On Communications through a Gaussian Noise Channel with an MMSE Disturbance Constraint

Alex Dytso, Ronit Bustin, Daniela Tuninetti, Natasha Devroye, H.Vincent Poor, Shlomo Shamai (Shitz)
Corresponding Author: Alex Dytso, University of Illinois at Chicago, USA, odytso2@uic.edu

***Abstract*—This paper considers a Gaussian channel with one transmitter and two receivers. The goal is to maximize the communication rate at the intended/primary receiver subject to a disturbance constraint at the unintended/secondary receiver. The disturbance is measured in terms of minimum mean square error (MMSE) of the interference that the transmission to the primary receiver inflicts on the secondary receiver.**

The paper presents a new upper bound for the problem of maximizing the mutual information subject to an MMSE constraint. The new bound holds for vector inputs of any length and recovers a previously known limiting (when the length for vector input tends to infinity) expression from the work of Bustin *et al.* The key technical novelty is a new upper bound on MMSE. This new bound allows one to bound the MMSE for all signal-to-noise ratio (SNR) values *below* a certain SNR at which the MMSE is known (which corresponds to the disturbance constraint). This new bound complements the 'single-crossing point property' of the MMSE that upper bounds the MMSE for all SNR values *above* a certain value at which the MMSE value is known. The new MMSE upper bound provides a refined characterization of the phase-transition phenomenon which manifests, in the limit as the length of the vector input goes to infinity, as a discontinuity of the MMSE for the problem at hand.

A matching lower bound, to within an additive gap of order $O\left(\log\log \frac{1}{\mathsf{MMSE}}\right)$ (where MMSE is the disturbance constraint), is shown by means of the mixed inputs recently introduced by Dytso *et al.*

I. Introduction

Consider a Gaussian noise channel with one transmitter and two receivers:

$$\mathbf{Y} = \sqrt{\mathsf{snr}}\mathbf{X} + \mathbf{Z}, \tag{1a}$$

$$\mathbf{Y}_{\mathsf{snr}_0} = \sqrt{\mathsf{snr}_0}\mathbf{X} + \mathbf{Z}_0, \tag{1b}$$

where $\mathbf{Z}, \mathbf{Z}_0, \mathbf{X}, \mathbf{Y}, \mathbf{Y}_{\mathsf{snr}_0} \in \mathbb{R}^n$, $\mathbf{Z}, \mathbf{Z}_0 \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ and $(\mathbf{Z}, \mathbf{Z}_0, \mathbf{X})$ are mutually independent. When it will be necessary to stress the SNR at $\mathbf{Y}$ we will denote it with $\mathbf{Y}_{\mathsf{snr}}$.

We denote the mutual information between input $\mathbf{X}$ and output $\mathbf{Y}$ as

$$I(\mathbf{X};\mathbf{Y}) = I(\mathbf{X}, \mathsf{snr}) := \mathbb{E}\left[\log\left(\frac{p_{\mathbf{Y}|\mathbf{X}}(\mathbf{Y}|\mathbf{X})}{p_{\mathbf{Y}}(\mathbf{Y})}\right)\right]. \tag{2}$$

We also denote the mutual information normalized by n as

$$I_n(\mathbf{X}, \mathsf{snr}) := \frac{1}{n} I(\mathbf{X}, \mathsf{snr}). \tag{3}$$

We denote the minimum mean squared error (MMSE) of estimating $\mathbf{X}$ from $\mathbf{Y}$ as

$$\mathrm{mmse}(\mathbf{X}|\mathbf{Y}) = \mathrm{mmse}(\mathbf{X}, \mathsf{snr}) := \frac{1}{n}\mathrm{Tr}\left(\mathbb{E}\left[\mathbf{Cov}(\mathbf{X}|\mathbf{Y})\right]\right), \tag{4}$$

where $\mathbf{Cov}(\mathbf{X}|\mathbf{Y})$ is the conditional covariance matrix of $\mathbf{X}$ given $\mathbf{Y}$ and is defined as

$$\mathbf{Cov}(\mathbf{X}|\mathbf{Y}) := \mathbb{E}\left[\left(\mathbf{X} - \mathbb{E}[\mathbf{X}|\mathbf{Y}]\right)\left(\mathbf{X} - \mathbb{E}[\mathbf{X}|\mathbf{Y}]\right)^{\mathrm{T}} |\mathbf{Y}\right].$$

Moreover, since the distribution of the noise is fixed, the quantities $I(\mathbf{X};\mathbf{Y})$ and $\mathrm{mmse}(\mathbf{X}|\mathbf{Y})$ are completely determined by $\mathbf{X}$ and snr and there is no ambiguity in using the notation $I(\mathbf{X}, \mathsf{snr})$ and $\mathrm{mmse}(\mathbf{X}, \mathsf{snr})$.

We consider a scenario in which a message, encoded as $\mathbf{X}$, must be decoded at the primary receiver $\mathbf{Y}_{\mathsf{snr}}$ while it is also seen at the unintended/secondary receiver $\mathbf{Y}_{\mathsf{snr}_0}$ for which it is an interferer. This scenario is motivated by the two-user Gaussian Interference Channel (G-IC), whose capacity is only known for some special cases. The following strategies are commonly used to manage interference in the G-IC:

1) *interference is treated as Gaussian noise*: in this approach the interference structure is neglected. It has been shown to be sum-capacity optimal in the so called very-weak interference regime [1].
2) *partial interference cancellation*: by using the Han-Kobayashi (HK) achievable scheme [2], part of the interfering message is decoded and subtracted off, and the remaining part is treated as Gaussian noise. This approach has been show to be capacity achieving in strong interference [3] and optimal within 1/2 bit per channel per user otherwise [4].
3) *soft-decoding / estimation*: the unintended receiver employs soft-decoding of part of the interference. This is enabled by using non-Gaussian inputs and designing the decoders that treat interference as noise by taking into account the correct (non-Gaussian) distribution of the interference. Such scenarios were considered in [5], [6] and [7], and shown to be optimal to within either a constant or a $O(\log\log(\mathsf{snr}))$ gap in [8].

In this paper we look at a somewhat simplified scenario as opposed to the G-IC as shown in Fig. 1. We assume that there is only one message for the primary receiver, and the primary user inflicts interference (disturbance) on a secondary receiver. The primary transmitter wishes to maximize its transmission rate, while subject to a constraint on the disturbance it inflicts on the secondary receiver. The disturbance is measured in terms of MMSE. Intuitively, the MMSE disturbance constraint quantifies the remaining interference after partial interference cancellation or soft-decoding have been performed [9], [10]. Formally, we aim to solve the following problem.

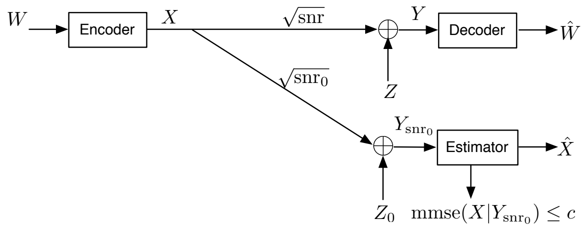


Fig. 1: Channel Model.

Definition 1. *(*max-I problem.*) For some* $\beta \in [0,1]$

$$\mathcal{C}_n(\mathsf{snr}, \mathsf{snr}_0, \beta) := \sup_{\mathbf{X}} I_n(\mathbf{X}, \mathsf{snr}), \tag{5a}$$

$$\textit{s.t.} \mathrm{Tr}\left(\mathbb{E}[\mathbf{X}\mathbf{X}^T]\right) \le n, \tag{5b}$$

$$\textit{and } \mathrm{mmse}(\mathbf{X}, \mathsf{snr}_0) \le \frac{\beta}{1+\beta \mathsf{snr}_0}. \tag{5c}$$

The subscript n in $\mathcal{C}_n(\mathsf{snr}, \mathsf{snr}_0, \beta)$ emphasizes that we seek to find bounds that hold for any input length n. Even though this model is somewhat simplified, compared to the G-IC, it can serve as an important building block, towards characterizing the capacity of G-IC [9], [10].

In [9] the capacity of the channel in Fig. 1 was properly defined and it was shown to be equal to $\lim_{n\to\infty} \mathcal{C}_n(\mathsf{snr}, \mathsf{snr}_0, \beta)$. The reason why the capacity does not have a 'single-letter' expression is because the MMSE constraint does not 'single-letterize'. Moreover, in [10, Sec. VI.3] and [9, Sec. VIII] it was conjectured that the optimal input for $\mathcal{C}_1(\mathsf{snr}, \mathsf{snr}_0, \beta)$ is discrete.

A. Contributions and Paper Outline

In Section II we position our work in the context of existing literature. In Section III we summarize our main results. In Section IV we develop bounds on the derivative of the MMSE, which we use to prove our main result in Theorem 1. Due to space limitations, most proofs are omitted and can be found in the extended version of the paper [11].

B. Notation

Throughout the paper we adopt the following notational conventions: deterministic scalar quantities are denoted by lower case letters and deterministic vector quantities are denoted by lower case bold letters; matrices are denoted by bold upper case letters; random variables are denoted by upper case letters and random vectors are denoted by bold uppercase letters; all logarithms are taken to be base e; we denote support of a random variable A by $\mathsf{supp}(A)$; $X \sim \mathsf{PAM}(N)$ denotes pulse-amplitude modulation (PAM) or the uniform probability mass function over a zero-mean constellation with $|\mathsf{supp}(X)| = N$ points, minimum distance $d_{\min(X)}$, and therefore average energy $\mathbb{E}[X^2] = d^2_{\min(X)} \frac{N^2-1}{12}$; we denote the Fisher information matrix of the random vector $\mathbf{A}$ by $\mathbf{J}(\mathbf{A})$; for $x \in \mathbb{R}$ we let $[x]^+ := \max(x, 0)$ and $\log^+(x) := [\log(x)]^+$; we use the Landau notation $f(x) = O(g(x))$ to mean that for some $c > 0$ there exists an x_0 such that $f(x) \le cg(x)$ for all $x \ge x_0$.

II. Past Work

The mutual information and the MMSE can be related, for any input $\mathbf{X}$, via the so called *I-MMSE relationship* [12, Theorem 1].

Proposition 1. *(*I-MMSE relationship *[12].) The I-MMSE is given by the derivative relationship*

$$\frac{d}{d\mathsf{snr}} I_n(\mathbf{X}, \mathsf{snr}) = \frac{1}{2} \mathrm{mmse}(\mathbf{X}, \mathsf{snr}), \tag{6a}$$

or the integral relationship [12, Eq.(47)]

$$I_n(\mathbf{X}, \mathsf{snr}) = \frac{1}{2} \int_0^{\mathsf{snr}} \mathrm{mmse}(\mathbf{X}, t) dt. \tag{6b}$$

In order to develop bounds on $\mathcal{C}_n(\mathsf{snr}, \mathsf{snr}_0, \beta)$ we require bounds on the MMSE. An important bound on the MMSE is the following *linear MMSE (LMMSE) upper bound*.

Proposition 2. *(*L-MMSE bound [12].*) For any* $\mathbf{X}$ *and* $\mathsf{snr} > 0$

$$\mathrm{mmse}(\mathbf{X}, \mathsf{snr}) \le \frac{1}{\mathsf{snr}}. \tag{7a}$$

If $\frac{1}{n} \mathrm{Tr}\left(\mathbb{E}[\mathbf{X}\mathbf{X}^T]\right) \le \sigma^2$*, then for any* $\mathsf{snr} \ge 0$

$$\mathrm{mmse}(\mathbf{X}, \mathsf{snr}) \le \frac{\sigma^2}{1+\sigma^2 \mathsf{snr}}, \tag{7b}$$

where equality in (7b) *is achieved iff* $\mathbf{X} \sim \mathcal{N}(0, \sigma^2 \mathbf{I})$.

Another important bound for the MMSE is the *single-crossing point property* bound developed in [13] for $n = 1$ and extended in [14] to any $n \ge 1$. Next the single-crossing point property is not stated in full generality, which can be found in [13, Proposition 16] and [14, Theorem 1].

Proposition 3. *(* Single-crossing point property [14].*) For any fixed* $\mathbf{X}$*, suppose that* $\mathrm{mmse}(\mathbf{X}, \mathsf{snr}_0) = \frac{\beta}{1+\beta\mathsf{snr}_0}$*, for some fixed* $\beta \ge 0$*. Then for all* $\mathsf{snr} \in [\mathsf{snr}_0, \infty)$ *we have that*

$$\mathrm{mmse}(\mathbf{X}, \mathsf{snr}) \le \frac{\beta}{1+\beta \mathsf{snr}}, \tag{8a}$$

and for all $\mathsf{snr} \in [0, \mathsf{snr}_0)$

$$\mathrm{mmse}(\mathbf{X}, \mathsf{snr}) \ge \frac{\beta}{1+\beta \mathsf{snr}}. \tag{8b}$$

In words, Proposition 3 means that if we know that the value of MMSE at snr_0 is given by $\mathrm{mmse}(\mathbf{X}, \mathsf{snr}) = \frac{\beta}{1+\beta\mathsf{snr}_0}$ then for all higher SNR values ($\mathsf{snr}_0 \le \mathsf{snr}$) we have the upper bound in (8a) and for all lower SNR values ($\mathsf{snr} \le \mathsf{snr}_0$) we have a lower bound in (8b). Unfortunately, Proposition 3 does not provide an upper bound on $\mathrm{mmse}(\mathbf{X}, \mathsf{snr})$ for $\mathsf{snr} \in [0, \mathsf{snr}_0)$ and one of the goals of this paper is to fill in this gap. Note that upper bounds on MMSE are useful, thanks to the I-MMSE relationship, as tools to derive converse results.

Motivated by the search of the complementary upper bound to Proposition 3 we define the following problem.

Definition 2. *(*max-MMSE problem.*) For some* $\beta \in [0,1]$

$$\mathrm{M}_n(\mathsf{snr},\mathsf{snr}_0,\beta) := \sup_{\mathbf{X}} \mathrm{mmse}(\mathbf{X},\mathsf{snr}), \tag{9a}$$

$$\textit{s.t. } \frac{1}{n}\mathrm{Tr}\left(\mathbb{E}[\mathbf{X}\mathbf{X}^T]\right) \le 1, \tag{9b}$$

$$\textit{and } \mathrm{mmse}(\mathbf{X},\mathsf{snr}_0) \le \frac{\beta}{1+\beta\mathsf{snr}_0}. \tag{9c}$$

Clearly, $\mathrm{M}_n(\mathsf{snr},\mathsf{snr}_0,\beta) \le \mathrm{M}_\infty(\mathsf{snr},\mathsf{snr}_0,\beta)$ for all finite n. Observe that the max-MMSE problem problems in (9) and the max-I problem in (5) have a different objective functions but have the same constraints.

Note that Proposition 3 gives a solution to the max-MMSE problem in (9) for $\mathsf{snr} \ge \mathsf{snr}_0$ and any $n \ge 1$ as follows:

$$\mathrm{M}_n(\mathsf{snr},\mathsf{snr}_0,\beta) = \frac{\beta}{1+\beta\mathsf{snr}}, \text{ for } \mathsf{snr} \ge \mathsf{snr}_0, \tag{10}$$

achieved by $\mathbf{X} \sim \mathcal{N}(0,\beta\mathbf{I})$. Therefore in the rest of the paper the treatment of max-MMSE problem will only focus on the regime $\mathsf{snr} \le \mathsf{snr}_0$.

The case $n = \infty$ of the max-MMSE problem in (9) was solved in [15, Section V-C] and [9, Theorem 2] as follows:

$$\mathrm{M}_\infty(\mathsf{snr},\mathsf{snr}_0,\beta) = \begin{cases} \frac{1}{1+\mathsf{snr}}, & \mathsf{snr} < \mathsf{snr}_0, \\ \frac{\beta}{1+\beta\mathsf{snr}}, & \mathsf{snr} \ge \mathsf{snr}_0, \end{cases}, \tag{11}$$

achieved by using superposition coding with Gaussian codebooks. Clearly there is a discontinuity in (11) at $\mathsf{snr} = \mathsf{snr}_0$ for $\beta < 1$. This fact is a well known property of MMSE referred to as a *phase transition* [15]. It is also well know that, for any finite n, $\mathrm{mmse}(\mathbf{X},\mathsf{snr})$ is a continuous function of snr [13]. Putting these two facts together we have that, for any finite n, the objective function $\mathrm{M}_n(\mathsf{snr},\mathsf{snr}_0,\beta)$ must be continuous in snr and converge to a function with a jump-discontinuity at snr_0 as $n \to \infty$. Therefore, $\mathrm{M}_n(\mathsf{snr},\mathsf{snr}_0,\beta)$ must be of the following form:

$$\begin{aligned} &\mathrm{M}_n(\mathsf{snr},\mathsf{snr}_0,\beta) \\ &= \begin{cases} \frac{1}{1+\mathsf{snr}}, & \mathsf{snr} \le \mathsf{snr}_L, \\ T_n(\mathsf{snr},\mathsf{snr}_0,\beta), & \mathsf{snr}_L \le \mathsf{snr} \le \mathsf{snr}_0, \\ \frac{\beta}{1+\beta\mathsf{snr}}, & \mathsf{snr}_0 \le \mathsf{snr}, \end{cases} \end{aligned} \tag{12}$$

for some snr_L. In this paper we seek to characterize $T_n(\mathsf{snr},\mathsf{snr}_0,\beta)$ and snr_L in (12) and give scaling bounds on the width of the phase transition region defined as

$$W(n) := \mathsf{snr}_0 - \mathsf{snr}_L. \tag{13}$$

Back to the max-I problem in (5). Clearly, we have the following relationship for every $n \ge 1$

$$\mathcal{C}_n(\mathsf{snr},\mathsf{snr}_0,\beta) \le \mathcal{C}_\infty(\mathsf{snr},\mathsf{snr}_0,\beta). \tag{14}$$

In [9, Theorem. 3] Bustin *et al.* proved

$$\begin{aligned} \mathcal{C}_\infty(\mathsf{snr},\mathsf{snr}_0,\beta) &= \lim_{n\to\infty} \mathcal{C}_n(\mathsf{snr},\mathsf{snr}_0,\beta), \\ &= \begin{cases} \frac{1}{2}\log(1+\mathsf{snr}), & \mathsf{snr} \le \mathsf{snr}_0, \\ \frac{1}{2}\log(1+\beta\mathsf{snr}) + \frac{1}{2}\log\left(1+\frac{\mathsf{snr}_0(1-\beta)}{1+\beta\mathsf{snr}_0}\right), & \mathsf{snr} \ge \mathsf{snr}_0, \end{cases} \\ &= \frac{1}{2}\log^+\left(\frac{1+\beta\mathsf{snr}}{1+\beta\mathsf{snr}_0}\right) + \frac{1}{2}\log\left(1+\min(\mathsf{snr},\mathsf{snr}_0)\right), \end{aligned} \tag{15}$$

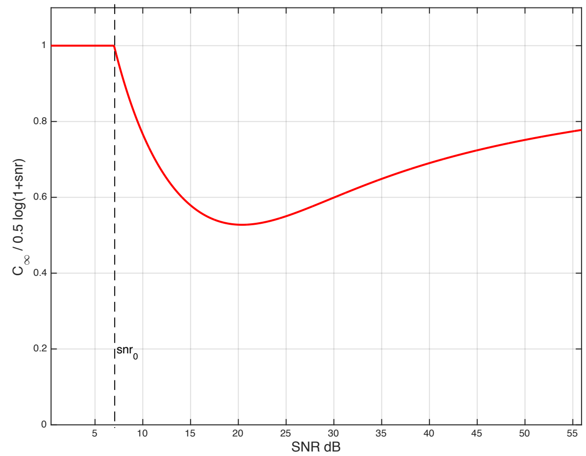


Fig. 2: Plot of $\frac{\mathcal{C}_\infty(\mathsf{snr},\mathsf{snr}_0,\beta)}{\frac{1}{2}\log(1+\mathsf{snr})}$ vs. snrdB, for $\beta = 0.01$, $\mathsf{snr}_0 = 5 = 6.989$ dB.

which is achieved by using superposition coding with Gaussian codebooks. Fig. 2 shows a plot of $\mathcal{C}_\infty(\mathsf{snr},\mathsf{snr}_0,\beta)$ normalized by the capacity of the point-to-point channel capacity $\frac{1}{2}\log(1+\mathsf{snr})$. The region $\mathsf{snr} \le \mathsf{snr}_0$ (flat part of the curve) is where the MMSE constraint is inactive since the channel with snr_0 can decode the interference and guarantee zero MMSE. The regime $\mathsf{snr} \ge \mathsf{snr}_0$ (curvy part of the curve) is where the receiver with snr_0 can no-longer decode the interference and the MMSE constraint becomes active, which in practice is the more interesting regime because the secondary receiver experiences 'weak interference' that can not be fully decoded (recall that in this regime superposition coding appears to be the best achievable strategy for G-IC but it is unknown whether it achieves capacity [4]).

The importance of studying models of communication systems with disturbance constraints has been recognized previously. For example, in [16] Bandemer *et al.* studied the following problem related to the max-I problem in (5).

Definition 3. *(*Bandemer *et al.* problem.*) For some* $R \ge 0$

$$\mathcal{I}_n(\mathsf{snr},\mathsf{snr}_0,R) := \max_{\mathbf{X}} I_n(\mathbf{X},\mathsf{snr}), \tag{16a}$$

$$\textit{s.t. } \frac{1}{n}\mathrm{Tr}\left(\mathbb{E}[\mathbf{X}\mathbf{X}^T]\right) \le 1, \tag{16b}$$

$$\textit{and } I_n(\mathbf{X},\mathsf{snr}_0) \le R. \tag{16c}$$

In [16] it was shown that the optimal solution for $\mathcal{I}_n(\mathsf{snr},\mathsf{snr}_0,R)$, for any n, is attained by $\mathbf{X} \sim \mathcal{N}(0,\alpha\mathbf{I})$ where $\alpha = \min\left(1, \frac{\mathrm{e}^{2R}-1}{\mathsf{snr}_0}\right)$; here α is such that the most stringent constraint between (16b) and (16c) is satisfied with equality. In other words, the optimal input is Gaussian with power reduced such that the disturbance constraint in (16c) is not violated.

Observe that the max-I problems in (5) and the one in (16) have the same objective function but have different constraints. The relationship between the constraints in (5c) and (16c) can be explained as follows. The constraint in (5c) imposes a maximum value on the function $\mathrm{mmse}(\mathbf{X}, \mathsf{snr})$ at $\mathsf{snr} = \mathsf{snr}_0$, while the constraint in (16c), via the integral I-MMSE relationship in (6), imposes a constraint on the area below the function $\mathrm{mmse}(\mathbf{X}, \mathsf{snr})$ in the range $\mathsf{snr} \in [0, \mathsf{snr}_0]$.

Measuring the disturbance with the mutual information as in (16), in contrast to measuring the disturbance with the MMSE as in (5), suggests that it is always optimal to use Gaussian codebooks with the reduced power without any rate splitting. Moreover, while the mutual information constraint in (16) limits the amount of information transmitted to the unintended receiver, it may not be the best choice when one models the interference, since any information that can be reliably decoded is not really interference. For this reason, it has been argued in [9] and [10] that the max-I problem with the MMSE disturbance constraint is a more suitable building block to study the G-IC and understand the key role of rate splitting in approaching capacity.

III. Main Results

A. max-MMSE problem: bounds on $\mathrm{M}_n(\mathsf{snr}, \mathsf{snr}_0, \beta)$

We start by giving bounds on the phase transition region of $\mathrm{M}_n(\mathsf{snr}, \mathsf{snr}_0, \beta)$ defined in (12).

Theorem 1. *(D-Bound.) For any $\mathbf{X}$ and $0 \le \mathsf{snr} \le \mathsf{snr}_0$, let $\mathrm{mmse}(\mathbf{X}, \mathsf{snr}_0) = \frac{\beta}{1+\beta\mathsf{snr}_0}$ for some $\beta \in [0, 1]$. Then*

$$\mathrm{mmse}(\mathbf{X}, \mathsf{snr}) \le \mathrm{mmse}(\mathbf{X}, \mathsf{snr}_0) + k_n \left(\frac{1}{\mathsf{snr}} - \frac{1}{\mathsf{snr}_0} \right) - \Delta, \tag{17a}$$

$$k_n \le n + 2, \ \ \Delta = 0. \tag{17b}$$

If $\mathbf{X}$ is such that $\frac{1}{n}\mathrm{Tr}\left(\mathbb{E}[\mathbf{X}\mathbf{X}^T]\right) \le 1$ then

$$\begin{aligned}\Delta := \Delta_{(17c)} &= 2\log\left(\frac{1+\mathsf{snr}_0}{1+\mathsf{snr}}\right) - 2\log\left(\frac{\mathsf{snr}_0}{\mathsf{snr}}\right) + \frac{1}{1+\mathsf{snr}} \\ &\quad - \frac{1}{1+\mathsf{snr}_0} + \frac{1}{\mathsf{snr}} - \frac{1}{\mathsf{snr}_0}. \end{aligned} \tag{17c}$$

The bound on $\mathrm{M}_n(\mathsf{snr}, \mathsf{snr}_0, \beta)$ in (17b) from Theorem 1 is depicted in Fig. 3a, where:

- The red line is the $\mathrm{M}_\infty(\mathsf{snr}, \mathsf{snr}_0, \beta)$ upper bound on $\mathrm{M}_1(\mathsf{snr}, \mathsf{snr}_0, \beta)$; and
- The blue line is the new upper bound on $\mathrm{M}_1(\mathsf{snr}, \mathsf{snr}_0, \beta)$ from Theorem 1.

Observe that the new bound provides a tighter and continuous upper bounds on $\mathrm{M}_1(\mathsf{snr}, \mathsf{snr}_0, \beta)$ than the trivial upper bound given by $\mathrm{M}_\infty(\mathsf{snr}, \mathsf{snr}_0, \beta)$.

We next show how fast the phase transition region shrinks with n as $n \to \infty$.

Proposition 4. *The bound in* (17a) *from Theorem 1 intersects the LMMSE bound in* (7a) *from Proposition 2 at*

$$\mathsf{snr}_L = \mathsf{snr}_0 \frac{1+\beta\mathsf{snr}_0}{\frac{k_n}{k_n-1} + \beta\mathsf{snr}_0} = O\left(\left(1 - \frac{1}{n}\right) \mathsf{snr}_0 \right). \tag{18a}$$

Thus, the width of the phase transition region is given by

$$W(n) = \frac{1}{k_n - 1} \frac{\mathsf{snr}_0}{\frac{k_n}{k_n-1} + \beta\mathsf{snr}_0} = O\left(\frac{1}{n}\right). \tag{18b}$$

In Proposition 4 we found the intersection between the LMMSE bound $\frac{1}{\mathsf{snr}}$ and the bound in (17a) from Theorem 1. Unfortunately, for the power constraint case, the intersection of the LMMSE bound $\frac{1}{1+\mathsf{snr}}$ and the bound in (17c) cannot be found analytically. However, the solution can be computed efficiently by using numerical methods. Moreover, the asymptotic behavior of the phase transition region is still given by $O\left(\frac{1}{n}\right)$. The bound in Theorem 1 for several values of n is shown in Fig. 3b where:

- The red line is the $\mathrm{M}_\infty(\mathsf{snr}, \mathsf{snr}_0, \beta)$ bound on $\mathrm{M}_n(\mathsf{snr}, \mathsf{snr}_0, \beta)$; and
- The blue line is the bound on $\mathrm{M}_n(\mathsf{snr}, \mathsf{snr}_0, \beta)$ from Theorem 1 for $n = 1, 3, 15$ and 70.

We observe that the new bound provides a refined characterization of the phase transition phenomenon for finite n and in particular it recovers the bound in (11) as $n \to \infty$.

B. max-I problem: bounds on $\mathcal{C}_n(\mathsf{snr}, \mathsf{snr}_0, \beta)$

Using the previous novel bound on $\mathrm{M}_n(\mathsf{snr}, \mathsf{snr}_0, \beta)$ in Theorem 1 we can find new upper bounds on $\mathcal{C}_n(\mathsf{snr}, \mathsf{snr}_0, \beta)$ by integration as follows:

$$\begin{aligned}\mathcal{C}_n(\mathsf{snr}, \mathsf{snr}_0, \beta) &\le \frac{1}{2}\int_0^{\mathsf{snr}} \mathrm{M}_n(t, \mathsf{snr}_0, \beta) dt \\ &= \frac{1}{2}\log(1+\mathsf{snr}_L) + \frac{1}{2}\int_{\mathsf{snr}_L}^{\mathsf{snr}_0} T_n(t, \mathsf{snr}_0, \beta) dt \\ &\quad + \frac{1}{2}\log\left(\frac{1+\beta\mathsf{snr}}{1+\beta\mathsf{snr}_0}\right), \ \text{for } \mathsf{snr}_0 \le \mathsf{snr}, \end{aligned} \tag{19}$$

and

$$\begin{aligned}&\mathcal{C}_n(\mathsf{snr}, \mathsf{snr}_0, \beta) \le \frac{1}{2}\int_0^{\mathsf{snr}} \mathrm{M}_n(t, \mathsf{snr}_0, \beta) dt \\ &\le \frac{1}{2}\log(1+\min(\mathsf{snr}_L, \mathsf{snr})) \\ &+ \frac{1}{2}\int_{\min(\mathsf{snr}_L, \mathsf{snr})}^{\mathsf{snr}} T_n(t, \mathsf{snr}_0, \beta) dt, \ \text{for } \mathsf{snr}_0 \ge \mathsf{snr}. \end{aligned} \tag{20}$$

By using Theorem 1 to bound $T_n(t, \mathsf{snr}_0, \beta)$ we get the following upper bounds on $\mathcal{C}_n(\mathsf{snr}, \mathsf{snr}_0, \beta)$.

Proposition 5. *For any $0 \le \mathsf{snr}_0$ and $\beta \in [0, 1]$ and snr_L given in Proposition 4 we have that for $\mathsf{snr}_0 \le \mathsf{snr}$*

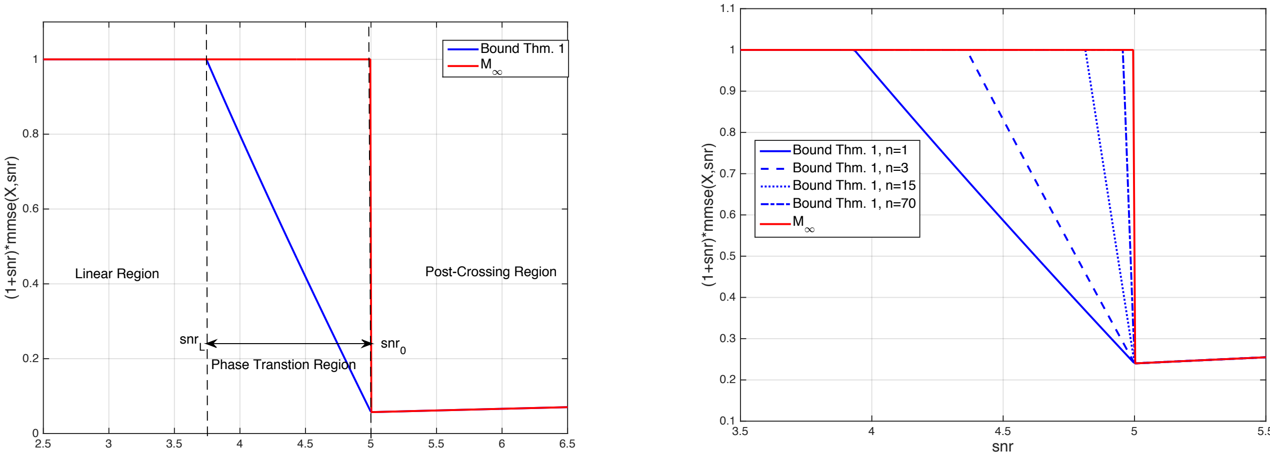
$$\mathcal{C}_n(\mathsf{snr}, \mathsf{snr}_0, \beta) \le \mathcal{C}_\infty(\mathsf{snr}, \mathsf{snr}_0, \beta) - \Delta_{(23)}, \tag{21}$$

and for $\mathsf{snr}_0 \ge \mathsf{snr}$

$$\mathcal{C}_n(\mathsf{snr}, \mathsf{snr}_0, \beta) \le \mathcal{C}_\infty(\mathsf{snr}, \mathsf{snr}_0, \beta) - \Delta_{(24)}, \tag{22}$$

where

$$\begin{aligned}0 \le \Delta_{(23)} &= \frac{1}{2}\log\left(\frac{1+\mathsf{snr}_0}{1+\mathsf{snr}_L}\right) - \frac{1}{2}\frac{\beta(\mathsf{snr}_0 - \mathsf{snr}_L)}{1+\beta\mathsf{snr}_0} \\ &\quad - \frac{(n+2)}{2}\log\left(\frac{\mathsf{snr}_0}{\mathsf{snr}_L}\right) + \frac{(n+2)(\mathsf{snr}_0 - \mathsf{snr}_L)}{2\mathsf{snr}_0}\end{aligned}$$



(a) For $n = 1$, $\mathsf{snr}_0 = 5$ and $\beta = 0.01$.

(b) For several values of n and where $\mathsf{snr}_0 = 5$ and $\beta = 0.05$.

Fig. 3: Bounds on $\mathrm{M}_n(\mathsf{snr}, \mathsf{snr}_0, \beta)$ vs. snr.

$$+\frac{1}{2}\left((2\mathsf{snr}_L+1)\log\left(\frac{\mathsf{snr}_0(1+\mathsf{snr}_L)}{\mathsf{snr}_L(1+\mathsf{snr}_0)}\right)-\frac{\mathsf{snr}_0-\mathsf{snr}_L}{1+\mathsf{snr}_0}\right.$$
$$\left.-\frac{\mathsf{snr}_0-\mathsf{snr}_L}{\mathsf{snr}_0}\right)=O\left(\frac{1}{n}\right), \tag{23}$$

and

$$0 \le \Delta_{(24)} = \frac{1}{2}\log\left(\frac{1+\mathsf{snr}}{1+\min(\mathsf{snr}_L,\mathsf{snr})}\right)$$
$$-\frac{\beta(\mathsf{snr}-\min(\mathsf{snr}_L,\mathsf{snr}))}{2(1+\beta\mathsf{snr}_0)}-\frac{(n+2)}{2}\log\left(\frac{\mathsf{snr}}{\min(\mathsf{snr}_L,\mathsf{snr})}\right)$$
$$+\frac{(n+2)(\mathsf{snr}-\min(\mathsf{snr}_L,\mathsf{snr}))}{2\mathsf{snr}_0}$$
$$+\frac{1}{2}\left((2\min(\mathsf{snr}_L,\mathsf{snr})+1)\log\left(\frac{1+\min(\mathsf{snr}_L,\mathsf{snr})}{\min(\mathsf{snr}_L,\mathsf{snr})}\right)\right.$$
$$-(2\mathsf{snr}+1)\log\left(\frac{1+\mathsf{snr}}{\mathsf{snr}}\right)$$
$$+2(\mathsf{snr}-\min(\mathsf{snr}_L,\mathsf{snr}))\log\left(\frac{1+\mathsf{snr}_0}{\mathsf{snr}_0}\right)$$
$$\left.-\frac{\mathsf{snr}-\min(\mathsf{snr}_L,\mathsf{snr})}{\mathsf{snr}_0}-\frac{\mathsf{snr}-\min(\mathsf{snr}_L,\mathsf{snr})}{1+\mathsf{snr}_0}\right)=O\left(\frac{1}{n}\right). \tag{24}$$

Fig. 4 compares the bounds on $\mathcal{C}_n(\mathsf{snr}, \mathsf{snr}_0, \beta)$ from Proposition 5 with $\mathcal{C}_\infty(\mathsf{snr}, \mathsf{snr}_0, \beta)$ for several values of n. The figure shows how the new bounds in Proposition 5 improves on the trivial $\mathcal{C}_\infty(\mathsf{snr}, \mathsf{snr}_0, \beta)$ bound.

C. *max-MMSE problem: achievability of* $\mathrm{M}_1(\mathsf{snr}, \mathsf{snr}_0, \beta)$

In this Section we propose an input that will be used in the achievable strategy for both the max-I problem and the max-MMSE problem with input length $n = 1$. This input is referred to as *mixed input* [8] and is defined as

$$X_{\text{mix}} := \sqrt{1-\delta}X_D + \sqrt{\delta}X_G,\ \ \delta \in [0,1], \tag{25}$$

where X_G and X_D are independent, $X_G \sim \mathcal{N}(0,1)$, $\mathbb{E}[X_D^2] \le 1$, and where the distribution of X_D and parameter δ are to

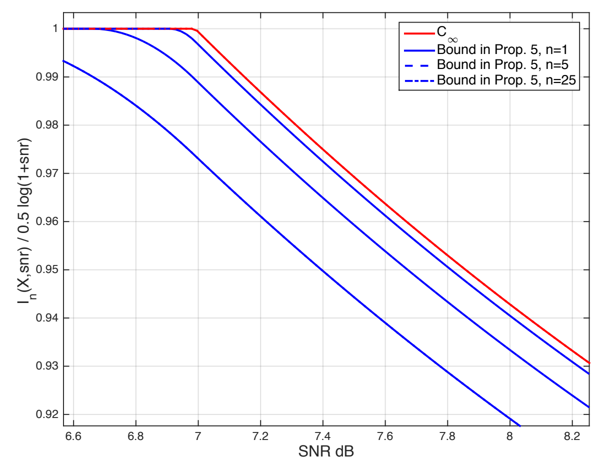


Fig. 4: Bounds on $\mathcal{C}_n(\mathsf{snr}, \mathsf{snr}_0, \beta)$ vs. snrdB, for $\beta = 0.1$ and $\mathsf{snr}_0 = 5 = 6.9897$ dB.

be optimized over. The input X_{mix} exhibits a decomposition property where the MMSE and the mutual information can be written as the sum of the MMSE and mutual information of the X_D and X_G components, albeit at different SNR values.

Proposition 6. *([11]) For* X_{mix} *defined in* (25) *we have that*

$$I(X_{\text{mix}}, \mathsf{snr}) = I\left(X_D, \frac{\mathsf{snr}(1-\delta)}{1+\delta\mathsf{snr}}\right) + I(X_G, \mathsf{snr}\ \delta), \tag{26a}$$

$$\mathrm{mmse}(X_{\text{mix}}, \mathsf{snr}) = \frac{1-\delta}{(1+\mathsf{snr}\delta)^2}\mathrm{mmse}\left(X_D, \frac{\mathsf{snr}(1-\delta)}{1+\delta\mathsf{snr}}\right)$$
$$+\delta\ \mathrm{mmse}(X_G, \mathsf{snr}\ \delta). \tag{26b}$$

Observe that Proposition 6, implies that in order for mixed inputs to comply with the MMSE constraints in (5c) and (9c) the MMSE of the X_D component must satisfy

$$\mathrm{mmse}\left(X_D, \frac{\mathsf{snr}_0(1-\delta)}{1+\delta\mathsf{snr}_0}\right) \le \frac{(\beta-\delta)(1+\delta\mathsf{snr}_0)}{(1-\delta)(1+\beta\mathsf{snr}_0)}. \tag{27}$$

The bound in (27) will be helpful in our choice of δ later on.

When X_D is a discrete random variable with $\mathsf{supp}(X_D) = N$ we use the following bounds from [17, App. C] and [8, Rem. 2].

Proposition 7. *([17], [8]) For discrete random variable X_D such that $p_i = \Pr(X_D = x_i)$ for $i \in [1:N]$ we have that*

$$\mathrm{mmse}(X_D, \mathsf{snr}) \le d_{\max}^2 \sum_{i=1}^{N} p_i \mathrm{e}^{-\frac{\mathsf{snr}}{8} d_i^2}, \tag{28a}$$

$$I(X_D, \mathsf{snr}) \ge H(X_D) - \frac{1}{2}\log\left(\frac{\pi}{6}\right) - \frac{1}{2}\log\left(1 + \frac{12}{d_{\min}^2}\mathrm{mmse}(X_D, \mathsf{snr})\right), \tag{28b}$$

where

$$d_\ell = \min_{x_i \in \mathsf{supp}(X_D): i \ne \ell} |x_\ell - x_i|, \tag{28c}$$

$$d_{\min} = \min_{\ell \in [1:N]} d_\ell, \tag{28d}$$

$$d_{\max} = \max_{x_k, x_i \in \mathsf{supp}(X_D)} |x_k - x_i|. \tag{28e}$$

Proposition 6 and Proposition 7 are particularly useful and allow us to design Gaussian and discrete components of mixed inputs independently.

Fig. 5 shows upper and lower bounds on $\mathrm{M}_1(\mathsf{snr}, \mathsf{snr}_0, \beta)$ where we show the following:

- The upper bound in (11) (solid red line);
- The upper bound from Theorem 1 (dashed cyan line);
- The Gaussian-only input lower bound (green line), with $X \sim \mathcal{N}(0, \beta)$, where the power has been reduced to meet the MMSE constraint;
- The mixed input lower bound (blue dashed line), with the input in (25). We used Proposition 6 where we optimized over X_D for $\delta = \beta \frac{\mathsf{snr}_0}{1+\mathsf{snr}_0}$. The choice of δ is motivated by the scaling property of MMSE that is $\delta \mathrm{mmse}(X_G, \mathsf{snr}\delta) = \mathrm{mmse}(\sqrt{\delta} X_G, \mathsf{snr})$ and the constraint on the discrete component in (27). That is, we chose δ such that the power of X_G is approximately β while the MMSE constraint on X_D in (27) is not equal to zero. The input X_D used in Fig. 5 was found by a local search algorithm on the space of distributions with $N = 3$, and resulted in $\mathcal{X}_D = [-1.8412, -1.7386, 0.5594]$ with $P_X = [0.1111, 0.1274, 0.7615]$, which we do not claim to be the optimal;
- The discrete-only input lower bound (brown dashed-dotted line), with $X_D = [-1.8412, -1.7386, 0.5594]$ with $P_X = [0.1111, 0.1274, 0.7615]$, that is, the same discrete part of the above mentioned mixed input. This is done for completeness, to see the performance of the MMSE of the discrete component of the mixed input, in order to emphasize the behavior with and without the Gaussian component; and
- The discrete-only input lower bound (dotted magenta line), with $X_D = [-1.4689, -1.1634, 0.7838]$ with $P_X = [0.1282, 0.2542, 0.6176]$, which was found by using a local search algorithm on the space of discrete-only distributions with $N = 3$ points.

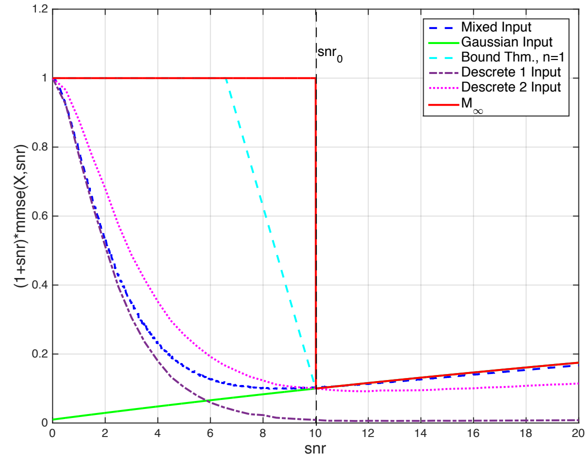


Fig. 5: Upper and lower bounds on $\mathrm{M}_1(\mathsf{snr}, \mathsf{snr}_0, \beta)$ vs. snr, for $\beta = 0.01$, $\mathsf{snr}_0 = 10$.

On the one hand, Fig. 5 shows that, for $\mathsf{snr} \ge \mathsf{snr}_0$, a Gaussian-only input with power reduced to β maximizes $\mathrm{M}_1(\mathsf{snr}, \mathsf{snr}_0, \beta)$ in agreement with the 'single-crossing point property' (green line). On the other hand, for $\mathsf{snr} \le \mathsf{snr}_0$, we see that a discrete-only input achieves higher MMSE than a Gaussian-only input with reduced power (brown dashed-dotted line). Interestingly, unlike Gaussian-only inputs, discrete-only inputs do not have to reduce power in order to meet the MMSE constraint. The reason discrete-only inputs can use full power, as per the power constraint only, is because their MMSE decreases fast enough (exponentially in SNR, as seen in (28a)) to comply with the MMSE constraint. However, for $\mathsf{snr} \ge \mathsf{snr}_0$, the behavior of the MMSE of discrete-only inputs prevents it from being optimal; this is due to their exponential tail behavior.

The mixed input (blue dashed line) gets the best of both (Gaussian-only and discrete-only) worlds: it has the behavior of Gaussian-only inputs for $\mathsf{snr} \ge \mathsf{snr}_0$ (without any reduction in power) and the behavior of discrete-only inputs for $\mathsf{snr} \le \mathsf{snr}_0$. This behavior of mixed inputs turns out to be important for the max-I problem, where we need to choose an input that has the largest area under the MMSE curve.

Finally, Fig. 5 shows the achievable MMSE with another discrete-only input (dotted magenta line) that achieves higher MMSE than the mixed input for $\mathsf{snr} \le \mathsf{snr}_0$ but lower than the mixed input for $\mathsf{snr} \ge \mathsf{snr}_0$.

The insight gained from analyzing different lower bounds on $\mathrm{M}_1(\mathsf{snr}, \mathsf{snr}_0, \beta)$ will be crucial to show an approximately optimal input for $\mathcal{C}_1(\mathsf{snr}, \mathsf{snr}_0, \beta)$, which we consider next.

D. max-I problem: achievability of $\mathcal{C}_1(\mathsf{snr}, \mathsf{snr}_0, \beta)$

In this Section we demonstrate that an inner bound on $\mathcal{C}_1(\mathsf{snr}, \mathsf{snr}_0, \beta)$ with the mixed input in (25) is to within an additive gap of the outer bound in Proposition 5.

Proposition 8. *A lower bound on $\mathcal{C}_1(\mathsf{snr}, \mathsf{snr}_0, \beta)$ with the mixed input in* (25)*, with $X_D \sim \mathsf{PAM}(N)$ and with input parameters as specified in Table I, is to within*

TABLE I: Parameters of mixed inputs in (25) used in the proof of Proposition 8.

Regime	Input Parameters
Weak Interference ($\mathsf{snr} \geq \mathsf{snr}_0$)	$N = \left\lfloor \sqrt{1 + c_1 \frac{(1-\delta)\mathsf{snr}_0}{1+\delta\mathsf{snr}_0}} \right\rfloor$, $c_1 = \frac{3}{2\log\left(\frac{12(1-\delta)(1+\beta\mathsf{snr}_0)}{(1+\mathsf{snr}_0\delta)(\beta-\delta)}\right)}$, $\delta = \beta\frac{\mathsf{snr}_0}{1+\mathsf{snr}_0}$.
Strong Interference ($\mathsf{snr} \leq \mathsf{snr}_0$)	$N = \left\lfloor \sqrt{1 + c_2\mathsf{snr}} \right\rfloor$, $c_2 = \frac{3}{2\log\left(\frac{12(1+\beta\mathsf{snr}_0)}{\beta}\right)}$, $\delta = 0$.

$O\left(\log\log(\frac{1}{\mathrm{mmse}(X,\mathsf{snr}_0)}\right)$ *of the outer bound in Proposition 5 with the exact gap value given by:*

$$\mathsf{snr} \geq \mathsf{snr}_0 \geq 1 : C_1(\mathsf{snr}, \mathsf{snr}_0, \beta) - I_1(X_{\text{mix}}, \mathsf{snr}) \leq \mathsf{gap}_1, \tag{29a}$$

$$\mathsf{snr}_0 \geq \mathsf{snr} \geq 1 : C_1(\mathsf{snr}, \mathsf{snr}_0, \beta) - I_1(X_{\text{mix}}, \mathsf{snr}) \leq \mathsf{gap}_2, \tag{29b}$$

$$\mathsf{snr} \leq 1 : C_1(\mathsf{snr}, \mathsf{snr}_0, \beta) - I_1(X_{\text{mix}}, \mathsf{snr}) \leq \frac{1}{2}\log(2), \tag{29c}$$

where

$$\mathsf{gap}_1 := \frac{1}{2}\log\left(\frac{2}{3}\log\left(\frac{24(1+(1-\beta)\mathsf{snr}_0)}{\beta}\right) + \frac{6\beta}{1+\beta\mathsf{snr}_0}\right) + \frac{1}{2}\log\left(\frac{4\pi}{3}\right) - \Delta_{(23)}, \tag{29d}$$

$$\mathsf{gap}_2 := \frac{1}{2}\log\left(1 + \frac{2}{3}\log\left(\frac{12(1+\beta\mathsf{snr}_0)}{\beta}\right)\right) + \frac{1}{2}\log\left(\frac{4\pi}{6}\right) - \Delta_{(24)}. \tag{29e}$$

Please note that the gap result in Proposition 8 is constant in snr (i.e., independent of snr) but not in snr_0.

Fig. 6 compares the inner bounds on $\mathcal{C}_1(\mathsf{snr}, \mathsf{snr}_0, \beta)$, normalized by the point-to-point capacity $1/2\log(1+\mathsf{snr})$, with mixed inputs (dashed magenta line) in Proposition 8 to:

- The upper bound in (15), (solid red line);
- The upper bound from Proposition 5 (dashed blue line);
- The inner bound with $X \sim \mathcal{N}(0, \beta)$, where the reduction in power is necessary to satisfy the MMSE constraint $\mathrm{mmse}(X, \mathsf{snr}_0) \leq \frac{\beta}{1+\beta\mathsf{snr}_0}$ (dotted green line).

Fig. 6 shows that Gaussian inputs are sub-optimal and that mixed-inputs achieve more degrees of freedom[1] compared to Gaussian inputs. Interestingly, in the regime $\mathsf{snr} \leq \mathsf{snr}_0$, it is approximately optimal to set $\delta = 0$, that is, only the discrete part of the input is used. This in particular supports the conjecture in [9] that discrete inputs may be optimal for $n = 1$ and $\mathsf{snr} \leq \mathsf{snr}_0$.

IV. Properties of the first derivative of MMSE

A key element in the proof of the 'single-crossing point property' in Proposition 3 was the characterization of the first derivative of the MMSE

$$-\mathrm{mmse}'(\mathbf{X}, \mathsf{snr}) = \frac{1}{n}\mathrm{Tr}\left(\mathbb{E}\left[\mathbf{Cov}^2(\mathbf{X}|\mathbf{Y})\right)\right) = \frac{1}{n}\mathrm{Tr}\left(\mathbb{E}\left[\mathbf{Cov}^2(\mathbf{X}, \mathsf{snr})\right]\right), \tag{30}$$

[1]The degrees of freedom, or pre-log, is defined as $d(X) := \lim_{\mathsf{snr}\to\infty} \frac{I(X,\mathsf{snr})}{0.5\log(1+\mathsf{snr})}$.

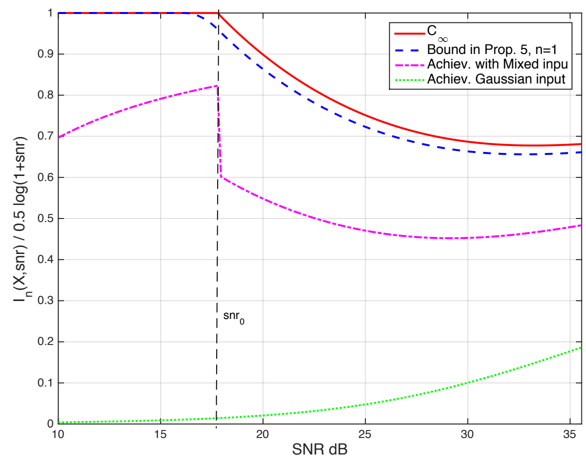


Fig. 6: Upper and lower bounds on $\mathcal{C}_{n=1}(\mathsf{snr}, \mathsf{snr}_0, \beta)$ vs. snrdB, for $\beta = 0.001$ and $\mathsf{snr}_0 = 60 = 17.6815$ dB.

which was given in [13, Proposition 9] for $n = 1$ and in [14, Lemma 3] for $n \geq 1$. The first derivative in (30) turns out to be instrumental in proving Theorem 1 as well.

For ease of presentation, in the rest of the section, instead of focusing on the derivative $\mathrm{mmse}'(\mathbf{X}, \mathsf{snr})$ we will focus on $\mathrm{Tr}\left(\mathbb{E}[\mathbf{Cov}^2(\mathbf{X}|\mathbf{Y})]\right)$. The quantity $\mathrm{Tr}\left(\mathbb{E}[\mathbf{Cov}^2(\mathbf{X}|\mathbf{Y})]\right)$ is well defined for any $\mathbf{X}$. Moreover, for the case of $n = 1$ it has been shown [13, Proposition 5]

$$\mathbb{E}\left[\mathrm{Cov}^2(X|Y)\right] \leq \frac{k_1}{\mathsf{snr}^2}, \text{ where } k_1 \leq 3 \cdot 2^4. \tag{31}$$

Before using (30) in the proof of Theorem 1, we will need to sharpen the existing constant for $n = 1$ in (31) (given by $k_1 \leq 3 \cdot 2^4$) and generalize it to any $n \geq 1$, which to the best of our knowledge has not been considered before.

Proposition 9. *For any* $\mathbf{X}$ *and* $\mathsf{snr} > 0$ *we have*

$$\frac{1}{n}\mathrm{Tr}\left(\mathbb{E}[\mathbf{Cov}^2(\mathbf{X}|\mathbf{Y})]\right) \leq \frac{k_n}{\mathsf{snr}^2}, \tag{32a}$$

where

$$k_n \leq \frac{n(n+2) - n\ \mathrm{mmse}(\mathbf{Z}\mathbf{Z}^T|\mathbf{Y}) - \mathrm{Tr}\left(\mathbf{J}^2(\mathbf{Y})\right)}{n} \leq n + 2. \tag{32b}$$

In Proposition 9 the bound on k_1 in (31) has been tightened from $k_1 \leq 3 \cdot 2^4$ in (31) to $k_1 \leq 3$. This improvement will result in tighter bounds in what follows.

The following tightens k_n for power constrained inputs.

Proposition 10. *If* $\mathbf{X}$ *such that* $\frac{1}{n}\mathrm{Tr}\left(\mathbb{E}\left[\mathbf{X}\mathbf{X}^T\right]\right) \le 1$ *then*

$$\mathrm{Tr}(\mathbf{J}^2(\mathbf{Y})) \ge \frac{n}{(1+\mathsf{snr})^2}. \tag{33}$$

Observe that, by using the bound in (32) from Proposition 9 together with the lower bound on the Fisher information in Proposition 10, we have that for power constrained inputs

$$k_n \le \frac{n(n+2) - \frac{n}{(1+\mathsf{snr})^2}}{n} = n+2 - \frac{1}{(1+\mathsf{snr})^2}. \tag{34}$$

We are now ready to prove our main result.

A. Proof of Theorem 1

The proof of Theorem 1 relies on the fact that MMSE is an infinitely differentiable function of snr [13, Proposition 7] and therefore we can write the difference of two MMSE functions using the fundamental theorem of calculus as

$$\begin{aligned}
&\mathrm{mmse}(\mathbf{X},\mathsf{snr}) - \mathrm{mmse}(\mathbf{X},\mathsf{snr}_0) = -\int_{\mathsf{snr}}^{\mathsf{snr}_0} \mathrm{mmse}'(\mathbf{X},\gamma) d\gamma \\
&\stackrel{a)}{=} \int_{\mathsf{snr}}^{\mathsf{snr}_0} \frac{1}{n}\mathrm{Tr}\left(\mathbb{E}[\mathbf{Cov}^2(\mathbf{X},\gamma)]\right) d\gamma \\
&\stackrel{b)}{\le} \int_{\mathsf{snr}}^{\mathsf{snr}_0} \frac{(n+2)}{\gamma^2} d\gamma = (n+2)\left(\frac{1}{\mathsf{snr}} - \frac{1}{\mathsf{snr}_0}\right),
\end{aligned}$$

where the (in)-equalities follow from: a) by using (30), and b) by using the bound in Proposition 9 with $k_n \le n+2$. If we further assume that $\mathbf{X}$ has finite power, instead of bounding $k_n \le n+2$, we can use (34),to obtain

$$\begin{aligned}
&\mathrm{mmse}(\mathbf{X},\mathsf{snr}) - \mathrm{mmse}(\mathbf{X},\mathsf{snr}_0) \le \int_{\mathsf{snr}}^{\mathsf{snr}_0} \frac{k_n}{\gamma^2} d\gamma \\
&\le \int_{\mathsf{snr}}^{\mathsf{snr}_0} \frac{n+2}{\gamma^2} d\gamma - \int_{\mathsf{snr}}^{\mathsf{snr}_0} \frac{1}{\gamma^2(1+\gamma)^2} d\gamma \\
&= (n+2)\left(\frac{1}{\mathsf{snr}} - \frac{1}{\mathsf{snr}_0}\right) - \Delta_{(17c)}
\end{aligned}$$

where

$$0 \le \Delta_{(17c)} = \int_{\mathsf{snr}}^{\mathsf{snr}_0} \frac{1}{\gamma^2(1+\gamma)^2} d\gamma.$$

This concludes the proof of Theorem 1.

V. Conclusion

In this paper we have considered a Gaussian channel with one transmitter and two receivers in which the maximization of the rate at the primary/intended receiver is subject to a disturbance constraint measured by the MMSE at the secondary/unintended receiver. We have derived new upper bounds on the capacity of this channel that hold for vector inputs of any length, and demonstrates a matching lower bound that is within an additive gap of the order $O\left(\log\log \frac{1}{\mathrm{mmse}(X,\mathsf{snr}_0)}\right)$ of the upper bound. At the heart of our proof is a new upper bound on the MMSE that complements the 'single-crossing point property' of the MMSE and maybe of independent interest.

Acknowledgment

The work of Alex Dytso, Daniela Tuninetti and Natasha Devroye was partially funded by NSF under award 1422511. The work of Ronit Bustin was supported in part by the Women Postdoctoral Sholarship of Israel's Council for Higher Education (VATAT) 2014-2015. The work of H. Vincent Poor and Ronit Bustin was partially supported by NSF under awards CCF-1420575 and ECCS-1343210. The work of Shlomo Shamai was supported by the Israel Science Foundation and the S. and N. Grand Research Fund. The contents of this article are solely the responsibility of the authors and do not necessarily represent the official views of the funding agencies.

References

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

[2] T. 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 Gaussian interference channel under strong interference," *IEEE Trans. Inf. Theory*, vol. IT-27, no. 6, pp. 786–788, Nov. 1981.

[4] R. Etkin, D. 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] A. Bennatan, S. Shamai, and A. Calderbank, "Soft-decoding-based strategies for relay and interference channels: Analysis and achievable rates using ldpc codes," *IEEE Trans. Inf. Theory*, vol. 60, no. 4, pp. 1977–2009, April 2014.

[6] K. Moshksar, A. Ghasemi, and A. Khandani, "An alternative to decoding interference or treating interference as Gaussian noise," *IEEE Trans. Inf. Theory*, vol. 61, no. 1, pp. 305–322, Jan 2015.

[7] A. Dytso, D. Tuninetti, and N. Devroye, "On the two-user interference channel with lack of knowledge of the interference codebook at one receiver," *IEEE Trans. Inf. Theory*, vol. 61, no. 3, pp. 1257–1276, March 2015.

[8] ——, "Interference as noise: Friend or foe?" Submitted to *IEEE Trans. Inf. Theory,* http://arxiv.org/abs/1506.02597, 2015.

[9] R. Bustin and S. Shamai, "MMSE of 'bad' codes," *IEEE Trans. Inf. Theory*, vol. 59, no. 2, pp. 733–743, Feb 2013.

[10] S. Shamai, "From constrained signaling to network interference alignment via an information-estimation perspective," *IEEE Information Theory Society Newsletter*, vol. 62, no. 7, pp. 6–24, September 2012.

[11] A. Dytso, R. Bustin, D. Tuninetti, N. Devroye, S. Shamai, and H. V. Poor, "New bounds on MMSE and applications to communication with the disturbance constraint," To be submitted to *IEEE Trans. Inf. Theory*, 2016.

[12] D. Guo, S. Shamai, and S. Verdú, "Mutual information and minimum mean-square error in Gaussian channels," *IEEE Trans. Inf. Theory*, vol. 51, no. 4, pp. 1261–1282, April 2005.

[13] D. Guo, Y. Wu, S. Shamai, and S. Verdú, "Estimation in Gaussian noise: Properties of the minimum mean-square error," *IEEE Trans. Inf. Theory*, vol. 57, no. 4, pp. 2371–2385, April 2011.

[14] R. Bustin, M. Payaró, D. P. Palomar, and S. Shamai, "On MMSE crossing properties and implications in parallel vector Gaussian channels," *IEEE Trans. Inf. Theory*, vol. 59, no. 2, pp. 818–844, Feb 2013.

[15] N. Merhav, D. Guo, and S. Shamai, "Statistical physics of signal estimation in Gaussian noise: Theory and examples of phase transitions," *IEEE Trans. Inf. Theory*, vol. 56, no. 3, pp. 1400–1416, March 2010.

[16] B. Bandemer and A. El Gamal, "Communication with disturbance constraints," *IEEE Trans. Inf. Theory*, vol. 60, no. 8, pp. 4488–4502, Aug 2014.

[17] A. Lozano, A. M. Tulino, and S. Verdú, "Optimum power allocation for parallel Gaussian channels with arbitrary input distributions," *IEEE Trans. Inf. Theory*, vol. 52, no. 7, pp. 3033–3051, July 2006.