On the Noisy Interference Regime of the MISO Gaussian Interference Channel

Bernd Bandemer, Aydin Sezgin, and Arogyaswami Paulraj

***Abstract*—In this work, we consider the two-user Gaussian MISO interference channel. We derive a criterion for sum-rate optimality of treating interference as noise at the receivers, which generalizes recent results for the SISO case [1]–[3]. In contrast to those results, however, the criterion is not in closed form. For the case where single-mode beamforming is applied at the transmitters, we propose a practical decision algorithm that will decide whether optimality holds for a given channel.**

I. Introduction

The information-theoretic interference channel has recently attracted increased research interest. In [4], [5], inner and outer bounds on the capacity region were found that are within a constant gap of each other. More recently, tighter upper bounds on the sum rate were developed for the two-user Gaussian single-input, single-output interference channel [1]–[3]. Interestingly, these new bounds match the sum rate achievable by treating the interference as noise at the receiver for sufficiently low interference, thus establishing the sum capacity in this regime ("noisy interference").

In this work, we apply similar bounding techniques to the two-user Gaussian interference channel where each link has multiple inputs and a single output (MISO). Specifically, we follow the ideas of [1] and extend them to the MISO case. We arrive at sufficient conditions for the optimality of treating interference as noise. For the special case where both transmitters use single-mode beamforming, we provide a way to decide sum-rate optimality algorithmically.

II. Channel model

A. Original MISO Channel

We consider the MISO Gaussian interference channel

$$\begin{aligned} Y_1 &= e_1^{\mathrm{T}} X_1 + h_2^{\mathrm{T}} X_2 + Z_1 \\ Y_2 &= h_3^{\mathrm{T}} X_1 + e_1^{\mathrm{T}} X_2 + Z_2, \end{aligned}$$

where $X_1, X_2 \in \mathbb{R}^M$ are the transmit signals, $h_i \in \mathbb{R}^M$ are the fixed channel coefficients, and $Z_1, Z_2 \sim \mathcal{N}(0,1)$ is independent additive Gaussian noise. We define the transmit covariance matrices $Q_i = \mathbb{E}(X_i X_i^{\mathrm{T}})$, and impose average transmit power constraints $\mathrm{tr}(Q_i) \leq P_i$, for $i = 1, 2$.

The authors are with the Information Systems Laboratory at Stanford University, 350 Serra Mall, Stanford, CA 94305. The corresponding author is Bernd Bandemer (bandemer@stanford.edu). The work of Bernd Bandemer is supported by an Eric and Illeana Benhamou Stanford Graduate Fellowship.

This work was also presented as a student poster at the Communications Theory Workshop (CTW'08), St. Croix, US Virgin Islands, May 2008.

In our model, the direct channels are $e_1 = [1, 0, 0, \dots]^{\mathrm{T}}$. This is without loss of generality, since arbitrary channel gains can always be rotated to e_1 through an orthogonal transformation and scaling [6], both of which can be absorbed into the transmit vectors. Note that the channel is completely characterized by the tuple (h_2, h_3, P_1, P_2).

B. Genie-aided MISO Channel

In order to arrive at computable sum capacity expressions, we supply the receivers with additional "genie" information. Specifically, receivers 1 and 2 will have access to side information G_1 and G_2, respectively, where

$$\begin{aligned} G_1 &= h_3^{\mathrm{T}} X_1 + \eta_1 W_1 \\ G_2 &= h_2^{\mathrm{T}} X_2 + \eta_2 W_2. \end{aligned}$$

The genie noise terms W_1, W_2 are i.i.d. with probability density function $\mathcal{N}(0,1)$. They are independent of the transmit signals, but correlated with the channel noises,

$$\mathbb{E}(W_1 Z_1) = \rho_1, \qquad \mathbb{E}(W_2 Z_2) = \rho_2.$$

This choice of genies is motivated by similar approaches for the SISO channel [1], [4].

The genie parameters $(\eta_1, \eta_2, \rho_1, \rho_2)$ fully specify the statistics of the side information. We will now determine these parameters such that the genie is helpful for our derivations.

III. Choosing the genie parameters

A. Sum rate of the genie-aided channel

We can upper bound the achievable rates in the genie-aided channel by

$$n(R_i - \varepsilon_n) \leq I(X_i^n; Y_i^n G_i^n), \quad i = 1, 2.$$

Using similar techniques as [1], it can be shown that the sum rate of genie-aided channels that satisfy

$$1 - \rho_2^2 \geq \eta_1^2, \qquad 1 - \rho_1^2 \geq \eta_2^2, \tag{1}$$

is bounded by

$$R_1 + R_2 - 2\varepsilon_n \leq \underbrace{I(\widetilde{X}_1; \widetilde{Y}_1 \widetilde{G}_1) + I(\widetilde{X}_2; \widetilde{Y}_2 \widetilde{G}_2)}_{U(Q_1, Q_2)}. \tag{2}$$

(A proof is given in Appendix A.) Here, the tilded variables are the ones induced by choosing the inputs X_1 and X_2 as Gaussian random variables with covariances Q_1 and Q_2, respectively. This upper bound $U(Q_1, Q_2)$ is thus achievable in the genie-aided channel. Because it involves only Gaussian random variables of known covariances, it can be written as

$$U(Q_1, Q_2, \eta_1, \eta_2, \rho_1, \rho_2) = \tfrac{1}{2} \log \left(\frac{\left(1+I_1+S_1-\frac{(e_1 Q_1 h_3+\eta_1\rho_1)^2}{I_2+\eta_1^2}\right)\left(\frac{1-\rho_2^2+I_2}{I_2+\eta_1^2}\right)\left(1+I_2+S_2-\frac{(e_1 Q_2 h_2+\eta_2\rho_2)^2}{I_1+\eta_2^2}\right)\left(\frac{1-\rho_1^2+I_1}{I_1+\eta_2^2}\right)}{\eta_1^2\eta_2^2} \right) \tag{3}$$

a closed form $\log$ expression which depends on the genie parameters η_i, ρ_i, shown in equation (3) at the top of the page, using notation from (9). For the original channel, $U(Q_1, Q_2)$ represents an upper bound on the sum rate.

B. Genies that do not increase the sum rate

We now consider the special case where the upper bound (2) is achievable in the *original* channel. This happens precisely when the genie side information does not increase the mutual information from each receiver to the corresponding transmitter, i.e., iff $I(\widetilde{X}_i; \widetilde{Y}_i \widetilde{G}_i) = I(\widetilde{X}_i; \widetilde{Y}_i)$, for $i = 1, 2$. This is equivalent to $\widetilde{G}_i - \widetilde{Y}_i - \widetilde{X}_i$ forming a Markov chain. Using the result from Appendix B, this happens iff

$$h_3^{\mathrm{T}} Q_1 = \frac{h_3^{\mathrm{T}} Q_1 e_1 e_1^{\mathrm{T}} Q_1 + \eta_1 \rho_1 e_1^{\mathrm{T}} Q_1}{e_1^{\mathrm{T}} Q_1 e_1 + h_2^{\mathrm{T}} Q_2 h_2 + 1} \tag{4}$$

$$h_2^{\mathrm{T}} Q_2 = \frac{h_2^{\mathrm{T}} Q_2 e_1 e_1^{\mathrm{T}} Q_2 + \eta_2 \rho_2 e_1^{\mathrm{T}} Q_2}{e_1^{\mathrm{T}} Q_2 e_1 + h_3^{\mathrm{T}} Q_1 h_3 + 1}. \tag{5}$$

If conditions (1), (4), and (5) hold simultaneously, it follows that the sum rate of the original channel is bounded by

$$\begin{aligned} & R_1 + R_2 - 2\varepsilon_n \\ & \quad \leq I(\widetilde{X}_1; \widetilde{Y}_1) + I(\widetilde{X}_2; \widetilde{Y}_2) \\ & \quad = \tfrac{1}{2} \log \left(\left(1 + \tfrac{e_1^{\mathrm{T}} Q_1 e_1}{1 + h_2^{\mathrm{T}} Q_2 h_2}\right) \left(1 + \tfrac{e_1^{\mathrm{T}} Q_2 e_1}{1 + h_3^{\mathrm{T}} Q_1 h_3}\right) \right). \end{aligned} \tag{6}$$

This is exactly the sum rate that can be supported by treating interference as noise at each receiver. We can draw a powerful conclusion: For a given original channel, if we can find genie parameters $(\eta_1, \eta_2, \rho_1, \rho_2)$ satisfying (1), (4), and (5) for the sum-rate optimal Q_1, Q_2, then treating the interference as noise achieves sum capacity. In practice, we do not know the sum-rate optimal Q_1, Q_2, but we can verify the conditions for *all* Q_1, Q_2 that could *potentially* be sum-rate optimal.

IV. THE SINGLE-MODE BEAMFORMING CASE

In the following, we focus on the case where the transmitters apply single-mode beamforming [7]. In practice, this may be motivated by transmitter complexity and cost constraints. It is equivalent to Q_1 and Q_2 having rank 1, i.e., $Q_i = v_i v_i^{\mathrm{T}},\ i = 1, 2$. (This assumption is often made, e.g., in [8]).

Then, conditions (4) and (5) may be restated as

$$\eta_1 = \frac{1}{\rho_1} \cdot \underbrace{\frac{h_3^{\mathrm{T}} v_1\ (1 + (v_2^{\mathrm{T}} h_2)^2)}{e_1^{\mathrm{T}} v_1}}_{D_1},$$

$$\eta_2 = \frac{1}{\rho_2} \cdot \underbrace{\frac{h_2^{\mathrm{T}} v_2\ (1 + (v_1^{\mathrm{T}} h_3)^2)}{e_1^{\mathrm{T}} v_2}}_{D_2}.$$

The question is then whether we can find $(\eta_1, \eta_2, \rho_1, \rho_2)$ satisfying these conditions and (1), i.e., such that

$$|D_1| \leq |\rho_1| \sqrt{1 - \rho_2^2}, \qquad |D_2| \leq |\rho_2| \sqrt{1 - \rho_1^2}.$$

Valid ρ_1 and ρ_2 (i.e., with magnitude not exceeding unity) can be found if and only if (following a similar argument in [1])

$$|D_1| + |D_2| \leq 1, \tag{7}$$

or, equivalently,

$$\sqrt{I_2/S_1}\ (1 + I_1) + \sqrt{I_1/S_2}\ (1 + I_2) \leq 1, \tag{8}$$

where

$$I_1 = h_2^{\mathrm{T}} Q_2 h_2 = (h_2^{\mathrm{T}} v_2)^2 \tag{9a}$$

$$I_2 = h_3^{\mathrm{T}} Q_1 h_3 = (h_3^{\mathrm{T}} v_1)^2 \tag{9b}$$

$$S_i = e_1^{\mathrm{T}} Q_i e_1 = (e_1^{\mathrm{T}} v_i)^2, \quad i = 1, 2, \tag{9c}$$

are the interference-to-noise and signal-to-noise ratios at receivers 1 and 2, respectively. As mentioned before, in order to conclude the sum-rate optimality of treating interference as noise, we need (8) to hold for the specific pair (v_1, v_2) that is globally sum-rate optimal.

A. Relaxed optimality criterion

Consider the following optimization problem with variables $Q_1, Q_2 \in \mathbb{R}^{M \times M}$.

$$\begin{aligned} \text{maximize} \quad & \sqrt{I_2/S_1}\ (1 + I_1) + \sqrt{I_1/S_2}\ (1 + I_2) \\ \text{subject to} \quad & \operatorname{tr}(Q_i) \leq P_i, \quad i = 1, 2 \\ & Q_i \succcurlyeq 0, \quad i = 1, 2 \\ & U(Q_1, Q_2) \geq R^\star. \end{aligned} \tag{10}$$

Here, I_i and S_i are linear functions of Q_1 and Q_2, as defined in equations (9). Our test for optimality of treating interference as noise will consist of checking whether the optimal value of (10) is less than unity, which if true implies (8) for the sum-rate optimal pair (v_1, v_2). In optimization terminology, (10) is a relaxation of (8) in two senses: First, it extends the feasible set of v_1, v_2 from including only the optimum pair to all pairs that could potentially be optimum. Secondly, it relaxes the rank-1 constraint on Q_1 and Q_2. We are effectively checking (8) not only for the globally sum-rate optimal pair (v_1, v_2), but instead for a larger class of (Q_1, Q_2), thus creating a sufficient but not necessary optimality criterion.[1]

The third constraint in (10) plays the role of making the set of (Q_1, Q_2) as small as possible while still containing the optimal pair. Here, $U(Q_1, Q_2)$ is the sum rate upper bound from (3), and $R^\star$ is any sum rate that is known to be achievable

[1] We note that (8) is already not necessary due to the genie-based proof. There may be cases where treating interference as noise is sum-rate optimal, but the genie-based approach fails to prove it.

with rank-1 matrices Q_1, Q_2. In other words, we rule out all pairs of (Q_1, Q_2) for which the upper bound is below $R^\star$.

In the implementation, it is impractical to enforce the third constraint in (10) with all possible bounds $U(Q_1, Q_2)$, since those are parameterized by $\eta_1, \eta_2, \rho_1, \rho_2$. As an approximation, we propose to use only tuples where (1) is met with equality, and sample the possible values of η_i on a uniform grid on $[0, 1]$. This approach yields a finite number of inequalities, all of which are convex constraints. A finer sampling grid will correspond to checking a smaller class of (Q_1, Q_2) (i.e., a less lenient relaxation of the optimization problem).

B. Implementation of optimization (10)

Optimization (10) is not a convex problem [9]. However, the feasible set is convex, and the objective function is of the form $f_1(\cdot)f_2(\cdot) + f_3(\cdot)f_4(\cdot)$ where the $f_i(\cdot)$ are quasiconcave. Thus, the feasibility problem

$$\begin{aligned} \text{find} \quad & Q_1, Q_2 \in \mathbb{R}^M && (11)\\ \text{subject to} \quad & \sqrt{I_2/S_1} \geq t_1, \quad 1 + I_1 \geq t_2 \\ & \sqrt{I_2/S_1} \geq t_3, \quad 1 + I_2 \geq t_4 \\ & \operatorname{tr}(Q_i) \leq P_i, \quad Q_i \succcurlyeq 0, \quad i = 1, 2 \\ & U(Q_1, Q_2) \geq R^\star \end{aligned}$$

is convex for all tuples $t = (t_1, t_2, t_3, t_4) \in \mathbb{R}^4_+$, and thus tractable. The question is now whether the sets

$$\mathcal{F} = \{t \in \mathbb{R}^4_+ \mid \text{(11) is feasible for } t\ \}$$
$$\mathcal{C} = \{t \in \mathbb{R}^4_+ \mid t_1 t_2 + t_3 t_4 \geq 1\}$$

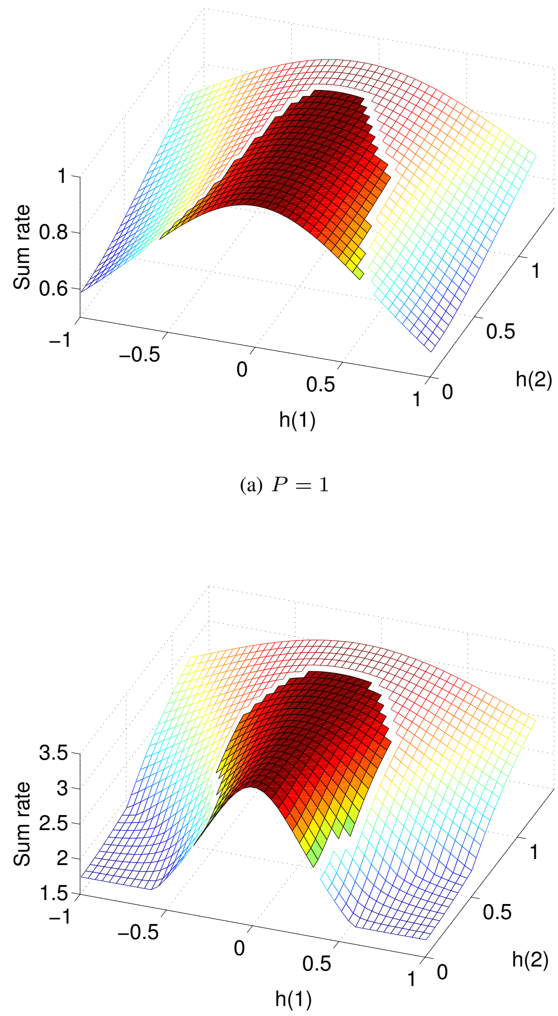
are disjoint or not. A generic algorithm for this type of disjointness problem is given in Appendix C. It requires the sets $\mathcal{F}$ and $\mathcal{C}$ to be specified by an indicator function $\mathbb{I}_\mathcal{F}$ and a projector function $\mathbb{P}_\mathcal{C}$, respectively. In our instance of the problem, the indicator function $\mathbb{I}_\mathcal{F}$ is evaluated by solving the convex feasibility problem (11). A valid projector $\mathbb{P}_\mathcal{C}$ is

$$\begin{aligned} \mathbb{P}_\mathcal{C}(t) &= t + \mathbf{1} \cdot (\min\{s \geq 0 \mid t + s\mathbf{1} \in \mathcal{C}\}) \\ &= t + s_0 \mathbf{1}, \quad \text{where} \\ s_0 &= -\frac{\mathbf{1}^\mathrm{T} t}{4} + \sqrt{\left(\frac{\mathbf{1}^\mathrm{T} t}{4}\right)^2 + 1 - t_1 t_2 - t_3 t_4}, \end{aligned}$$

where $\mathbf{1} \in \mathbb{R}^4$ is the all-one vector. It is easily verified that our problem instance of $\mathcal{F}$ and $\mathcal{C}$ satisfy the technical requirements of the algorithm.

C. Achievable sum rate by treating interference as noise

One way to obtain an achievable $R^\star$ for the optimization (10) is to maximize (6) over rank-1 transmit covariance matrices. This problem itself is also non-convex, and thus potentially hard. However, its structure permits a transformation to a relatively simple equivalent optimization in $\mathbb{R}^4$, independently of M, which can be tackled by standard non-linear optimization techniques.



(a) $P = 1$

(b) $P = 10$

Fig. 1. Achievable sum rate by treating interference as noise over the cross channel $h \in \mathbb{R}^2$. The solid part of the surface shows where this scheme was established to be optimal, under the assumption of single-mode beamforming at the transmitters.

D. Example: Symmetric, two-antenna channel

To demonstrate the feasibility of the proposed method, the algorithm has been implemented and run for a symmetric channel where $P_1 = P_2 = P$, $h_2 = h_3 = h \in \mathbb{R}^2$.

The results are shown in Figure 1. For fixed transmit power $P = 1$ or $P = 10$, the figures show the sum rate achievable by treating interference as noise over the cross channel h plane. In addition, the solid part of the surfaces indicates where our method proves that this sum rate is optimal (given that single-mode beamforming is used).

V. Conclusions

In this work, we considered conditions under which treating interference as noise in the MISO Gaussian interference channel is sum-rate optimal. In contrast to the SISO case, closed form conditions cannot be easily obtained. For the case where the transmitters employ single-mode beamforming, we instead developed an algorithm that decides sum-rate optimality. As in the SISO case, the condition is sufficient but not necessary, meaning that treating interference as noise may still be optimal even in cases where it is not proved by the algorithm.

APPENDIX

A. Proof of the upper bound (2)

This proof follows exactly along the same lines of that given in [1] for the SISO case. From Fano's inequality, we conclude

$$\begin{aligned} n(R_1 - \varepsilon_n) &\le I(X_1^n; Y_1^n G_1^n) \\ &= I(X_1^n; G_1^n) + I(X_1^n; Y_1^n | G_1^n) \\ &= h(G_1^n) - h(G_1^n | X_1^n) + h(Y_1^n | G_1^n) \\ &\quad - h(Y_1^n | G_1^n X_1^n). \end{aligned} \tag{12}$$

The second term can be directly evaluated to $h(G_1^n|X_1^n) = nh(\eta_1 W_1)$, and the third term can be bounded as follows:

$$\begin{aligned} h(Y_1^n|G_1^n) &\overset{(a)}{\le} \sum_{i=1}^n h(Y_{1i}|G_{1i}) \\ &= \sum_{i=1}^n h(e_1^{\mathrm T} X_{1i} + h_2^{\mathrm T} X_{2i} + Z_{1i} \mid h_3^{\mathrm T} X_{1i} + \eta_1 W_{1i}) \\ &\overset{(b)}{\le} \sum_{i=1}^n \frac{1}{2} \log\Bigg(e_1^{\mathrm T} Q_{1i} e_1 + h_2^{\mathrm T} Q_{2i} h_2 + 1 - \\ &\qquad\qquad \frac{(e_1^{\mathrm T} Q_{1i} h_3 + \eta_1\rho_1)^2}{a_1^{\mathrm T} Q_{1i} h_3 + \eta_1^2} \Bigg) \\ &\overset{(c)}{\le} \frac{n}{2} \log\Bigg(e_1^{\mathrm T} Q_1 e_1 + h_2^{\mathrm T} Q_2 h_2 + 1 - \\ &\qquad\qquad \frac{(e_1^{\mathrm T} Q_1 h_3 + \eta_1\rho_1)^2}{a_1^{\mathrm T} Q_1 h_3 + \eta_1^2} \Bigg) \\ &\overset{(d)}{=} h(\widetilde{Y}_1^n | \widetilde{G}_1^n), \end{aligned}$$

where

- (a) stems from the chain rule, and equality holds if and only if the X_{1i} and X_{2i} are independent across the block.
- (b) upper bounds the entropy using the error variance of the best linear estimator. This bound is met with equality if the involved signals are Gaussian random variables. Here, the matrices $Q_{1i} = \mathbb{E}(X_{1i} X_{1i}^{\mathrm T})$ and $Q_{2i} = \mathbb{E}(X_{2i} X_{2i}^{\mathrm T})$ are the covariance matrices of the ith symbol in the block (the expectations are over messages).
- For (c), Jensen's inequality was used based on concavity of the given function in the Qs. From their definition, Q_1 and Q_2 are the arithmetic means of Q_{1i} and Q_{2i} over $i = 1, \dots, n$. Equality in (c) is met if the Q_{1i} and Q_{2i} are constant for all i.
- In (d), the case where (a)–(c) hold with equality is captured. The tilded variables are the ones induced by choosing X_{1i} and X_{2i} as independent Gaussian random variables of constant covariance Q_1 and Q_2, respectively.

In parallel to (12), we may also write

$$\begin{aligned} n(R_2 - \varepsilon_n) &\le h(G_2^n) - h(G_2^n | X_2^n) + h(Y_2^n | G_2^n) \\ &\quad - h(Y_2^n | G_2^n X_2^n), \end{aligned} \tag{13}$$

where two of the terms are as easily obtained as for (12). By summing (12) and (13), we arrive at an expression for the sum rate $R_1 + R_2$. The total of four unevaluated terms can be regrouped into sums of the form

$$\begin{aligned} & h(G_1^n) - h(Y_2^n | G_2^n X_2^n) \\ &= h(h_3^{\mathrm T} X_1^n + \eta_1 W_1^n) \\ &\quad - h(h_3^{\mathrm T} X_1^n + e_1^{\mathrm T} X_2^n + Z_2^n | a_2^{\mathrm T} X_2^n + \eta_2 W_2^n, X_2^n) \\ &= h(h_3^{\mathrm T} X_1^n + \eta_1 W_1^n) - h(h_3^{\mathrm T} X_1^n + Z_2^n | W_2^n) \\ &= h(h_3^{\mathrm T} X_1^n + \underbrace{\eta_1 W_1^n}_{\mathcal{N}(0,\eta_1^2)}) - h(h_3^{\mathrm T} X_1^n + \underbrace{V_2^n}_{\mathcal{N}(0,1-\rho_2^2)}), \end{aligned}$$

where the components of V_2^n in the last line are white Gaussian noise with density $\mathcal{N}(0, 1 - \rho_2^2)$. Note also that the components of the noise term in the first summand have density $\mathcal{N}(0, \eta_1^2)$.

Assuming $1 - \rho_2^2 \ge \eta_1^2$, we may equivalently assume $V_{2i} = \eta_1 W_{1i} + V_i$, where $V_i \sim \mathcal{N}(0, 1 - \eta_1^2 - \rho_2^2)$, temporally white and independent of W_{1i}. Then we have

$$\begin{aligned} & h(G_1^n) - h(Y_2^n | G_2^n X_2^n) \\ &= h(h_3^{\mathrm T} X_1^n + \eta_1 W_1^n) - h(h_3^{\mathrm T} X_1^n + \eta_1 W_1^n + V^n) \\ &= h(h_3^{\mathrm T} X_1^n + \eta_1 W_1^n + V^n | V^n) - h(h_3^{\mathrm T} X_1^n \eta_1 W_1^n + V^n) \\ &= -I(V^n; h_3^{\mathrm T} X_1^n + \eta_1 W_1^n + V^n). \end{aligned}$$

The last expression contains the capacity of an additive noise channel with input V and noise $h_3^{\mathrm T} X_1 + \eta_1 W_1$ of given *average* variance $h_3^{\mathrm T} Q_1 h_3 + \eta_1^2$. By convexity arguments, it is minimized if the noise, and thus $h_3^{\mathrm T} X_1$, is Gaussian with constant covariance over the entire block (worst-case noise result, [10]). One way to achieve that is by letting X_1 itself be Gaussian with covariance not changing over the block (i.e., $\widetilde{X}_1$). Thus we conclude

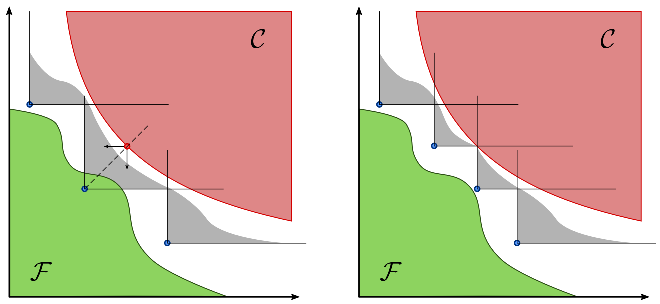
$$h(G_1^n) - h(Y_2^n | G_2^n X_2^n) \le h(\widetilde{G}_1^n) - h(\widetilde{Y}_2^n | \widetilde{G}_2^n \widetilde{X}_2^n), \tag{14}$$

where again, the variables with tildes are the ones induced by indendent Gaussian codebooks of covariances that are not changing over the block.

Using the same argument on $h(G_2^n) - h(Y_1^n | G_1^n X_1^n)$, and combining the previous results, we conclude that if $1 - \rho_2^2 \ge \eta_1^2$ and $1 - \rho_1^2 \ge \eta_2^2$, then the sum rate of the genie-aided channel can be upper bounded by

$$R_1 + R_2 - 2\varepsilon_n \le nI(\widetilde{X}_1; \widetilde{Y}_1 \widetilde{G}_1) + nI(\widetilde{X}_2; \widetilde{Y}_2 \widetilde{G}_2). \tag{15}$$

Since the upper bound is achieved by using i.i.d. Gaussian codebooks, such a codebook is thus sum-rate optimal in genie-aided channels for which (1) holds.



(a) Before refinement step. (b) After refinement step.

Fig. 2. Illustration of a single refinement step in ISDISJOINT.

B. Markov chain condition for Gaussian random variables

We will prove the following necessary and sufficient condition for Markovity. Let X, Y, Z be jointly Gaussian vector random variables. They form a Markov chain $X - Y - Z$ if and only if

$$\mathbb{E}(XZ^{\mathrm{T}}) = \mathbb{E}(XY^{\mathrm{T}})\mathbb{E}(YY^{\mathrm{T}})^{-1}\mathbb{E}(YZ^{\mathrm{T}}).$$

Proof: Consider the best estimator of X given Y. It is

$$\hat{X}_{|Y} = \mathbb{E}(XY^{\mathrm{T}})\mathbb{E}(YY^{\mathrm{T}})^{-1}Y.$$

The estimation error of X given Y is thus given as

$$E_{X|Y} = X - \hat{X}_{|Y} = X - \mathbb{E}(XY^{\mathrm{T}})\mathbb{E}(YY^{\mathrm{T}})^{-1}Y.$$

Equally, the estimation error of Z given Y can be written as

$$E_{Z|Y} = Z - \hat{Z}_{|Y} = Z - \mathbb{E}(ZY^{\mathrm{T}})\mathbb{E}(YY^{\mathrm{T}})^{-1}Y.$$

The Markov chain $X - Y - Z$ holds iff X and Z are conditionally independent given Y. This is equivalent to independence between the two estimation errors computed above. Hence we have the necessary and sufficient condition

$$\mathbb{E}(E_{X|Y}E_{Z|Y}^{\mathrm{T}}) = 0,$$

which evaluates to the condition given above. ■

C. An algorithm to decide disjointness of sets

We describe an algorithm that decides whether two given closed sets $\mathcal{F}, \mathcal{C} \subseteq \mathbb{R}^n$ have non-zero intersection. We make the following assumptions on the sets:

$$\begin{aligned} b \in \delta^\star\mathcal{F} \quad &\Rightarrow \quad b + \mathbb{R}_+^n \cap \mathcal{F} = \emptyset \\ b \in \delta^\star\mathcal{F} \quad &\Rightarrow \quad (b - \mathbb{R}_+^n) \cap \mathbb{R}_+^n \subseteq \mathcal{F} \\ b \in \delta\mathcal{C} \quad &\Rightarrow \quad b - \mathbb{R}_+^n \cap \mathcal{C} = \emptyset \\ & \quad\quad 0 \notin \mathcal{C}, \quad 0 \in \mathcal{F}, \end{aligned}$$

where $\delta\mathcal{X}$ is the boundary of a set $\mathcal{X}$, and $\delta^\star\mathcal{F}$ is the set of boundary points of $\mathcal{F}$ for which no coordinate is zero.

The sets are specified to the algorithm via two functions

$$\mathbb{I}_{\mathcal{F}}(x) = \begin{cases} 1 & \text{if } x \in \mathcal{F} \\ 0 & \text{otherwise} \end{cases}$$

$$\mathbb{P}_{\mathcal{C}}(x) : \mathbb{R}^n \setminus \mathcal{C} \to \delta\mathcal{C} \cap (x + \mathbb{R}_+^n),$$

where $\mathbb{I}_{\mathcal{F}}$ is an indicator function and the projector function $\mathbb{P}_{\mathcal{C}}$ maps any outside $\mathcal{C}$ to a larger point on the boundary $\delta\mathcal{C}$.

The algorithm is given by the following pseudocode.

function ISDISJOINT($\mathbb{I}_{\mathcal{F}}, \mathbb{P}_{\mathcal{C}}, \varepsilon$)
  $\mathcal{S} \leftarrow \{0\}$
  loop
    Pick a point $p \in \mathcal{S}$.
    $y \leftarrow \mathbb{P}_{\mathcal{C}}(p)$
    if $\mathbb{I}_{\mathcal{F}}(y)$ **then** Return "Not disjoint"
    if $\|y - p\| \leq \varepsilon$ **then** Return "Undecided"
    for all $i = 1, \ldots, n$ **do**
      $q^{(i)} \leftarrow p + (y_i - p_i)e_i$
      if $\mathbb{I}_{\mathcal{F}}(q^{(i)})$ **then** $\mathcal{S} \leftarrow \mathcal{S} \cup \{q^{(i)}\}$
    $\mathcal{S} \leftarrow \mathcal{S} \setminus \{p\}$
    if $|\mathcal{S}| = 0$ **then** Return "Disjoint"

The algorithm works by spanning the set $\mathcal{C}$ with a finite set of sample points $\mathcal{R} = \{r_i\}$ such that $\mathcal{C} \subseteq \bigcup_{r \in \mathcal{R}} \left(r + \mathbb{R}_+^n\right)$. Initially, we choose $\mathcal{R} = \{0\}$. While maintaining the property that $\mathcal{R}$ spans $\mathcal{C}$, we attempt to iteratively refine $\mathcal{R}$ until it is disjoint with $\mathcal{F}$, since the existence of such an $\mathcal{R}$ proves that $\mathcal{F} \cap \mathcal{C} = \emptyset$. In each step, a point from $\mathcal{R}$ that is in $\mathcal{F}$ is replaced by n points that are closer to the boundary of $\mathcal{C}$, and thus more likely to be outside $\mathcal{F}$. In the implementation given above, we store only those elements of $\mathcal{R}$ that are still inside $\mathcal{F}$. This subset of $\mathcal{R}$ is named $\mathcal{S}$.

If the boundaries of $\mathcal{F}$ and $\mathcal{C}$ are too close, the algorithm terminates without a decision. Thus the algorithm is not a decision algorithm in the theoretical computer science sense. However, whenever the algorithm returns a decision, correctness is guaranteed. A single refinement step of the algorithm is illustrated for $\mathcal{F}, \mathcal{C} \in \mathbb{R}^2$ in Fig. 2.

REFERENCES

[1] V. S. Annapureddy and V. Veeravalli, "Sum Capacity of the Gaussian Interference Channel in the Low Interference Regime," in *Proceedings of ITA Workshop*, San Diego, CA, USA, Jan. 2008.

[2] X. Shang, G. Kramer, and B. Chen, "A New Outer Bound and the Noisy-Interference Sum-Rate Capacity for Gaussian Interference Channels," *IEEE Transactions on Information Theory*, submitted for publication.

[3] A. S. Motahari and A. K. Khandani, "Capacity Bounds for the Gaussian Interference Channel," *IEEE Transactions on Information Theory*, submitted for publication.

[4] R. Etkin, D. Tse, and H. Wang, "Gaussian Interference Channel Capacity to Within One Bit," *IEEE Transactions on Information Theory*, submitted for publication.

[5] E. Telatar and D. Tse, "Bounds on the capacity region of a class of interference channels," in *Proceedings of the International Symposium on Information Theory (ISIT 2007)*, Nice, France, June 2007.

[6] S. Vishwanath and S. A. Jafar, "On the Capacity of Vector Gaussian Interference Channels," in *Proceedings of the Information Theory Workshop (ITW 2004)*, San Antonio, TX, USA, Oct. 2004.

[7] A. Goldsmith, S. A. Jafar, N. Jindal, and S. Vishwanath, "Capacity Limits of MIMO Channels," *IEEE Journal on Selected Areas in Communications*, vol. 21, no. 5, pp. 684–702, June 2003.

[8] E. A. Jorswieck and E. G. Larsson, "The MISO Interference Channel from a Game-Theoretic perspective: A Combination of Selfishness and Altruism Achieves Pareto Optimality," in *Proceedings of ICASSP 2008*, Las Vegas, NV, USA, Apr. 2008, pp. 5364–5367.

[9] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.

[10] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. Hoboken, New Jersey: John Wiley and Sons, 2006.