbf{b} &= ((\mathbf{d} - \mu h \mathbf{c})(\mathbf{d} - \mu h \mathbf{c})^\top + h^2 \mathbf{c}\mathbf{c}^\top)\mathbf{b} \\ &= (\mathbf{b}^\top \mathbf{d} - \mu h \mathbf{b}^\top \mathbf{c})(\mathbf{d} - \mu h \mathbf{c}) + h^2 \mathbf{b}^\top \mathbf{c}\mathbf{c} \\ &= (\mathbf{b}^\top \mathbf{d} - \mu h \mathbf{b}^\top \mathbf{c})\mathbf{d} - ((\mathbf{b}^\top \mathbf{d} - \mu h \mathbf{b}^\top \mathbf{c})\mu h - h^2 \mathbf{b}^\top \mathbf{c})\mathbf{c} \\ &\stackrel{(a)}{=} (\mathbf{b}^\top \mathbf{d} - \mu h \mathbf{b}^\top \mathbf{c})\mathbf{d} - h^2 P \mathbf{b}^\top \mathbf{c} (\mathbf{b}^\top \mathbf{d} - \mu h \mathbf{b}^\top \mathbf{c}) \|J^{-1} \mathbf{d}\| \mathbf{c} \\ &= (\mathbf{b}^\top \mathbf{d} - \mu h \mathbf{b}^\top \mathbf{c}) \|J^{-1} \mathbf{d}\| \left( \frac{\mathbf{d}}{\|J^{-1} \mathbf{d}\|} - h^2 P \mathbf{b}^\top \mathbf{c}\mathbf{c} \right) \\ &\stackrel{(b)}{=} (\mathbf{b}^\top \mathbf{d} - \mu h \mathbf{b}^\top \mathbf{c}) \|J^{-1} \mathbf{d}\| \mathbf{b}. \end{aligned}$$

where step (a) follows from Claim 4, and step (b) follows from Claim 1 in Appendix E. From Lemma 11, it follows that  $\mathbf{b}$  corresponds to the largest eigenvalue of the matrix  $M = (\mathbf{d} - \mu h \mathbf{c})(\mathbf{d} - \mu h \mathbf{c})^\top + h^2 \mathbf{c}\mathbf{c}^\top$ . The conditions required in Lemma 11 are shown below

$$\begin{aligned} (\mathbf{d} - \mu h \mathbf{c})^\top \mathbf{c} &= \cos\theta - \mu h \\ &\stackrel{(a)}{>} 0. \end{aligned}$$

Also

$$\begin{aligned} \underline{b}^\top \underline{c} &= \frac{\underline{c}^\top J^{-1} \underline{d}}{\|J^{-1} \underline{d}\|} \\ &\stackrel{(b)}{=} \frac{\cos \theta}{(1 + h^2 P) \|J^{-1} \underline{d}\|} \\ &\stackrel{(c)}{>} 0 \end{aligned}$$

and

$$\begin{aligned} (\underline{d} - \mu h \underline{c})^\top \underline{b} &= \frac{(\underline{d} - \mu h \underline{c})^\top J^{-1} \underline{d}}{\|J^{-1} \underline{d}\|} \\ &\stackrel{(d)}{=} \frac{1 + h^2 P \sin^2 \theta - \mu h \cos \theta}{(1 + h^2 P) \|J^{-1} \underline{d}\|} \\ &\stackrel{(e)}{>} \frac{1 + h^2 P \sin^2 \theta - \cos^2 \theta}{(1 + h^2 P) \|J^{-1} \underline{d}\|} \\ &= \frac{\sin^2 \theta + h^2 P \sin^2 \theta}{(1 + h^2 P) \|J^{-1} \underline{d}\|} \\ &= \frac{\sin^2 \theta}{\|J^{-1} \underline{d}\|} \\ &\stackrel{(f)}{>} 0 \end{aligned}$$

where steps (a) and (e) follow from the hypothesis (56), steps (b) and (d) follow from (49), steps (c) and (f) follow because of our assumption that  $\theta \in (0, \frac{\pi}{2})$ .

#### H. Proof of Lemma 5

The proof is very similar to the proof of Lemma 3 in Appendix B. The only difference is steps (46) and (47), where we showed that  $h(\underline{S}_1^n) - h(\underline{Y}_2^n | \underline{S}_2^n, \underline{X}_2^n)$  and  $h(\underline{S}_2^n) - h(\underline{Y}_1^n | \underline{S}_1^n, \underline{X}_1^n)$  are maximized by  $X_{1G}^n$  and  $X_{2G}^n$ , respectively. Here we show that the condition (12) is not necessary and a weaker condition

$$(\underline{c}^\top \Sigma^{-1} \underline{c})^{-1} \leq (\underline{c}^\top (I - A \Sigma^{-1} A^\top)^{-1} \underline{c})^{-1}$$

is sufficient for (46) and (47) to be true. We give the proof assuming  $\underline{c} = [1 \ 0]^\top$  to simplify the presentation, but can be extended to arbitrary  $\underline{c}$  using the arguments in Appendix D. Following (46), we have

$$\begin{aligned} h(\underline{S}_1^n) - h(\underline{Y}_2^n | \underline{S}_2^n, \underline{X}_2^n) \\ = h(h \underline{c} X_1^n + \underline{W}_1^n) - h(h \underline{c} X_1^n + \underline{V}_1^n) \end{aligned}$$

where  $\underline{W}_1 \sim \mathcal{N}(\underline{0}, \Sigma)$  and  $\underline{V}_1 \sim \mathcal{N}(\underline{0}, I - A \Sigma^{-1} A^\top)$ . Now

$$\begin{aligned} h(h \underline{c} X_1^n + \underline{W}_1^n) &= h(h X_1^n + W_{11}^n, W_{12}^n) \\ &= h(h X_1^n + W_{11}^n | W_{12}^n) + h(W_{12}^n) \\ &= h(h X_1^n + \hat{W}_1^n) + nh(W_{12}) \end{aligned}$$

where  $\hat{W}_1 \sim \mathcal{N}(0, \text{Cov}(W_{11} | W_{12}))$ . Let

$$\Sigma = \text{Cov}(\underline{W}_1) = \begin{bmatrix} a_1 & b \\ b & a_2 \end{bmatrix}$$

then

$$\Sigma^{-1} = \frac{1}{a_1 a_2 - b^2} \begin{bmatrix} a_2 & -b \\ -b & a_1 \end{bmatrix}$$

and since  $\underline{c} = [1 \ 0]^\top$ , we have

$$\begin{aligned} (\underline{c}^\top \Sigma^{-1} \underline{c})^{-1} &= \frac{a_1 a_2 - b^2}{a_2} \\ &= a_1 - \frac{b^2}{a_2} \\ &= \text{Cov}(W_{11} | W_{12}) = \text{Cov}(\hat{W}_1). \end{aligned}$$

Similarly

$$h(h \underline{c} X_1^n + \underline{V}_1^n) = h(h X_1^n + \hat{V}_1^n) + nh(V_{12})$$

where

$$\hat{V}_1 \sim \mathcal{N}(0, (\underline{c}^\top (I - A \Sigma^{-1} A^\top)^{-1} \underline{c})^{-1}).$$

Therefore

$$\begin{aligned} h(\underline{S}_1^n) - h(\underline{Y}_2^n | \underline{S}_2^n, X_2^n) \\ = h(h \underline{c} X_1^n + \underline{W}_1^n) - h(h \underline{c} X_1^n + \underline{V}_1^n) \\ = h(h X_1^n + \hat{W}_1^n) - h(h X_1^n + \hat{V}_1^n) \\ + nh(W_{12}) - nh(V_{12}) \\ \leq nh(h X_{1G} + \hat{W}_1) - nh(h X_{1G} + \hat{V}_1) \\ + nh(W_{12}) - nh(V_{12}). \end{aligned}$$

We have thus shown that  $h(\underline{S}_1^n) - h(\underline{Y}_2^n | \underline{S}_2^n, X_2^n)$  is maximized by  $X_{1G}^n$  and similarly we can show that  $h(\underline{S}_2^n) - h(\underline{Y}_1^n | \underline{S}_1^n, X_1^n)$  is maximized by  $X_{2G}^n$ .

#### I. Existence of SIMO Genie Parameters

Here we show that if  $h \leq h_0(\theta, P)$  where  $h_0(\theta, P)$  is the positive solution of

$$h^2 - \sin^2 \theta = \left( \frac{\cos \theta}{1 + h^2 P} - h \right)_+^2.$$

then there exist  $k$ ,  $\eta$  and  $\underline{v}$  satisfying the conditions

$$\begin{aligned} \eta &\geq k^2 \\ \eta &\leq \frac{\eta - k^2}{\eta - k^2 (1 - (\underline{c}^\top \underline{v})^2)} \\ h &= k \underline{v}^\top (I + h^2 P \underline{c} \underline{c}^\top)^{-1} \underline{d}. \end{aligned} \quad (59)$$

First, we eliminate  $\eta$ , and obtain conditions on  $k$  and  $\underline{v}$ . Let  $t = \eta - k^2$ , then the second condition in (59) is equivalent to

$$\begin{aligned} t + k^2 &\leq \frac{t}{t + k^2 (\underline{c}^\top \underline{v})^2} \\ &\Rightarrow t^2 + t(k^2 (1 + (\underline{c}^\top \underline{v})^2) - 1) + k^4 (\underline{c}^\top \underline{v})^2 \leq 0. \end{aligned}$$

The following claim can be easily verified.

*Claim 5:* A non-negative solution ( $t \geq 0$ ) to the inequality  $t^2 + bt + c \leq 0$ , with  $c \geq 0$ , exists iff  $b \leq -2\sqrt{c}$ .

Applying the above claim, a valid  $t$  exists iff  $k$  and  $\underline{v}$  satisfy

$$\begin{aligned} k^2(1 + (\underline{c}^\top \underline{v})^2) - 1 &\leq -2|\underline{c}^\top \underline{v}| k^2 \\ \Rightarrow k^2(1 + |\underline{c}^\top \underline{v}|)^2 &\leq 1 \\ \Rightarrow k(1 + |\underline{c}^\top \underline{v}|) &\leq 1. \end{aligned}$$

Now, using the last condition in (59) and eliminating  $k$ , we get

$$h(1 + |\underline{c}^\top \underline{v}|) \leq \underline{v}^\top (I + h^2 P \underline{c} \underline{c}^\top)^{-1} \underline{d}.$$

Without any loss of generality, let  $\underline{v} = [\alpha \beta]^\top$  such that  $\alpha^2 + \beta^2 = 1$ . Using the assumptions that  $\underline{c} = [1 \ 0]^\top$  and  $\underline{d} = [\cos \theta \ \sin \theta]^\top$ , the above inequality is equivalent to

$$h(1 + |\alpha|) \leq \alpha \frac{\cos \theta}{1 + h^2 P} + \beta \sin \theta.$$

Observe that we can restrict ourselves to only positive values for  $\alpha$ , because if a negative  $\alpha$  works then the corresponding positive  $\alpha$  works as well. Thus, we reduced the problem to finding the conditions on  $h$ ,  $\theta$  and  $P$ , so that there exist  $\alpha > 0$  and  $\beta$  exist satisfying  $\alpha^2 + \beta^2 = 1$

$$h \leq \alpha \left( \frac{\cos \theta}{1 + h^2 P} - h \right) + \beta \sin \theta.$$

Now, observe that if  $\frac{\cos \theta}{1 + h^2 P} - h < 0$ , then the best is to set  $\alpha = 0$  and  $\beta = 1$ , giving rise to the condition

$$h \leq \sin \theta.$$

Otherwise, if  $\frac{\cos \theta}{1 + h^2 P} - h \geq 0$ , then the best is choose  $\alpha$  and  $\beta$  proportional to  $\frac{\cos \theta}{1 + h^2 P} - h$  and  $\sin \theta$ , respectively, giving rise to the condition

$$h^2 \leq \left( \frac{\cos \theta}{1 + h^2 P} - h \right)^2 + \sin^2 \theta.$$

Thus, we proved that the genie parameters  $\eta, k, \underline{v}$  exist satisfying (59) iff  $h, \theta$  and  $P$  satisfy

$$h^2 - \sin^2 \theta \leq \left( \frac{\cos \theta}{1 + h^2 P} - h \right)_+^2.$$

It is easy to verify that the above is a threshold condition on  $h$ :  $h \leq h_0(\theta, P)$ .

#### ACKNOWLEDGMENT

The authors would like to thank S. Vishwanath for helping formulate the MIMO interference channel problem, particularly in the symmetric MISO setting. They would also like to thank A. Khisti for useful discussions.

#### REFERENCES

- [1] A. B. Carleial, "A case where interference does not reduce capacity," *IEEE Trans. Inf. Theory*, vol. IT-21, no. 1, pp. 569–570, Sep. 1975.
- [2] T. S. Han and K. Kobayashi, "A new achievable rate region for the interference channel," *IEEE Trans. Inf. Theory*, vol. IT-27, no. 1, pp. 49–60, Jan. 1981.
- [3] H. Sato, "The capacity of the Gaussian interference channel under strong interference," *IEEE Trans. Inf. Theory*, vol. IT-27, no. 6, pp. 786–788, Nov. 1981.
- [4] 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.
- [5] X. Shang, G. Kramer, and B. Chen, "A new outer bound and noisy-interference sum-rate capacity for the Gaussian interference channels," *IEEE Trans. Inf. Theory*, vol. 55, no. 2, pp. 689–699, Feb. 2009.
- [6] 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.
- [7] 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. 6, pp. 3032–3050, Jun. 2009.
- [8] S. Vishwanath and S. A. Jafar, "On the capacity of vector Gaussian interference channels," in *Proc. IEEE Information Theory Workshop (ITW)*, 2004, pp. 365–369.
- [9] X. Shang, B. Chen, and M. J. Gans, "On the achievable sum rate for MIMO interference channels," *IEEE Trans. Inf. Theory*, vol. 52, no. 9, pp. 4313–4320, Sep. 2006.
- [10] E. Telatar and D. Tse, "Bounds on the capacity region of a class of interference channels," in *Proc. IEEE Int. Symp. Information Theory (ISIT)*, Nice, France, Jun. 2007, pp. 2871–2874.
- [11] X. Shang, B. Chen, G. Kramer, and H. V. Poor, "On the capacity of MIMO interference channels," presented at the 46th Annu. Allerton Conf. Commun. Cont. Comp., Oct. 2008, Univ. Illinois.
- [12] X. Shang, B. Chen, G. Kramer, and H. V. Poor, "Capacity regions and sum-rate capacities of vector Gaussian interference channels," *IEEE Trans. Inf. Theory*, vol. 56, no. 10, pp. 5030–5044, Oct. 2010.
- [13] X. Shang, B. Chen, G. Kramer, and H. V. Poor, "Noisy-interference sum-rate capacity of parallel Gaussian interference channels," *IEEE Trans. Inf. Theory*, vol. 57, no. 1, pp. 210–226, Jan. 2011.
- [14] B. Bandemer, A. Sezgin, and A. Paulraj, "On the noisy interference regime of the MISO Gaussian interference channel," in *Proc. 42nd Asilomar Conf. Signals, Systems and Computers*, 2008, pp. 1098–1102.
- [15] X. Shang, B. Chen, and H. V. Poor, "Multi-user MISO interference channels with single-user detection: Optimality of beamforming and the achievable rate region," *IEEE Trans. Inf. Theory* [Online]. Available: <http://arxiv.org/abs/0907.0505>, to be published
- [16] R. Horn and C. R. Johnson, *Matrix Analysis*. New York: Cambridge Univ. Press, 1985.
- [17] S. Boyd and L. Vandenberghe, *Convex Optimization*. New York: Cambridge Univ. Press, 2006.
- [18] E. A. Jorswieck, E. G. Larsson, and D. Danev, "Complete characterization of the pareto boundary for the MISO interference channel," *IEEE Trans. Signal Process.*, vol. 56, no. 10, pp. 5292–5296, Oct. 2008.
- [19] R. Zakhour and D. Gesbert, "Coordination on the MISO interference channel using the virtual SINR framework," presented at the Int. ITG Workshop on Smart Antennas, 2009.
- [20] J. A. Thomas, "Feedback can at most double Gaussian multiple access channel capacity," *IEEE Trans. Inf. Theory*, vol. 33, no. 5, pp. 711–716, Sep. 1987.

**V. Sreekanth Annapureddy** (S'08) received the B.Tech. degree in electrical engineering from the Indian Institute of Technology, Madras, in 2006, and the M.S. degree in electrical and computer engineering from University of Illinois at Urbana-Champaign, Urbana, in 2008, where he is currently pursuing the Ph.D. degree.

He was an intern at Ittiam Systems, Texas Instruments, Qualcomm, and Apple. His research interests include information theory and wireless communications.

Mr. Annapureddy was a member of the team that won the 2010 Qualcomm Cognitive Radio contest. He was a recipient of the James Henderson fellowship and the Vodafone graduate fellowship from the University of Illinois and the Roberto Padovani scholarship from Qualcomm.



**Venugopal V. Veeravalli** (S'86–M'92–SM'98–F'06) received the B.Tech. degree in 1985 from the Indian Institute of Technology, Bombay (Silver Medal Honors), the M.S. degree in 1987 from Carnegie Mellon University, Pittsburgh, PA, and the Ph.D. degree in 1992 from the University of Illinois at Urbana-Champaign (UIUC), Urbana, all in electrical engineering.

He joined UIUC in 2000, where he is currently a Professor in the Department of Electrical and Computer Engineering, a Research Professor in the Coordinated Science Laboratory, and the Director of the Illinois Center for Wireless Systems (ICWS). He served as a program director for communications research at the U.S. National Science Foundation in Arlington, VA, from 2003 to 2005. He was with Cornell University, Ithaca, NY, before he joined UIUC, and has been on sabbatical at MIT, IISc Bangalore, and Qualcomm, Inc. His research

interests include distributed sensor systems and networks, wireless communications, detection and estimation theory, and information theory.

Dr. Veeravalli is a Distinguished Lecturer for the IEEE Signal Processing Society for 2010–2011. He is a Fellow of the IEEE and has been on the Board of Governors of the IEEE Information Theory Society. He has also served as an Associate Editor for the IEEE TRANSACTIONS ON INFORMATION THEORY and the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS. Among the awards he has received for research and teaching are the IEEE Browder J. Thompson Best Paper Award, the National Science Foundation CAREER Award, and the Presidential Early Career Award for Scientists and Engineers (PECASE).