Discrete Inputs in Gaussian Interference Networks: Performance Analysis and Approximate Optimality

BY

ALEX DYTSO
B.S., University of Illinois at Chicago, 2011

THESIS

Submitted as partial fulfillment of the requirements
for the degree of Doctor of Philosophy as Electrical Engineering and Computer Engineering
in the Graduate College of the
University of Illinois at Chicago, 2016

Chicago, Illinois

Defense Committee:

Daniela Tuninetti, Chair and Advisor
Natasha Devroye, Advisor
H. Vincent Poor, Princeton University
Hulya Seferoglu
Dan Schonfeld
Rashid Ansari

ACKNOWLEDGMENTS

My academic journey would not have been possible without technical guidance, immense personal support, and patience of my Ph.D. advisors Prof. Daniela Tuninetti and Prof. Natasha Devroye. I am very grateful to them for taking in an unexperienced undergrad and shaping me into a person that I am today. I would like to thank my committee members Prof. H. Vincent Poor, Prof. Rashid Ansari, Prof. Hulya Seferoglu and Prof. Dan Schonfeld for providing their valuable time.

I would also like to extend my respect and gratitude to the staff and faculty of the Department of Electrical and Computer Engineering at University of Illinois at Chicago, especially, to Dr. Vladimir Goncharoff for recommending and introducing me to my Ph.D. advisors.

Many thanks also go to my friends and Ph.D. companions Zhiyu Cheng, Diana Maamari, Yanying Chen and Tang Liu.

Last but not least, this journey would not have been possible without enormous personal sacrifice, support and encouragement of my girlfriend and family, and to whom I dedicate this thesis.

AD

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

CHAPTER **PAGE**

TABLE OF CONTENTS (Continued)

LIST OF TABLES

LIST OF FIGURES

LIST OF FIGURES (Continued)

FIGURE **PAGE**

SUMMARY

The last two decades have witnessed an enormous growth in the telecommunication industry. Much of the technological progress, undeniably, is a consequence of the theory put forth by Claude Shannon in his landmark paper “A Mathematical Theory of Communications” (1), which also established the field of *information theory*. The information theoretic notion of *channel capacity*, developed by Shannon, has become the benchmark parameter for the design of many of the communication systems and protocols used by telecommunication engineers.

Shannon’s theory has been extended from the point-to-point channel to a variety of multi-user channels such as multiple access and broadcast channels. Most of the results assume a channel models where transmitters and receivers have almost inexhaustible memory and computing power, are fully synchronized, and have global knowledge of all the communication protocols, number of users and state of the channel, etc. Clearly, in real world applications, these assumptions are too idealistic and it is of interest to understand the fundamental limits of more practical models where these assumptions are relaxed.

It has become apparent over the years that classical techniques are no longer sufficient to study such channel models. For example, random Gaussian codes, that are optimal for many Gaussian noise channels do not perform well in competitive multi-user scenarios where users share spectrum resources. In asynchronous multi-user channels, some examples have shown that discrete inputs may outperform Gaussian signaling. Regrettably, such conclusions have been either drawn from numerical evaluations, which are difficult to generalize analytically, or existence proofs that show that discrete inputs are optimal but give little information about the optimal input distribution. Therefore, it is of interest to develop a theoretical framework and tools for establishing the optimal performance in such multi-user settings.

Broadly this thesis consists of two parts. In the first part, the necessary machinery required for evaluating the performance of discrete inputs is developed. The framework is rooted in novel connections between *information theory*, *additive combinatorics*, *number theory* and *estimation theory*. Many of the tools developed are of interest in and of themselves and lead to interesting connections between the aforementioned fields that are worthy of further exploration.

In the second part of this thesis, we apply the developed tools to evaluate the performance of discrete inputs in several important channel models and scenarios.

Our first focus is the capacity region of the *interference channel with partial codebook knowledge*. A quite surprising result is shown: that systems with limited codebook knowledge are almost as efficient as systems with full global codebook knowledge for the practically relevant *additive white Gaussian Noise (AWGN)* channels.

Next we focus on a transmission strategy dubbed as *treating interference as noise (TIN)*. TIN is a very simple and commonly used strategy. In the past this achievable scheme was evaluated by using the optimal (nearest neighbor) decoding rule in AWGN. This sub-optimal detection rule is equivalent to assuming the worst (in terms of achievable rates) noise distribution: Gaussian noise. One consequence of adapting TIN with Gaussian inputs as a strategy is an overly pessimistic view of interference which suggests that networks are inference limited, and that user orthogonalization is needed; this is the design paradigm of all commercially available networks. With the tools developed in the first part of the thesis,we show that using the correct distribution of noise+interference, TIN can be approximately optimal.

SUMMARY (Continued)

Because the TIN achievable rate region applies to the block asynchronous interference channel we can make similar conclusions on optimality there as well. Therefore, our result suggests that accounting for the correct distribution of the noise can bring considerable gain. Moreover, we show that the gain is not vanishing as SNR$\to \infty$ but correspond to a degrees of freedom gain (i.e., grow in dB scale with SNR).

Next we consider 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 the minimum mean square error (MMSE), of the interference that the transmission to the primary receiver inflicts on the secondary receiver. This is one of the simplest models in which one can study the effect of interference and interference mitigating strategies. Moreover, even though this model is somewhat simplistic compared to the Gaussian Interference Channel (G-IC), it may serve as an important building block towards characterizing the capacity of the G-IC. The advantage of this estimation theoretic perspective is that it gives a very natural explanation of the nature of the gains that may be attributed to the discrete inputs.

CHAPTER 1

INTRODUCTION

Part of this chapter has been previously published in (2; 3; 4; 5; 6; 7; 8; 9). ©[2013] IEEE. Reprinted, with permission, from (2). ©[2014] IEEE. Reprinted, with permission, from (3) and (4). ©[2015] IEEE. Reprinted, with permission, from (5), (6) and (7). ©[2016] IEEE. Reprinted, with permission from (8) and (9).

Many of the current channel models in network information theory use assumptions that are too idealistic: users have global knowledge of the network protocols, are fully synchronized, and have almost infinite computing power. Moreover, works that tried to relax these idealistic assumptions, often due to the lack of analytical tools, have mainly focussed on numerical results. As today's networks are constantly increasing in size and demanding higher and higher communication rates, there is an urgent need to develop tools which may help us analytically address the capacity of more realistic network models that relax many of these commonly used, but idealistic, assumptions.

One example of a classical, but perhaps unrealistic, assumption in multi-user information theory is that each node in the network possesses knowledge of the *codebooks* used by every other node. Knowledge of the codebook implies that the node is aware of the transmission schemes of the all the neighboring nodes and thus can mimic those protocols to decode and remove interference from neighboring nodes. However, such an assumption might not be practical in heterogeneous, cognitive, distributed or dynamic networks. For example, in very large ad-hoc networks, where nodes enter and leave at will, it might not be practical for new nodes to learn the codebooks of old nodes and vice-versa. In cognitive radio scenarios, where new cognitive systems coexist with legacy systems, requiring the legacy systems to know the codebooks of the new cognitive systems might not be viable. This motivates the study of networks where each node possesses only a subset of the codebooks used in the network. We refer to such systems as networks with *partial codebook knowledge* and to nodes with only knowledge of a subset of the codebooks as *oblivious* nodes.

In has also become apparent that, in multi-user systems, Gaussian random coding is no longer sufficient when studying the capacity regions of many multi-user channels. Recently, in some studies it has been suggested that codes based on discrete random variables instead of Gaussian ones may be effective and lead to higher overall rates. However, due to the lack of required tools for the analytical study of different input pdfs, the majority of results have so far been based on numerical simulations.

This work focuses on developing the necessary tools for evaluating the performance (in terms of achievable rates) when using discrete inputs in multi-user scenarios. With these new tools, we address several issues that are of theoretical and practical relevance. We assess the potential gains when using discrete inputs in channels with partial codebook knowledge. We also demonstrate the approximate optimality of a very robust transmission strategy referred to as treating interference as noise (TIN), which sheds light on the capacity of channels with partial codebook knowledge and some block asynchronous channels. Finally, we apply the developed tools to study communications with a minimum mean squared error (MMSE) disturbance constraint. This is one of the simplest models in which one can study the effect of interference and interference mitigating strategies. Even though this model is somewhat simplistic when compared to a classical Gaussian interference channel, G-IC, it can serve as an important building block towards characterizing the capacity of the G-IC.

The main contributions and structure of the thesis are as follows. In Chapter 1.4.1 we present all the relevant notation used through the thesis.

1.1 Contributions: Tools for Evaluating the Performance of Discrete Inputs

In Chapter 2 we develop the necessary tools for enabling an analytical handling of discrete inputs. The developed tools are of interest in and of themselves, and use techniques from many fields such as information theory, additive combinatorics, number theory and estimation theory.

We are not the first to consider discrete inputs for Gaussian noise channels. In (10) the authors considered the point-to-point power-constrained Gaussian noise channel and derived lower bounds on the achievable rate when the input is constrained to be an equally spaced Pulse Amplitude Modulation (PAM) in which each point is used with equal probability; such an input was shown to be optimal to within 0.41 bits per channel use (10, eq.(9)). As pointed out in (11), already in 1948 Claude Shannon in the unpublished work (12) argued the asymptotically optimality of a PAM input for the point-to-point power-constrained Gaussian noise channel.

In (13, Theorems 6 and 7), the authors *asymptotically* characterized the optimal input distribution over N masses at high and low signal to noise ratio (SNR), respectively, for a point-to-point power-constrained Gaussian noise channel by assuming that N is not dependent on SNR. For the purpose of analytically characterizing the capacity of networks under discrete inputs, these bounds cannot be used, as 1) these bounds are optimized for a specific SNR while we shall need to lower bound the rate achievable by a discrete input at multiple receivers each characterized by a different SNR; 2) we need a *firm* bound that holds at *all* finite SNR; and 3) we need to properly choose N as a function of SNR, a question posed but left open in (13).

The sub-optimality of Gaussian inputs for Gaussian noise channels has been observed before. Past work on the asynchronous IC (14) and (15) showed that non-Gaussian inputs may outperform i.i.d. Gaussian inputs by using local perturbations of an i.i.d. Gaussian input: (14, Lemma 3) considers a fourth order approximation of mutual information, while (15, Theorem 4) uses perturbations in the direction of Hermite polynomials of order larger than three. In both cases the input distribution is assumed to have a density, though (14, Fig. 1) numerically shows the performance of a ternary PAM input as well. For the cases considered in (14) and (15), the improvement over i.i.d. Gaussian inputs shows in the decimal digits of the achievable rates; it is hence not clear that perturbed continuous Gaussian inputs as in (14) and (15) can actually provide degrees of freedom (DoF) gains over Gaussian inputs (note that a strict DoF gain implies an unbounded rate gain as SNR increases) which we seek in this work. In a way this work follows the philosophy of (16): the main idea is to use sub-optimal point-to-point codes in which the reduction in achievable rates is more than compensated for by the decrease in the interference created at the other users.

Finally, it is worth mentioning that there is large body of work, initiated by (17), that demonstrates that discrete inputs are optimal for Gaussian noise channels, with constraints other than the power constraint, such as: amplitude constraint (17), duty cycle constraint (18). Regrettably, the employed technique uses a proof of by contradiction and only demonstrates that discrete inputs are optimal without specifying the nature of the optimal input distribution.

The main contributions of this part are:

1. In Section 2.1, Proposition 2.1.1 presents a generalization of a lower bound from (10) on the mutual information attained by a discrete input on a point-to-point additive noise channel and compares its performance with other lower bounds available in the literature.

2. In Section 2.1.1, Proposition 2.1.2 and Proposition 2.1.3 present new bounds on the cardinality and minimum distance of the sum of two discrete constellations. In multi-user communication, similarly to a point to point communication, the minimum distance between the points of the received constellation will play a key role. However, in multi-user scenarios due to supperposition of several discrete constellations the minimum distance of the aggregate received constellation can have very complicated behavior, for example it can be dependent on whether channel gains take rational or irrational values. Therefore, in order to provide bounds on the minimum distance and the size of the aggregate constellation we require the use of sum-set and number theories.

3. Section 2.1.2 serves as an example of how we intend to use the developed tools. First, we show that discrete inputs are approximately optimal on a point-to-point Gaussian channel. Second, via an example of a point-to-point channel with discrete state (interference), we show that discrete interference acts almost as if there is no interference. This is one of the main ideas of the thesis upon which we will build the majority of our results.

The above mentioned contributions have in part been published in (3; 4; 5; 6; 7; 8) and (9).

1.2 Contributions: The Two User Interference Channel with Lack of Knowledge of the Interference Codebook at One Receiver

In Chapter 3 we study the capacity of an interference channel with one oblivious receiver (IC-OR). To the best of our knowledge, systems with oblivious terminals were first introduced in (19). In (19) lack of codebook knowledge was modeled by using *codebook indices*, which index the random encoding function that maps the messages to the codewords. If a node has codebook knowledge it knows the index (or instance) of the random encoding function used; else it does not and the codewords essentially look like the symbols were produced in an independent identically distributed (i.i.d.) fashion from a given distribution. In (20) and (21) this concept of partial codebook knowledge was extended to model oblivious relays and capacity results were derived. However, as pointed out in (20, Section III.A) and (21, Remark 5), these capacity regions are "non-computable" in the sense that it is not known how to find the optimal input distribution in general. In particular, the capacity achieving input distribution for the practically relevant Gaussian noise channel remains an open problem.

The main contributions of this part are:

1. In Section 3.3, Theorem 3.3.2 derives a novel outer bound that incorporates this partial codebook knowledge explicitly. In this bound, the single rate bounds are valid for a general memoryless IC-OR while the sum-rate bound is valid for the *injective semi-deterministic interference channel with one oblivious receiver (ISD-IC-OR)* only.

2. In Section 3.4 we demonstrate a series of capacity and approximate capacity results for various regimes and classes of IC-OR. Specifically, by using the achievable region in Proposition 3.4.1 we prove:

   (a) In Theorem 3.4.2 we obtain the capacity region for the general memoryless IC-OR in very strong interference at the non-oblivious receiver;

(b) In Theorem 3.4.3 we demonstrate the capacity region to within a gap for the injective semi deterministic interference channel with oblivious receiver (ISD-IC-OR); and

(c) In Corollary 3.4.4 we show that for the injective fully deterministic IC-OR the gap is zero.

3. In Section 3.4.4, we look at the practically relevant Gaussian IC-OR (G-IC-OR) and its corresponding Linear Deterministic Approximation (LDA-IC-OR) in the spirit of (22), which models the G-IC-OR at high SNR. Surprisingly, for the LDA-IC-OR we numerically demonstrate that for the proposed achievable scheme in Proposition 3.4.1, i.i.d. Bernoulli(1/2) input bits (known to be optimal for the LDA-IC with full codebook knowledge (23)) are outperformed by other (correlated and non-uniform) input distributions.

4. In Section 3.4.5, for the G-IC-OR, we show in Corollary 3.4.5 that our inner and outer bounds are to within $1/2$ bit (per channel use per user) of one another. However, similarly to prior work on oblivious models, we are not able to find the set of input distributions that exhaust the outer bound in Theorem 3.3.2, in particular we cannot argue whether i.i.d. Gaussian inputs exhaust the outer bound. Inspired by the results for the LDA-IC-OR, we numerically show that a larger sum-capacity is attainable by using a discrete input at the non-oblivious transmitter than by selecting i.i.d. Gaussian inputs, or using time-division, or treating interference as Gaussian noise (TIN-G), in the strong interference regime at high SNR. This suggests that the penalty for the lack of codebook knowledge is not as severe as one might initially expect.

5. For the remainder of the chapter we consider the G-IC-OR, and by building on the intuition from Section 3.4.5, we demonstrate that even with partial codebook knowledge we are able to achieve to within $\frac{1}{2}\log(12\pi\mathrm{e}) \approx 3.34$ bits per channel use of the *symmetric capacity region* of the G-IC

with full codebook knowledge through the use of *mixed inputs* (i.e., superposition of Gaussian and discrete random variables).

With the tools from Chapter 2, in Section 3.5 in Theorems 3.5.1 and 3.5.2, we evaluate the achievable rate region presented in Proposition 3.4.1 for the G-IC-OR when the non-oblivious transmitter uses either a PAM input or a mixed input that comprises a Gaussian component and a PAM component. Corollaries 3.6.1 and 3.6.2 provide the generalized degrees of freedom (gDoF) characterization of the achievable regions in Theorems 3.5.1 and 3.5.2.

6. In past work on networks with oblivious nodes no performance guarantees were provided as the capacity regions could not be evaluated. In Section 3.6 we study the gDoF achievable with mixed inputs. In Theorem 3.6.3, we show that mixed inputs achieve the gDoF of the classical G-IC, hence implying that there is no loss in performance due to lack of codebooks in a gDoF sense / at high SNR. This is quite surprising considering that the oblivious receiver cannot perform joint decoding of the two messages, which is optimal for the classical G-IC in the strong and very strong interference regimes.

7. Finally, in Section 3.7 we turn our attention to the finite SNR regime and in Theorem 3.7.1 we show that the capacity of the symmetric G-IC-OR is within $\frac{1}{2}\log\left(12\pi\mathrm{e}\right) \approx 3.34$ bits per channel use of the outer bound to the capacity region of the classical symmetric G-IC with full codebook knowledge. To the best of our knowledge. this is the first approximated capacity result that has performance guarantees for systems with partial codebook knowledge.

The above mentioned contributions have in part been published in (2; 3) and (5).

1.3 Contributions: Approximate Optimality of Treating Interference as Noise

In Chapter 4 we look at the performance of TIN in classical G-IC, and its extension to both codebook oblivious and block asynchronous channel models.

Recently there has been lots of interest in characterizing when TIN, with or without time sharing (TS), is approximately optimal. For example, in (24) "It is shown that in the K-user interference channel, if for each user the desired signal strength is no less than the sum of the strengths of the strongest interference from this user and the strongest interference to this user (all values in dB scale), then the simple scheme of using point to point Gaussian codebooks with appropriate power levels at each transmitter and TIN at every receiver achieves all points in the capacity region to within a constant gap. The generalized degrees of freedom (gDoF) region under this condition is a polyhedron, which is shown to be fully achieved by the same scheme, without the need for time-sharing." In this thesis we aim to show that one can always use treating interference as noise with no time sharing (TINnoTS) and be optimal to within an additive gap in all parameter regimes, and not just in the very weak interference regime identified in (24). The key is to use more "friendly" codebooks than Gaussian codebooks. We note that for an input constrained additive-noise channel where the noise distribution is arbitrary, Gaussian inputs are known to be optimal to within 1/2 bit (25); what our work shows is that the same is not true in general in a multi-user competitive scenario.

The main contribution of this part are:

1. In Section 4.2, Proposition 4.2.1 presents an inner bound obtained by evaluating the TINnoTS region with our proposed mixed inputs, whose performance will then be compared to the outer bound in Proposition 4.2.2.

2. Section 4.3 focuses on the symmetric G-IC. Theorem 4.3.1 shows that TINnoTS with mixed inputs is to within $O(1)$, or $O\left(\log\left(\frac{\ln(\min(\mathsf{snr},\mathsf{inr}))}{\gamma}\right)\right)$ except for a set of Lebesgue measure γ for any $\gamma \in (0,1]$, of the outer bound in Proposition 4.2.2, and where snr and inr are the largest singnal to noise ratio and the largest interference to noise ratio, respectively. From this result we infer that:

   (a) The discrete part of the mixed input behaves as a "common message" whose contribution can be removed from the channel output of the non-intended receiver, even though explicit joint decoding of the interference is not employed in TINnoTS;

   (b) The continuous part of the mixed input behaves as a "private message" whose power should be chosen such that it is either received below the noise floor of the non-intended receiver (26), or to have a rate that is approximately half the target rate; and

   (c) Time-sharing may be mimicked by varying the number of points in the discrete part of the mixed inputs.

3. In Section 4.4 we extend the gap result of Theorem 4.3.1 to some general asymmetric G-IC's. The channel parameter regime covered in Theorem 4.4.1 is such that bounds of the form $2R_1 + R_2$ or $R_1 + 2R_2$ are not active in the outer bound in Proposition 4.2.2. The excluded regime, roughly speaking, is such that the sum of the crosslink gains is upper bounded by the sum of the direct link gains and lower bounded by the minimum of the direct link gains, all quantities expressed in dB scale. Numerical experiments suggest that the insights gained in the symmetric case (see

above item 2) hold for the asymmetric case as well and that the proposed TINnoTS with mixed inputs is approximately optimal for the general asymmetric G-IC.

4. In Section 4.5, Theorem 4.5.1 shows that TINnoTS with mixed inputs is gDoF optimal almost everywhere (a.e.), that is, for all channel gains except for an *outage set* of zero measure.

5. In Section 4.6 shows that our approximate optimality results hold for a variety of channels, such as for example the *block-asynchronous* G-IC and the *codebook oblivious* G-IC, thereby demonstrating that lack of codeword synchronism or of codebook knowledge at the receivers results in penalty of at most $O(1)$, or $O\left(\log\left(\frac{\ln(\min(\mathsf{snr},\mathsf{inr}))}{\gamma}\right)\right)$, compared to the classical G-IC. This section generalizes the result of Chapter 3 where we considered only one oblivious receivers to the case when both receivers are oblivious.

6. In Section 4.7 we discuss some practical implications of our TINnoTS with mixed inputs achievability scheme, such as

   (a) in Section 4.7.1 we discuss an approximate maximum a posteriori probability (MAP) decoder for the very strong interference regime that is very simple to implement with TINnoTS;

   (b) in Section 4.7.2 we show through numerical evaluations that our gap results are very conservative and that in practice the achievable rates are much closer to capacity than predicted by our analytical results; and

(c) in Section 4.7.3 we show that a gap result can be obtained by using as inputs purely discrete random variables, i.e., to within an additive gap the Gaussian part of the mixed inputs can be replaced by another PAM input.

The above mentioned contributions have in part been published in (3; 4; 6; 7) and (8).

1.4 Contributions: Communication under an MMSE Disturbance Constraint

In this Chapter we look at a somewhat simplified scenario compared to the G-IC. 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 communication 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 (27; 11). The goal here is to give an estimation theoretic explanation for the near optimal performance of mixed inputs.

The importance of studying models of communication systems with disturbance constraints has been recognized previously. For example, in (28) authors studied a similar scenario where the disturbance was measured in terms of the mutual information at the secondary user. However, as will be explained in Chapter 5, such a disturbance measured is not very suitable for modeling the interference.

Results of this Chapter will also focus on deriving non-asymptotic results, i.e. coding is done over blocks of length n. The advantage of this approach is two fold: 1) asymptotic results can be recovered by taking $n \to \infty$; 2) almost all information theoretic and estimation theoretic quantiles experience a loss of analyticity as $n \to \infty$ (i.e. their derivative with respect to SNR becomes discontinuous). The value of SNR for which the mutual information loses analyticity is referred to as a *phase transition*. By looking at a finite n, considerable insight can be gained on the rate of convergence to the phase transition, in terms of n, which is important for practical systems.

The main contributions of this Chapter are as follows:

1. In Section 5.2 we summarize our main results:
   (a) Theorem 5.2.1, our main technical result, provides new upper bounds for the max-MMSE problem for arbitrary n that complement the single-crossing point property (SCPP) bound.
   (b) Proposition 5.2.2 provides a lower bound on the width of the phase transition region of the order of $\frac{1}{n}$.
   (c) Proposition 5.2.3 provides a new upper bound for the max-I problem for arbitrary n.
   (d) Proposition 5.2.6 shows that, for the case of $n = 1$ *mixed input* inputs achieves the proposed upper bound on the max-I problem from Proposition 5.2.3 to within an additive gap of order $O\left(\log\log \frac{1}{\mathsf{MMSE}}\right)$ (where MMSE is the disturbance constraint).
2. In Section 5.3 we develop bounds on the derivative of MMSE, which we use to prove the main result in this Chapter. In particular, in Proposition 5.3.1 we considerably refines existing bounds on the derivative of MMSE for $n = 1$ and generalizes them to any n.
3. In Section 5.4 we explore some interesting relationships between the MMSE constraint and the power constraint.

The above mentioned contributions have in part been published in (9) and (29) and submitted to (30).

1.4.1 Notation

Throughout the thesis we adopt the following notation convention:

- Lower case variables are instances of upper case random variables which take on values in calligraphic alphabets.
- $\log(\cdot)$ denotes logarithms in base 2 and $\ln(\cdot)$ in base e.
- $[n_1 : n_2]$ is the set of integers from n_1 to $n_2 \geq n_1$.
- Y^j is a vector of length j with components $(Y_1, \ldots, Y_j)$.
- Whenever vector manipulation will be required (Chapter 5) we will also denote random vectors by bold bold uppercase letters.
- Ordering notation $\mathbf{A} \succeq \mathbf{B}$ implies that $\mathbf{A} - \mathbf{B}$ is a positive semidefinite matrix;
- If A is a random variable we denote its support by $\mathsf{supp}(A)$.
- The symbol $|\cdot|$ may denote different things: $|\mathcal{A}|$ is the cardinality of the set $\mathcal{A}$, $|X|$ is the cardinality of $\mathsf{supp}(X)$ of the r.v. X , or $|x|$ is the absolute value of the real-valued x.
- For $x \in \mathbb{R}$ we let $\lfloor x \rfloor$ denote the largest integer not greater than x.
- For $x \in \mathbb{R}$ we let $[x]^+ := \max(x, 0)$ and $\log^+(x) := [\log(x)]^+$.
- $d_{\min(\mathcal{S})} := \min_{i \neq j : s_i, s_j \in \mathcal{S}} |s_i - s_j|$ denotes the minimum distance among the points in the set $\mathcal{S}$. With some abuse of notation we also use $d_{\min(X)}$ to denote $d_{\min(\mathsf{supp}(X))}$ for a r.v. X.
- Let $f(x), g(x)$ be two real-valued functions. We use the Landau notation $f(x) = O(g(x))$ to mean that for *some* $c > 0$ there exists an x_0 such that $f(x) \leq c\, g(x)$ for all $x \geq x_0$.

- Operator $\mathsf{co}(\cdot)$ will refer to convex hull operation.

- $X \sim \mathcal{N}(\mu, \sigma^2)$ denotes the density of a real-valued Gaussian r.v. X with mean μ and variance σ^2.

- $X \sim \mathsf{PAM}\left(N, d_{\min(X)}\right)$ denotes the uniform probability mass function over a zero-mean PAM constellation with $|\mathsf{supp}(X)| = N$ points, minimum distance $d_{\min(X)}$, and therefore average energy $\mathbb{E}[X^2] = d^2_{\min(X)} \frac{N^2-1}{12}$.

- $m(\mathcal{S})$ denotes Lebesgue measure of the set $\mathcal{S}$.

- We let

$$\mathsf{I_g}(x) := \frac{1}{2}\log(1+x), \tag{1.1}$$

$$\mathsf{I_d}(X) := \left[H(X) - \frac{1}{2}\log\left(\frac{2\pi\mathrm{e}}{12}\right) - \frac{1}{2}\log\left(1 + \frac{12}{d^2_{\min(X)}}\right)\right]^+, \tag{1.2}$$

$$\mathsf{N_d}(x) := \left\lfloor \sqrt{1+x} \right\rfloor, \tag{1.3}$$

$$\mathsf{I_d}\left(N, x\right) := \left[\mathsf{I_g}\left(\min\left(N^2-1, x\right)\right) - \frac{1}{2}\log\left(\frac{\pi\mathrm{e}}{3}\right)\right]^+ \tag{1.4}$$

  where the subscript d reminds the reader that discrete inputs are involved, while g that Gaussian inputs are involved.

- Here $H(X)$ is the entropy of the discrete random variable X, while $h(X)$ is the differential entropy of the absolutely continuous random variable X.

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

$$I(\mathbf{X};\mathbf{Y}) := \mathbb{E}\left[\log\left(\frac{p_{\mathbf{Y}|\mathbf{X}}(\mathbf{Y}|\mathbf{X})}{p_{\mathbf{Y}}(\mathbf{Y})}\right)\right]. \tag{1.5}$$

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

$$\mathrm{mmse}(\mathbf{X}|\mathbf{Y}) := \frac{1}{n}\mathrm{Tr}\left(\mathbb{E}\left[\mathbf{Cov}(\mathbf{X}|\mathbf{Y})\right]\right), \tag{1.6}$$

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

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

- When $\mathbf{X}$ and $\mathbf{Y}$ are related through $\mathbf{Y}=\sqrt{\mathsf{snr}}\mathbf{X}+\mathbf{Z}$ where $\mathbf{Z}\sim\mathcal{N}(0,\mathbf{I})$ and independent of $\mathbf{X}$ we will use

$$I(\mathbf{X};\mathbf{Y}) = I(\mathbf{X},\mathsf{snr}) \tag{1.7}$$

$$\mathrm{mmse}(\mathbf{X}|\mathbf{Y}) = \mathrm{mmse}(\mathbf{X},\mathsf{snr}) \tag{1.8}$$

  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 denote the Fisher information matrix of the random vector $\mathbf{A}$ by $\mathbf{J}(\mathbf{A})$.

1.4.2 Abbreviations

AWGN additive white Gaussian Noise

DoF degrees of freedom

G-IC Gaussian Interference Channel

G-IC-OR Gaussian IC-OR

gDoF generalized degrees of freedom

i.i.d. independent identically distributed

IC-OR interference channel with one oblivious receiver

ISD-IC-OR injective semi-deterministic interference channel with one oblivious receiver

LDA-IC-OR Linear Deterministic Approximation

MAP maximum a posteriori probability

MMSE minimum mean square error

PAM Pulse Amplitude Modulation

SCPP single-crossing point property

SNR signal to noise ratio

TIN treating interference as noise

TIN-G treating interference as Gaussian noise

TINnoTS treating interference as noise with no time sharing

TS time sharing

CHAPTER 2

MAIN TOOLS

In this section we present a set of tools that we will use to evaluate performance of achievable schemes with discrete (or mixed inputs). Moreover, these tools can be of interest on their own since they bridge a gap between the seemingly unrelated fields of information theory, additive combinatorics, number theory and estimation theory.

2.1 Generalization of the Ozarow-Wyner Bound

At the core of our proofs is the following lower bound on the rate achieved by a discrete input on a point-to-point additive noise channel. The important point here is to derive firm bounds that are valid for *any* discrete constellation at *any* SNR, as opposed to bounds that are either optimized for a fixed SNR, or hold asymptotically in the low or high SNR regimes.

Proposition 2.1.1 (Ozarow-Wyner-B bound)**.** *Let X_D be a discrete random variable with minimum distance $d_{\min(X_D)} > 0$. Let Z be a zero-mean unit-variance random variable independent of X_D (not necessarily Gaussian). Then*

$$\mathsf{I_d}(X_D) := \left[H(X_D) - \mathsf{G}_{\mathsf{d}(\text{Eq. 2.1})}\right]^+ \le I(X_D; X_D + Z) \le H(X_D), \tag{2.1a}$$

$$\mathsf{G}_{\mathsf{d}(\text{Eq. 2.1})} := \frac{1}{2}\log\left(\frac{2\pi \mathrm{e}}{12}\right) + \frac{1}{2}\log\left(1 + \frac{12}{d^2_{\min(X_D)}}\right). \tag{2.1b}$$

Proof. The upper bound in (Eq. 2.1a) is trivial. The lower bound follows by the approach used in (10, Part b)), where the assumption that X_D is a PAM is not necessary.

For the lower bound, let $\widetilde{X} := X_D + U$ with U uniformly distributed on $[-d_{\min(X_D)}/2, +d_{\min(X_D)}/2]$ and independent of X_D and Z, and let $Y := X_D + Z$. Following the approach of (10, eq(15)) via the data processing inequality for $\widetilde{X} \to X_D \to Y$ we know that

$$I(X_D; Y) \geq I(\widetilde{X}; Y) = h(\widetilde{X}) - h(\widetilde{X}|Y). \tag{2.2}$$

The assumption that X_D is a PAM used in (10) is not needed and we write (10, eq(16)) as

$$h(\widetilde{X}) = H(X_D) + \log(d_{\min(X_D)}). \tag{2.3}$$

Therefore, it remains to upper bound $h(\widetilde{X}|Y)$. The bound follows by the same argument that leads to (10, eq.(19)), which holds under the assumptions of the proposition, i.e., no need to assume a PAM input or a Gaussian noise, and states that for any s^2 and k:

$$h(\tilde{X}|Y) \leq \frac{1}{2}\log\left(2\pi s^2\right) + \frac{\log(\mathrm{e})}{2s^2} E[(\tilde{X} - kY)^2]. \tag{2.4}$$

Thus, by using $s^2 = \mathbb{E}[(\widetilde{X} - kY)^2]$ with $k = \frac{\mathbb{E}[\widetilde{X}Y]}{\mathbb{E}[Y^2]}$ we write (Eq. 2.4) as

$$h(\widetilde{X}|Y) \leq \frac{1}{2}\log\left[2\pi\mathrm{e}\left(\frac{d^2_{\min(X_D)}}{12} + \frac{\mathbb{E}[X_D^2]}{\mathbb{E}[X_D^2] + 1}\right)\right]. \tag{2.5}$$

Combining this, by the non-negativity of mutual information, and since $\frac{\mathbb{E}[X_D^2]}{\mathbb{E}[X_D^2]+1} \leq 1$, the lower bound in (Eq. 2.1a) with the gap expression in (Eq. 2.1b) follows immediately. □

Remark 1. *The proof of Proposition 2.1.1 holds for any continuous U such that $\mathsf{supp}(U) \subseteq [-d_{\min(X_D)}/2, +d_{\min(X_D)}/2]$. In this case $\log(d_{\min(X_D)})$ must be replaced by $h(U)$ in* (Eq. 2.3)*, and $\frac{d^2_{\min(X_D)}}{12}$ must be replaced by the variance of U in* (Eq. 2.5)*. However, for this more general case, it may not be easy to analytically express the entropy as a function of the variance, and to relate them to the bound on the size of the support of the distribution given by $d_{\min(X_D)}$.*

Remark 2. *In the proof of Proposition 2.1.1 we can express $\frac{\mathbb{E}[X_D^2]}{\mathbb{E}[X_D^2]+1} = \mathrm{lmmse}(X_D|Y)$, that is, the linear minimum mean square error of estimating X_D from observation $Y = X_D + Z$. This term can be tightened and replaced by the minimum mean squared error (MMSE) by using the following relationship between conditional differentiable entropy and the MMSE, from (31, Thm. 8.6.6)*

$$\mathrm{mmse}(X|Y) \geq \frac{1}{2\pi \mathrm{e}} 2^{2h(X|Y)} \Rightarrow h(X|Y) \leq \frac{1}{2}\log\left((2\pi\mathrm{e})\ \mathrm{mmse}(X|Y)\right), \tag{2.6}$$

and the expression in (Eq. 2.5) *can be tightened to*

$$h(\widetilde{X}|Y) \leq \frac{1}{2}\log\left[2\pi\mathrm{e}\left(\frac{d^2_{\min(X_D)}}{12} + \mathrm{mmse}(X_D|Y)\right)\right]. \tag{2.7}$$

The bound in (Eq. 2.7) *would lead to a smaller gap than in* (Eq. 2.1b) *given by*

$$\mathsf{G}_{\mathsf{d}(\text{Eq. 2.8})} := \frac{1}{2}\log\left(\frac{2\pi\mathrm{e}}{12}\right) + \frac{1}{2}\log\left(1 + \frac{12\ \mathrm{mmse}(X_D|Y)}{d^2_{\min(X_D)}}\right). \tag{2.8}$$

Remark 3. *If in Proposition 2.1.1 we set $Z = Z_G \sim \mathcal{N}(0,1)$, then we can tighten the upper bound in* (Eq. 2.1a) *to*

$$\mathsf{I_d}(X_D) \leq I(X_D; X_D + Z_G) \leq \min\left(H(X_D), \mathsf{I_g}(\mathbb{E}[X_D^2])\right), \tag{2.9}$$

since a Gaussian input is capacity achieving for the power-constrained point-to-point Gaussian noise channel. Moreover, for $|\mathsf{supp}(X_D)| = N$, the mutual information bounds in (Eq. 2.1a) *are the largest (i.e. maximizes $\mathsf{I_d}(X_D)$) for a PAM constellation. This follows from the fact that PAM is uniformly distributed and satisfies with equality the general inequality $H(X_D) \leq \log(N)$, and we have that*

$$\mathsf{I_d}\left(N, \mathbb{E}[X_D^2]\right) \leq I(X_D;\ X_D + Z_G) \tag{2.10}$$

$$\leq \mathsf{I_g}\left(\min\left(N^2 - 1, \mathbb{E}[X_D^2]\right)\right). \tag{2.11}$$

The lower bound in (Eq. 2.10) *follows by letting $x_{\min} := \min(N^2 - 1, \mathbb{E}[X_D^2])$ and $x_{\max} := \max(N^2 - 1, \mathbb{E}[X_D^2])$. We have*

$$\begin{aligned}
&I(X_D; X_D + Z_G) \\
&\overset{\text{from Proposition 2.1.1}}{\geq} \frac{1}{2}\log\left(1 + (N^2 - 1)\right) - \frac{1}{2}\log\left(1 + \frac{N^2 - 1}{1 + \mathbb{E}[X_D^2]}\right) - \frac{1}{2}\log\left(\frac{\pi e}{6}\right) \\
&= \mathsf{I_g}\left(x_{\min}\right) + \mathsf{I_g}\left(x_{\max}\right) - \mathsf{I_g}\left(x_{\min} + x_{\max}\right) - \frac{1}{2}\log\left(\frac{\pi e}{6}\right) \\
&= \mathsf{I_g}\left(x_{\min}\right) - \mathsf{I_g}\left(\frac{x_{\min}}{1 + x_{\max}}\right) - \frac{1}{2}\log\left(\frac{\pi e}{6}\right) \\
&\geq \mathsf{I_g}\left(x_{\min}\right) - \frac{1}{2}\log\left(\frac{\pi e}{3}\right),
\end{aligned}$$

since $\frac{x_{\min}}{1+x_{\max}} \in [0,1]$. *This, combined with non-negativity of mutual information, gives the lower bound in* (Eq. 2.10).

We next compare the Ozarow-Wyner-B lower bound in Proposition 2.1.1 to bounds available in the literature.

Ozarow-Wyner-A, or Fano-based, bound

Proposition 2.1.1 generalizes the approach of (10, Part b)). Had we generalized (10, Part a)), we would have obtained the following lower bound valid for Gaussian noise only

$$\left[H(X_D) - \mathsf{G}_{\mathsf{d}\,\text{(Eq. 2.12)}}\right]^+ \le I(X_D; X_D + Z_G), \tag{2.12a}$$

$$\mathsf{G}_{\mathsf{d}\,\text{(Eq. 2.12)}} := \xi \log \frac{1}{\xi} + (1-\xi) \log \frac{1}{1-\xi} + \xi \log(N-1) \;, \tag{2.12b}$$

$$\xi := 2Q\left(\frac{d_{\min(X_D)}}{2}\right), \tag{2.12c}$$

where ξ is the union-of-events upper bound on the probability of symbol error for a minimum-distance symbol-by-symbol detector in Gaussian noise from Fano's inequality. We note that a similar Fano-based bounding technique was also used in (32, Theorem 3).

In the following we are interested in showing that certain upper and lower bounds are to within a constant gap of one another, regardless of the channel parameters. For bounds as in (Eq. 2.1), the quantity "G_{d}" upper bounds the difference between the upper and lower bounds. The gap in (Eq. 2.12) (that generalizes (10, Part a)) to any discrete input on the Gaussian noise channel) is bounded if the term

$\xi \log(N-1)$ is bounded; by using the Chernoff's bound for the Q-function, i.e., $Q(x) \le \frac{1}{2}\mathrm{e}^{-x^2/2}$ and by imposing $\xi \log(N-1) \le 1$,we get

$$\text{bounded gap in (Eq. 2.12)} \Longleftrightarrow \log(N-1) \le \mathrm{e}^{d^2_{\min(X_D)}/8}$$

$$\Longleftrightarrow d^2_{\min(X_D)} \ge 8\ln(\log(N-1)),$$

in other words, the minimum distance squared must be of the order of $\ln(\log(N))$ for the gap in (Eq. 2.12) to be bounded. On the other hand, the gap in (Eq. 2.1) (that generalizes (10, Part b)) to any discrete input on any additive noise channel) is bounded as long as the minimum distance is lower bounded by a constant; for example

$$\text{bounded gap in (Eq. 2.1), say } \mathsf{G}_{\mathsf{d}(\text{Eq. 2.1b})} \le \frac{1}{2}\log(6\pi\mathrm{e}) \approx 2.047 \text{ bits}$$

$$\Longleftrightarrow d_{\min(X_D)} \ge 2,$$

that is, the minimum distance does not need to grow in a particular way with the number of points of the constellation, but it is required to be bounded by a constant from below.

DTD-ITA'14 bound

By taking a different approach via Jensen's inequality, we derived the following lower bound for the mutual information with a discrete input on a Gaussian noise channel. As before, let the noise

$Z_G \sim \mathcal{N}(0,1)$ be independent of the discrete input X_D, and let $\Pr[X_D = s_j] = p_j > 0,\ j \in [1:N]$ such that $\sum_{j\in[1:N]} p_j = 1$. We have – the proof can be found in Appendix A:

$$\left[\log(N) - \mathsf{G}_{\mathsf{d}(\text{Eq. 2.13})}\right]^+ \tag{2.13a}$$

$$\leq \left[-\log\left(\sum_{(i,j)\in[1:N]^2} p_i p_j \frac{1}{\sqrt{4\pi}} \mathrm{e}^{-\frac{(s_i-s_j)^2}{4}}\right) - \frac{1}{2}\log\left(2\pi\mathrm{e}\right)\right]^+ \leq I(X_D; X_D + Z_G), \tag{2.13b}$$

$$\mathsf{G}_{\mathsf{d}(\text{Eq. 2.13})} := \frac{1}{2}\log\left(\frac{\mathrm{e}}{2}\right) + \log\left(1 + (N-1)\mathrm{e}^{-d^2_{\min(X_D)}/4}\right). \tag{2.13c}$$

The advantage of the bound in (Eq. 2.13a) (referred to in the following as 'simple DTD-ITA'14 bound') is its simplicity: it only depends on the constellation through the number of points and the minimum distance. The bound in (Eq. 2.13b) (referred to in the following as 'full DTD-ITA'14 bound') is in general tighter than the one in (Eq. 2.13a) but requires the knowledge of the whole "distance spectrum" (all pair-wise distances among constellation points) as well as the "shaping" of the constellation (the a priori probability of each constellation point), which does not make it amenable for closed form analytical computations in general.

Again aiming at a bounded gap, we have

$$\begin{aligned}\text{bounded gap in (Eq. 2.13)} &\Longleftrightarrow (N-1)\mathrm{e}^{-\frac{d^2_{\min(X_D)}}{4}} \leq 1 \\ &\Longleftrightarrow d^2_{\min(X_D)} \geq 4\ln(N-1),\end{aligned}$$

in other words, the minimum distance squared must be of the order of $\log(N)$ for the gap in (Eq. 2.13c) to be bounded. Because of this 'strong' requirement on the minimum distance, in (4) and (3) we could show that a mixed input achieves the capacity region of the classical G-IC to within an additive gap of the order of $O\left(\log\left(\frac{\ln(\min(\mathsf{snr},\mathsf{inr}))}{\gamma}\right)\right)$, rather than a constant gap; but it was nonetheless sufficient to show that TINnoTS with mixed inputs achieves the sum gDoF of the classical G-IC for all channel gains up to a set of zero measure.

Numerical Comparisons

We conclude this subsection by numerically comparing the lower bounds in (Eq. 2.1), (Eq. 2.12) and (Eq. 2.13) for the Gaussian noise channel with a PAM input, which is asymptotically capacity achieving at high SNR (10). In Fig. Figure 1 we plot bounds on $I(X_D; \sqrt{\mathsf{snr}}\, X_D + Z_G)$ vs. snr in dB; here snr represents the SNR at the receiver, $Z_G \sim \mathcal{N}(0,1)$ is the noise, and $X_D \sim \mathsf{PAM}\left(N, \sqrt{\frac{12}{N^2-1}}\right)$ is the input with $N = \mathsf{N_d}(\mathsf{snr}) = \left\lfloor \sqrt{1+\mathsf{snr}} \right\rfloor \approx \mathsf{snr}^{\frac{1}{2}}$. In Fig. 1a we plot the rate bounds while in Fig. 1b the gap to capacity, i.e., the difference between the channel capacity and the different lower bounds. In both figures we show:

1. The black curve is the channel capacity $\mathsf{I_g}(\mathsf{snr})$.

2. The blue curve is the Ozarow-Wyner-B bound in (Eq. 2.1a). From Fig. 1b this bound is asymptotically (for $\mathsf{snr} \geq 30\text{dB}$) to within 0.754 bits of capacity, which is much better than the analytic worst case gap of $\frac{1}{2}\log(6\pi\mathrm{e}) = 2.8395$ bits shown before.

3. The magenta curve is the Ozarow-Wyner-A bound in (Eq. 2.12a). This bound is to within $O(\log(\mathsf{snr}))$ of capacity (i.e., straight line as a function of $\mathsf{snr}|_{\text{dB}}$).

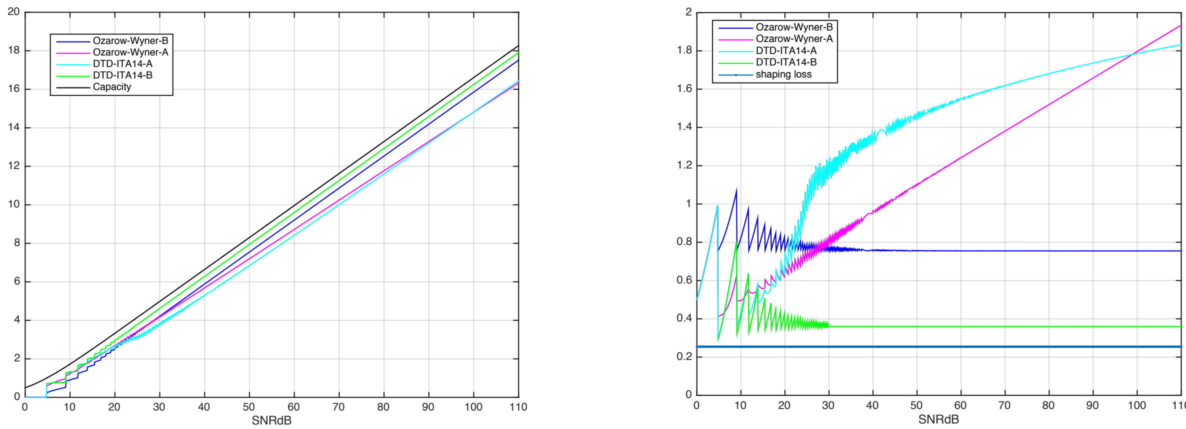
4. The cyan curve is the simple DTD-ITA'14 bound in (Eq. 2.13a). Here we used $N = \mathsf{N_d}(\mathsf{snr}^{1-\epsilon}) \approx \mathsf{snr}^{\frac{1-\epsilon}{2}}$ with $\epsilon = \max\left(0, \frac{\log(\frac{1}{6}\ln(\mathsf{snr}))}{\log(\mathsf{snr})}\right)$. This choice of ϵ was derived in (3, Theorem 3) in

order to have a $O(\log\log(\mathsf{snr}))$ gap to capacity. Had we chosen $\epsilon = 0$ then we could only achieve a 'gap' of $O(\log(\mathsf{snr}))$. Similarly, for the Ozarow-Wyner-A, had we choose the same $\epsilon = \max\left(0, \frac{\log(\frac{1}{6}\ln(\mathsf{snr}))}{\log(\mathsf{snr})}\right)$ a similar $O(\log\log(\mathsf{snr}))$ gap would have been observed.

5. The green curve is the full DTD-ITA'14 bound in (Eq. 2.13b), which from Fig. 1b achieves asymptotically (for $\mathsf{snr} \geq 30$dB) to within 0.36 bits of capacity.

The quantity $\frac{1}{2}\log\left(\frac{\pi e}{6}\right)$ is also shown for reference in Fig. 1b; this is the "shaping loss" for a one-dimensional infinite lattice and is the limiting gap if the number of points N grows faster than $\mathsf{snr}^{1/2}$. The "zig-zag" behavior of the curves at low SNR is due to the floor operation in $N = \left\lfloor\sqrt{1+\mathsf{snr}}\right\rfloor$.

We observe that the relative ranking among the bounds at low SNR (roughly less than 27 dB) is different than at high SNR. In particular we observe a qualitatively different behavior at high SNR: the Ozarow-Wyner-B bound in (Eq. 2.1a) (blue curve) and the full DTD-ITA'14 bound in (Eq. 2.13b) (green curve) result in a constant gap, while the Ozarow-Wyner-A bound in (Eq. 2.12a) (magenta curve) and the simple DTD-ITA'14 bound in (Eq. 2.13a) (cyan curve) result in a gap that grows with SNR; this is in agreement with the previous discussion that points out that for a constant gap in the latter two cases the number of points N must grow slower than $\mathsf{snr}^{1/2}$. The smallest gap at high SNR for $N \approxeq \mathsf{snr}^{1/2}$ is given by our full DTD-ITA'14 bound in (Eq. 2.13b) (green curve); as pointed out earlier, this bound is unfortunately not amenable for closed form analytical evaluations, so in the following we shall use the Ozarow-Wyner-B bound in (Eq. 2.1a) (blue curve) from Proposition 2.1.1 whose simplicity comes at the cost of a larger gap.



(a) Bounds on $I(X_D; \sqrt{\mathsf{snr}}X_D + Z_G)$ vs snr.

(b) Gap to capacity vs snr.

Figure 1: Comparison of different bounds for a PAM input on a Gaussian noise channel.

2.1.1 Cardinality and Minimum Distance Bounds for Sum-Sets

In multi-user settings, we may wish to select one user's input as Gaussian, another as discrete, or both mixtures of discrete and Gaussian. To handle such scenarios, we need bounds on the cardinality and minimum distance of sums of discrete constellations. If X and Y are two sets, we denote the *sum-set* as

$$X + Y := \{x + y | x \in X, y \in Y\}.$$

Tight bounds on the cardinality and the minimum distance of $X + Y$, for general X and Y, are an open problem in the area of additive combinatorics and number theory (33). The following set of sufficient conditions for the sum-set obtained with two PAM constellations (actually the probability with which each point is used does not matter as long as it is strictly positive) will play an important role in evaluating our inner bound.

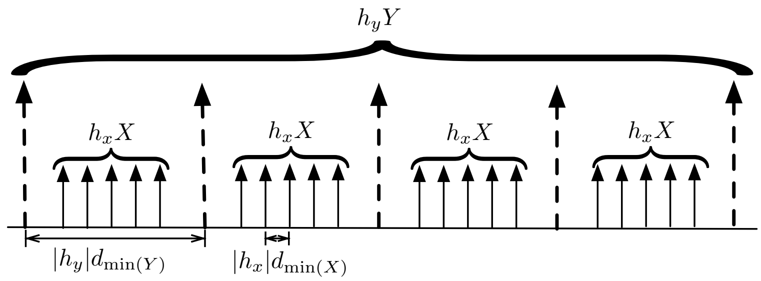


Figure 2: Structure of the sum-set under the conditions in Proposition 2.1.2.

Proposition 2.1.2. *Let* $(h_x, h_y) \in \mathbb{R}^2$ *be two constants such that* $h_x \cdot h_y \neq 0$.

Let $X \sim \mathsf{PAM}(|X|, d_{\min(X)})$ *and* $Y \sim \mathsf{PAM}(|Y|, d_{\min(Y)})$. *Then*

$$|h_xX + h_yY| = |X||Y|, \tag{2.14}$$

$$d_{\min(h_xX+h_yY)} = \min\left(|h_x|d_{\min(X)}, |h_y|d_{\min(Y)}\right), \tag{2.15}$$

under the following conditions

$$\textit{either}\ \ |Y||h_y|d_{\min(Y)} \leq |h_x|d_{\min(X)}, \tag{2.16a}$$

$$\textit{or}\ \ |X||h_x|d_{\min(X)} \leq |h_y|d_{\min(Y)}. \tag{2.16b}$$

Proof. The condition in (Eq. 2.16) is such that one PAM constellation is completely contained within two points of the other PAM constellation, see Fig. Figure 2 for a visual illustration. □

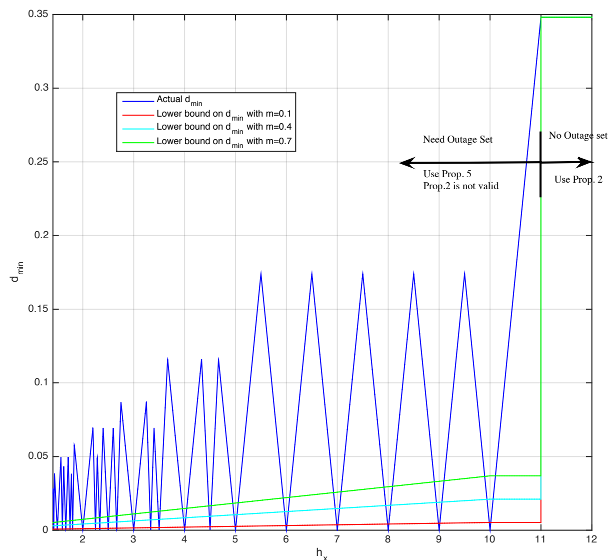


Figure 3: Minimum distance (blue line) for the sum-set $h_x X + h_y Y$ as a function of h_x for fixed $h_y = 1$ and for $X \sim Y \sim \mathsf{PAM}\,(10, 1)$. On the right of the vertical green line Proposition 2.1.2 is valid. On the left of the vertical green line Proposition 2.1.3 must be used; in this case, the minimum distance lower bound in (Eq. 2.19a) holds for set of h_x's for which the blue line is above the red / cyan / green line, where the red, cyan and green lines represent a different value for the measure of the outage set.

We will refer to the condition in (Eq. 2.16) as the *non-overlap* condition. Unfortunately, Proposition 2.1.2 is not sufficient for our purposes because it restricts the set of channel parameters for which we can compute the minimum distance to those cases where the non-overlap condition holds. When the non-overlap condition in (Eq. 2.16) is not satisfied, the minimum distance is very sensitive to the fractional values of h_x and h_y. Fig. Figure 3 shows, in solid blue line, the minimum distance for the sum-set $h_x X + h_y Y$ as a function of h_x for fixed $h_y = 1$ and where X and Y are the same $\mathsf{PAM}(10, 1)$ constellation. It can be observed that there are channel gains for which the minimum distance is zero; those occur on the left of the vertical green line, which separates the values of h_x for which Proposition 2.1.2 is valid (right side) for those where it is not (left side).

Remark 4. *To bound the cardinality and the minimum distance when the condition in* (Eq. 2.16) *is not satisfied we use the approach of (34) and (35). In (34; 35) it was observed that capacity of G-IC is sensitive to the fractional values of the channel gains, and the G-IC input out relationship can be described by*

$$Y_1 = h_{11x} \lceil h_{11} \rceil X_1 + h_{12x} \lceil h_{12} \rceil X_2 + Z_1 \tag{2.17a}$$

$$Y_2 = h_{21x} \lceil h_{21} \rceil X_1 + h_{22x} \lceil h_{22} \rceil X_2 + Z_2 \tag{2.17b}$$

where h_{ijx} and $\lceil h_{ij} \rceil$ are the fractional and integer parts of the channel gain h_{ij}, respectively. In this representation, the integer part $\lceil h_{ij} \rceil$ captures the magnitude and coarse structure of the channel gain, and the fractional part h_{ijx} is thought to capture the finer structure of the channel gain.

The variations shown on Fig. Figure 3 mainly depend on the fractional part of the channel gains. Following the approach of (34; 35), we define the "outage set"- that is the set of fractional channel gains (for fixed integer part) for which the minimum distance falls below a given target. Moreover, the size of the outage set and the target minimum distance are tunable parameters.

Finally, we remark that for $X \sim \mathsf{PAM}(|X|, d_{\min(X)})$ *and* $Y \sim \mathsf{PAM}(|Y|, d_{\min(Y)})$ *the resulting sum-set given by* $h_x X + h_y Y$ *can always be restated as*

$$h_x X + h_y Y = h_{xx} \lceil h_x \rceil X + h_{yy} \lceil h_y \rceil Y = h_{xx} \hat{X} + h_y \hat{Y} \tag{2.17c}$$

where $(h_{xx}, h_y) \in [0,1]^2$ *and where* $\hat{X} \sim \mathsf{PAM}(|X|, d_{\min(\lceil h_x \rceil X)})$ *and* $\hat{Y} \sim \mathsf{PAM}(|Y|, d_{\min(\lceil h_y \rceil Y)})$*. Therefore, for the remainder of the thesis, we assume that the integer parts* $\lceil h_x \rceil, \lceil h_y \rceil$ *are fixed and we consider Lebesgue measure over the fractional parts* $(h_{xx}, h_y) \in [0,1]^2$*.*

Applications of discrete and mixed inputs to theG-IC will be investigate in detail in Chapter 4.

We will use the following result to bound the cardinality and the minimum distance when the condition in (Eq. 2.16) is not satisfied.

Proposition 2.1.3. *Let* $X \sim \mathsf{PAM}(|X|, d_{\min(X)})$ *and* $Y \sim \mathsf{PAM}(|Y|, d_{\min(Y)})$*. Then for* $(h_{xx}, h_{yy}) \in [0,1]^2$

$$|h_{xx} X + h_{yy} Y| = |X||Y| \text{ almost everywhere (a.e.),} \tag{2.18}$$

and for any $\gamma > 0$ there exists a set $E \subseteq [0,1]^2$ such that for all $(h_{xx}, h_{yy}) \in E$

$$d_{\min(h_{xx}X+h_{yy}Y)} \geq \kappa_{\gamma,|X|,|Y|} \cdot \min\left(|h_{xx}|\, d_{\min(X)}, |h_{yy}|\, d_{\min(Y)}, \Upsilon_{|h_{xx}|,|h_{yy}|,|X|,|Y|}\right), \tag{2.19a}$$

$$\kappa_{\gamma,|X|,|Y|} := \frac{\gamma/2}{1+\ln(\max(|X|,|Y|))}, \tag{2.19b}$$

$$\Upsilon_{|h_{xx}|,|h_{yy}|,|X|,|Y|} := \max\left(\frac{|h_{xx}|\, d_{\min(X)}}{|Y|}, \frac{|h_{yy}|\, d_{\min(Y)}}{|X|}\right), \tag{2.19c}$$

where the Lebesgue measure of the complement of the set E ($E^c = [0,1]^2 \backslash E$ is referred to as the outage set) satisfies $m(E^c) \leq \gamma$.

Proof. □

The reason we need to introduce an outage set in Proposition 2.1.3 is that there are values of (h_x, h_y) for which the minimum distance is zero, as it can be seen from Fig. Figure 3. In computing the gap later on, we want to exclude the set of channel gains for which the minimum distance is too close to zero; the measure of this set can be controlled through the parameter γ. The green, cyan, and red lines in Fig. Figure 3 represent lower bounds on the minimum distance that are valid everywhere except for a set of measure no greater than $\gamma = 0.1, 0.3$ and 0.7, respectively. It is important to notice that the set of channel gains for which the minimum distance is exactly zero satisfies:

Proposition 2.1.4. *Let $X \sim \mathsf{PAM}(|X|, d_{\min(X)})$ and $Y \sim \mathsf{PAM}(|Y|, d_{\min(Y)})$. Then the set of $(h_{xx}, h_{yy}) \in [0,1]^2$ such that $d_{\min(h_{xx}X+h_{yy}Y)} = 0$ has measure zero.*

Proof. The proof follows by observing that the set of channel gains for which $d_{\min(h_{xx}X+h_{yy}Y)} = 0$ and $|h_x X + h_y Y| \neq |X||Y|$ are equivalent and given by eq.(Eq. B.1) in Appendix B. The rest of the proof is similar to that of Proposition 2.1.3. $\square$

Remark 5. *Different minimum distance bounds for sum-sets based on Diophantine approximations were used in (36). For example, consider the sum-set $h_1 X + h_2 X$, i.e., both transmitters use the same PAM constellation X, where $h_1^2 = h_S^2 \mathsf{snr}$ and $h_2^2 = h_I^2 \mathsf{snr}^\alpha$ for some fixed $(h_S, h_I) \in \mathbb{R}^2$ and $\alpha > 0$. The authors of (36) focused on the degrees of freedom (DoF) for the case when $\alpha = 1$; in this case the minimum distance can be lower bounded as follows*

$$
\begin{aligned}
d_{\min(h_1 X + h_2 X)} &= \min_{x_{1i}, x_{2i} \in X} |h_1 x_{1i} - h_2 x_{2i}| \\
&= \min_{z_{1i}, z_{2i} \in [-\frac{N}{2} : \frac{N}{2}],} |h_S \sqrt{\mathsf{snr}} d_{\min(X)}\ z_{1i} - h_I \sqrt{\mathsf{snr}} d_{\min(X)}\ z_{2i}| \\
&= \sqrt{\mathsf{snr}}\, d_{\min(X)} \min_{z_{1i}, z_{2i} \in [-\frac{N}{2} : \frac{N}{2}],} |h_S z_{1i} - h_I z_{2i}| && \text{(2.20a)} \\
&\geq \kappa_\epsilon \frac{2^\epsilon}{N^\epsilon} \sqrt{\mathsf{snr}}\, d_{\min(X)}, && \text{(2.20b)}
\end{aligned}
$$

where the inequality in (Eq. 2.20b) *comes from Diophantine approximation results, specifically from the Khintchine-Groshev theorem, and says that for almost all real numbers (h_S, h_I) and for any $\epsilon > 0$ there exists a constant $\kappa_\epsilon > 0$, whose analytical expression is not known, such that the bound in* (Eq. 2.20b) *holds.*

Unfortunately, bounds such as (Eq. 2.20b) *are only well suited for the derivation of DoF (i.e $\alpha = 1$ but not for gDoF (i.e. $\alpha \neq 1$), which is of interest here. The fundamental problem is that for $\alpha \neq 1$,* *the factorization in* (Eq. 2.20a) *is no longer possible and κ_ϵ may end up being a function of* snr *and* α*. Moreover, the fact that we have auxiliary constants ϵ and κ_ϵ in* (Eq. 2.20b)*, and where κ_ϵ is essentially not known in closed form, makes derivation of closed form gap results very difficult.*

2.1.2 Examples

In this Section we give an example of how we intend to use discrete inputs for the G-IC by considering the familiar point-to-point power-constrained additive white Gaussian noise channel. The goal is to derive some properties / results for a simple setting that we shall use often in the subsequent sections. Specifically, we aim to show that the unit-energy discrete input X_D with a properly chosen number of points $N = |\mathsf{supp}(X_D)|$ as a function of snr achieves, roughly speaking ($\approx$)

$$I(X_D; \sqrt{\mathsf{snr}} X_D + Z_G) \approx \log(N), \text{ where } Z_G \sim \mathcal{N}(0,1), \tag{2.21}$$

$$I(X_G; \sqrt{\mathsf{snr}} X_G + X_D + Z_G) \approx \mathsf{I_g}(\mathsf{snr}), \text{ where } X_G \sim \mathcal{N}(0,1), \tag{2.22}$$

that is, the discrete input X_D is a “good” input and a “good” interference. To put it more clearly, when we use a discrete constellation as input, as in (Eq. 2.21), the mutual information is roughly equal to the entropy of the constellation, which is highly desirable. On the other hand, if the interference, unknown to transmitter and receiver, is from a discrete constellation as in (Eq. 2.22), the mutual information is roughly as if there was no interference, which is again highly desirable. In contrast, a Gaussian input instead of X_D would be the “best” input for (Eq. 2.21) but the “worst” interference/noise in (Eq. 2.22). We next formalize the approximate statements in (Eq. 2.21) and (Eq. 2.22).

Gaussian Channel

Consider the point-to-point power-constrained Gaussian noise channel

$$Y = \sqrt{\mathsf{snr}}\, X + Z_G, \tag{2.23a}$$

$$\mathbb{E}[X^2] \le 1,\ Z_G \sim \mathcal{N}(0,1), \tag{2.23b}$$

where X is the information carrying signal, independent of the noise Z_G. The capacity of this channel, as a function of the SNR snr, is $C(\mathsf{snr}) = \mathsf{I_g}(\mathsf{snr})$ and is achieved by $X \sim \mathcal{N}(0,1)$ for every snr. Consider now the input $X = X_D \sim \mathsf{PAM}\left(N, \sqrt{\frac{12}{N^2-1}}\right)$ on the channel in (Eq. 2.23). By Proposition 2.1.1 and Remark 3

$$\left[\log(N) - \frac{1}{2}\log\left(\frac{2\pi \mathrm{e}}{12}\right) - \frac{1}{2}\log\left(1 + \frac{N^2-1}{\mathsf{snr}}\right)\right]^+ \le I(X_D; \sqrt{\mathsf{snr}}X_D + Z_G) \le \mathsf{I_g}(\mathsf{snr}). \tag{2.24}$$

By observing the bounds in (Eq. 2.24), we see that for a PAM input to be optimal to within a constant gap we need that $\log(N) \approx \mathsf{I_g}(\mathsf{snr})$ and that $\frac{N^2-1}{\mathsf{snr}}$ is upper bounded by a constant. By choosing $N = \left\lfloor \sqrt{1+\mathsf{snr}} \right\rfloor =: \mathsf{N_d}(\mathsf{snr})$ it is easy to see that a PAM input can achieve the capacity $\mathsf{I_g}(\mathsf{snr})$ to within $\frac{1}{2}\log\left(\frac{2\pi \mathrm{e}}{3}\right) \approx 1.25$ bits, where the maximum gap is for $\mathsf{snr} = 3 - \epsilon$ for some $0 < \epsilon \ll 1$.

Note that, had we kept the term $\frac{\mathbb{E}[X_D^2]}{\mathbb{E}[X_D^2]+1}$ in (Eq. 2.5), the bound in (Eq. 2.24) would have had $\frac{N^2-1}{\mathsf{snr}+1}$ in place of $\frac{N^2-1}{\mathsf{snr}}$ and would have resulted in a gap of at most $\frac{1}{2}\log\left(\frac{\pi \mathrm{e}}{2}\right) \approx 1.047$ bits. As always, bounds which allow for expressions that are easier to manipulate analytically come at the expense of a larger gap.

Gaussian Channel with States

The above example showed that a discrete input with $\log(N) \approx \mathsf{I_g}(\mathsf{snr})$ is a "good" input in the sense alluded to by (Eq. 2.21). We now show that a discrete interference is a "good" interference in the sense alluded to by (Eq. 2.22). We study an extension of the channel in (Eq. 2.23) by considering an *additive state T available neither at the encoder nor at the decoder*. The input-output relationship is

$$Y = \sqrt{\mathsf{snr}}\, X + h\, T + Z_G : \tag{2.25a}$$

$$\mathbb{E}[X^2] \le 1,\ Z_G \sim \mathcal{N}(0,1), \tag{2.25b}$$

$$T \text{ discrete with finite power.} \tag{2.25c}$$

It is well known (37, Section 7.4) that the capacity of the channel with random state in (Eq. 2.25) is

$$C = \max_{P_X} I(X;Y) \le \max_{P_X} I(X;Y|T) = \mathsf{I_g}(\mathsf{snr}). \tag{2.26}$$

From (25) we know that $X = X_G \sim \mathcal{N}(0,1)$ is at most 1/2 bit from the capacity C, but the value of the capacity is unknown. In particular it is not know whether the gap to the interference free capacity $\mathsf{I_g}(\mathsf{snr}) - C$ is a bounded function of snr.

Assume we use the input $X = X_G \sim \mathcal{N}(0,1)$, as a Gaussian input is not too bad for an additive noise channel (25); assume also that $d_{\min(T)} > 0$; then the capacity C is lower bounded as:

$$C \geq \mathsf{I_g}(\mathsf{snr}) - \mathsf{G}_{\mathsf{d}(\text{Eq. 2.27})}, \tag{2.27a}$$

$$\mathsf{G}_{\mathsf{d}(\text{Eq. 2.27})} := \frac{1}{2}\log\left(\frac{2\pi \mathrm{e}}{12}\right) + \frac{1}{2}\log\left(1 + \frac{12}{d^2_{\min(T)}}\right), \tag{2.27b}$$

since

$$\begin{aligned} I(X_G;Y) &= I(X_G; \sqrt{\mathsf{snr}}\, X_G + h\,T + Z_G) \\ &= \underbrace{h(\sqrt{\mathsf{snr}}\, X_G + h\,T + Z_G) - h(\sqrt{\mathsf{snr}}\, X_G + Z_G)}_{\geq \mathsf{I_d}\left(\frac{h}{\sqrt{1+\mathsf{snr}}}T\right) \geq H(T) - \mathsf{G}_{\mathsf{d}(\text{Eq. 2.27})}} \\ &- \underbrace{\left(h(h\,T + Z_G) - h(Z_G)\right)}_{\leq H(T)} \\ &+ \underbrace{\left(h(\sqrt{\mathsf{snr}}\, X_G + Z_G) + h(Z_G)\right)}_{= \mathsf{I_g}(\mathsf{snr})}. \end{aligned}$$

Thus, as long as $d_{\min(T)}$ is lower-bounded by a constant, it is possible to achieve the interference-free capacity to within the constant gap in (Eq. 2.27b) even when the state is unknown to both the transmitter and the receiver.

The expression in (Eq. 2.27) can be readily used to obtained inner bounds on the capacity region of a G-IC where one user has a Gaussian input and the other a discrete input and where the discrete input is treated as noise, as we shall do in the next sections.

CHAPTER 3

ON THE TWO-USER INTERFERENCE CHANNEL WITH LACK OF KNOWLEDGE OF THE INTERFERENCE CODEBOOK AT ONE RECEIVER

3.1 Introduction

A classical assumption in multi-user information theory is that each node in the network possesses knowledge of the codebooks used by every other node. However, such an assumption might not be practical in heterogeneous, cognitive, distributed or dynamic networks. For example, in very large ad-hoc networks, where nodes enter and leave at will, it might not be practical for new nodes to learn the codebooks of old nodes and vice-versa. In cognitive radio scenarios, where new cognitive systems coexist with legacy systems, requiring the legacy systems to know the codebooks of the new cognitive systems might not be viable. This motivates the study of networks where each node possesses only a subset of the codebooks used in the network. We will refer to such systems as networks with *partial codebook knowledge* and to nodes with only knowledge of a subset of the codebooks as *oblivious* nodes.

We make progress on this front by demonstrating that certain rates are achievable for the Gaussian noise interference channel with oblivious receivers (G-IC-OR) through the evaluation of a simplified Han-Kobayashi scheme (38) in which joint decoding of the intended and interfering messages is not required at the oblivious receiver. The major contribution of this work is the realization that Gaussian

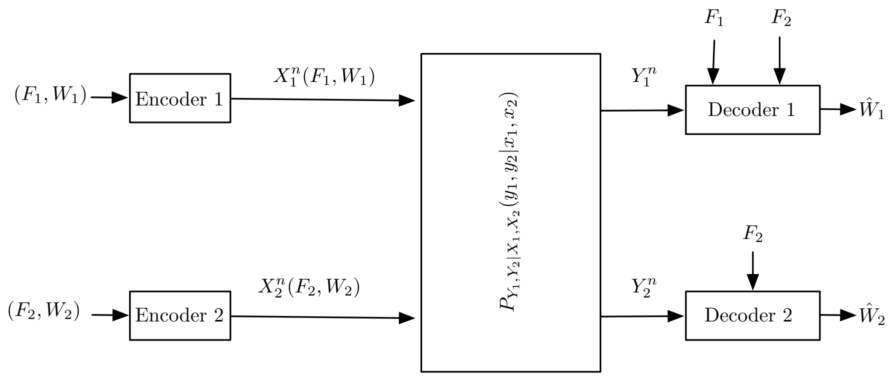


Figure 4: The IC-OR, where F_1 and F_2 represent codebook indices known to one or both receivers.

inputs perform poorly in the proposed achievable region. We therefore propose to use a class of inputs that we termed *mixed inputs*. A mixed input is a random variable that is a mixture of a continuous and a discrete part, such as for example a Gaussian random variable and a uniformly distributed random variable on an equally spaced set of discrete points. We then properly design the distribution of the mixed input as a function of the channel parameters.

3.2 Channel Model

3.2.1 General Memoryless IC-OR

An IC-OR consists of the two-user memoryless interference channel $(\mathcal{X}_1, \mathcal{X}_2, P_{Y_1 Y_2|X_1 X_2}, \mathcal{Y}_1, \mathcal{Y}_2)$ where receiver 2 is oblivious of transmitter 1's codebook. We use the terminology "codebook" to denote a set of codewords and the (one-to-one) mapping of the messages to these codewords. We model lack of codebook knowledge as in (19), where transmitters use randomized encoding functions, which are indexed by a message index and a "codebook index" (F_1 and F_2 in Fig. Figure 4). An oblivious receiver is unaware of the "codebook index" (F_1 is not given to decoder 2 in Fig. Figure 4) and hence does not know how codewords are mapped to messages. The basic modeling assumption is that without the knowledge of the codebook index a codeword looks unstructured. More formally, by extending (20, Definition 2), a $(2^{nR_1}, 2^{nR_2}, n)$ code for the IC-OR with enabled time sharing is a six-tuple $(P_{F_1|Q^n}, \sigma_1^n, \phi_1^n, P_{F_2|Q^n}, \sigma_2^n, \phi_2^n)$, where the distribution $P_{F_i|Q^n}$, $i \in [1:2]$, is over a finite alphabet $\mathcal{F}_i$ conditioned on the time-sharing sequences q^n from some finite alphabet $\mathcal{Q}$, and where the encoders σ_i^n and the decoders ϕ_i^n, are mappings

$$\begin{aligned}
&\sigma_1^n : [1:2^{nR_1}] \times [1:|\mathcal{F}_1|] \to \mathcal{X}_1^n, \\
&\sigma_2^n : [1:2^{nR_2}] \times [1:|\mathcal{F}_2|] \to \mathcal{X}_2^n, \\
&\phi_1^n : [1:|\mathcal{F}_1|] \times [1:|\mathcal{F}_2|] \times \mathcal{Y}_1^n \to [1:2^{nR_1}], \\
&\phi_2^n : [1:|\mathcal{F}_2|] \times \mathcal{Y}_2^n \to [1:2^{nR_2}].
\end{aligned}$$

Moreover, when user 1's codebook index is unknown at decoder 2, the encoder σ_1^n and the distribution $P_{F_1|Q^n}$ must satisfy

$$\begin{aligned}
&\mathbb{P}[X_1^n = x_1^n | Q^n = q^n] \\
&= \sum_{w_1=1}^{2^{nR_1}} \sum_{f_1=1}^{|\mathcal{F}_1|} P_{F_1|Q^n}(f_1|q^n)\, 2^{-nR_1}\, \delta\big(x_1^n - \sigma_1^n(w_1, f_1)\big) \\
&= \prod_{t \in [1:n]} P_{X_1|Q}(x_{1t}|q_t), \qquad (3.1)
\end{aligned}$$

according to some distribution $P_{X_1|Q}$. In other words, when averaged over the probability of selecting a given codebook and over a uniform distribution on the message set, the transmitted codeword conditioned on any time sharing sequence has an i.i.d. distribution according to some distribution $P_{X_1|Q}$. We refer the reader to (20, Remark 1) for further justifications of the condition in (Eq. 3.1).

A non-negative rate pair (R_1, R_2) is said to be achievable if there exist a sequence of encoding functions $\sigma_1^n(W_1, F_1),\ \sigma_2^n(W_2, F_2),$ and decoding functions $\phi_1^n(Y_1^n, F_1, F_2),\ \phi_2^n(Y_2^n, F_2)$, such that the average probability of error satisfies $\max_{i\in[1:2]} \mathbb{P}[\widehat{W}_i \neq W_i] \to 0$ as $n \to +\infty$. The capacity region is defined as the convex closure of all achievable rate pairs (R_1, R_2) (37).

Remark 6. *One of the key features of our model is that the codebook index may change from codeword to codeword. In particular, one can show that the number of codebooks is given by $|F| = |X|^{n2^{nR}}$ (20). Therefore, communicating the index of the codebook - before the transmission of every codeword - incurs a non vanishing overhead. For more discussion on which communication schemes are permitted and which are not we refer reader to (19; 20).*

3.2.2 Injective Semi-Deterministic IC-OR

For a general memoryless IC-OR, no restrictions are imposed on the transition probability $P_{Y_1Y_2|X_1X_2}$. The ISD-IC-OR is a special IC-OR with transition probability

$$P_{Y_1Y_2|X_1X_2}(y_1, y_2|x_1, x_2) = \sum_{t_1,t_2} P_{T_1|X_1}(t_1|x_1) P_{T_2|X_2}(t_2|x_2) \cdot \delta\big(y_1 - g_1(x_1, t_2)\big)\, \delta\big(y_2 - g_2(x_2, t_1)\big), \tag{3.2}$$

for some memoryless transition probabilities $P_{T_1|X_1}$ and $P_{T_2|X_2}$, and some deterministic functions $g_1(\cdot,\cdot)$ and $g_2(\cdot,\cdot)$ that are injective when their first argument is held fixed (39). The ISD property implies that

$$H(Y_1|X_1) = H(T_2) \text{ and } H(Y_2|X_2) = H(T_1), \quad \forall P_{X_1X_2} = P_{X_1}P_{X_2}, \tag{3.3}$$

or in other words that the T_u is a deterministic function of the pair (Y_u, X_u), $u \in [1:2]$. For channels with continuous alphabets, the summation in (Eq. 3.2) should be replaced with an integral and the discrete entropies in (Eq. 3.3) with the differential entropies.

3.3 Outer Bounds

In this section we present novel outer bounds for the IC-OR. In particular, we derive the single rate bounds that are valid for a general memoryless IC-OR and a sum-rate bound that is valid for the ISD-IC-OR only.

We begin by proving a property of the output distributions that is key to deriving single-letter expressions in our outer bounds; this property holds for a general memoryless IC-OR.

Proposition 3.3.1. *The output of the oblivious decoder has a product distribution conditioned on the signal whose codebook is known, that is,*

$$P_{Y_2^n|X_2^n,F_2}(y_2^n|x_2^n,f_2) = \prod_{i=1}^{n} P_{Y_{2i}|X_{2i}}(y_{2i}|x_{2i}).$$

which implies

$$H(Y_2^n|X_2^n,F_2) = \sum_{i=1}^{n} H(Y_{2i}|X_{2i})$$
$$\stackrel{\text{f}orISD-IC-OR}{=} \sum_{i=1}^{n} H(T_{1i}).$$

Proof of Proposition 3.3.1. Starting from the joint distribution of Y_2^n, X_1^n conditioned on X_2^n, F_2 we have that

$$\begin{aligned}
&P_{Y_2^n,X_1^n|X_2^n,F_2}(y_2^n,x_1^n|x_2^n,f_2)\\
&\stackrel{\text{a)}}{=} P_{X_1^n}(x_1^n)\prod_{i=1}^{n} P_{Y_{2i}|X_{1i},X_{2i}}(y_{2i}|x_{1i},x_{2i})\\
&\stackrel{\text{b)}}{=} \prod_{i=1}^{n} P_{X_{1i}}(x_{1i})\prod_{i=1}^{n} P_{Y_{2i}|X_{1i},X_{2i}}(y_{2i}|x_{1i},x_{2i})\\
&\stackrel{\text{c)}}{=} \prod_{i=1}^{n} P_{Y_{2i},X_{1i}|X_{2i}}(y_{2i},x_{1i}|x_{2i})
\end{aligned}$$

where the equalities follows from: a) the inputs are independent and the channel is memoryless; b) the assumption that X_1^n has a product distribution if not conditioned on F_1 as in (Eq. 3.1); and c) the inputs are independent. By marginalizing with respect to X_1^n yields

$$P_{Y_2^n|X_2^n,F_2}(y_2^n|x_2^n,f_2) = \prod_{i=1}^{n}\sum_{x_{1i}} P_{Y_{2i},X_{1i}|X_{2i}}(y_{2i},x_{1i}|x_{2i}) = \prod_{i=1}^{n} P_{Y_{2i}|X_{2i}}(y_{2i}|x_{2i}),$$

as claimed.

□

The main result of the section is the following upper bound:

Theorem 3.3.2. *Any achievable rate pair* (R_1, R_2) *for the IC-OR must satisfy*

$$R_1 \leq I(Y_1; X_1 | X_2, Q), \qquad \textit{(memoryless IC-OR)} \tag{3.4a}$$

$$R_2 \leq I(Y_2; X_2 | Q), \qquad \textit{(memoryless IC-OR)} \tag{3.4b}$$

$$\begin{aligned} R_1 + R_2 &\leq H(Y_1|Q) + H(Y_2|U_2, Q) \\ &\quad - H(T_2|X_2, Q) - H(T_1|Q) \qquad \textit{(memoryless ISD-IC-OR)} \\ &= I(Y_1; X_1, X_2|Q) + I(Y_2; X_2|U_2, Q), \end{aligned} \tag{3.4c}$$

for some input distribution that factors as

$$P_{Q,X_1,X_2,U_2}(q, x_1, x_2, u_2) = P_Q(q) P_{X_1|Q}(x_1|q) P_{X_2|Q}(x_2|q) P_{T_2|X_2}(u_2|x_2), \tag{3.4d}$$

and with $|\mathcal{Q}| \leq 2$*. We denote the region in* (Eq. 3.4) *as* $\mathcal{R}_{\text{out}}$*.*

Proof of Theorem 3.3.2. By Fano's inequality $H(W_1|Y_1^n, F_1, F_2) \leq n\epsilon_n$ and $H(W_2|Y_2^n, F_2) \leq n\epsilon_n$ for some $\epsilon_n \to 0$ as $n \to \infty$.

We begin with the R_1-bound (non-oblivious receiver) in (Eq. 3.4a):

$$
\begin{aligned}
n(R_1 - \epsilon_n) &\overset{\text{a)}}{\leq} I(W_1; Y_1^n, F_1, F_2) \\
&\overset{\text{b)}}{\leq} I(W_1; Y_1^n | F_1, F_2, W_2) \\
&\overset{\text{c)}}{\leq} I(X_1^n; Y_1^n | F_1, F_2, X_2^n) \\
&\overset{\text{d)}}{=} H(Y_1^n | F_1, F_2, X_2^n) - \sum_{i=1}^{n} H(Y_{1i} | X_{1i}, X_{2i}) \\
&\overset{\text{e)}}{\leq} \sum_{i=1}^{n} H(Y_{1i} | X_{2i}) - \sum_{i=1}^{n} H(Y_{1i} | X_{1i}, X_{2i}) \\
&= \sum_{i=1}^{n} I(X_{1i}; Y_{1i} | X_{2i}),
\end{aligned}
$$

where the (in)-equalities follow from: a) Fano's inequality, b) giving W_2 as side information and using the fact that F_1, F_2, W_1 and W_2 are mutually independent; c) data processing $(F_i, W_i) \rightarrow X_i^n \rightarrow$

Y_1^n, for $i \in [1:2]$; d) because the channel is memoryless; and e) by chain rule of entropy and by "conditioning reduces entropy". For the R_2-bound (oblivious receiver) in (Eq. 3.4b) we have:

$$\begin{aligned}
n(R_2 - \epsilon_n) &\overset{\text{a)}}{\le} I(W_2; Y_2^n, F_2) \\
&\overset{\text{b)}}{\le} I(W_2; Y_2^n | F_2) \\
&\overset{\text{c)}}{\le} I(X_2^n; Y_2^n | F_2) \\
&\overset{\text{d)}}{=} H(Y_2^n | F_2) - \sum_{i=1}^{n} H(Y_{2i} | X_{2i}) \\
&\overset{\text{e)}}{\le} \sum_{i=1}^{n} H(Y_{2i}) - \sum_{i=1}^{n} H(Y_{2i} | X_{2i}) \\
&= \sum_{i=1}^{n} I(X_{2i}; Y_{2i}),
\end{aligned}$$

where the (in)-equalities follow from: a) Fano's inequality; b) the fact that F_2 and W_2 are independent; c) data processing $(F_i, W_i) \to X_i^n \to Y_1^n$, for $i \in [1:2]$; d) by Proposition 3.3.1; and e) from chain rule of entropy and "conditioning reduces entropy".

Next, by providing U_2 as side information to receiver 2 (oblivious receiver) similarly to (39)[1], where U_2 is jointly distributed with the inputs according to (Eq. 3.4d), we have:

$$\begin{aligned}
n(R_1 + R_2 - 2\epsilon_n) &\overset{a)}{\leq} I(X_1^n; Y_1^n | F_1, F_2) + I(X_2^n; Y_2^n, U_2^n | F_2) \\
&= H(Y_1^n | F_1, F_2) - H(Y_1^n | F_1, F_2, X_1^n) \\
&+ H(U_2^n | F_2) - H(U_2^n | F_2, X_2^n) \\
&+ H(Y_2^n | F_2, U_2^n) - H(Y_2^n | F_2, X_2^n, U_2^n) \\
&\overset{b)}{=} H(Y_1^n | F_1, F_2) - H(T_2^n | F_1, F_2) \\
&+ H(U_2^n | F_2) - H(U_2^n | F_2, X_2^n) \\
&+ H(Y_2^n | F_2, U_2^n) - H(T_1^n) \\
&\overset{c)}{=} H(Y_1^n | F_1, F_2) - H(T_2^n | F_1, F_2) \\
&+ H(T_2^n | F_2) - H(T_2^n | F_2, X_2^n) \\
&+ H(Y_2^n | F_2, U_2^n) - H(T_1^n) \\
&\overset{d)}{=} H(Y_1^n | F_1, F_2) + H(Y_2^n | F_2, U_2^n) - H(T_2^n | X_2^n) - H(T_1^n) \\
&\overset{e)}{\leq} \sum_{i=1}^{n} H(Y_{1i} | F_1, F_2) + H(Y_{2i} | F_2, U_{2i}) - H(T_{2i} | X_{2i}) - H(T_{1i}), \\
&\overset{f)}{\leq} \sum_{i=1}^{n} H(Y_{1i}) + H(Y_{2i} | U_{2i}) - H(T_{2i} | X_{2i}) - H(T_{1i}),
\end{aligned}$$

---

[1] Random variable U_2 is obtained by passing X_2 through an auxiliary channel described by $P_{T_2|X_2}$. Intuitively, U_2 represents interference caused by X_2 plus noise at the output Y_1. The idea is that providing a noisy version of X_2 as side information will result in a tighter bound than for example giving just X_2.

where (in)-equalities follow from: a) by Fano's inequality and by giving U_2 as side information and by proceeding as done for the single rate bounds up to step labeled "c)"; b) by the injective property in (Eq. 3.2) and the independence of (X_1^n, T_1^n) and X_2^n; c) by definition of U_2 in (Eq. 3.4d) we have $H(U_2^n|F_2) = H(T_2^n|F_2)$; d) by independence of the messages we have $H(T_2^n|F_1, F_2) - H(T_2^n|F_2) = 0$; e) since the channel is memoryless and thus $H(T_2^n|F_2, X_2^n) = H(T_2^n|X_2^n) = \sum_{i=1}^n H(T_{2i}|X_{2i})$ and since $H(T_1^n) = H(Y_2^n|X_2^n)$ can be single-letterized by using Proposition 3.3.1; and f) by conditioning reduces entropy.

The introduction of a time-sharing random variable $Q \sim \mathsf{Unif}[1:n]$ yields the bounds in (Eq. 3.4). The Fenchel-Eggleston-Caratheodory theorem (40, Chapter 14) guarantees that we may restrict attention to $|Q| \leq 2$ without loss of optimality.

Finally, the equality in (Eq. 3.4c) follows from the injective property in (Eq. 3.2), the independence of the inputs and the memoryless property of the channel, i.e.,

$$
\begin{aligned}
&H(T_2|X_2) = H(T_2|X_1, X_2) = H(Y_1|X_1, X_2, Q),\\
&H(T_1|Q) = H(T_1|U_2, Q, X_2).
\end{aligned}
$$

This concludes the proof. □

3.4 Capacity Results

In this section we prove that the outer bound in (Eq. 3.4) is (approximately) tight in certain regimes or for certain classes of channels. To start, we propose an achievable rate region based on a simplified

Han-Kobayashi scheme (38) in which joint decoding of the intended and interfering messages is not required at receiver 2 (the oblivious receiver) and in which every node uses an i.i.d. codebook.

3.4.1 Inner Bound

Consider an achievability scheme where encoder 1 transmits using an i.i.d. codebook, while encoder 2, corresponding to the oblivious receiver, rate-splits as in the Han and Kobayashi achievability scheme for the classical IC (38). It may then be shown that the following rates are achievable,

Proposition 3.4.1. *The set of non-negative rate pairs* (R_1, R_2) *satisfying*

$$R_1 \leq I(Y_1; X_1 | U_2, Q), \tag{3.5a}$$

$$R_2 \leq I(Y_2; X_2 | Q), \tag{3.5b}$$

$$R_1 + R_2 \leq I(Y_1; X_1, U_2 | Q) + I(Y_2; X_2 | U_2, Q), \tag{3.5c}$$

is achievable for every input distribution that factorizes as

$$P_{Q,X_1,X_2,U_2} = P_Q P_{X_1|Q} P_{X_2|Q} P_{U_2|X_2Q}, \tag{3.5d}$$

and where $|Q| \leq 8$ *from (41). We denote the region in* (Eq. 3.5) *as* $\mathcal{R}_{\text{in}}$*, which is achievable for any memoryless IC-OR.*

Proof of Proposition 3.4.1. The proof follows by setting the auxiliary r.v. U_1 in the Han and Kobayashi rate region in (37, Section 6.5) to $U_1 = \emptyset$. Note, that this modified version of the Han and Kobayashi scheme employs joint decoding (of desired and undesired messages) only at receiver 1 (the non-oblivious receiver) and hence knowledge of the codebook of transmitter 1 is not needed at receiver 2 (the oblivious receiver). □

Remark 7. *By comparing the outer bound region $\mathcal{R}_{\rm out}$ in Theorem 3.3.2 to the inner bound region $\mathcal{R}_{\rm in}$ in Proposition 3.4.1 we notice the following differences: 1) in* (Eq. 3.4d) *the side information random variable U_2 is distributed as T_2 conditioned on X_2, while in* (Eq. 3.5d) *the auxiliary random variable U_2 can have any distribution conditioned on X_2; 2) the mutual information terms involving Y_1 have X_2 in the outer bound, but U_2 in the inner bound; and 3) the mutual information terms involving Y_2 are the same in both regions.*

3.4.2 Capacity in Very Strong Interference at the Non-oblivious Receiver for the General Memoryless IC-OR

In this section we show that under special channel conditions, akin to the very strong interference regime for the classical IC, the outer bound region in Theorem 3.3.2 is tight.

A general memoryless IC-OR for which

$$I(X_2;Y_2|X_1) \le I(X_2;Y_1), \quad \forall P_{X_1,X_2} = P_{X_1}P_{X_2}, \tag{3.6}$$

is said to have *very strong interference at the non-oblivious receiver* (receiver 1). Intuitively, when the condition in (Eq. 3.6) holds, the non-oblivious receiver should be able to first decode the interfering signal by treating its own signal as noise and then decode its own intended signal free of interference. This should “de-activate” the sum-rate bound in (Eq. 3.4c). Next we formalize this intuition.

Theorem 3.4.2. *When the condition in* (Eq. 3.6) *holds the capacity region of the IC-OR is given by*

$$R_1 \leq I(X_1; Y_1 | X_2, Q), \tag{3.7a}$$

$$R_2 \leq I(X_2; Y_2 | Q), \tag{3.7b}$$

taken over the union of all input distributions that factor as $P_{Q,X_1,X_2} = P_Q P_{X_1|Q} P_{X_2|Q}$ *and where* $|Q| \leq 2$*.*

Proof of Theorem 3.4.2. By dropping the sum-rate outer bound in (Eq. 3.4c) we see that the region in (Eq. 3.7) is an outer bound for a general memoryless IC-OR. By setting $U_2 = X_2$ in the achievable region in (Eq. 3.5), the region

$$R_1 \leq I(X_1; Y_1 | X_2, Q), \tag{3.8a}$$

$$R_2 \leq I(X_2; Y_2 | Q), \tag{3.8b}$$

$$R_1 + R_2 \leq I(X_1, X_2; Y_1 | Q), \tag{3.8c}$$

taken over the union of all $P_{Q,X_1,X_2} = P_Q P_{X_1|Q} P_{X_2|Q}$, is achievable. We see that the single rate bounds in (Eq. 3.8) match the upper bounds in (Eq. 3.7). We next intend to show that when the condition in (Eq. 3.6) holds, the sum-rate bound in (Eq. 3.8c) is redundant. By summing (Eq. 3.8a) and (Eq. 3.8b)

$$
\begin{aligned}
R_1 + R_2 &\leq I(X_1; Y_1|X_2, Q) + I(X_2; Y_2|Q) \\
&\overset{a)}{\leq} I(X_1; Y_1|X_2, Q) + I(X_2; Y_2, X_1|Q) \\
&\overset{b)}{=} I(X_1; Y_1|X_2, Q) + I(X_2; Y_2|X_1, Q) \\
&\overset{c)}{\leq} I(X_1; Y_1|X_2, Q) + I(X_2; Y_1|Q) \\
&= I(X_1, X_2; Y_1|Q) = \text{eq.(Eq. 3.8c)},
\end{aligned}
$$

where in a) we loosened the achievable sum-rate by adding X_1 as "side information" to receiver 2; in b) we used the independence of the inputs; and in c) the condition in (Eq. 3.6). Therefore, the sum-rate bound in (Eq. 3.8c) can be dropped without affecting the achievable rate region. This shows that the outer bound in (Eq. 3.7) is achievable thereby proving the claimed capacity result. □

Remark 8. *For the classical IC, the very strong interference regime is defined as*

$$
\begin{aligned}
I(X_1; Y_1|X_2) &\leq I(X_1; Y_2), \\
I(X_2; Y_2|X_1) &\leq I(X_2; Y_1),
\end{aligned}
$$

for all product input distributions; under these pair of conditions capacity can be shown. For the IC-OR, the very strong interference constraint at receiver 2 (oblivious receiver) is not needed in order to *show capacity. Therefore, the very strong interference condition for the IC-OR is less stringent than that for the classical IC. We believe this is so because the oblivious receiver (receiver 2) cannot decode the message of user 1 as per the modeling assumption. Indeed, we feel that the "lack of codebook knowledge" as originally proposed in (19) actually models the inability of a receiver to jointly decode its message along with unintended ones, as the mapping between the messages and codewords is not known.*

3.4.3 Capacity to within a Constant Gap for the ISD-IC-OR

We now show that $\mathcal{R}_{in}$ in Proposition 3.4.1 lies to within a gap of the outer bound $\mathcal{R}_{out}$ in Theorem 3.3.2 for the general ISD-IC-OR. We have

Theorem 3.4.3. *For the ISD-IC-OR, if* $(R_1, R_2) \in \mathcal{R}_{\text{out}}$ *then* $([R_1 - I(X_2; T_2|U_2, Q)]^+, R_2) \in \mathcal{R}_{\text{in}}$.

Proof of Theorem 3.4.3. The proof is as in (39). First, we define a new outer bound region $\bar{\mathcal{R}}_{out}$ by replacing X_2 with U_2 in all *positive* entropy terms of region $\mathcal{R}_{\text{out}}$, which is permitted as $H(Y_2|X_2) \leq H(Y_2|U_2)$ by the data processing inequality. We conclude that $\mathcal{R}_{\text{out}} \subseteq \bar{\mathcal{R}}_{\text{out}}$. We next compare $\bar{\mathcal{R}}_{\text{out}}$ and $\mathcal{R}_{\text{in}}$ term by term (we only need to compare the mutual informations invoking Y_1 as those involving Y_2 are the same in both bounds, see Remark 7, thus implying a zero gap for rate R_2): the difference is that $\bar{\mathcal{R}}_{out}$ has $-H(Y_1|X_1, X_2)$ where $\mathcal{R}_{\text{in}}$ has $-H(T_2|U_2, Q)$; thus the gap is

$$-H(Y_1|X_1, X_2) + H(T_2|U_2, Q) = -H(T_2|X_2) + H(T_2|U_2, Q) = I(X_2; T_2|U_2, Q).$$

This concludes the proof. $\square$

Remark 9. *Note that*

$$I(X_2;T_2|U_2,Q) = H(T_2|U_2,Q) - H(T_2|X_2) \leq H(T_2) - H(T_2|X_2) \leq \max_{p_{X_2}} I(T_2;X_2),$$

so the gap is finite/constant for all channel $P_{T_2|X_2}$ with finite capacity.

We next give an example of constant gap characterization in Section 3.4.5 after having discussed in Section 3.4.4 a special class of ISD-IC-OR for which the gap to capacity is zero.

3.4.4 Exact Capacity for the Injective Fully Deterministic IC-OR

We now specialize Theorem 3.4.3 to the class of injective *fully deterministic* ICs (42). For this class of channels the mappings T_1 and T_2 in (Eq. 3.2) are deterministic functions of X_1 and X_2, respectively. We have

Corollary 3.4.4. *For the injective fully deterministic IC-OR the outer bound in Theorem 3.3.2 is tight.*

Proof of Corollary 3.4.4. The injective fully deterministic IC-OR has $T_2 = U_2$ and therefore $I(X_2;T_2|U_2,Q) = 0$ in Theorem 3.4.3. □

As an application of Corollary 3.4.4 we consider next the Linear Deterministic Approximation (LDA) of the Gaussian IC-OR at high SNR, whose classical counterpart (where all codebooks are known) was first proposed in (22). The LDA-IC-OR has input/output relationship

$$Y_1 = \mathbf{S}^{q-n_{11}} X_1 + \mathbf{S}^{q-n_{12}} X_2, \quad T_2 = \mathbf{S}^{q-n_{12}} X_2, \tag{3.9a}$$

$$Y_2 = \mathbf{S}^{q-n_{21}} X_1 + \mathbf{S}^{q-n_{22}} X_2, \quad T_1 = \mathbf{S}^{q-n_{21}} X_1, \tag{3.9b}$$

where inputs and outputs are binary-valued vectors of length q, $\mathbf{S}$ is the $q \times q$ shift matrix (22), $(n_{11}, n_{12}, n_{21}, n_{22})$ are non-negative integers and $q := \max\{n_{11}, n_{12}, n_{21}, n_{22}\}$. Summations and multiplications are bit-wise over the binary field.

For simplicity, we next evaluate the *symmetric* sum-capacity of the LDA-IC-OR. The symmetric LDA-IC-OR has parameters $n_{11} = n_{22} = n_{\mathrm{S}}$ and $n_{12} = n_{21} = n_{\mathrm{I}} := n_{\mathrm{S}}\,\alpha$ for some non-negative α. The maximum symmetric rate, or symmetric sum-capacity normalized by the sum-capacity of an interference-free channel, is defined as

$$d(\alpha) := \frac{\max\{R_1 + R_2\}}{2\,n_{\mathrm{S}}}, \tag{3.10}$$

where the maximization is over all achievable rate pairs (R_1, R_2) satisfying Theorem 3.3.2, which is the capacity region by Corollary 3.4.4. Since we may provide the oblivious receiver in the LDA-IC-OR with the additional codebook index so as to obtain the classical LDA-IC with full codebook knowledge, we immediately have

$$d(\alpha) \le d^{(W)}(\alpha) = \min\left(1, \max\left(\frac{\alpha}{2}, 1 - \frac{\alpha}{2}\right), \max\left(\alpha, 1 - \alpha\right)\right), \tag{3.11}$$

where $d^{(W)}(\alpha)$, the so-called W-curve (26), is the maximum symmetric rate of the classical LDA-IC. In (23) it was shown that i.i.d. $\mathsf{Bernoulli}(1/2)$ input bits in the Han and Kobayashi region yield $d^{(W)}(\alpha)$.

Although Theorem 3.3.2 gives the exact capacity region of the LDA-IC-OR, it is not immediately clear which input distribution achieves the maximum symmetric rate. Instead of analytically deriving the sum-capacity, we proceeded to numerically evaluate Theorem 3.3.2 for $|Q| = 1$, which is not necessarily optimal. We observe the surprising result that even with $|Q| = 1$ i.e., without time sharing, some of the points on the normalized sum-capacity of the LDA-IC-OR are equal to $d^{(W)}(\alpha)$, see Fig. Figure 5 and Table I. Although we lack a formal proof that we can achieve the whole W-curve with a non i.i.d. $\mathsf{Bernoulli}(1/2)$ input we do, however, conjecture that it is indeed possible with the scheme in Proposition 3.4.1. If true, this implies that partial codebook knowledge at one receiver does not impact the sum-rate of the symmetric LDA-IC-OR at these points. This is quite unexpected, especially in the strong interference regime ($\alpha \geq 1$) where the optimal strategy for the classical LDA-IC is to jointly decode the interfering message along with the intended message—a strategy that seems to be precluded by the lack of codebook knowledge at one receiver. This might suggest a more general principle: there is no loss of optimality in lack of codebook knowledge as long as the oblivious receiver can remove the interfering codeword, regardless of whether or not it can decode the message carried by this codeword.

Another interesting observation is that i.i.d. $\mathsf{Bernoulli}(1/2)$ input bits may no longer be optimal (though we do not show their strict sub-optimality). In Table I we report, for some values of α and $n_{\rm S}, n_{\rm I}$, the input distributions to be used in $\mathcal{R}_{\rm out}$ in Theorem 3.3.2. We notice that, at least when evaluating the region in Theorem 3.3.2 for $|Q| = 1$ only, that the region exhausting inputs are now *correlated*. For example, Table I shows that, for $\alpha = 4/3$ the inputs X_1 and X_2 are binary vectors of length $\log(16) = 4$ bits; out of the 16 different possible bit sequences, only 4 are actually used at each transmitter with strictly positive probability to achieve $d^{(W)}(4/3) = 4/6$. By using i.i.d. $\mathsf{Bernoulli}(1/2)$ input bits in Theorem 3.3.2 for $|Q| = 1$ we would obtain a normalized sum-rate of $1/2 = 3/6$, the same as achieved by time division (23).

TABLE I: LDA-IC-OR: EXAMPLES OF SUM-RATE OPTIMAL INPUT DISTRIBUTIONS FOR THE CAPACITY REGION IN THEOREM **??**.

α, $(n_{\mathrm{S}}, n_{\mathrm{I}})$	Probability mass function with $\lvert Q \rvert = 1$
$\frac{1}{2}$, $(2, 1)$	$P_{X_1} = [0.5, 0, 0.5, 0]$ $P_{X_2} = [0, 0.5, 0, 0.5]$
$\frac{2}{3}$, $(3, 2)$	$P_{X_1} = [0, 0, 0.25, 0.25, 0, 0, 0.25, 0.25]$ $P_{X_2} = [0, 0, 0.25, 0.25, 0, 0, 0.25, 0.25]$
1, $(2, 2)$	$P_{X_1} = [0, 0, 0.5, 0.5]$ $P_{X_2} = [0, 0.5, 0, 0.5]$
$\frac{4}{3}$, (3,4)	$P_{X_1} = [0, 0, 0, 0, 0, 0.25, 0, 0.25, 0, 0, 0, 0, 0, 0.25, 0, 0.25]$ $P_{X_2} = [0, 0, 0, 0.25, 0, 0.25, 0, 0, 0, 0, 0, 0, 0, 0.25, 0, 0.25]$
2, $(2, 1)$	$P_{X_1} = [0, 0.5, 0, 0.5]$ $P_{X_2} = [0, 0.5, 0, 0.5]$

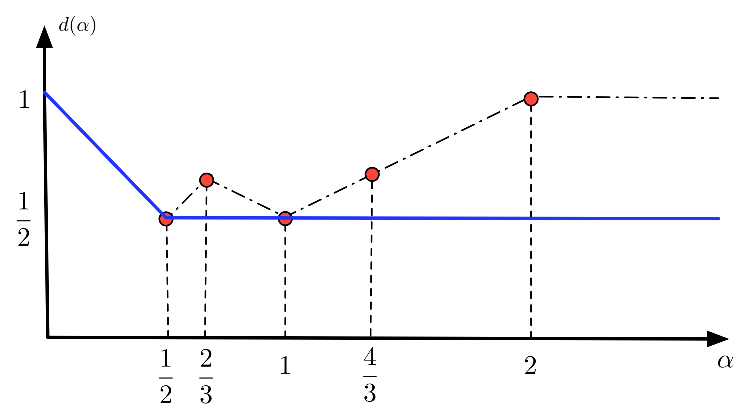


Figure 5: The normalized sum-capacity, or maximum symmetric rate, for the classical LDA-IC (dash-dotted black line). Normalized sum-rates achieved by the input distributions in Table I (red dots) for the LDA-IC-OR. The normalized sum-rate achieved by i.i.d. Bernoulli$(1/2)$ inputs and $|Q| = 1$ (solid blue line) in the capacity region in Theorem 3.3.2 for the LDA-IC-OR.

Also, i.i.d. $\mathsf{Bernoulli}(1/2)$ inputs in the LDA model usually are translated to i.i.d. Gaussian inputs in the Gaussian noise model. This intuition is reinforced, in the next section, by showing that i.i.d. Gaussian are also suboptimal for the Gaussian noise model for $|Q| = 1$. Also, the fact that there exist other, non i.i.d $\mathsf{Bernoulli}(1/2)$, input distributions that are capacity achieving for the LDA stimulates search for non-Gaussian inputs that might be capacity achieving for a Gaussian noise channel. In fact the rest of the thesis tries to use intuition gained in this section to construct non-Gaussian inputs that will be capacity or constant gap capacity approaching.

3.4.5 The Gaussian Noise IC-OR

We now consider the practically relevant real-valued single-antenna power-constrained Gaussian noise channel, whose input/output relationship is

$$Y_1 = h_{11}X_1 + h_{12}X_2 + Z_1 = h_{11}X_1 + T_2, \quad T_2 = h_{12}X_2 + Z_1, \tag{3.12a}$$

$$Y_2 = h_{21}X_1 + h_{22}X_2 + Z_2 = h_{22}X_2 + T_1, \quad T_1 = h_{21}X_1 + Z_2, \tag{3.12b}$$

where h_{ij} are the real-valued channel coefficients for $(i, j) \in [1 : 2]^2$ assumed constant and known to all nodes, the input $X_i \in \mathbb{R}$ is subject to per block power constraints $\frac{1}{n}\sum_{i=1}^{n} X_i^2 \leq 1,\ i \in [1 : 2]$, and the noise $Z_i,\ i \in [1 : 2]$, is a unit-variance zero-mean Gaussian r.v.

By specializing the result of Theorem 3.4.3 to the G-IC-OR we may show the following:

Corollary 3.4.5. *For the G-IC-OR the gap is at most* $1/2$ *bit per channel use.*

Proof of Corollary 3.4.5. For the G-IC-OR $T_2 = h_{12}X_2 + Z_1$, and thus we set U_2 in Theorem 3.3.2 to $U_2 = h_{12}X_2 + Z_1^*$, where $Z_1 \sim Z_1^*$ and mutually independent. We thus have

$$\begin{aligned} I(X_2; T_2|U_2, Q) &= h(T_2|U_2, Q) - h(Z_2) \\ &\leq h(T_2 - U_2) - h(Z_1) \\ &= h(Z_1 - Z_1^*) - h(Z_1) = \frac{1}{2}\log(2), \end{aligned}$$

as claimed. □

In the classical G-IC with full codebook knowledge, Gaussian inputs exhaust known outer bounds, which are achievable to within $1/2$ bit per channel use (26). From the rate expression in Theorem 3.3.2 it is not clear whether Gaussian inputs are optimal for $\mathcal{R}_{out}$. The following discussion shows that in general the answer is in the negative. For simplicity we focus on the achievable generalized Degrees of Freedom (gDoF) for the symmetric G-IC-OR. The symmetric G-IC-OR has $|h_{11}|^2 = |h_{22}|^2 = \mathsf{snr}$ and $|h_{12}|^2 = |h_{21}|^2 = \mathsf{inr}$, with $\mathsf{inr} = \mathsf{snr}^\alpha$ for some non-negative α. The sum-gDoF is defined as

$$d(\alpha) := \lim_{\mathsf{snr}\to+\infty} \frac{\max\{R_1 + R_2\}}{2 \cdot \frac{1}{2}\log(1 + \mathsf{snr})}, \tag{3.13}$$

where the maximization is over all possible achievable rate pairs. By using the classical G-IC as a trivial upper bound, we have $d(\alpha) \leq d^{(W)}(\alpha)$ where $d^{(W)}(\alpha)$ is given in (Eq. 3.11).

By evaluating Theorem 3.3.2 for independent Gaussian inputs and $|Q| = 1$ (which we do *not* claim to be optimal, but which gives us an achievable rate up to $1/2$ bit) we obtain

$$(R_1 + R_2)^{(GG)} = \min \left\{ \mathsf{I_g}\left(\mathsf{snr}\right) + \mathsf{I_g}\left(\frac{\mathsf{snr}}{1+\mathsf{inr}}\right), \mathsf{I_g}\left(\frac{\mathsf{snr}}{\mathsf{inr}+1}\right) + \mathsf{I_g}\left(\mathsf{inr} + \frac{\mathsf{snr}}{1+\mathsf{inr}}\right) \right\},$$
$$\iff \quad d^{(GG)}(\alpha) = \frac{1}{2} + \left[\frac{1}{2} - \alpha\right]^+,$$

the superscript “GG” indicates that both transmitters use a Gaussian input. For future reference, with Time Division (TD) and Gaussian codebooks we can achieve

$$(R_1 + R_2)^{(TD)} = \frac{1}{2}\log\left(1 + 2\,\mathsf{snr}\right) \quad \iff \quad d^{(TD)}(\alpha) = \frac{1}{2}.$$

We plot the achievable gDoF vs. α in Fig. Figure 5, together with the gDoF of the classical G-IC given by $d^{(W)}(\alpha)$ (26). We note that Gaussian inputs are indeed optimal for $0 \leq \alpha \leq 1/2$, i.e., $d^{(GG)}(\alpha) = d^{(W)}(\alpha)$, where interference is treated as noise even for the classical G-IC (which is also achievable by the G-IC-OR). For $\alpha > 1/2$ we have $d^{(GG)}(\alpha) = d^{(TD)}(\alpha)$, that is, Gaussian inputs achieve the same rates as time division. Interestingly, Gaussian inputs are sub-optimal in our achievable region in general as we show next.

Consider $\alpha = 4/3$. With Gaussian inputs we only achieve $d^{(GG)}(4/3) = d^{(TD)}(4/3) = 1/2$. Notice the similarity with the LDA-IC-OR: the input distribution that is optimal for the non-oblivious IC performs as time division for the G-IC-OR. Inspired by the LDA-IC-OR we explore now the possibility of using non-Gaussian inputs. By following (19, Section VI.A), which demonstrated that binary signaling outperforms Gaussian signaling for a fixed finite snr, we consider a uniform PAM constellation with

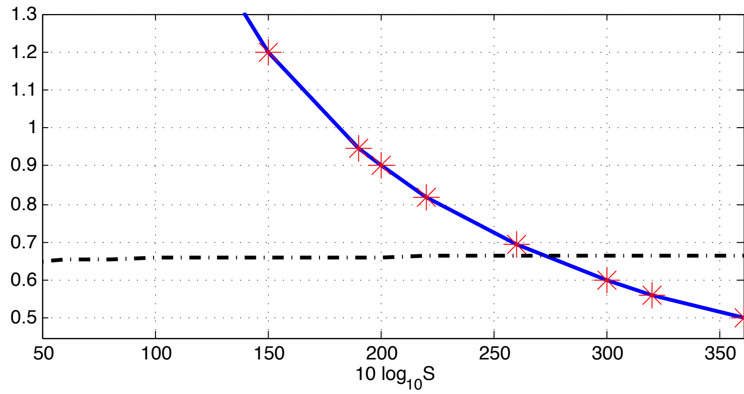


Figure 6: Achievable normalized sum-rate for the symmetric G-IC-OR with $\alpha = 4/3$ vs snr in dB. Legend: time division in solid blue line; Gaussian inputs at both transmitters in red stars; X_1 is a uniform PAM with $N = \lfloor \mathsf{snr}^{\frac{1}{6}} \rfloor$ points and X_2 is Gaussian in dash-dotted black line.

N points. Fig. Figure 6 shows the achievable normalized sum-rate $\frac{R_1+R_2}{2\cdot\frac{1}{2}\log(1+\mathsf{snr})}$ as a function of snr for the case where X_1 (the input of the non-oblivious pair) is a PAM constellation with $N = \lfloor \mathsf{snr}^{1/6} \rfloor$ points and X_2 (the input of the oblivious pair) is Gaussian; we refer to the achievable gDoF of this inputs as $d^{(DG)}(\alpha)$. Notice that the number of points in the discrete input is a function of snr. We also report the achievable normalized sum-rate with time division and Gaussian inputs. Fig. Figure 6 shows that, for sufficiently large snr, using a discrete input outperforms time division; moreover, for the range of simulated snr, it seems that the proposed discrete input achieves a gDoF of $d^{(DG)}(\alpha) = \alpha/2 = 4/6$ as for the classical G-IC with full codebook knowledge. In the sections that follow we analytically show that using discrete input (or mixed) at the non-oblivious transmitter indeed achieves the full gDoF and symmetric capacity region to within a constant gap.

3.5 Achievable Regions for the G-IC-OR

We now analyze the G-IC-OR by using Remark 3 (i.e., bounds on the mutual information achievable by a PAM input on a point-to-point power-constrained Gaussian noise channel) and the insight on the nature of the gap due to a PAM input from Remark **??**. We first present a scheme (an achievable rate region evaluated using a mixed input) that will prove to be useful in strong and very strong interference, and then present a more involved scheme that will be useful in the somewhat trickier weak and moderate interference regimes. Although the second scheme includes the first as a special case, we start with a simpler scheme to highlight the important steps of the derivation without getting caught up in excessive technical details.

3.5.1 Achievable Scheme I

We first derive an achievable rate region from Proposition 3.4.1 with inputs

$$\text{Scheme I:} \quad X_{1D} \sim \mathsf{PAM}\left(N\right), \ N \in \mathbb{N}, \ \text{independent of} \tag{3.14a}$$

$$X_{2G} \sim \mathcal{N}(0,1), \tag{3.14b}$$

$$X_1 = X_{1D}, \ X_2 = X_{2G}, \tag{3.14c}$$

$$U_2 = X_2, \ Q = \emptyset. \tag{3.14d}$$

which we will show in the next sections to be gDoF optimal and to within a constant gap of the symmetric capacity of the classical G-IC in the strong and very strong interference regimes. Such results may not be shown by using i.i.d. Gaussian inputs in the same achievable scheme in Proposition 3.4.1. The achievable region is derived for a general G-IC-OR and later on specialized to the symmetric case.

Theorem 3.5.1. *For the G-IC-OR the following rate region is achievable by the input in* (Eq. 3.14)

$$R_1 \le \mathsf{I_d}\left(N, |h_{11}|^2\right), \tag{3.15a}$$

$$R_2 \le \mathsf{I_d}\left(N, \frac{|h_{21}|^2}{1+|h_{22}|^2}\right) + \mathsf{I_g}\left(|h_{22}|^2\right) - \mathsf{I_g}\left(\min\left(N^2-1, |h_{21}|^2\right)\right), \tag{3.15b}$$

$$R_1 + R_2 \le \mathsf{I_d}\left(N, \frac{|h_{11}|^2}{1+|h_{12}|^2}\right) + \mathsf{I_g}\left(|h_{12}|^2\right). \tag{3.15c}$$

Proof of Theorem 3.5.1. We proceed to evaluate the rate region in Proposition 3.4.1 with the inputs in (Eq. 3.14), that is, the achievable region in (Eq. 3.8) with $|Q| = 1$.

The rate of the user 1 is bounded by $R_1 \le I(X_1; Y_1|X_2) = I(X_{1D}; h_{11}X_{1D} + Z_1)$, where $I(X_{1D}; h_{11}X_{1D} + Z_1)$ can be further lower bounded by using (Eq. 2.10) from Remark 3 with $\mathsf{snr} = |h_{11}|^2$; by doing so we obtain the bound in (Eq. 3.15a).

The rate of the user 2 is bounded by

$$
\begin{aligned}
R_2 &\leq I(X_2; Y_2) \\
&= h(h_{21}X_{1D} + \underbrace{h_{22}X_{2G} + Z_2}_{\sim\mathcal{N}(0,1+|h_{22}|^2)}) - h(h_{21}X_{1D} + Z_2) \\
&= \underbrace{\left(h\left(\frac{h_{21}}{\sqrt{1+|h_{22}|^2}}X_{1D} + Z_2\right) - h(Z_2)\right)}_{\geq \mathsf{I}_\mathsf{d}\left(N, \frac{|h_{21}|^2}{1+|h_{22}|^2}\right) \text{ from (Eq. 2.10)}} + \frac{1}{2}\log(1+|h_{22}|^2) \\
&- \underbrace{\left(h\left(h_{21}X_{1D} + Z_2\right) - h(Z_2)\right)}_{\leq \mathsf{I}_\mathsf{g}(\min(N^2-1, |h_{21}|^2)) \text{ from (Eq. 2.11)}},
\end{aligned}
$$

from which we conclude that the achievable rate for user 2 is lower bounded as in (Eq. 3.15b).

The sum-rate is bounded by $R_1 + R_2 \leq I(X_1, X_2; Y_1) = I(X_1; Y_1) + I(X_2; Y_1|X_1)$, where $I(X_1; Y_1)$ can be lower bounded by means of Remark 3 with $\mathsf{snr} = \frac{|h_{11}|^2}{1+|h_{12}|^2}$ and where $I(X_2; Y_1|X_1) = I(X_{2G}; h_{12}X_{2G} + Z_1) = \mathsf{I}_\mathsf{g}(|h_{12}|^2)$; by combining the two terms we obtain the bound in (Eq. 3.15c).

□

3.5.2 Achievable Scheme II

The input in (Eq. 3.14) might not be optimal in general and may be generalized as follows. Consider the rate region in Proposition 3.4.1 with inputs

$$\text{Scheme II:} \quad X_{1D}, X_{1G}, X_{2Gc}, X_{2Gp} \text{ independent and distributed as} \tag{3.16a}$$

$$X_{1D} \sim \mathsf{PAM}(N), \ N \in \mathbb{N}, \tag{3.16b}$$

$$\text{all the others are } \mathcal{N}(0,1), \tag{3.16c}$$

$$X_1 = \sqrt{1-\delta_1} X_{1D} + \sqrt{\delta_1} X_{1G}, \ \delta_1 \in [0,1], \tag{3.16d}$$

$$X_2 = \sqrt{1-\delta_2} X_{2Gc} + \sqrt{\delta_2} X_{2Gp}, \ \delta_2 \in [0,1]. \tag{3.16e}$$

$$U_2 = X_{2Gc}, \ Q = \emptyset. \tag{3.16f}$$

In Scheme II, X_{2Gc} encodes a “common” message, and X_{2Gp} and X_{1G} encode the “private” messages as in the classical Han-Kobayashi scheme (38). We shall also interpret X_{1D} as encoding a “common” message even if X_{1D} cannot be decoded at receiver 2 (the oblivious receiver) as receiver 2 lacks knowledge of the codebook(s) used by transmitter 1. The main message of this Chapter is in fact that, even with lack of codebook knowledge, if there would-be-common message is from a discrete alphabet then its effect on the rate region—up to a constant gap—is as if the message could indeed be jointly decoded. We believe this is because lack of codebook knowledge may be translated as lack of knowledge of the mapping of the codewords to the messages, but does not preclude a receiver's ability to perform *symbol-by-symbol estimation* of the symbols in the interfering codeword (rather than decoding the messages carried by the codeword). Correctly estimating and subtracting off the interfering symbols is as effective as decoding the actual interfering codeword, as the message carried by the codeword is not desired anyhow. A similar intuition was pointed out in (19) where the authors write "We indeed see that BPSK signaling outperforms Gaussian signaling. This is because demodulation is some form of primitive decoding, which is not possible for the Gaussian signaling."

In the next sections we will show that Proposition 3.4.1 with the inputs in (Eq. 3.16) is gDoF optimal and is to within a constant gap of a capacity outer bound for the classical G-IC in the weak and moderate interference regimes. Also note that with $\delta_1 = \delta_2 = 0$ Scheme II in (Eq. 3.16) reduces to Scheme I in (Eq. 3.14).

The achievable region is derived for a general G-IC-OR and later on specialized to the symmetric case. The rate region achievable by Scheme II is:

Theorem 3.5.2. *For the G-IC-OR the following rate region is achievable with inputs as in* (Eq. 3.16)

$$
\begin{aligned}
R_1 &\leq \mathsf{I_d}\left(N, \frac{|h_{11}|^2(1-\delta_1)}{1+|h_{11}|^2\delta_1+|h_{12}|^2\delta_2}\right) + \mathsf{I_g}\left(\frac{|h_{11}|^2\delta_1}{1+|h_{12}|^2\delta_2}\right), && \text{(3.17a)}\\
R_2 &\leq \mathsf{I_d}\left(N, \frac{|h_{21}|^2(1-\delta_1)}{1+|h_{21}|^2\delta_1+|h_{22}|^2}\right) + \mathsf{I_g}\left(\frac{|h_{22}|^2}{1+|h_{21}|^2\delta_1}\right)\\
&\quad - \mathsf{I_g}\left(\min\left(N^2-1, \frac{|h_{21}|^2(1-\delta_1)}{1+|h_{21}|^2\delta_1}\right)\right), && \text{(3.17b)}\\
R_1+R_2 &\leq \mathsf{I_d}\left(N, \frac{|h_{11}|^2(1-\delta_1)}{1+|h_{11}|^2\delta_1+|h_{12}|^2}\right) + \mathsf{I_g}\left(|h_{11}|^2\delta_1+|h_{12}|^2\right) - \mathsf{I_g}\left(|h_{12}|^2\delta_2\right)\\
&\quad + \mathsf{I_d}\left(N, \frac{|h_{21}|^2(1-\delta_1)}{1+|h_{21}|^2\delta_1+|h_{22}|^2\delta_2}\right) + \mathsf{I_g}\left(\frac{|h_{22}|^2\delta_2}{1+|h_{21}|^2\delta_1}\right)\\
&\quad - \mathsf{I_g}\left(\min\left(N^2-1, \frac{|h_{21}|^2(1-\delta_1)}{1+|h_{21}|^2\delta_1}\right)\right). && \text{(3.17c)}
\end{aligned}
$$

Proof of Theorem 3.5.2. The proof can be found in Appendix C and follows similarly to the proof of Theorem 3.5.1. □

3.6 High SNR Performance

We now analyze the performance of the schemes in Theorems 3.5.1 and 3.5.2 for the symmetric G-IC-OR at high-SNR by using the gDoF region as performance metric. The notion of gDoF has been introduced in (26) and has become an important metric that sheds lights on the behavior of the capacity when exact capacity results are not available. The gDoF region is formally defined as follows. For an achievable pair (R_1, R_2), let

$$\mathcal{D}(\alpha) := \left\{ (d_1, d_2) \in \mathbb{R}_+^2 : d_i := \lim_{\substack{\mathsf{inr} = \mathsf{snr}^\alpha, \\ \mathsf{snr} \to \infty}} \frac{R_i}{\frac{1}{2}\log(1+\mathsf{snr})}, i \in [1:2],\ (R_1, R_2) \text{ is achievable} \right\}. \tag{3.18}$$

Let $\mathcal{D}^{\text{G-IC}}(\alpha)$ and $\mathcal{D}^{\text{G-IC-OR}}(\alpha)$ be the gDoF region of the classical G-IC and of the G-IC-OR, respectively.

We first present two different achievable gDoF regions based on Theorems 3.5.1 and 3.5.2, which we will compare to $\mathcal{D}^{\text{G-IC}}(\alpha)$ given by (26)

$$\mathcal{D}^{\text{G-IC}}(\alpha)\ :\ d_1 \le 1, \tag{3.19a}$$

$$d_2 \le 1, \tag{3.19b}$$

$$d_1 + d_2 \le \max(\alpha, 2-\alpha), \tag{3.19c}$$

$$d_1 + d_2 \le \max(2\alpha, 2-2\alpha), \tag{3.19d}$$

$$2d_1 + d_2 \le 2,\ \text{only for } \alpha \in [1/2, 1], \tag{3.19e}$$

$$d_1 + 2d_2 \le 2,\ \text{only for } \alpha \in [1/2, 1]. \tag{3.19f}$$

Corollary 3.6.1 (gDoF region from achievable Scheme I)**.** *Let* $N = \mathsf{N_d}(\mathsf{snr}^\beta)$ *and*

$$\mathcal{D}^{\text{I}}(\alpha,\beta)\ :\ d_1 \le \min(\beta, 1), \tag{3.20a}$$

$$d_2 \le \min(\beta, [\alpha-1]^+) + 1 - \min(\beta, \alpha), \tag{3.20b}$$

$$d_1 + d_2 \le \min(\beta, [1-\alpha]^+) + \alpha. \tag{3.20c}$$

for any $\beta \ge 0$*. By Theorem 3.5.1, the gDoF region* $\mathcal{D}^{\text{I}}(\alpha,\beta)$ *is achievable.*

Proof of Corollary 3.6.1. We prove the bound in (Eq. 3.20b) only as the other bounds follow similarly. With $\mathsf{inr} = \mathsf{snr}^{\alpha}$ and $N = \mathsf{N_d}(\mathsf{snr}^{\beta})$ we have

$$\lim_{\mathsf{snr}\to\infty} \frac{\log(N^2)}{\log(1+\mathsf{snr})} = \beta,$$
$$\lim_{\mathsf{snr}\to\infty} \frac{\log(1+\mathsf{inr})}{\log(1+\mathsf{snr})} = \alpha.$$

Therefore d_2 can be bounded as

$$\begin{aligned} d_2 &= \lim_{\mathsf{snr}\to\infty} \frac{\text{left hand side of eq.(Eq. 3.15b)}}{\frac{1}{2}\log(1+\mathsf{snr})} \\ &= \min(\beta, [\alpha-1]^+) + 1 - \min(\beta,\alpha), \end{aligned}$$

thus proving (Eq. 3.20b). $\square$

Next, by using Theorem 3.5.2 with the power split as in (26) we show yet another achievable gDoF region.

Corollary 3.6.2 (gDoF region from achievable Scheme II)**.** *Let* $N = \mathsf{N_d}(\mathsf{snr}^{\beta})$ *and*

$$\mathcal{D}^{\mathrm{II}}(\alpha,\beta) \;:\; d_1 \le \min(\beta, 1+\alpha-\max(1,\alpha)) + [1-\alpha]^+, \tag{3.21a}$$

$$d_2 \le \min(\beta, [\alpha-1]^+) + 1 - \min(\beta,\alpha), \tag{3.21b}$$

$$\begin{aligned} d_1 + d_2 \le{}& \min(\beta, [1+\alpha-\max(1,2\alpha)]^+) + \max(\alpha, 1-\alpha) + \\ &+ \min(\beta, [2\alpha-\max(1,\alpha)]^+) + [1-\alpha]^+ - \min(\beta,\alpha). \end{aligned} \tag{3.21c}$$

for any $\beta \geq 0$. By Theorem 3.5.2, the gDoF region $\mathcal{D}^{II}(\alpha, \beta)$ is achievable.

Proof of Corollary 3.6.2. Let $\mathsf{inr} = \mathsf{snr}^{\alpha}$, $N = \mathsf{N_d}(\mathsf{snr}^{\beta})$, and $\delta_1 = \delta_2 = \frac{1}{1+\mathsf{inr}}$ in Theorem 3.5.2 (see the region in (Eq. C.2) in Appendix C) and take limits similarly to the proof of Corollary 3.6.1. $\square$

We are now ready to prove the main result of this section:

Theorem 3.6.3. *For the G-IC-OR there is no loss in gDoF compared to the classical G-IC, i.e.,*

$$\mathcal{D}^{\textit{G-IC}}(\alpha) = \mathcal{D}^{\textit{G-IC-OR}}(\alpha).$$

Proof of Theorem 3.6.3. We consider several regimes:

3.6.1 Very Strong Interference Regime $\alpha \geq 2$

In this regime the gDoF region outer bound $\mathcal{D}^{\text{G-IC}}(\alpha)$ is characterized by (Eq. 3.19a) and (Eq. 3.19b). For achievability we consider Corollary 3.6.1 with $\beta = 1$, that is,

$$\begin{aligned}
\mathcal{D}^{\mathrm{I}}(\alpha, 1)\ :\ & d_1 \leq \min(1, 1) = 1, \\
& d_2 \leq \min(1, [\alpha - 1]^+) + 1 - \min(1, \alpha) = 1, \\
& d_1 + d_2 \leq \min(1, [1 - \alpha]^+) + \alpha = \alpha \text{ (redundant because } \alpha \geq 2).
\end{aligned}$$

Since the sum-gDoF is redundant, we get that

$$\mathcal{D}^{\mathrm{I}}(\alpha, \beta = 1) = \{d_i \in [0, 1],\ i \in [1:2]\} = \mathcal{D}^{\text{G-IC-OR}}(\alpha) = \mathcal{D}^{\text{G-IC}}(\alpha).$$

Fig. 7a illustrates the region $\mathcal{D}^{\mathrm{I}}(\alpha, \beta = 1)$.

3.6.2 Strong Interference Regime $1 \leq \alpha < 2$

In this regime the gDoF region outer bound $\mathcal{D}^{\text{G-IC}}(\alpha)$ is characterized by (Eq. 3.19a)-(Eq. 3.19c) and has two dominant corner points: $(d_1, d_2) = (1, \alpha - 1)$ and $(d_1, d_2) = (\alpha - 1, 1)$. For achievability we consider the following achievable gDoF regions

$$\begin{aligned} \mathcal{D}^{\mathrm{I}}(\alpha, 1) \ : \ d_1 &\leq 1, \\ d_2 &\leq \alpha - 1, \\ d_1 + d_2 &\leq \alpha \text{ (redundant)}. \end{aligned}$$

and

$$\begin{aligned} \mathcal{D}^{\mathrm{I}}(\alpha, \alpha - 1) \ : \ d_1 &\leq \alpha - 1, \\ d_2 &\leq 1, \\ d_1 + d_2 &\leq \alpha \text{ (redundant)}, \end{aligned}$$

Fig. 7b illustrates that

$$\mathsf{co}\left(\mathcal{D}^{\mathrm{I}}(\alpha, 1) \cup \mathcal{D}^{\mathrm{I}}(\alpha, \alpha - 1)\right) = \mathcal{D}^{\text{G-IC}}(\alpha) = \mathcal{D}^{\text{G-IC-OR}}(\alpha).$$

3.6.3 Moderately Weak Interference Regime $\frac{1}{2} < \alpha < 1$

In this regime the gDoF region outer bound $\mathcal{D}^{\text{G-IC}}(\alpha)$ is characterized by all the constraints in (Eq. 3.19) and has four corner points: $(d_1, d_2) = (1, 0)$, $(d_1, d_2) = (0, 1)$, and $(d_1, d_2) = (\min(4\alpha - 2, \alpha), 2 - 2\alpha)$ and $(d_1, d_2) = (2 - 2\alpha, \min(4\alpha - 2, \alpha))$. The gDoF pair $(d_1, d_2) = (1, 0)$ is trivially achievable by silencing user 2, and similarly $(d_1, d_2) = (0, 1)$ by silencing user 1. For achievability of the remaining two corner points, we consider the following achievable gDoF regions

$$
\begin{aligned}
\mathcal{D}^{\text{II}}(\alpha, 2\alpha - 1) \ : \ d_1 &\le \min(2\alpha - 1, 1 + \alpha - 1) + 1 - \alpha = \alpha, \\
d_2 &\le \min(2\alpha - 1, 0) + 1 - \min(2\alpha - 1, \alpha) = 2 - 2\alpha, \\
d_1 + d_2 &\le \min(2\alpha - 1, [1 + \alpha - \max(1, 2\alpha)]^+) + \max(\alpha, 1 - \alpha) + \\
&\quad + \min(2\alpha - 1, [2\alpha - 1]^+) + 1 - \alpha - \min(2\alpha - 1, \alpha) \\
&= \min(2\alpha, 2 - \alpha), \quad \text{(redundant for } \alpha \in [2/3, 1]).
\end{aligned}
$$

and

$$
\begin{aligned}
\mathcal{D}^{\text{II}}(\alpha, 1 - \alpha) \ : \ d_1 &\le \min(1 - \alpha, 1 + \alpha - 1) + 1 - \alpha = 2 - 2\alpha, \\
d_2 &\le \min(1 - \alpha, 0) + 1 - \min(1 - \alpha, \alpha) = \alpha, \\
d_1 + d_2 &\le \min(1 - \alpha, [1 + \alpha - \max(1, 2\alpha)]^+) + \max(\alpha, 1 - \alpha) + \\
&\quad + \min(1 - \alpha, [2\alpha - 1]^+) + 1 - \alpha - \min(1 - \alpha, \alpha) \\
&= \min(2\alpha, 2 - \alpha), \quad \text{(redundant for } \alpha \in [2/3, 1]).
\end{aligned}
$$

Fig. 7c (for $\alpha \in [2/3, 1]$) and Fig. 7d (for $\alpha \in [1/2, 2/3]$) illustrate that

$$\mathsf{co}\left(\left\{(d_1, d_2) = (1, 0)\right\} \cup \left\{(d_1, d_2) = (0, 1)\right\} \cup \mathcal{D}^{\mathrm{II}}(\alpha, 2\alpha - 1) \cup \mathcal{D}^{\mathrm{II}}(\alpha, 1 - \alpha)\right) = \mathcal{D}^{\text{G-IC}}(\alpha)$$
$$= \mathcal{D}^{\text{G-IC-OR}}(\alpha).$$

3.6.4 Noisy Interference $0 \leq \alpha \leq \frac{1}{2}$

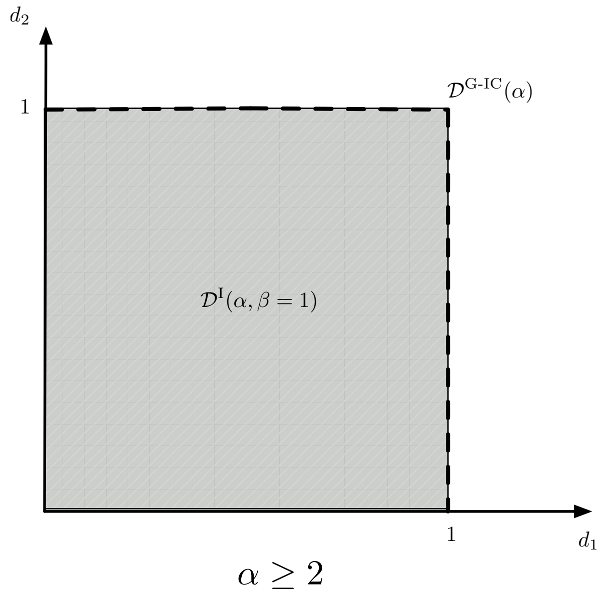
In this regime one may achieve the whole optimal G-IC gDoF region by using Gaussian inputs, treating interference as noise, and power control. Since this strategy is feasible for the G-IC-OR, the G-IC gDoF region is achievable also for the G-IC-OR.

This concludes our proof. □

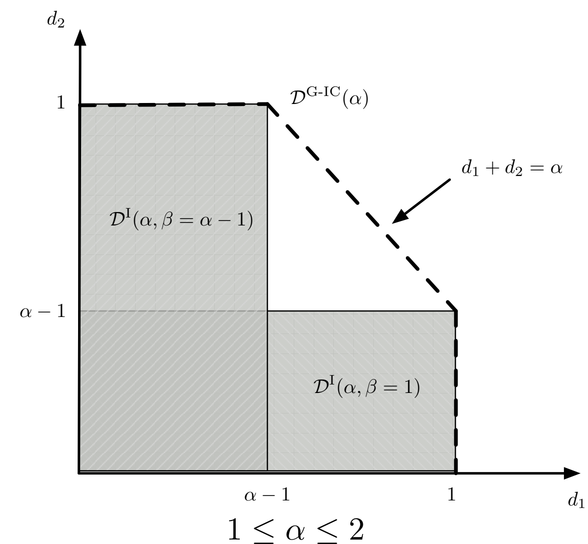
The result of Theorem 3.6.3 is quite surprising, namely, that for the G-IC-OR we can achieve the gDoF region of the classical G-IC in all regimes. This is especially surprising in the strong and very strong interference regimes where joint decoding of intended and interfering messages is optimal for the classical G-IC—recall that joint decoding appears to be precluded by the absence of codebook knowledge in the G-IC-OR. This seems to suggest that while decoding of the undesired messages is not possible, one may still estimate (i.e., symbol-by-symbol demodulate) the codeword symbols corresponding to the undesired messages.

3.7 Finite SNR Performance

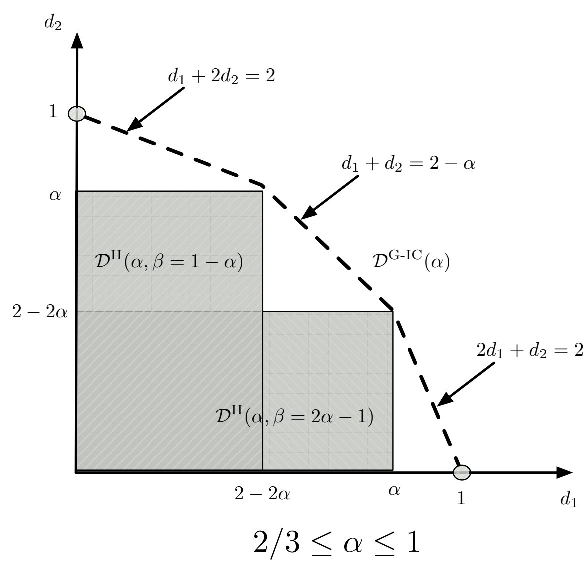
In the previous section we showed that the gDoF region of the classical G-IC can be achieved even when one receiver lacks knowledge of the interfering codebook. One may then ask whether it is possible to achieve the capacity, possibly up to a constant gap, of the classical G-IC at all finite SNRs. We next



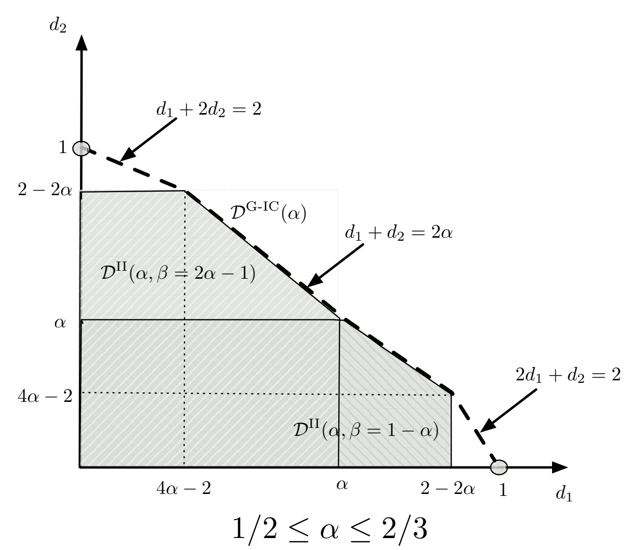
(a) Very strong interference.



(b) Strong interference.



(c) Moderately weak interference 1.



(d) Moderately weak interference 2.

Figure 7: How to achieve the gDoF region for the G-IC-OR in different parameter regimes.

show that this is indeed possible. For future use, the capacity region of the classical G-IC is outer bounded by (26)

$$
\begin{aligned}
\mathcal{R}_{out}^{\text{(G-IC)}}: \quad & R_1 \le \mathsf{I_g}\left(\mathsf{snr}\right), && (3.22a)\\
& R_2 \le \mathsf{I_g}\left(\mathsf{snr}\right), && (3.22b)\\
R_1 + R_2 &\le \left[\mathsf{I_g}\left(\mathsf{snr}\right) - \mathsf{I_g}\left(\mathsf{inr}\right)\right]^+ + \mathsf{I_g}\left(\mathsf{snr}+\mathsf{inr}\right), && (3.22c)\\
R_1 + R_2 &\le 2\mathsf{I_g}\left(\mathsf{inr} + \frac{\mathsf{snr}}{1+\mathsf{inr}}\right), && (3.22d)\\
2R_1 + R_2 &\le \left[\mathsf{I_g}\left(\mathsf{snr}\right) - \mathsf{I_g}\left(\mathsf{inr}\right)\right]^+ + \mathsf{I_g}\left(\mathsf{snr}+\mathsf{inr}\right) + \mathsf{I_g}\left(\mathsf{inr} + \frac{\mathsf{snr}}{1+\mathsf{inr}}\right), && (3.22e)\\
R_1 + 2R_2 &\le \left[\mathsf{I_g}\left(\mathsf{snr}\right) - \mathsf{I_g}\left(\mathsf{inr}\right)\right]^+ + \mathsf{I_g}\left(\mathsf{snr}+\mathsf{inr}\right) + \mathsf{I_g}\left(\mathsf{inr} + \frac{\mathsf{snr}}{1+\mathsf{inr}}\right), && (3.22f)
\end{aligned}
$$

which is tight for $\mathsf{snr} \le \mathsf{inr}$ and optimal to with $1/2$ bit (per channel use per user) otherwise.

The main result of this section is:

Theorem 3.7.1. *For the G-IC-OR it is possible to achieve the outer bound region in* (Eq. 3.22) *to within* $\frac{1}{2}\log\left(12\pi\mathrm{e}\right) \approx 3.34$ *bits per channel use per user.*

Proof of Theorem 3.7.1. We consider different regimes separately.

3.7.1 Very Strong Interference $\mathsf{snr}(1+\mathsf{snr}) \le \mathsf{inr}$

In the regime the capacity region of the classical G-IC is given by (Eq. 3.22a) and (Eq. 3.22b). For achievability we consider the achievable region in Theorem 3.5.1 with

$$N = \mathsf{N_d}(\mathsf{snr}) \quad \text{(equivalent of } \beta = 1) \Longrightarrow N^2 - 1 \le \mathsf{snr} \le \frac{\mathsf{inr}}{1+\mathsf{snr}} \le \mathsf{inr}. \tag{3.23}$$

Recall that the achievable region in Theorem 3.5.1 is the region in (Eq. 3.8) with the inputs as in (Eq. 3.14); the sum-rate in Theorem 3.5.1 is redundant if $I(X_1;Y_1|X_2)+I(X_2;Y_2)\leq I(X_1,X_2;Y_1)$, that is, if $I(X_2;Y_2)\leq I(X_2;Y_1)$, for all input distributions in (Eq. 3.14). With a Gaussian X_2 as in (Eq. 3.14):

$$I(X_2;Y_2)\leq I(X_2;Y_2|X_1)=I(X_{2G};\sqrt{\mathsf{snr}}\,X_{2G}+Z_2)=\mathsf{I_g}(\mathsf{snr}),$$

and

$$I(X_2;Y_1)=I(X_{2G};\sqrt{\mathsf{inr}}\,X_{2G}+\sqrt{\mathsf{snr}}\,X_{1D}+Z_2)\geq \mathsf{I_g}\left(\frac{\mathsf{inr}}{1+\mathsf{snr}}\right),$$

because a Gaussian noise is the worst noise for a Gaussian input. Since in very strong interference we have $\mathsf{I_g}(\mathsf{snr})\leq\mathsf{I_g}\left(\frac{\mathsf{inr}}{1+\mathsf{snr}}\right)$, the condition $I(X_2;Y_2)\leq I(X_2;Y_1)$ is verified for all inputs in (Eq. 3.14) and hence we can drop the sum-rate constraint in (Eq. 3.15c) from Theorem 3.5.1. Therefore, in this regime the following rates are achievable

$$\mathcal{R}_{\text{in}}^{\text{(G-IC-OR very strong)}}:\quad R_1\leq \mathsf{I_g}(\mathsf{snr})-\Delta_1, \tag{3.24a}$$

$$R_2\leq \mathsf{I_g}(\mathsf{snr})-\Delta_2, \tag{3.24b}$$

where for $N = \mathsf{N_d}(\mathsf{snr})$

$$
\begin{aligned}
\Delta_1 &:= \mathsf{I_g}(\mathsf{snr}) - \mathsf{I_d}\left(N, \mathsf{snr}\right) \\
&\leq \frac{1}{2}\log\left(\frac{4\pi\mathrm{e}}{3}\right) \text{ for the reasoning leading to eq.(\textbf{??})}, \qquad (3.24c)\\
\Delta_2 &:= \mathsf{I_g}\left(\min\left(N^2-1, \mathsf{inr}\right)\right) - \mathsf{I_d}\left(N, \frac{\mathsf{inr}}{1+\mathsf{snr}}\right) \\
&= \log(N) - \left[\log(N) - \frac{1}{2}\log\left(\frac{\pi\mathrm{e}}{3}\right)\right]^+ \leq \frac{1}{2}\log\left(\frac{\pi\mathrm{e}}{3}\right), \qquad (3.24d)
\end{aligned}
$$

where the equality in (Eq. 3.24d) is a consequence of the relationships in (Eq. 3.23).

It is immediate to see that (Eq. 3.24c) is the gap for R_1 and that (Eq. 3.24d) is the gap for R_2. Therefore in this regime the gap is at most $\frac{1}{2}\log\left(\frac{4\pi\mathrm{e}}{3}\right)$ per channel use per user, and it is due to shaping loss and integer penalty.

3.7.2 Strong Interference $\mathsf{snr} \leq \mathsf{inr} < \mathsf{snr}(1+\mathsf{snr})$

In this regime the capacity region of the classical G-IC is given by (Eq. 3.22a)-(Eq. 3.22c), and has two dominant corner points

$$
\mathcal{R}_{out}^{\text{(G-IC strong P1)}}: \quad (R_1, R_2) = \left(\mathsf{I_g}\left(\mathsf{snr}\right), \mathsf{I_g}\left(\frac{\mathsf{inr}}{1+\mathsf{snr}}\right)\right), \qquad (3.25a)
$$

and

$$
\mathcal{R}_{out}^{\text{(G-IC strong P2)}}: \quad (R_1, R_2) = \left(\mathsf{I_g}\left(\frac{\mathsf{inr}}{1+\mathsf{snr}}\right), \mathsf{I_g}\left(\mathsf{snr}\right)\right). \qquad (3.25b)
$$

The other two corner points are $(R_1, R_2) = (\mathsf{I_g}(\mathsf{snr}), 0)$ and $(R_1, R_2) = (0, \mathsf{I_g}(\mathsf{snr}))$ that can be exactly achieved by silencing one of the users.

For achievability we mimic the proof of the gDoF region in the same regime (see Fig. 7b), that is, we show the achievability to within a constant gap of the corner points in (Eq. 3.25a) and (Eq. 3.25b) by choosing two different values of N in Theorem 3.5.1. For the corner point in (Eq. 3.25a) we consider the achievable region in Theorem 3.5.1 with

$$N = \mathsf{N_d}(\mathsf{snr}) \quad \text{(equivalent of } \beta = 1) \Longrightarrow N^2 - 1 \le \mathsf{snr} \le \mathsf{inr} \le \mathsf{snr}(1+\mathsf{snr}), \tag{3.26a}$$

and for the corner point (Eq. 3.25b) we consider the achievable region in Theorem 3.5.1 with

$$N = \mathsf{N_d}\left(\frac{\mathsf{inr}}{1+\mathsf{snr}}\right) \quad \text{(equivalent of } \beta = \alpha - 1) \Longrightarrow N^2 - 1 \le \frac{\mathsf{inr}}{1+\mathsf{snr}} \le \mathsf{snr} \le \mathsf{inr}. \tag{3.26b}$$

For the choice of N in (Eq. 3.26a) the achievable region in Theorem 3.5.1 can be written as

$$\begin{aligned}
R_1 &\le \mathsf{I_d}\left(N, \mathsf{snr}\right) \\
&= \left[\log\left(N\right) - \frac{1}{2}\log\left(\frac{\pi \mathrm{e}}{3}\right)\right]^+, \\
R_2 &\le \mathsf{I_d}\left(N, \frac{\mathsf{inr}}{1+\mathsf{snr}}\right) + \mathsf{I_g}\left(\mathsf{snr}\right) - \mathsf{I_g}\left(\min\left(N^2-1, \mathsf{inr}\right)\right) \\
&= \left[\mathsf{I_g}\left(\frac{\mathsf{inr}}{1+\mathsf{snr}}\right) - \frac{1}{2}\log\left(\frac{\pi \mathrm{e}}{3}\right)\right]^+ + \mathsf{I_g}(\mathsf{snr}) - \log(N), \\
R_1 + R_2 &\le \mathsf{I_d}\left(N, \frac{\mathsf{snr}}{1+\mathsf{inr}}\right) + \mathsf{I_g}\left(\mathsf{inr}\right) \\
&= \left[\mathsf{I_g}\left(\frac{\mathsf{snr}}{1+\mathsf{inr}}\right) - \frac{1}{2}\log\left(\frac{\pi \mathrm{e}}{3}\right)\right]^+ + \mathsf{I_g}(\mathsf{inr}),
\end{aligned}$$

which can further be lower bounded as

$$
\begin{aligned}
\mathcal{R}_{in}^{\text{(G-IC strong P1)}}\ :\ R_1 &\leq \log\left(N\right) - \frac{1}{2}\log\left(\frac{\pi \mathrm{e}}{3}\right) \\
&= \mathsf{I_g}\left(\mathsf{snr}\right) - \Delta_1, && \text{(3.27a)} \\
R_2 &\leq \mathsf{I_g}\left(\mathsf{snr}+\mathsf{inr}\right) - \log(N) - \frac{1}{2}\log\left(\frac{\pi \mathrm{e}}{3}\right) \\
&= \mathsf{I_g}\left(\frac{\mathsf{inr}}{1+\mathsf{snr}}\right) - \Delta_2, && \text{(3.27b)} \\
R_1 + R_2 &\leq \mathsf{I_g}\left(\mathsf{snr}+\mathsf{inr}\right) - \frac{1}{2}\log\left(\frac{\pi \mathrm{e}}{3}\right) \\
&= \left(\mathsf{I_g}\left(\mathsf{snr}\right) - \Delta_1\right) + \left(\mathsf{I_g}\left(\frac{\mathsf{inr}}{1+\mathsf{snr}}\right) - \Delta_2\right) + \frac{1}{2}\log\left(\frac{\pi \mathrm{e}}{3}\right), && \text{(3.27c)}
\end{aligned}
$$

where the sum-rate bound is clearly redundant and where

$$\Delta_1 := \mathsf{I_g}\left(\mathsf{snr}\right) - \log(N) + \frac{1}{2}\log\left(\frac{\pi \mathrm{e}}{3}\right) \leq \frac{1}{2}\log\left(\frac{4\pi \mathrm{e}}{3}\right), \tag{3.28a}$$

$$\Delta_2 := \log(N) - \mathsf{I_g}\left(\mathsf{snr}\right) + \frac{1}{2}\log\left(\frac{\pi \mathrm{e}}{3}\right) \leq \frac{1}{2}\log\left(\frac{\pi \mathrm{e}}{3}\right). \tag{3.28b}$$

Therefore, with N as in (Eq. 3.26a) in Theorem 3.5.1, the gap to the corner point in (Eq. 3.25a) is at most $\frac{1}{2}\log\left(\frac{4\pi \mathrm{e}}{3}\right)$ per channel use per user, as for the very strong interference regime.

By following similar steps, for the choice of N in (Eq. 3.26b) in Theorem 3.5.1, the gap to the corner point in (Eq. 3.25b) is still given by (Eq. 3.28), that is, the gap is at most $\frac{1}{2}\log\left(\frac{4\pi \mathrm{e}}{3}\right)$ per channel use per user, as for the very strong interference regime.

3.7.3 Moderately Weak Interference $\mathsf{inr} \leq \mathsf{snr} \leq \mathsf{inr}(1+\mathsf{inr})$

In this regime the capacity of the G-IC is outer bounded by (Eq. 3.22).

As we did for the gDoF region (see Figs. 7c and 7d), we show here that we can achieve, up to a constant gap, all dominant corner points of (Eq. 3.22). By silencing one of the users, we can achieve $(R_1, R_2) = (\mathsf{I_g}(\mathsf{snr}), 0)$ and $(R_1, R_2) = (0, \mathsf{I_g}(\mathsf{snr}))$; these rate points are to within 1 bit of the corner points of (Eq. 3.22) given by $(R_1, R_2) = (A, \mathsf{I_g}(\mathsf{snr}))$ and $(R_1, R_2) = (\mathsf{I_g}(\mathsf{snr}), A)$ where

$$\begin{aligned}
A &:= \mathsf{I_g}(\mathsf{snr} + \mathsf{inr}) + \mathsf{I_g}\left(\mathsf{inr} + \frac{\mathsf{snr}}{1 + \mathsf{inr}}\right) - \mathsf{I_g}(\mathsf{snr}) - \mathsf{I_g}(\mathsf{inr}) \\
&= \mathsf{I_g}\left(\frac{\mathsf{inr}}{1 + \mathsf{snr}}\right) + \mathsf{I_g}\left(\frac{\mathsf{snr}}{(1 + \mathsf{inr})^2}\right) \\
&\le \mathsf{I_g}\left(\frac{\mathsf{snr}}{1 + \mathsf{snr}}\right) + \mathsf{I_g}\left(\frac{\mathsf{inr}}{1 + \mathsf{inr}}\right) \le 2 \cdot \frac{1}{2}\log(2) = 1.
\end{aligned}$$

We therefore have to show the achievability of the remaining two corner points obtained by the intersection of the sum-rate outer bound (given by $\min$(eq.(Eq. 3.22c),eq.(Eq. 3.22d))) with either (Eq. 3.22e) or (Eq. 3.22f). For these corner points, the gDoF-optimal choices of β were $2\alpha - 1$ and $1 - \alpha$, which we mimic here by choosing the following values of N in the region in (Eq. C.2) (a simplified achievable region from Theorem 3.5.2)

$$\begin{aligned}
&N = \mathsf{N_d}\left(\frac{\mathsf{inr}^2}{1 + \mathsf{snr} + 2\mathsf{inr}}\right) \quad \text{(equivalent of } \beta = 2\alpha - 1\text{)} \\
&\Longrightarrow N^2 - 1 \le \frac{\mathsf{inr}^2}{1 + \mathsf{snr} + 2\mathsf{inr}} \le \min\left(\frac{\mathsf{inr}^2}{1 + 2\mathsf{inr}}, \frac{\mathsf{inr} \cdot \mathsf{snr}}{1 + \mathsf{snr} + 2\mathsf{inr}}\right),
\end{aligned} \tag{3.29}$$

because $\mathsf{inr} \leq \mathsf{snr}$, and

$$N = \mathsf{N_d}\left(\frac{\mathsf{snr} \cdot \mathsf{inr}}{(1+\mathsf{inr})^2 + \mathsf{snr}}\right) \quad \text{(equivalent of } \beta = 1-\alpha\text{)}$$
$$\Longrightarrow N^2 - 1 \leq \frac{\mathsf{snr} \cdot \mathsf{inr}}{(1+\mathsf{inr})^2 + \mathsf{snr}} \leq \min\left(\frac{\mathsf{inr}^2}{1+2\mathsf{inr}}, \frac{\mathsf{inr} \cdot \mathsf{snr}}{1+\mathsf{snr}+2\mathsf{inr}}\right), \tag{3.30}$$

because $\mathsf{snr} \leq \mathsf{inr}(1+\mathsf{inr})$. In the regime $\mathsf{inr} \leq \mathsf{snr} \leq \mathsf{inr}(1+\mathsf{inr})$ we also have

$$\frac{\mathsf{inr}^2}{(1+\mathsf{inr})(1+\mathsf{snr})+\mathsf{inr}} \leq \frac{\mathsf{inr}^2}{(1+\mathsf{inr})^2+\mathsf{inr}} \leq 1 \leq N^2 - 1, \quad \forall N \geq 2. \tag{3.31}$$

With (Eq. 3.29)-(Eq. 3.31), and by recalling that $\mathsf{I_g}(x) - \frac{1}{2}\log(4) \le \log(\mathsf{N_d}(x)) \le \mathsf{I_g}(x),\ x \ge 0$, the region in (Eq. C.2) can be further lower bounded as follows[1]

$$\begin{aligned}
\mathcal{R}_{in}^{\text{(G-IC-OR weak)}}: \quad R_1 &\le \mathsf{I_g}(x) - \frac{1}{2}\log(4) \\
&\quad - \frac{1}{2}\log\left(\frac{\pi \mathrm{e}}{3}\right) + \mathsf{I_g}\left(\frac{\mathsf{snr}}{1+2\mathsf{inr}}\right), && \text{(3.32a)}\\
R_2 &\le \mathsf{I_g}\left(\frac{\mathsf{inr}^2}{(1+\mathsf{inr})(1+\mathsf{snr})+\mathsf{inr}}\right) - \frac{1}{2}\log\left(\frac{\pi \mathrm{e}}{3}\right) \\
&\quad + \mathsf{I_g}\left(\frac{\mathsf{snr}}{2}\right) - \mathsf{I_g}(x), && \text{(3.32b)}\\
R_1 + R_2 &\le \mathsf{I_g}\left(\min\left(\frac{\mathsf{inr}^2}{1+\mathsf{snr}+2\mathsf{inr}}, \frac{\mathsf{snr}\cdot\mathsf{inr}}{(1+\mathsf{inr})^2+\mathsf{snr}}\right)\right) - \frac{1}{2}\log(4) \\
&\quad + \mathsf{I_g}\left(\mathsf{inr} + \frac{\mathsf{snr}}{1+\mathsf{inr}}\right) - \mathsf{I_g}\left(\frac{\mathsf{inr}}{1+\mathsf{inr}}\right) + \mathsf{I_g}\left(\frac{\mathsf{snr}}{1+2\mathsf{inr}}\right) \\
&\quad - 2\cdot\frac{1}{2}\log\left(\frac{\pi \mathrm{e}}{3}\right), && \text{(3.32c)}
\end{aligned}$$

where

$$x := \frac{\mathsf{inr}^2}{1+\mathsf{snr}+2\mathsf{inr}} \text{ if } N \text{ as in (Eq. 3.29), or} \quad \text{(3.32d)}$$

$$x := \frac{\mathsf{snr}\cdot\mathsf{inr}}{(1+\mathsf{inr})^2+\mathsf{snr}} \text{ if } N \text{ as in (Eq. 3.30).} \quad \text{(3.32e)}$$

---

[1] In order to get the sum-rate, let $n = N^2 - 1 \in \mathbb{N}$ and consider either $N = \mathsf{N_d}(a) : n_a := \mathsf{N_d}(a)^2 - 1 \le a \in \mathbb{R}^+$ or $N = \mathsf{N_d}(b) : n_b := \mathsf{N_d}(a)^2 - 1 \le b \in \mathbb{R}^+$ in the expression $y(n) := \mathsf{I_g}(\min(n,a)) + \mathsf{I_g}(\min(n,b)) - \mathsf{I_g}(n)$ that appears in the sum-rate. It follows easily that for $N = \mathsf{N_d}(a)\ :\ y = \mathsf{I_g}(\min(n_a,b)) \ge \mathsf{I_g}(\min(n_a,n_b)) \ge \mathsf{I_g}(\min(a,b)) - \frac{1}{2}\log(4)$, and for $N = \mathsf{N_d}(b)\ :\ y = \mathsf{I_g}(\min(a,n_b)) \ge \mathsf{I_g}(\min(n_a,n_b)) \ge \mathsf{I_g}(\min(a,b)) - \frac{1}{2}\log(4)$, where the term $\frac{1}{2}\log(4)$ is due to the "integer penalty".

In Appendix D we show that region in (Eq. 3.32) achieves the classical G-IC outer bound to within $\frac{1}{2}\log(12\pi\mathrm{e}) \approx 3.34$ bits (per channel user per user).

3.7.4 Noisy Interference $\mathsf{inr}(1+\mathsf{inr}) \le \mathsf{snr}$

In this regime Gaussian inputs, treating interference as noise, and power control is optimal to within $1/2$ bit (per channel use per user) for the classical G-IC; since this scheme does not require codebook knowledge / joint decoding, the gap result applies to the G-IC-OR as well.

This concludes the proof. □

3.8 Conclusion

In this part of the thesis we derived capacity results for the interference channel where one of the receivers lacks knowledge of the interfering codebook, in contrast to a classical model where both receivers possess full codebook knowledge. For the class of injective semi-deterministic interference channels with one oblivious receiver, we derived a capacity result to within a constant gap; the gap is zero for fully deterministic channels, thereby providing an exact capacity characterization. We also derived the exact capacity region for a general memoryless interference channel with one oblivious receiver in the regime where the non-oblivious receiver experiences very strong interference.

We next proceeded to the Gaussian noise channel, where, unlike past work on oblivious receivers, we were able to demonstrate performance guarantees. For the symmetric case we derived the gDoF region and the capacity region to within a constant gap of $\frac{1}{2}\log(12\pi\mathrm{e}) \approx 3.34$ bits (per channel use per user). Surprisingly, this lack of codebook knowledge at one receiver does not impact the gDoF at all, and only the Gaussian capacity region to within a constant gap, compared to having full codebook knowledge. We believe this is because even though the mapping from codewords to messages may not be known, this does not prevent the receiver from estimating (for example by symbol-by-symbol demodulation) and removing the effect of the interfering codeword itself.

An interesting direction is to consider a generalization with lack of interfering codebook knowledge at both receivers, where one might surmise that both inputs would have discrete components. However, this generalization is highly non-trivial and significantly more mathematically challenging, and was left as an open problem in (20). The major issue that arises when both users employ discrete inputs is the need to compute the cardinality and minimum distance of the sum of two discrete sets. These quantities are not only difficult to compute in general, but are also very sensitive to whether channel gains are rational or irrational (this is an open problem in additive combinatorics). We will make considerable progress on this problem in Chapter 4 and give an approximate optimality results.

We studied the performance of mixed inputs on the G-IC. Its application to oblivious and asynchronous ICs somewhat surprisingly implies that much less "global coordination" between nodes is needed than one might initially expect: synchronism and codebook knowledge might not be critical if one is happy with "approximate" capacity results. Why discrete inputs are able to resolve these issues might be because even simple expressions such as $I(X_1;Y_1) + I(X_2;Y_2)$, which do not appear to employ joint decoding, may still capture some form of "interference estimation". This estimation theoretic explanation will be revisited in detail in Chapter 5.

CHAPTER 4

APPROXIMATE OPTIMALITY OF TREATING INTERFERENCE AS NOISE

4.1 Introduction

Part of this chapter has been previously published in (8).

Consider the two-user memoryless real-valued additive white Gaussian noise interference channel (G-IC) with input-output relationship

$$Y_1^n = h_{11}X_1^n + h_{12}X_2^n + Z_1^n, \tag{4.1a}$$

$$Y_2^n = h_{21}X_1^n + h_{22}X_2^n + Z_2^n, \tag{4.1b}$$

where $X_j^n := (X_{j1}, \cdots X_{jn})$ and $Y_j^n := (Y_{j1}, \cdots Y_{jn})$ are the length-n vector input and output, respectively, for user $j \in [1:2]$, the noise vector Z_j^n has i.i.d. zero-mean unit-variance Gaussian components, the input X_j^n is subject to a per-block power constraint $\frac{1}{n}\sum_{i=1}^n X_{ji}^2 \leq 1$, and the channel gains $(h_{11}, h_{12}, h_{21}, h_{22})$ are fixed and known to all nodes. The input X_j^n, $j \in [1:2]$, carries the independent message W_j that is uniformly distributed on $[1:2^{nR_j}]$, where R_j is the rate and n the block-length. Receiver $j \in [1:2]$ wishes to recover W_j from the channel output Y_j^n with arbitrarily small probability of error. Achievable rates and capacity region are defined in the usual way (37). We shall denote the capacity region by $\mathcal{C}$.

For simplicity we will focus primarily on the *symmetric* G-IC defined by

$$|h_{11}|^2 = |h_{22}|^2 = \mathsf{snr} \geq 0, \tag{4.2a}$$

$$|h_{12}|^2 = |h_{21}|^2 = \mathsf{inr} \geq 0, \tag{4.2b}$$

and we will discuss how the results for the symmetric G-IC extend to the general asymmetric setting.

The general *discrete memoryless* IC was introduced in (43) where it was shown that the capacity region of an *information stable IC* (44) is given by

$$\mathcal{C} = \lim_{n\to\infty} \mathsf{co}\left(\bigcup_{P_{X_1^n X_2^n} = P_{X_1^n} P_{X_2^n}} \left\{ \begin{array}{l} 0 \leq R_1 \leq \frac{1}{n} I(X_1^n; Y_1^n) \\ 0 \leq R_2 \leq \frac{1}{n} I(X_2^n; Y_2^n) \end{array} \right\} \right). \tag{4.3}$$

For the G-IC in (Eq. 4.1), the maximization in (Eq. 4.3) is further restricted to inputs satisfying the power constraint $\frac{1}{n}\sum_{i=1}^n X_{i,j}^2 \leq 1,\ j \in [1:2]$.

An inner bound to the capacity region in (Eq. 4.3) can be obtained by considering i.i.d. inputs in (Eq. 4.3) thus giving

$$\mathcal{R}_{\text{in}}^{\text{TIN+TS}} = \mathsf{co}\left(\bigcup_{P_{X_1 X_2} = P_{X_1} P_{X_2}} \left\{ \begin{array}{l} 0 \leq R_1 \leq I(X_1; Y_1) \\ 0 \leq R_2 \leq I(X_2; Y_2) \end{array} \right\} \right), \tag{4.4}$$

where the superscript "TIN+TS" reminds the reader that the region is achieved by Treating Interference as Noise (TIN)[1] and with Time Sharing (TS), where TS is enabled by the convex hull operation (37). By further removing the convex hull operation in (Eq. 4.4) we arrive at

$$\mathcal{R}_{\text{in}}^{\text{TINnoTS}} = \bigcup_{P_{X_1X_2}=P_{X_1}P_{X_2}} \left\{ \begin{array}{c} 0 \le R_1 \le I(X_1;Y_1) \\ \\ 0 \le R_2 \le I(X_2;Y_2) \end{array} \right\}. \tag{4.5}$$

The region in (Eq. 4.5) does not allow the users to time-share. For the G-IC the maximization in (Eq. 4.4) and (Eq. 4.5) is further restricted to inputs satisfy average power constraint $\frac{1}{n}\sum_{i=1}^{n} X_{i,j}^2 \le 1,\ j \in [1 : 2]$.

Obviously

$$\mathcal{R}_{\text{in}}^{\text{TINnoTS}} \subseteq \mathcal{R}_{\text{in}}^{\text{TIN+TS}} \subseteq \mathcal{C}.$$

---

[1]We use the terminology "treating interference as noise" to denote the rates obtained when evaluating expressions for the interference channel of the form

$$\text{Desired rate} \le I(\text{desired input}; \text{output}),$$

without any other rate expressions, mutual information terms, or explicit rate splits. When evaluated with independent and identically distributed (i.i.d.) Gaussian inputs in the interference channel in (Eq. 4.1), these rate expressions look like those in which the interference is indeed treated as noise, i.e.,

$$0 \le R_i \le \frac{1}{2}\log\left(1 + \frac{\mathsf{snr}}{1+\mathsf{inr}}\right),\ i \in [1:2],$$

where the 'effective noise' (at the denominator within the log) looks like the true noise power plus all the interferer's power. Whether this expression has the same "treating interference as noise" interpretation when using non-Gaussian inputs is open to interpretation, and is one of the focusses of this work. We will however continue to use this terminology.

The question of interest in this chapter is how $\mathcal{R}_{\text{in}}^{\text{TINnoTS}}$ fares compared to $\mathcal{C}$. Note that there are many advantages in using TINnoTS in practice. For example, TINnoTS does not require codeword synchronization, as for example for joint decoding or interference cancellation, and does not require much coordination between users, thereby reducing communications overhead. The goal of this chapter is to show that despite its simplicity, TINnoTS approximately achieves the capacity $\mathcal{C}$.

Next, we review past work relevant to our investigation. We refer the interested reader to (37) for a comprehensive literature survey on general discrete memoryless ICs.

4.1.1 Past Work

In general, little is known about the optimizing input distribution in (Eq. 4.3) for the G-IC (or in (Eq. 4.4) and in (Eq. 4.5)) and only some special cases have been solved. In (45) it was shown that i.i.d. Gaussian inputs maximize the sum-capacity in (Eq. 4.3) for $\sqrt{\frac{\mathsf{inr}}{\mathsf{snr}}}(1+\mathsf{inr}) \leq \frac{1}{2}$ in the symmetric case. In contrast, the authors of (46) showed that in general multivariate Gaussian inputs do not exhaust regions of the form in (Eq. 4.3). The difficulty arises from the competitive nature of the problem (15): for example, say X_2 is i.i.d. Gaussian, taking X_1 to be Gaussian increases $I(X_1;Y_1)$ but simultaneously decreases $I(X_2;Y_2)$, as Gaussians are known to be the "best inputs" for Gaussian point-to-point power-constrained channels, but are also the "worst noise" (or interference, if it is treated as noise) for a Gaussian input.

In Chapter 3, for the G-IC with one oblivious receiver, we showed that a properly chosen discrete input has a somewhat different behavior than a Gaussian input: a discrete X_2 may yield a "good" $I(X_1;Y_1)$ while keeping $I(X_2;Y_2)$ relatively unchanged compared to a Gaussian input, thus substantially improving the rates compared to Gaussian inputs in the same achievable region expression. Moreover, in Chapter 3 we showed that treating interference as noise at the oblivious receiver and joint decoding at the other receiver is to within an additive gap of 3.34 bits of the capacity. In this work we seek to *analytically* evaluate the lower bound in (Eq. 4.5) for a special class of *mixed inputs* by generalizing the approach taken in Chapter 3 and show that using TINnoTS at both receivers is to within an additive gap of the capacity. In a way this work follows the philosophy of (16): the main idea is to use sub-optimal point-to-point codes in which the reduction in achievable rate for the intended receiver is more than compensated by the decrease in the interference created at the other receiver, which results in an overall rate region improvement.

We remark that the gDoF optimality of TINnoTS for all channel parameters for the G-IC was pointed out in (37, Remark 6.12). The proof follows since TINnoTS is always optimal for the Linear Deterministic Approximation (LDA) of the G-IC at high-SNR (47). Moreover, a scheme for the LDA can be translated into a scheme for the real-valued G-IC that is optimal to within at most 18.6 bits (23, Theorem 2). This line of reasoning based on a universal gap between the LDA and the G-IC, thus giving a constant gap result, does not provide a concrete practical construction of an approximately optimal scheme. The idea of 'lifting' an LDA optimal scheme to the G-IC has been used in (48) where a $O(\log\log(\mathsf{snr}))$, rather than a constant, gap result was proved for the symmetric sum-capacity. Our proof here extends our original approach in (4) and provides, in closed form, the optimal number of points in the discrete part of the mixed inputs, as well as of the optimal power split among the discrete and continuous parts of the mixed inputs. Moreover, our derived gap is in general smaller than 18.6 bits (this is so because the log-log function grows very slowly in its argument).

In a conference version of this thesis (4) we demonstrated that TINnoTS is gDoF optimal and can achieve to within an additive gap the symmetric sum-capacity of the classical G-IC. In (48, Theorem 3.), the authors showed that the sum-capacity result of (4) can be achieved by an input with purely discrete marginals, i.e., the Gaussian part of our mixed inputs can be replaced by a discrete random variable.

We conclude this overview of relevant past work by pointing out that in practice it is well known that a non-Gaussian interference should not be treated as a Gaussian noise. The optimal detector for an additive non-Gaussian noise channel may however be far more complex than a classical minimum-distance decoder. Nonetheless, since the performance increase can be substantial for a moderate complexity increase, Network-Assisted Interference Cancellation and Suppression (NAICS) receivers, which account for the discrete and coded nature of the interference, were adopted in the Long Term Evolution (LTE) Advanced Release 12 (49; 50; 51). The boost in performance of NAICS-type detectors may be understood as follows. As we pointed out in (5), with TIN the mapping of the codewords to the messages is unknown but the codeword symbols may be known through soft symbol-by-symbol estimation as remarked in (19), where the authors write “We indeed see that BPSK signaling outperforms Gaussian signaling. This is because demodulation is some form of primitive decoding, which is not possible for the Gaussian signaling.”

4.2 TINnoTS with Mixed Inputs Achievable Rate Region and an Outer Bound for the G-IC

For the G-IC in (Eq. 4.1) we now evaluate the TINnoTS region in (Eq. 4.5) with inputs

$$X_i = \sqrt{1-\delta_i}\, X_{iD} + \sqrt{\delta_i}\, X_{iG},\ i \in [1:2]: \tag{4.6a}$$

$$X_{iD} \sim \mathsf{PAM}\left(N_i, \sqrt{\frac{12}{N_i^2-1}}\right), \tag{4.6b}$$

$$X_{iG} \sim \mathcal{N}(0,1), \tag{4.6c}$$

$$\mathbf{p} := [N_1, N_2, \delta_1, \delta_2] \in \mathbb{N} \times \mathbb{N} \times [0,1] \times [0,1], \tag{4.6d}$$

where the random variables X_{ij} are independent for $i \in [1:2]$ and $j \in \{D, G\}$. The input in (Eq. 4.6) has four parameters, collected in the vector $\mathbf{p}$, namely: the number of points $N_i \in \mathbb{N}$ and the power split $\delta_i \in [0,1]$, for $i \in [1:2]$, which must be chosen carefully in order to match a given outer bound.

Proposition 4.2.1. *For the G-IC the TINnoTS region in* (Eq. 4.5) *contains the region* $\mathcal{R}_{in}$ *defined as*

$$\mathcal{R}_{in} := \bigcup \left\{ \begin{array}{l} 0 \le R_1 \le \mathsf{I_d}\left(S_1\right) + \mathsf{I_g}\left(\frac{|h_{11}|^2\delta_1}{1+|h_{12}|^2\delta_2}\right) - \min\left(\log(N_2), \mathsf{I_g}\left(\frac{|h_{12}|^2(1-\delta_2)}{1+|h_{12}|^2\delta_2}\right)\right) \\ 0 \le R_2 \le \mathsf{I_d}\left(S_2\right) + \mathsf{I_g}\left(\frac{|h_{22}|^2\delta_2}{1+|h_{21}|^2\delta_1}\right) - \min\left(\log(N_1), \mathsf{I_g}\left(\frac{|h_{21}|^2(1-\delta_1)}{1+|h_{21}|^2\delta_1}\right)\right) \end{array} \right\}, \tag{4.7}$$

where the union is over all possible parameters $[N_1, N_2, \delta_1, \delta_2] \in \mathbb{N}^2 \times [0,1]^2$ *for the mixed inputs in* (Eq. 4.6) *and where the equivalent discrete constellations seen at the receivers are*

$$S_1 := \frac{1}{\sqrt{1+|h_{11}|^2\delta_1+|h_{12}|^2\delta_2}}(\sqrt{1-\delta_1}h_{11}X_{1D} + \sqrt{1-\delta_2}h_{12}X_{2D}), \tag{4.8a}$$

$$S_2 := \frac{1}{\sqrt{1+|h_{21}|^2\delta_1+|h_{22}|^2\delta_2}}(\sqrt{1-\delta_1}h_{21}X_{1D} + \sqrt{1-\delta_2}h_{22}X_{2D}). \tag{4.8b}$$

Proof. Due to the symmetry of the problem we derive a lower bound on $I(X_2;Y_2)$ only by following steps similar to those in (Eq. 2.27); a lower bound on $I(X_1;Y_1)$ follows by swapping the role of the users. Let $Z_G \sim \mathcal{N}(0,1)$. Then:

$$\begin{aligned}
I(X_2;Y_2) &= I(X_2; h_{21}X_1 + h_{22}X_2 + Z_G) \\
&= \underbrace{\left[h\left(\frac{\sqrt{1-\delta_1}h_{21}X_{1D} + \sqrt{1-\delta_2}h_{22}X_{2D}}{\sqrt{1+|h_{21}|^2\delta_1+|h_{22}|^2\delta_2}} + Z_G\right) - h(Z_G)\right]}_{\geq \mathsf{I_d}(S_2)\text{ by Proposition 2.1.1}} \\
&- \underbrace{\left[h\left(\frac{\sqrt{1-\delta_1}}{\sqrt{1+|h_{21}|^2\delta_1}}h_{21}X_{1D} + Z_G\right) - h(Z_G)\right]}_{\leq \min\left(\log(N_1), \frac{1}{2}\log\left(1+\frac{|h_{21}|^2(1-\delta_1)}{1+|h_{21}|^2\delta_1}\right)\right)\text{ by Remark 3}} \\
&+ \frac{1}{2}\log\left(1+|h_{21}|^2\delta_1+|h_{22}|^2\delta_2\right) - \frac{1}{2}\log(1+|h_{21}|^2\delta_1).
\end{aligned}$$

By considering the union over all possible choices of parameters for the mixed inputs we obtain the region in (Eq. 4.7), which is contained within the achievable region in (Eq. 4.5) and hence forms a lower bound to the capacity region. □

In the following sections we shall show that our TINnoTS region with mixed inputs in Proposition 4.2.1 is to within an additive gap of the outer bound region given by:

Proposition 4.2.2. *The capacity region of the G-IC is contained in*

$$
\begin{aligned}
\mathcal{R}_{out} = \Big\{ \quad & R_1 \le \mathsf{I_g}\left(|h_{11}|^2\right), \text{ \textit{cut-set bound},} && (4.9a)\\
& R_2 \le \mathsf{I_g}\left(|h_{22}|^2\right), \text{ \textit{cut-set bound},} && (4.9b)\\
R_1 + R_2 &\le \Big[\mathsf{I_g}\left(|h_{11}|^2\right) - \mathsf{I_g}\left(|h_{21}|^2\right)\Big]^+ + \mathsf{I_g}(|h_{21}|^2 + |h_{22}|^2), \text{ \textit{from (52)},} && (4.9c)\\
R_1 + R_2 &\le \Big[\mathsf{I_g}\left(|h_{22}|^2\right) - \mathsf{I_g}\left(|h_{12}|^2\right)\Big]^+ + \mathsf{I_g}(|h_{11}|^2 + |h_{12}|^2), \text{ \textit{from (52)},} && (4.9d)\\
R_1 + R_2 &\le \mathsf{I_g}\left(|h_{12}|^2 + \frac{|h_{11}|^2}{1+|h_{21}|^2}\right) + \mathsf{I_g}\left(|h_{21}|^2 + \frac{|h_{22}|^2}{1+|h_{12}|^2}\right), \text{ \textit{from (26)},} && (4.9e)\\
2R_1 + R_2 &\le \mathsf{I_g}(|h_{11}|^2 + |h_{12}|^2) + \mathsf{I_g}\left(|h_{21}|^2 + \frac{|h_{22}|^2}{1+|h_{12}|^2}\right)\\
&\quad + \Big[\mathsf{I_g}\left(|h_{11}|^2\right) - \mathsf{I_g}\left(|h_{21}|^2\right)\Big]^+, \text{ \textit{from (26)},} && (4.9f)\\
R_1 + 2R_2 &\le \mathsf{I_g}(|h_{21}|^2 + |h_{22}|^2) + \mathsf{I_g}\left(|h_{12}|^2 + \frac{|h_{11}|^2}{1+|h_{21}|^2}\right)\\
&\quad + \Big[\mathsf{I_g}\left(|h_{22}|^2\right) - \mathsf{I_g}\left(|h_{12}|^2\right)\Big]^+, \text{ \textit{from (26)}} \quad \Big\}. && (4.9g)
\end{aligned}
$$

For the classical G-IC where all nodes are synchronous and possess full codebook knowledge, this outer bound is tight in strong interference $\{|h_{21}|^2 \ge |h_{11}|^2,\ |h_{12}|^2 \ge |h_{22}|^2\}$ *(53) and achievable to within* $1/2$ *bit otherwise (26).*

The key step to match, to within an additive gap, the outer bound region $\mathcal{R}_{\text{out}}$ in Proposition 4.2.2 to our TINnoTS achievable region with mixed inputs $\mathcal{R}_{\text{in}}$ in Proposition 4.2.1 is to carefully choose the mixed input parameter vector $[N_1, N_2, \delta_1, \delta_2]$. This 'carefully picking of the mixed input parameters' is the objective of Section 4.3.

4.3 Symmetric Capacity Region to within a Gap

The main result of this chapter is:

Theorem 4.3.1. *For the symmetric G-IC, as defined in* (Eq. 4.2)*, the TINnoTS achievable region in* (Eq. 4.7)*, with the parameters for the mixed inputs chosen as indicated in Table II, and the outer bound in* (Eq. 4.9) *are to within a gap of:*

- *Very Weak Interference:* $\mathsf{snr} \geq \mathsf{inr}(1+\mathsf{inr})$*:*

$$\mathsf{G_d} \leq \frac{1}{2} \textit{ bits},$$

- *Moderately Weak Interference Type2:* $\mathsf{inr} \leq \mathsf{snr} \leq \mathsf{inr}(1+\mathsf{inr})$, $\frac{1+\mathsf{snr}}{1+\mathsf{inr}+\frac{\mathsf{snr}}{1+\mathsf{inr}}} > \frac{1+\mathsf{inr}+\frac{\mathsf{snr}}{1+\mathsf{inr}}}{1+\frac{\mathsf{snr}}{1+\mathsf{inr}}}$*:*

$$\mathsf{G_d} \leq \frac{1}{2} \log\left(\frac{608\ \pi\mathrm{e}}{27}\right) \approx 3.79 \textit{ bits},$$

- *Moderately Weak Interference Type1:* $\mathsf{inr} \leq \mathsf{snr} \leq \mathsf{inr}(1+\mathsf{inr})$, $\frac{1+\mathsf{snr}}{1+\mathsf{inr}+\frac{\mathsf{snr}}{1+\mathsf{inr}}} \leq \frac{1+\mathsf{inr}+\frac{\mathsf{snr}}{1+\mathsf{inr}}}{1+\frac{\mathsf{snr}}{1+\mathsf{inr}}}$*:*

$$\mathsf{G_d} \leq \frac{1}{2} \log\left(\frac{16\pi\mathrm{e}}{3}\right) + \frac{1}{2} \log\left(1 + 45 \cdot \frac{(1+1/2\ln(1+\min(\mathsf{inr},\mathsf{snr})))^2}{\gamma^2}\right) \textit{ bits},$$

  except for a set of measure γ for any $\gamma \in (0,1]$,

- *Strong Interference:* $\mathsf{snr} < \mathsf{inr} < \mathsf{snr}(1+\mathsf{snr})$*:*

$$\mathsf{G_d} \leq \frac{1}{2} \log\left(\frac{2\pi\mathrm{e}}{3}\right) + \frac{1}{2} \log\left(1 + 8 \cdot \frac{(1+1/2\ln(1+\min(\mathsf{inr},\mathsf{snr})))^2}{\gamma^2}\right) \textit{ bits},$$

except for a set of measure γ *for any* $\gamma \in (0, 1]$*,*

- *Very Strong Interference:* $\mathsf{inr} \geq \mathsf{snr}(1 + \mathsf{snr})$*:*

$$\mathsf{G_d} \leq \frac{1}{2} \log \left(\frac{2\pi \mathrm{e}}{3} \right) \approx 1.25 \textit{ bits}.$$

Before we move to the proof of Theorem 4.3.1, we would like to offer our thoughts on why a $O\left(\log\left(\frac{\ln(\min(\mathsf{snr},\mathsf{inr}))}{\gamma}\right)\right)$ gap is obtained in some regimes up to an outage set of controllable measure (the larger the measure of the channel gains for which the derived gap does not hold, the lower the gap). We start by noticing that, for the symmetric G-IC, whenever the TINnoTS region with our mixed input is optimal to within a *constant gap* then the gap result holds for *all channel gains*. Otherwise, the optimality is to within a $O\left(\log\left(\frac{\ln(\min(\mathsf{snr},\mathsf{inr}))}{\gamma}\right)\right)$ *gap* and holds for all channel gains *up to an outage set*. We found a $O\left(\log\left(\frac{\ln(\min(\mathsf{snr},\mathsf{inr}))}{\gamma}\right)\right)$ gap up to an outage set whenever the sum-rate upper bound $\min\big($eq.(Eq. 4.9c), eq.(Eq. 4.9d)$\big)$ is active, which in gDoF corresponds to the regime $\alpha \in (2/3, 2)$ meaning that the interference is neither very weak nor very strong. It is thus natural to ask: (a) whether the $O\left(\log\left(\frac{\ln(\min(\mathsf{snr},\mathsf{inr}))}{\gamma}\right)\right)$ gap and/or the 'up to an outage set' condition are necessary (not a consequence of the achievable scheme used), and (b) whether a $O\left(\log\left(\frac{\ln(\min(\mathsf{snr},\mathsf{inr}))}{\gamma}\right)\right)$ gap and the 'up to an outage set' condition are necessarily always together. We do not have answers to these questions, but we provide our perspective next.

The sum-rate bounds in (Eq. 4.9c) and (Eq. 4.9d) were originally derived for the classical two-user IC in Gaussian noise in (52) and then extended to any memoryless two-user IC with source cooperation / generalized feedback in (54), and then to any memoryless cooperative two-user IC (where each node can have an input and an output to the channel) in (55) – see also K-user extensions in (56; 57). In (55) it was noted that surprisingly these bounds hold for a broad class of two-user IC-type channels, which includes for example cognitive ICs and certain ICs with cooperation. The difference is that the mutual information optimization is over all product input distributions for the classical IC, while it is over all joint input distributions for the cooperative or cognitive IC. The ability to correlate inputs is well known to only increase the rates by a constant number of bits; thus, up to a constant gap, channel models from the basic classical IC to the intricate cognitive IC have the same sum-rate upper bound in some regimes. Note that for the real-valued cognitive G-IC for example, the sum-rate bound is achievable to within 1/2 bit for all channel gains by using Dirty Paper Coding. It is not clear at this point whether the $O\left(\log\left(\frac{\ln(\min(\mathsf{snr},\mathsf{inr}))}{\gamma}\right)\right)$ gap up to an outage set for the classical G-IC is thus a fundamental consequence of the fact that the upper bound can be achieved to within a constant gap with sophisticated coding techniques (such as Dirty Paper Coding for the cognitive G-IC) but not with simpler ones (essentially rate splitting and superposition coding as in the Han-Kobayashi scheme) allowed for the classical G-IC.

Another intriguing observation is that these bounds also determine the optimality of "everybody gets half the cake"-DoF result for the K-user G-IC (58; 36). For the K-user G-IC with fixed channel gains it is well known that the DoF are discontinuous at rational channel gains (59). This seems to suggest, at least for $\alpha = 1$, that a gap result up to an outage set is actually *fundamental* and not a consequence of the achievable scheme used. Whether the converse result of (59) for $\alpha = 1$ can be extended to the whole regime $\alpha \in (2/3, 2)$ is an open question. We also note that a constant (not $O\left(\log\left(\frac{\ln(\min(\mathsf{snr},\mathsf{inr}))}{\gamma}\right)\right)$) gap result up to an outage set for the whole regime $\alpha \in (2/3, 2)$ was found in (60); in this case the

TABLE II: PARAMETERS FOR THE MIXED INPUTS IN (**??**), AS USED IN THE PROOF OF THEOREM **??**. NOTATION: FOR $\mathbf{p} = [N_1, N_2, \delta_1, \delta_2]$ WE LET $\mathbf{p}' = [N_2, N_1, \delta_2, \delta_1]$. THE PARAMETER α MEASURES THE LEVEL OF THE INTERFERENCE WHEN $\mathsf{snr} \to \infty$ ACCORDING TO THE PARAMETRIZATION $\mathsf{inr} = \mathsf{snr}^\alpha$.

Regime	Input Parameter $\mathbf{p}$ in (Eq. 4.6)	Gap(bits)
$\mathsf{snr} \geq \mathsf{inr}(1+\mathsf{inr})$, $\alpha \in [0, 1/2]$ (very weak)	$\mathbf{p}_t \cup \mathbf{p}'_t$, for all $t \in [0,1]$; $\mathbf{p}_t := [1, 1, t, 1]$;	constant gap $\mathsf{G}_{\mathsf{d}\text{(Eq. 4.49)}} \leq 1/2$
$\mathsf{snr} < \mathsf{inr}(1+\mathsf{inr})$, $\frac{1+\mathsf{snr}}{1+\mathsf{inr}+\frac{\mathsf{snr}}{1+\mathsf{inr}}} > \frac{1+\mathsf{inr}+\frac{\mathsf{snr}}{1+\mathsf{inr}}}{1+\frac{\mathsf{snr}}{1+\mathsf{inr}}}$, $\alpha \in (1/2, 2/3)$ (moderately weak 2)	$\mathbf{p}_{1,t} \cup \mathbf{p}_{2,t} \cup \mathbf{p}'_{2,t}$, for all $t \in [0,1]$; $\mathbf{p}_{1,t}$: values can be found in (Eq. I.1); $\mathbf{p}_{2,t}$: values can be found in (Eq. I.7);	constant gap $\mathsf{G}_{\mathsf{d}\text{(Eq. 4.47)}} \approx 3.79$
$\mathsf{inr} \leq \mathsf{snr}$, $\frac{1+\mathsf{snr}}{1+\mathsf{inr}+\frac{\mathsf{snr}}{1+\mathsf{inr}}} \leq \frac{1+\mathsf{inr}+\frac{\mathsf{snr}}{1+\mathsf{inr}}}{1+\frac{\mathsf{snr}}{1+\mathsf{inr}}}$, $\alpha \in [2/3, 1]$ (moderately weak 1)	$\mathbf{p}_{1,t} \cup \mathbf{p}_{2,t} \cup \mathbf{p}'_{2,t}$, for all $t \in [0,1]$; $\mathbf{p}_{1,t}$ values can be found in (Eq. 4.36); $\mathbf{p}_{2,t}$ values can be found in (Eq. G.5);	log-log gap $\mathsf{G}_{\mathsf{d}\text{(Eq. 4.40)}}$
$\mathsf{snr} < \mathsf{inr} < \mathsf{snr}(1+\mathsf{snr})$, $\alpha \in (1, 2)$ (strong)	$\mathbf{p}_t$, for all $t \in [0,1]$; $\mathbf{p}_t$: values can be found in (Eq. 4.22);	log-log gap $\mathsf{G}_{\mathsf{d}\text{(Eq. 4.21)}}$
$\mathsf{inr} \geq \mathsf{snr}(1+\mathsf{snr})$, $\alpha \in [2, \infty)$ (very strong)	$\mathbf{p} = [\mathsf{N}_\mathsf{d}(\mathsf{snr}), \mathsf{N}_\mathsf{d}(\mathsf{snr}), 0, 0]$;	constant gap $\mathsf{G}_{\mathsf{d}\text{(Eq. 4.15)}} \approx 1.25$

achievable region was based on a *multi-letter* scheme inspired by compute-and-forward. It is not clear at this point whether *single-letter* schemes, such as out TINnoTS, are fundamentally suboptimal compared to multi-letter ones.

Proof. The parameters of the mixed inputs in (Eq. 4.6) are chosen as indicated in Table II depending on the regime of operation. We now analyze each regime separately.

4.3.1 Very Strong Interference, i.e., $\mathsf{inr} \geq \mathsf{snr}(1+\mathsf{snr})$

Outer Bound

In the very strong interference regime the capacity of the classical G-IC is given by

$$\mathcal{R}_{\text{out}}^{(4.3.1)} = \left\{ \begin{array}{c} 0 \leq R_1 \leq \mathsf{I_g}\left(\mathsf{snr}\right) \\ 0 \leq R_2 \leq \mathsf{I_g}\left(\mathsf{snr}\right) \end{array} \right\}. \tag{4.10}$$

Inner Bound

The capacity of the classical G-IC in this regime is achieved by sending only common messages from Gaussian codebooks; a receiver first decodes the interfering message, strips it from the received signal, and then decodes the intended message in an equivalent interference-free channel. Even though joint decoding is not allowed in our TINnoTS region, we shall see that the discrete part of the input behaves as a common message (as if it could be decoded at the non-intended destination). We therefore do not send the Gaussian portion of the input (as Gaussian inputs treated as noise increase the noise floor of the receiver) and in (Eq. 4.6) we set

$$N_1 = N_2 = N = \mathsf{N_d}\left(\mathsf{snr}\right), \tag{4.11a}$$

$$\delta_1 = \delta_2 = \delta = 0, \tag{4.11b}$$

resulting in

$$S_1 \sim S_2 \sim S, \quad S := \sqrt{\mathsf{snr}} X_{1D} + \sqrt{\mathsf{inr}} X_{2D}, \tag{4.12}$$

for the received constellations in (Eq. 4.8). The number of points and the minimum distance for the constellation S in (Eq. 4.12) can be computed from Proposition 2.1.2 as follows. If we identify $|h_x|^2 = \mathsf{snr}$, $|h_y|^2 = \mathsf{inr}$, $|X| = |Y| = N$, $d^2_{\min(X)} = d^2_{\min(Y)} = \frac{12}{N^2-1}$, then the condition in (Eq. 2.16) reads $N^2\mathsf{snr} \leq \mathsf{inr}$, which is readily verified since $N^2\mathsf{snr} \leq (1+\mathsf{snr})\mathsf{snr}$ by definition of N in (Eq. 4.11a), and $(1+\mathsf{snr})\mathsf{snr} \leq \mathsf{inr}$ by the definition of the very strong interference regime. We therefore have

$$|S| = N^2, \quad \text{with equally likely points,} \tag{4.13}$$

$$\frac{d^2_{\min(S)}}{12} = \min\{\mathsf{snr}, \mathsf{inr}\}\frac{1}{N^2-1} = \frac{\mathsf{snr}}{N^2-1}. \tag{4.14}$$

By plugging these values in Proposition 4.2.1, an achievable rate region is

$$\mathcal{R}_{\text{in}}^{(4.3.1)} = \begin{Bmatrix} 0 \leq R_1 \leq r_0 \\ \\ 0 \leq R_2 \leq r_0 \end{Bmatrix} \text{ such that} \tag{4.15a}$$

$$\begin{aligned} r_0 &\geq \mathsf{I_d}\left(S\right) - \min\left(\log(N), \mathsf{I_g}\left(\mathsf{inr}\right)\right) \\ &\geq \left[\log(N^2) - \frac{1}{2}\log\left(\frac{2\pi \mathrm{e}}{12}\right) - \frac{1}{2}\log\left(1 + \frac{N^2-1}{\mathsf{snr}}\right)\right]^+ - \log(N) \\ &\geq \mathsf{I_g}\left(\mathsf{snr}\right) - \mathsf{G}_{\mathsf{d}(\text{Eq. 4.15})}, \end{aligned} \tag{4.15b}$$

$$\mathsf{G}_{\mathsf{d}(\text{Eq. 4.15})} := \frac{1}{2}\log\left(\frac{2\pi \mathrm{e}}{3}\right) \approx 1.25 \text{ bits}, \tag{4.15c}$$

where the gap in (Eq. 4.15c) is as for the point-to-point Gaussian channel without states in Section 2.1.2.

Gap

It is immediate to see that the achievable region in (Eq. 4.15) and the upper bound in (Eq. 4.10) are at most to within $\mathsf{G}_{\mathsf{d}(\text{Eq. 4.15})}$ bits of one another, where $\mathsf{G}_{\mathsf{d}(\text{Eq. 4.15})}$ is given in (Eq. 4.15c).

4.3.2 Strong (but not Very Strong) Interference, i.e., $\mathsf{snr} < \mathsf{inr} < \mathsf{snr}(1+\mathsf{snr})$

Outer Bound

The capacity region of the G-IC in this regime is

$$\mathcal{R}_{\text{out}}^{(4.3.2)} = \left\{ \begin{array}{l} 0 \le R_1 \le \mathsf{I_g}\left(\mathsf{snr}\right) \\ 0 \le R_2 \le \mathsf{I_g}\left(\mathsf{snr}\right) \\ R_1 + R_2 \le \mathsf{I_g}(\mathsf{snr}+\mathsf{inr}) \end{array} \right\}$$

$$= \bigcup_{t\in[0,1]} \left\{ \begin{array}{ll} 0 \le R_1 & \le \frac{1-t}{2}\log\left(1+\frac{\mathsf{inr}}{1+\mathsf{snr}}\right) + \frac{t}{2}\log\left(1+\mathsf{snr}\right) \\ & =: \mathsf{I_g}\left(\mathsf{snr}_{0,a,t}\right) \\ 0 \le R_2 & \le \frac{1-t}{2}\log\left(1+\mathsf{snr}\right) + \frac{t}{2}\log\left(1+\frac{\mathsf{inr}}{1+\mathsf{snr}}\right) \\ & =: \mathsf{I_g}\left(\mathsf{snr}_{0,b,t}\right) \end{array} \right\}, \tag{4.16}$$

where $t \in [0,1]$ is the time-sharing parameter (i.e., by varying t we obtain all points on the dominant face of the capacity region described by $R_1 + R_2 = \mathsf{I_g}(\mathsf{snr}+\mathsf{inr})$).

Inner Bound

The capacity of the classical G-IC in this regime is achieved by sending only common messages from Gaussian codebooks, and by performing joint decoding of the intended and interfering messages at both receivers. Similarly to the very strong interference regime, we do not send the Gaussian portion of the mixed inputs (i.e., $\delta_1 = \delta_2 = 0$). Differently from the very strong interference regime, here we do not set the number of points of the discrete part of the inputs to be the same for the two users since the corner point of (Eq. 4.16) for a fixed t has $R_1 \neq R_2$. Moreover, we lower bound the minimum distance of the sum-set constellations S_1 and S_2 in (Eq. 4.8) by using Proposition 2.1.3 as follows

$$\frac{d^2_{\min(S_1)}}{12} \geq \kappa^2_{\gamma,N_1,N_2} \min\left(\frac{\mathsf{snr}}{N_1^2-1}, \frac{\mathsf{inr}}{N_2^2-1}, \max\left(\frac{\mathsf{inr}}{N_1^2(N_2^2-1)}, \frac{\mathsf{snr}}{N_2^2(N_1^2-1)}\right)\right), \tag{4.17}$$

$$\frac{d^2_{\min(S_2)}}{12} \geq \kappa^2_{\gamma,N_1,N_2} \min\left(\frac{\mathsf{inr}}{N_1^2-1}, \frac{\mathsf{snr}}{N_2^2-1}, \max\left(\frac{\mathsf{inr}}{N_1^2(N_2^2-1)}, \frac{\mathsf{snr}}{N_2^2(N_1^2-1)}\right)\right), \tag{4.18}$$

$$\kappa_{\gamma,N_1,N_2} := \frac{\gamma/2}{1+\ln(\max(N_1,N_2))}, \tag{4.19}$$

where the minimum distance lower bounds in (Eq. 4.17) and (Eq. 4.18) hold for all channel gains up to an outage set of Lebesgue measure less than γ for any $\gamma \in (0,1]$.

By combining the bounds in (Eq. 4.17) and (Eq. 4.18) we obtain

$$\begin{aligned}\min_{i\in[1:2]} \frac{d^2_{\min(S_i)}}{12} &\geq \kappa^2_{\gamma,N_1,N_2} \min\left(\frac{\min(\mathsf{snr},\mathsf{inr})}{\max(N_1^2,N_2^2)-1}, \frac{\max(\mathsf{snr},\mathsf{inr})}{N_1^2N_2^2-1}\right) \\ &\overset{\text{for } \mathsf{snr}\leq\mathsf{inr}}{=} \kappa^2_{\gamma,N_1,N_2} \min\left(\frac{\mathsf{snr}}{\max(N_1^2,N_2^2)-1}, \frac{\mathsf{inr}}{N_1^2N_2^2-1}\right).\end{aligned} \tag{4.20}$$

With (Eq. 4.20), it can be easily seen that the achievable region in Proposition 4.2.1 can be written as the union over all (N_1, N_2) of the region

$$\mathcal{R}_{\text{in}}^{(4.3.2)}\left([N_1, N_2, 0, 0]\right) = \left\{ \begin{array}{c} 0 \le R_1 \le r_1 \\ \\ 0 \le R_2 \le r_2 \end{array} \right\} \text{ such that} \tag{4.21a}$$

$$r_1 \ge \mathsf{I_d}\left(S_1\right) - \min\Big(\log(N_2), \mathsf{I_g}\left(\mathsf{inr}\right)\Big) \ge \log(N_1) + \log(2) - \mathsf{G}_{\mathsf{d}(\text{Eq. } 4.21)}, \tag{4.21b}$$

$$r_2 \ge \mathsf{I_d}\left(S_2\right) - \min\Big(\log(N_1), \mathsf{I_g}\left(\mathsf{inr}\right)\Big) \ge \log(N_2) + \log(2) - \mathsf{G}_{\mathsf{d}(\text{Eq. } 4.21)}, \tag{4.21c}$$

$$\begin{aligned} \mathsf{G}_{\mathsf{d}(\text{Eq. } 4.21)} &\le \log(2) + \frac{1}{2}\log\left(\frac{2\pi \mathrm{e}}{12}\right) \\ &+ \frac{1}{2}\log\left(1 + \frac{1}{\kappa^2_{\gamma,N_1,N_2}}\max\left(\frac{\max(N_1^2, N_2^2) - 1}{\mathsf{snr}}, \frac{N_1^2 N_2^2 - 1}{\mathsf{inr}}\right)\right), \end{aligned} \tag{4.21d}$$

where the expression for $\mathsf{G}_{\mathsf{d}(\text{Eq. } 4.21)}$ comes from the minimum distance expression in (Eq. 4.20).

We next need to pick N_1 and N_2 in (Eq. 4.21). Our choice is guided by the expression of the 'compound MAC' capacity region in this regime given by (Eq. 4.16). In our TINnoTS region, time-sharing is not allowed, but varying the number of points of the discrete constellations is; we therefore mimic time-sharing in (Eq. 4.16) by choosing as number of points in the discrete part of the mixed inputs as follows: for some fixed t we let

$$N_1 = \mathsf{N_d}\left(\mathsf{snr}_{0,a,t}\right), \ \mathsf{snr}_{0,a,t} := \left(1 + \frac{\mathsf{inr}}{1 + \mathsf{snr}}\right)^{1-t} (1 + \mathsf{snr})^t - 1, \tag{4.22a}$$

$$N_2 = \mathsf{N_d}\left(\mathsf{snr}_{0,b,t}\right), \ \mathsf{snr}_{0,b,t} := \left(1 + \frac{\mathsf{inr}}{1 + \mathsf{snr}}\right)^{t} (1 + \mathsf{snr})^{1-t} - 1. \tag{4.22b}$$

The whole TINnoTS achievable region is obtained by taking union over $t \in [0,1]$ of the region in (Eq. 4.21) with the number of points as in (Eq. 4.22).

Gap

Since $\frac{\mathsf{inr}}{1+\mathsf{snr}} \leq \mathsf{snr} \leq \mathsf{inr}$ by the definition of the strong interference regime, we immediately have that in (Eq. 4.22) the equivalent SNRs satisfy $\max(\mathsf{snr}_{0,a,t}, \mathsf{snr}_{0,b,t}) \leq \mathsf{snr}$ for all $t \in [0,1]$. Thus, for the minimum distance expression in (Eq. 4.20), we have

$$\max(N_1^2, N_2^2) - 1 \leq \max(\mathsf{snr}_{0,a,t}, \mathsf{snr}_{0,b,t}) \leq \mathsf{snr} = \min(\mathsf{snr}, \mathsf{inr}), \tag{4.23}$$

$$N_1^2 N_2^2 - 1 \leq (\mathsf{snr}_{0,a,t} + 1)(\mathsf{snr}_{0,b,t} + 1) - 1 = \mathsf{snr} + \mathsf{inr} \leq 2\mathsf{inr}. \tag{4.24}$$

Finally, since $\mathsf{I_g}(x) \leq \log(\mathsf{N_d}(x)) + \log(2)$, the inner bound in (Eq. 4.21) is at most $\mathsf{G}_{\mathsf{d}(\text{Eq. 4.21})}$ bits from the outer bound in (Eq. 4.16), uniformly over all $t \in [0,1]$, where $\mathsf{G}_{\mathsf{d}(\text{Eq. 4.21})}$ in (Eq. 4.21d) can be further upper bounded thanks to (Eq. 4.23)-(Eq. 4.24) as

$$\begin{aligned} \mathsf{G}_{\mathsf{d}(\text{Eq. 4.21})} &\leq \frac{1}{2}\log\left(\frac{2\pi e}{3}\left(1 + \frac{\max(1,2)}{\kappa^2_{\gamma,N_1,N_2}}\right)\right) \\ &\leq \frac{1}{2}\log\left(\frac{2\pi e}{3}\left(1 + 8 \cdot \frac{\left(1 + 1/2\ln(1+\min(\mathsf{snr},\mathsf{inr}))\right)^2}{\gamma^2}\right)\right) \text{ bits,} \end{aligned} \tag{4.25}$$

where γ is the Lebesgue measure of the outage set over which the lower bounds on the minimum distance in (Eq. 4.20) does not apply. Recall that γ is a tunable parameter that represents a tradeoff between gap and set of channel gains for which the gap result holds, i.e., by increasing the measure of the outage set we can reduce the gap, and vice-versa. A similar behavior was pointed out already in (60).

Remark 10. *Note that, had we been able to use Proposition 2.1.2 instead of Proposition 2.1.3 to bound the minimum distance of the received constellations, we would have obtained a constant gap result instead of a* $O\left(\log\left(\frac{\ln(\min(\mathsf{snr},\mathsf{inr}))}{\gamma}\right)\right)$ *gap result. It turns out that in this regime the condition of Proposition 2.1.2 is not satisfied – the proof is very tedious and is not reported here for sake of space.*

4.3.3 Moderately Weak Interference, i.e. $\mathsf{inr} \leq \mathsf{snr} \leq (1+\mathsf{inr})\mathsf{inr}$: General Setup

The weak interference regime is notoriously more involved to analyze than the other regimes. In this subsection we aim to derive a general framework to deal with the weak (but not very weak) interference regime. Before we move into the gap derivation for this regime, let us summarize the key trick we developed in the strong interference regime to obtain a capacity result to within a gap: write the closure of the capacity outer bound in parametric form so as to get insight on how to choose the number of points of the discrete part of the mixed inputs. In the weak interference regime we will follow the same approach but the computations will be more involved because the capacity region outer bound in weak interference has three dominant faces (and not just one dominant face as in strong interference).

Outer Bound

In this regime, we express the upper bound in Proposition 4.2.2 as the convex closure of its corner points, that is

$$\mathcal{R}_{\text{out}}^{(4.3.3)} = \mathsf{co}\Big\{ (R_{1A}, R_{2A}) := (\mathsf{I_g}(\mathsf{snr}), c)\,, \tag{4.26a}$$

$$(R_{1B}, R_{2B}) := (b-a, 2a-b)\,, \tag{4.26b}$$

$$(R_{1C}, R_{2C}) := (2a-b, b-a)\,, \tag{4.26c}$$

$$(R_{1D}, R_{2D}) := (c, \mathsf{I_g}(\mathsf{snr}))\ \Big\}, \tag{4.26d}$$

where

$$a := \min\left(\mathsf{I_g}(\mathsf{inr}+\mathsf{snr}) + \mathsf{I_g}\left(\mathsf{snr}\right) - \mathsf{I_g}\left(\mathsf{inr}\right), 2\,\mathsf{I_g}\left(\mathsf{inr} + \frac{\mathsf{snr}}{1+\mathsf{inr}}\right)\right), \tag{4.26e}$$

$$b := \mathsf{I_g}\left(\mathsf{inr} + \frac{\mathsf{snr}}{1+\mathsf{inr}}\right) + \mathsf{I_g}(\mathsf{snr}+\mathsf{inr}) + \mathsf{I_g}\left(\mathsf{snr}\right) - \mathsf{I_g}\left(\mathsf{inr}\right), \tag{4.26f}$$

$$\begin{aligned} c &:= \mathsf{I_g}(\mathsf{inr}+\mathsf{snr}) - \mathsf{I_g}\left(\mathsf{snr}\right) + \mathsf{I_g}\left(\mathsf{inr} + \frac{\mathsf{snr}}{1+\mathsf{inr}}\right) - \mathsf{I_g}\left(\mathsf{inr}\right) \\ &= \mathsf{I_g}\left(\frac{\mathsf{inr}}{1+\mathsf{snr}}\right) + \mathsf{I_g}\left(\frac{\mathsf{snr}}{(1+\mathsf{inr})^2}\right) \\ &\leq \mathsf{I_g}\left(\frac{\mathsf{snr}}{1+\mathsf{snr}}\right) + \mathsf{I_g}\left(\frac{\mathsf{inr}}{1+\mathsf{inr}}\right) \leq \log(2). \end{aligned} \tag{4.26g}$$

Under the constraint $\frac{\mathsf{snr}}{1+\mathsf{inr}} \leq \mathsf{inr}$ it can be verified numerically that actually $c \leq 0.5537$ bits (rather than $c \leq 1$ bit) attained for $\mathsf{inr} = \sqrt{3}+1$; however, for notational convenience we will use in the following $c \leq 1$ bit.

An explicit expression for $\mathcal{R}_{\text{out}}^{(4.3.3)}$ obtained by time-sharing between the corner points in (Eq. 4.26) is

$$\mathcal{R}_{\text{out}}^{(4.3.3)} = \mathcal{R}_{2R_1+R_2}^{(4.3.3)} \cup \mathcal{R}_{R_1+R_2}^{(4.3.3)} \cup \mathcal{R}_{R_1+2R_2}^{(4.3.3)}, \quad \text{where} \tag{4.27a}$$

$$\mathcal{R}_{2R_1+R_2}^{(4.3.3)} = \bigcup_{t\in[0,1]} \left\{ \begin{array}{c} R_1 \leq tR_{1A} + (1-t)R_{1B} \\ \\ R_2 \leq tR_{2A} + (1-t)R_{2B} \end{array} \right\}, \tag{4.27b}$$

$$\mathcal{R}_{R_1+R_2}^{(4.3.3)} = \bigcup_{t\in[0,1]} \left\{ \begin{array}{c} R_1 \leq tR_{1B} + (1-t)R_{1C} \\ \\ R_2 \leq tR_{2B} + (1-t)R_{2C} \end{array} \right\}, \tag{4.27c}$$

$$\mathcal{R}_{R_1+2R_2}^{(4.3.3)} = \bigcup_{t\in[0,1]} \left\{ \begin{array}{c} R_1 \leq tR_{1C} + (1-t)R_{1D} \\ \\ R_2 \leq tR_{2C} + (1-t)R_{2D} \end{array} \right\}. \tag{4.27d}$$

Because the sum-rate upper bound in (Eq. 4.26) is in the form

$$R_1 + R_2 \leq \text{eq.(Eq. 4.26e)} = \min(\text{eq.(Eq. 4.9d)}, \text{eq.(Eq. 4.9e)}),$$

we will distinguish between two cases: when the constraint in (Eq. 4.9d) is active, referred to as *Weak1*, and when the constraint in (Eq. 4.9e) is active, referred to as *Weak2*, that is, within $\mathsf{inr} \leq \mathsf{snr} \leq \mathsf{inr}(1+\mathsf{inr})$ we further distinguish between

$$\text{Weak1:} \quad \frac{1+\mathsf{snr}}{1+\mathsf{inr}+\frac{\mathsf{snr}}{1+\mathsf{inr}}} \leq \frac{1+\mathsf{inr}+\frac{\mathsf{snr}}{1+\mathsf{inr}}}{1+\frac{\mathsf{snr}}{1+\mathsf{inr}}}, \tag{4.28}$$

$$\text{Weak2:} \quad \frac{1+\mathsf{snr}}{1+\mathsf{inr}+\frac{\mathsf{snr}}{1+\mathsf{inr}}} > \frac{1+\mathsf{inr}+\frac{\mathsf{snr}}{1+\mathsf{inr}}}{1+\frac{\mathsf{snr}}{1+\mathsf{inr}}}. \tag{4.29}$$

Inner Bound

For the G-IC in weak interference the best know strategy is to send common and private messages from Gaussian codebooks, and for each of the receivers to jointly decode both common messages and the desired private message while treating the private message of the interferer as noise. Unlike in the strong and very strong interference regimes, in this case we will use the Gaussian portion of the mixed inputs by setting δ_1 and δ_2 to be non-zero. Moreover, we will vary (δ_1, δ_2) jointly with (N_1, N_2) to mimic time sharing and power control.

In this regime, we further simplify the achievable rate region in (Eq. 4.7) from Proposition 4.2.1 as follows

$$\mathcal{R}_{\text{in}}^{(4.3.3)} = \bigcup_{\substack{[N_1, N_2, \delta_1, \delta_2] \in \mathbb{N}^2 \times [0,1]^2 : \\ \max(\delta_1, \delta_2) \leq \frac{1}{1+\mathsf{inr}}}} \mathcal{R}_{\text{in}}^{(4.3.3)}\left([N_1, N_2, \delta_1, \delta_2]\right), \quad \text{where}$$

$$\mathcal{R}_{\text{in}}^{(4.3.3)}\left([N_1, N_2, \delta_1, \delta_2]\right) := \left\{ \begin{array}{l} 0 \leq R_1 \leq \log(N_1) + \mathsf{I_g}\left(\mathsf{snr}\delta_1\right) - \Delta_{\text{(Eq. 4.30)}} \\ 0 \leq R_2 \leq \log(N_2) + \mathsf{I_g}\left(\mathsf{snr}\delta_2\right) - \Delta_{\text{(Eq. 4.30)}} \\ \Delta_{\text{(Eq. 4.30)}} = \frac{1}{2}\log\left(\frac{\pi \mathrm{e}}{3}\right) + \frac{1}{2}\log\left(1 + \frac{12}{\min_{i \in [1:2]} d^2_{\min(S_i)}}\right) \end{array} \right\}, \quad (4.30)$$

where the received constellations S_1 and S_2 are given in (Eq. 4.8). Note that, inspired by (26), we restricted the power splits between the continuous and discrete parts of the mixed inputs to satisfy $\max(\delta_1, \delta_2) \leq \frac{1}{1+\mathsf{inr}}$. The simplified form of the TINnoTS region with mixed inputs in (Eq. 4.30) is obtained from (Eq. 4.7) as follows. For the achievable rate R_1 we have

$$\begin{aligned} R_1 &\geq \mathsf{I_d}\left(S_1\right) + \mathsf{I_g}\left(\frac{\mathsf{snr}\delta_1}{1+\mathsf{inr}\delta_2}\right) - \min\left(\log(N_2), \mathsf{I_g}\left(\frac{\mathsf{inr}(1-\delta_2)}{1+\mathsf{inr}\delta_2}\right)\right) \\ &\overset{(a)}{=} \left[\log(N_1 N_2) - \frac{1}{2}\log\left(\frac{2\pi \mathrm{e}}{12}\right) - \frac{1}{2}\log\left(1 + \frac{12}{d^2_{\min(S_1)}}\right)\right]^+ \\ &\quad + \mathsf{I_g}\left(\frac{\mathsf{snr}\delta_1}{1+\mathsf{inr}\delta_2}\right) - \min\left(\log(N_2), \mathsf{I_g}\left(\frac{\mathsf{inr}(1-\delta_2)}{1+\mathsf{inr}\delta_2}\right)\right) \\ &\overset{(b)}{\geq} \log(N_1 N_2) - \frac{1}{2}\log\left(\frac{2\pi \mathrm{e}}{12}\right) - \frac{1}{2}\log\left(1 + \frac{12}{d^2_{\min(S_1)}}\right) + \mathsf{I_g}\left(\frac{\mathsf{snr}\delta_1}{1+\mathsf{inr}\delta_2}\right) - \log(N_2) \\ &\overset{(c)}{\geq} \log(N_1) + \mathsf{I_g}\left(\frac{\mathsf{snr}\delta_1}{2}\right) + \frac{1}{2}\log(2) - \Delta_{\text{(Eq. 4.30)}} \\ &\overset{(d)}{\geq} \log(N_1) + \mathsf{I_g}\left(\mathsf{snr}\delta_1\right) - \Delta_{\text{(Eq. 4.30)}}, \end{aligned}$$

where the (in)-equalities are due to: (a) because regardless of whether we use Proposition 2.1.2 or Proposition 2.1.3 to compute the minimum distance for the received sum-set constellations S_1 and S_2 in (Eq. 4.8), these constellations always comprise $|S_1| = |S_2| = N_1N_2$ equally likely points either exactly or almost surely; (b) because $[x]^+ \geq x$ and $\min(x, y) \leq x$; (c) because we imposed $\max(\delta_1, \delta_2) \leq \frac{1}{1+\mathsf{inr}}$ and by definition of $\Delta_{\text{(Eq. 4.30)}}$ in (Eq. 4.30); and (d) because $\log(1 + x/2) \geq \log(1 + x) - \log(2)$. The rate expression for user 2 follows similarly.

For the evaluation of $\Delta_{\text{(Eq. 4.30)}}$, the minimum distance of the received constellations S_1 and S_2 defined in (Eq. 4.8) will be computed with either Proposition 2.1.2 or Proposition 2.1.3. By using Proposition 2.1.3, which is valid for all channel gains up to a set of controllable Lebesgue measure less than γ, for any $\gamma > 0$, we have

$$\frac{d^2_{\min(S_1)}}{12} \geq \kappa^2_{\gamma,N_1,N_2} \frac{\min\left(\frac{(1-\delta_1)\mathsf{snr}}{N_1^2-1}, \frac{(1-\delta_2)\mathsf{inr}}{N_2^2-1}, \max\left(\frac{(1-\delta_2)\mathsf{inr}}{N_1^2(N_2^2-1)}, \frac{(1-\delta_1)\mathsf{snr}}{N_2^2(N_1^2-1)}\right)\right)}{1 + \mathsf{snr}\delta_1 + \mathsf{inr}\delta_2} \tag{4.31a}$$

$$\geq \kappa^2_{\gamma,N_1,N_2} \frac{1 - \max(\delta_1, \delta_2)}{1 + \mathsf{snr}\delta_1 + \mathsf{inr}\delta_2} \min\left(\frac{\mathsf{snr}}{N_1^2 - 1}, \frac{\mathsf{inr}}{N_2^2 - 1}, \frac{\max(\mathsf{snr}, \mathsf{inr})}{N_1^2N_2^2 - 1}\right) \tag{4.31b}$$

$$\overset{\text{for } \mathsf{inr} \leq \mathsf{snr}}{=} \kappa^2_{\gamma,N_1,N_2} \frac{1 - \max(\delta_1, \delta_2)}{1 + \mathsf{snr}\delta_1 + \mathsf{inr}\delta_2} \min\left(\frac{\mathsf{inr}}{N_2^2 - 1}, \frac{\mathsf{snr}}{N_1^2N_2^2 - 1}\right), \tag{4.31c}$$

$$\frac{d^2_{\min(S_2)}}{12} \geq \kappa^2_{\gamma,N_1,N_2} \frac{1 - \max(\delta_1, \delta_2)}{1 + \mathsf{snr}\delta_2 + \mathsf{inr}\delta_1} \min\left(\frac{\mathsf{inr}}{N_1^2 - 1}, \frac{\mathsf{snr}}{N_1^2N_2^2 - 1}\right), \tag{4.31d}$$

$$\kappa_{\gamma,N_1,N_2} = \frac{\gamma/2}{1 + 1/2\ln(\max(N_1^2, N_2^2))}. \tag{4.31e}$$

If instead we use Proposition 2.1.2 we have

$$\min_{i\in[1:2]} \frac{d^2_{\min(S_i)}}{12} = \min_{(i,i')\in\{(1,2),(2,1)\}} \frac{1}{1 + \mathsf{snr}\delta_i + \mathsf{inr}\delta_{i'}} \min\left(\frac{(1 - \delta_i)\mathsf{snr}}{N_i^2 - 1}, \frac{(1 - \delta_{i'})\mathsf{inr}}{N_{i'}^2 - 1}\right), \tag{4.32a}$$

which holds if

$$\mathsf{inr}(1-\delta_{i'})\frac{N_{i'}^2}{N_{i'}^2-1} \le \frac{\mathsf{snr}(1-\delta_i)}{N_i^2-1} \quad \forall (i,i') \in \{(1,2),(2,1)\}. \tag{4.32b}$$

We observe that in (Eq. 4.30) each achievable rate is bounded by the sum of two terms: one that depends on the number of points of the discrete part of the mixed inputs, and the other that depends on the continuous part of the mixed inputs through the power splits. This is reminiscent of rate-splitting in the Han-Kobayashi achievable scheme, where each rate is written as the sum of the common-message rate and the private-message rate. The simplified Han-Kobayashi achievable region in (26) is known to achieve the outer bound in Proposition 4.2.2 to within 1/2 bit; however, to the best of our knowledge, it is not known how much information should be conveyed through the private messages and how much through the common messages for a general rate-pair (R_1, R_2) on the convex closure of the outer bound in Proposition 4.2.2 and for a general set of channel parameters. Next we will identify the (to within 1/2 bit) optimal rate splits and use the found analytical closed-form expressions for the common-message and private-message rates to come up with an educated guess for the values of the parameters of our mixed inputs.

Let $R_u = R_{u,p} + R_{u,c}$, where $R_{u,p}$ is the rate of the private message and $R_{u,c}$ is the rate of the common message for user $u \in [1:2]$. From the analysis of the symmetric LDA in (23, Lemma 4), which gives the optimal gDoF region for the symmetric G-IC before Fourier-Motzkin elimination, it is not difficult to see that it is always optimal to set

$$R_{u,p} \approxeq \min\left(\mathsf{I_g}\left(\frac{\mathsf{snr}}{1+\mathsf{inr}}\right), \frac{R_u}{2}\right),\ u \in [1:2], \tag{4.33}$$

where with $\approxeq$ we mean equality up to an additive term that grows slower than $\log(\mathsf{snr})$ when $\mathsf{snr} \to \infty$. We found that, with the exception of the sum-capacity for $\alpha \in (1/2, 2/3)$, the optimal 'rate splits' are unique and are given by (Eq. 4.33). These 'rate splits' shed light on the interplay between private and common messages, which was not immediately obvious from the outer bound in (Eq. 4.9).

In the following it will turn out to be convenient to think of the discrete part of a mixed input (contributing to the rate with the term $\log(N_i), i \in [1:2]$) as a 'common message' and of the continuous part of a mixed input (contributing to the rate with the term $\mathsf{I_g}(S\delta_i), i \in [1:2]$) as a 'private message'. We shall refer to this 'mapping' of our TINnoTS scheme to the Han-Kobayashi scheme as the *discrete→common map*. Note that there is a fundamental difference between a common message in the Han-Kobayashi achievable scheme and the discrete part of the mixed input in our scheme. In our scheme the interfering signal is treated as noise while in Han-Kobayashi achievable scheme the common message is jointly decoded, albeit non-uniquely, with the intended signals at the non-intended receiver. The *discrete→common map* is thus just intended to provide an educated guess on how to pick the parameters of our mixed input in the following analysis. We do not claim here that the *discrete→common map* is the only possible way to 'match' our TINnoTS scheme to the Han-Kobayashi scheme. In fact, we will give an example later on where with the proposed *discrete→common map* we obtain a $O\left(\log\left(\frac{\ln(\min(\mathsf{snr},\mathsf{inr}))}{\gamma}\right)\right)$ gap, but with a *discrete→private map* we obtain a constant gap. Although finding the smallest possible gap in each regime would be desirable, here for sake of simplicity we consistently use the *discrete→common map*.

With the inner and outer bounds defined, as well as the 'rate splits', we are ready to determine an optimal (to within a gap) choice of parameters for the mixed inputs in the weak interference regime. Next, we will focus on the regime in (Eq. 4.28) and the regime in (Eq. 4.29) separately and for each regime we will match each point on the closure of the outer bound in (Eq. 4.26) with an achievable region as in (Eq. 4.30).

4.3.4 Moderately Weak Interference, subregime Weak1

The regime of interest here is the subset of $\mathsf{inr} \leq \mathsf{snr} \leq \mathsf{inr}(1+\mathsf{inr})$ for which (Eq. 4.28) holds. For convenience, we analyze the regime $\mathsf{inr} \leq \mathsf{snr} \leq 1+\mathsf{inr}$ in Appendix F and focus next on the subset of $(1+\mathsf{inr}) \leq \mathsf{snr} \leq \mathsf{inr}(1+\mathsf{inr})$ for which (Eq. 4.28) holds. The condition $1+\mathsf{inr} \leq \mathsf{snr}$ allows us to state $\frac{1+\mathsf{snr}}{1+\mathsf{inr}+\frac{\mathsf{snr}}{1+\mathsf{inr}}} \geq 1$ in the following.

Outer Bound Corner Points and Rate Splits

Whenever the condition in (Eq. 4.28) holds, the outer bound in (Eq. 4.9) is given by all the constraints in (Eq. 4.9) except for the one in (Eq. 4.9e) – in the symmetric case the constraints in (Eq. 4.9c) and (Eq. 4.9d) are the same. The corner points for the outer bound region in (Eq. 4.27) are thus

$$\text{eq.(Eq. 4.9a)} = \text{eq.(Eq. 4.9f)} \Rightarrow (R_{1A}, R_{2A}) = (\mathsf{I_g}(\mathsf{snr}), \tag{4.34a}$$

$$\mathsf{I_g}\left(\frac{\mathsf{snr}}{1+\mathsf{inr}}\right) + \mathsf{I_g}\left(\mathsf{inr} + \frac{\mathsf{snr}}{1+\mathsf{inr}}\right) - \mathsf{I_g}(\mathsf{snr})\bigg)\,; \tag{4.34b}$$

$$\text{eq.(Eq. 4.9f)} = \text{eq.(Eq. 4.9c)} \Rightarrow (R_{1B}, R_{2B}) = \left(\mathsf{I_g}\left(\mathsf{inr} + \frac{\mathsf{snr}}{1+\mathsf{inr}}\right)\right., \tag{4.34c}$$

$$\left.\mathsf{I_g}(\mathsf{snr}) + \mathsf{I_g}\left(\frac{\mathsf{snr}}{1+\mathsf{inr}}\right) - \mathsf{I_g}\left(\mathsf{inr} + \frac{\mathsf{snr}}{1+\mathsf{inr}}\right)\right)\,; \tag{4.34d}$$

$$\text{eq.(Eq. 4.9f)} = \text{eq.(Eq. 4.9c)} \Rightarrow (R_{1C}, R_{2C}) = \left(\mathsf{I_g}(\mathsf{snr}) + \mathsf{I_g}\left(\frac{\mathsf{snr}}{1+\mathsf{inr}}\right) - \mathsf{I_g}\left(\mathsf{inr} + \frac{\mathsf{snr}}{1+\mathsf{inr}}\right)\right., \tag{4.34e}$$

$$\left.\mathsf{I_g}\left(\mathsf{inr} + \frac{\mathsf{snr}}{1+\mathsf{inr}}\right)\right)\,; \tag{4.34f}$$

$$\text{eq.(Eq. 4.9b)} = \text{eq.(Eq. 4.9g)} \Rightarrow (R_{1D}, R_{2D}) = \left(\mathsf{I_g}\left(\frac{\mathsf{snr}}{1+\mathsf{inr}}\right) + \mathsf{I_g}\left(\mathsf{inr} + \frac{\mathsf{snr}}{1+\mathsf{inr}}\right) - \mathsf{I_g}(\mathsf{snr}),\right. \tag{4.34g}$$

$$\left.\mathsf{I_g}(\mathsf{snr})\right). \tag{4.34h}$$

As explained before, inspired by the proposed *discrete→common map*, we choose to 'split' the rates as:

1. for the sum-rate face / region $\mathcal{R}_{R_1+R_2}$: we set $R_{1,p} = R_{2,p} \approx \mathsf{I_g}\left(\frac{\mathsf{snr}}{1+\mathsf{inr}}\right)$.
2. for the other dominant face / region $\mathcal{R}_{2R_1+R_2}$: we set $R_{1p} \approx \mathsf{I_g}\left(\frac{\mathsf{snr}}{1+\mathsf{inr}}\right)$ and $R_{2p} \approx \frac{R_2}{2}$;
3. we will not explicitly consider the remaining dominant face / region $\mathcal{R}_{R_1+2R_2}$ because a gap result can be obtained by proceeding as for $\mathcal{R}_{2R_1+R_2}$ but with the role of the users swapped.

Outer Bound $\mathcal{R}_{R_1+R_2}$

With the corner point expressions in (Eq. 4.34) we write the outer bound sum-rate face in (Eq. 4.27c) as

$$\mathcal{R}^{(4.3.4)}_{R_1+R_2} = \bigcup_{t\in[0,1]} \left\{ \begin{array}{l} R_1 \leq \frac{t}{2}\log\left(\frac{1+\mathsf{inr}+\frac{\mathsf{snr}}{1+\mathsf{inr}}}{1+\frac{\mathsf{snr}}{1+\mathsf{inr}}}\right) + \frac{1-t}{2}\log\left(\frac{1+\mathsf{snr}}{1+\mathsf{inr}+\frac{\mathsf{snr}}{1+\mathsf{inr}}}\right) + \frac{1}{2}\log\left(1+\frac{\mathsf{snr}}{1+\mathsf{inr}}\right) \\ \qquad =: \mathsf{I_g}(\mathsf{snr}_{1,a,t}) + \mathsf{I_g}\left(\frac{\mathsf{snr}}{1+\mathsf{inr}}\right) \\ R_2 \leq \frac{1-t}{2}\log\left(\frac{1+\mathsf{inr}+\frac{\mathsf{snr}}{1+\mathsf{inr}}}{1+\frac{\mathsf{snr}}{1+\mathsf{inr}}}\right) + \frac{t}{2}\log\left(\frac{1+\mathsf{snr}}{1+\mathsf{inr}+\frac{\mathsf{snr}}{1+\mathsf{inr}}}\right) + \frac{1}{2}\log\left(1+\frac{\mathsf{snr}}{1+\mathsf{inr}}\right) \\ \qquad =: \mathsf{I_g}(\mathsf{snr}_{1,b,t}) + \mathsf{I_g}\left(\frac{\mathsf{snr}}{1+\mathsf{inr}}\right) \end{array} \right\}. \tag{4.35}$$

Inner Bound for $\mathcal{R}_{R_1+R_2}$

In order to approximately achieve the points in (Eq. 4.35), we pick

$$N_1 = \mathsf{N_d}\left(\mathsf{snr}_{1,a,t}\right), \ \mathsf{snr}_{1,a,t} := \left(\frac{1+\mathsf{inr}+\frac{\mathsf{snr}}{1+\mathsf{inr}}}{1+\frac{\mathsf{snr}}{1+\mathsf{inr}}}\right)^{t}\left(\frac{1+\mathsf{snr}}{1+\mathsf{inr}+\frac{\mathsf{snr}}{1+\mathsf{inr}}}\right)^{1-t} - 1, \tag{4.36a}$$

$$N_2 = \mathsf{N_d}\left(\mathsf{snr}_{1,b,t}\right), \ \mathsf{snr}_{1,b,t} := \left(\frac{1+\mathsf{inr}+\frac{\mathsf{snr}}{1+\mathsf{inr}}}{1+\frac{\mathsf{snr}}{1+\mathsf{inr}}}\right)^{1-t}\left(\frac{1+\mathsf{snr}}{1+\mathsf{inr}+\frac{\mathsf{snr}}{1+\mathsf{inr}}}\right)^{t} - 1, \tag{4.36b}$$

$$\delta_1 = \frac{1}{1+\mathsf{inr}}, \tag{4.36c}$$

$$\delta_2 = \frac{1}{1+\mathsf{inr}}. \tag{4.36d}$$

Gap for $\mathcal{R}_{R_1+R_2}$

The gap between the outer bound region in (Eq. 4.35) and the achievable rate region in (Eq. 4.30) with the parameters as in (Eq. 4.36) is

$$\begin{aligned}\Delta_{R_1} &= \mathsf{I_g}(\mathsf{snr}_{1,a,t}) + \mathsf{I_g}\left(\frac{\mathsf{snr}}{1+\mathsf{inr}}\right) - \log(\mathsf{N_d}(\mathsf{snr}_{1,a,t})) - \mathsf{I_g}\left(\frac{\mathsf{snr}}{1+\mathsf{inr}}\right) + \Delta_{\text{(Eq. 4.30)}} \\ &\leq \log(2) + \Delta_{\text{(Eq. 4.30)}},\end{aligned}$$

where the term $\log(2)$ is the "integrality gap" $\log(\mathsf{N_d}(x)) + \log(2) \geq \mathsf{I_g}(x)$; similarly, we have

$$\Delta_{R_2} \leq \log(2) + \Delta_{\text{(Eq. 4.30)}}.$$

We are thus left with bounding $\Delta_{\text{(Eq. 4.30)}}$ in (Eq. 4.30), which is related to the minimum distance of the received constellations S_1 and S_2 defined in (Eq. 4.8). In Appendix G.0.1 we show that

$$\min_{i\in[1:2]} \frac{d^2_{\min(S_i)}}{12} \geq \kappa^2_{\gamma,N_1,N_2} \cdot \frac{3}{8}, \tag{4.37}$$

where κ_{γ,N_1,N_2} is given in (Eq. 4.31e), and $\max(N_1^2, N_2^2) - 1 \leq \mathsf{inr} = \min(\mathsf{snr}, \mathsf{inr})$. With this, the gap for this face is bounded by

$$\begin{aligned}
\mathsf{G}_{\mathsf{d}\text{(Eq. 4.38)}} &\leq \max(\Delta_{R_1}, \Delta_{R_2}) = \log(2) + \Delta_{\text{(Eq. 4.30)}} \\
&\leq \frac{1}{2}\log\left(\frac{4\pi\mathrm{e}}{3}\right) + \frac{1}{2}\log\left(1 + \frac{8}{3}\cdot\frac{1}{\kappa^2_{\gamma,N_1,N_2}}\right) \\
&\leq \frac{1}{2}\log\left(\frac{4\pi\mathrm{e}}{3}\right) + \frac{1}{2}\log\left(1 + \frac{32}{3}\cdot\frac{(1+1/2\ln(1+\min(\mathsf{snr},\mathsf{inr})))^2}{\gamma^2}\right) \text{ bits.}
\end{aligned} \tag{4.38}$$

Gap for $\mathcal{R}_{2R_1+R_2}$

The derivation of the gap and other results for $\mathcal{R}_{2R_1+R_2}$ are delegated to Appendix G.0.2. The resulting gap is

$$\mathsf{G}_{\mathsf{d}\text{(Eq. 4.39)}} \leq \frac{1}{2}\log\left(\frac{16\pi\mathrm{e}}{3}\right) + \frac{1}{2}\log\left(1 + 45\cdot\frac{(1+1/2\ln(1+\min(\mathsf{snr},\mathsf{inr})))^2}{\gamma^2}\right) \text{ bits.} \tag{4.39}$$

Overall Gap for Weak1

To conclude the proof for this sub-regime, the gap is the maximum between the gaps of the different faces and is given by

$$\mathsf{G}_{\mathsf{d}(\text{Eq. 4.40})} \leq \max\left(\mathsf{G}_{\mathsf{d}(\text{Eq. 4.38})}, \mathsf{G}_{\mathsf{d}(\text{Eq. 4.39})}\right) = \mathsf{G}_{\mathsf{d}(\text{Eq. 4.39})}. \tag{4.40}$$

4.3.5 Moderately Weak Interference, subregime Weak2

We focus here on the subset of $\mathsf{inr} \leq \mathsf{snr} \leq \mathsf{inr}(1 + \mathsf{inr})$ for which (Eq. 4.29) holds.

Outer Bound Corner Points and Rate Splits

Under the condition in (Eq. 4.29), the outer bound in (Eq. 4.9) is given by all the constraints except for the ones in (Eq. 4.9c) and (Eq. 4.9d).

The corner points are thus

$$\text{eq.(Eq. 4.9a)} = \text{eq.(Eq. 4.9f)} \Rightarrow (R_{1A}, R_{2A}) = (\mathsf{I_g}(\mathsf{snr}), \tag{4.41a}$$

$$\mathsf{I_g}\left(\frac{\mathsf{snr}}{1+\mathsf{inr}}\right) + \mathsf{I_g}\left(\mathsf{inr} + \frac{\mathsf{snr}}{1+\mathsf{inr}}\right) - \mathsf{I_g}(\mathsf{snr})\Big);$$

$$\tag{4.41b}$$

$$\text{eq.(Eq. 4.9f)} = \text{eq.(Eq. 4.9e)} \Rightarrow (R_{1B}, R_{2B}) = \left(\mathsf{I_g}\left(\frac{\mathsf{snr}}{1+\mathsf{inr}}\right) + \mathsf{I_g}(\mathsf{snr}) - \mathsf{I_g}\left(\mathsf{inr} + \frac{\mathsf{snr}}{1+\mathsf{inr}}\right),\right.$$

$$\tag{4.41c}$$

$$\left.3\mathsf{I_g}\left(\mathsf{inr} + \frac{\mathsf{snr}}{1+\mathsf{inr}}\right) - \mathsf{I_g}(\mathsf{snr}) - \mathsf{I_g}\left(\frac{\mathsf{snr}}{1+\mathsf{inr}}\right)\right);$$

$$\tag{4.41d}$$

$$\text{eq.(Eq. 4.9e)} = \text{eq.(Eq. 4.9g)} \Rightarrow (R_{1C}, R_{2C}) = \left(3\mathsf{I_g}\left(\mathsf{inr} + \frac{\mathsf{snr}}{1+\mathsf{inr}}\right) - \mathsf{I_g}(\mathsf{snr}) - \mathsf{I_g}\left(\frac{\mathsf{snr}}{1+\mathsf{inr}}\right),\right.$$

$$\tag{4.41e}$$

$$\left.\mathsf{I_g}\left(\frac{\mathsf{snr}}{1+\mathsf{inr}}\right) + \mathsf{I_g}(\mathsf{snr}) - \mathsf{I_g}\left(\mathsf{inr} + \frac{\mathsf{snr}}{1+\mathsf{inr}}\right)\right);$$

$$\tag{4.41f}$$

$$\text{eq.(Eq. 4.9b)} = \text{eq.(Eq. 4.9g)} \Rightarrow (R_{1D}, R_{2D}) = \left(\mathsf{I_g}\left(\frac{\mathsf{snr}}{1+\mathsf{inr}}\right) + \mathsf{I_g}\left(\mathsf{inr} + \frac{\mathsf{snr}}{1+\mathsf{inr}}\right) - \mathsf{I_g}(\mathsf{snr}),\right.$$

$$\tag{4.41g}$$

$$\mathsf{I_g}(\mathsf{snr})\Big). \tag{4.41h}$$

As explained before, inspired by the proposed *discrete→common map*, we choose to 'split' the rates as:

1. for the sum-rate face / region $\mathcal{R}_{R_1+R_2}$: we set $R_{1,p} \approxeq \frac{R_1}{2}$ and $R_{2,p} \approxeq \frac{R_2}{2}$.

2. for the other dominant face / region $\mathcal{R}_{2R_1+R_2}$: we set $R_{1,p} \approxeq \mathsf{I_g}\left(\frac{\mathsf{snr}}{1+\mathsf{inr}}\right)$ and $R_{2,p} \approxeq \frac{R_2}{2}$.

3. we will not explicitly consider the remaining dominant face / region $\mathcal{R}_{R_1+2R_2}$ because a gap result can be obtained by proceeding as for $\mathcal{R}_{2R_1+R_2}$ but with the role of the users swapped.

Outer Bound $\mathcal{R}_{R_1+R_2}$

With the corner point expressions in (Eq. 4.41) we write the outer bound sum-rate face in (Eq. 4.27c) as

$$\mathcal{R}^{(4.3.5)}_{R_1+R_2} = \bigcup_{t\in[0,1]} \left\{ \begin{array}{l} R_1 \leq \frac{1-t}{2}\log\left(\frac{\left(1+\frac{\mathsf{snr}}{1+\mathsf{inr}}\right)(1+\mathsf{snr})}{1+\mathsf{inr}+\frac{\mathsf{snr}}{1+\mathsf{inr}}}\right) + \frac{t}{2}\log\left(\frac{\left(1+\mathsf{inr}+\frac{\mathsf{snr}}{1+\mathsf{inr}}\right)^3}{\left(1+\frac{\mathsf{snr}}{1+\mathsf{inr}}\right)(1+\mathsf{snr})}\right) \\ \qquad =: 2\cdot\mathsf{I_g}\left(\mathsf{snr}_{3,a,t}\right) \\ R_1 \leq \frac{t}{2}\log\left(\frac{\left(1+\frac{\mathsf{snr}}{1+\mathsf{inr}}\right)(1+\mathsf{snr})}{1+\mathsf{inr}+\frac{\mathsf{snr}}{1+\mathsf{inr}}}\right) + \frac{1-t}{2}\log\left(\frac{\left(1+\mathsf{inr}+\frac{\mathsf{snr}}{1+\mathsf{inr}}\right)^3}{\left(1+\frac{\mathsf{snr}}{1+\mathsf{inr}}\right)(1+\mathsf{snr})}\right) \\ \qquad =: 2\cdot\mathsf{I_g}\left(\mathsf{snr}_{3,b,t}\right) \end{array} \right\}. \tag{4.42}$$

Inner Bound for $\mathcal{R}_{R_1+R_2}$

In order to approximately achieve the points in $\mathcal{R}^{(4.3.5)}_{R_1+R_2}$ in (Eq. 4.42) we pick

$$N_1 = \mathsf{N_d}\left(\mathsf{snr}_{3,a,t}\right), \ \mathsf{snr}_{3,a,t} := \left(\frac{\left(1+\frac{\mathsf{snr}}{1+\mathsf{inr}}\right)(1+\mathsf{snr})}{1+\mathsf{inr}+\frac{\mathsf{snr}}{1+\mathsf{inr}}}\right)^{\frac{1-t}{2}} \left(\frac{\left(1+\mathsf{inr}+\frac{\mathsf{snr}}{1+\mathsf{inr}}\right)^3}{\left(1+\frac{\mathsf{snr}}{1+\mathsf{inr}}\right)(1+\mathsf{snr})}\right)^{\frac{t}{2}} - 1, \tag{4.43a}$$

$$N_2 = \mathsf{N_d}\left(\mathsf{snr}_{3,b,t}\right), \ \mathsf{snr}_{3,b,t} := \left(\frac{\left(1+\frac{\mathsf{snr}}{1+\mathsf{inr}}\right)(1+\mathsf{snr})}{1+\mathsf{inr}+\frac{\mathsf{snr}}{1+\mathsf{inr}}}\right)^{\frac{t}{2}} \left(\frac{\left(1+\mathsf{inr}+\frac{\mathsf{snr}}{1+\mathsf{inr}}\right)^3}{\left(1+\frac{\mathsf{snr}}{1+\mathsf{inr}}\right)(1+\mathsf{snr})}\right)^{\frac{1-t}{2}} - 1, \tag{4.43b}$$

$$\delta_1 : \mathsf{I_g}\left(\mathsf{snr}\delta_1\right) = \mathsf{I_g}\left(\mathsf{snr}_{3,a,t}\right) \Longleftrightarrow \delta_1 = \frac{\mathsf{snr}_{3,a,t}}{\mathsf{snr}}, \tag{4.43c}$$

$$\delta_2 : \mathsf{I_g}\left(\mathsf{snr}\delta_2\right) = \mathsf{I_g}\left(\mathsf{snr}_{3,b,t}\right) \Longleftrightarrow \delta_2 = \frac{\mathsf{snr}_{3,b,t}}{\mathsf{snr}}, \tag{4.43d}$$

where

$$\max(\delta_1, \delta_2) = \frac{\max(\mathsf{snr}_{3,a,t}, \mathsf{snr}_{3,b,t})}{\mathsf{snr}} \leq \frac{1}{1+\mathsf{inr}},$$

as required for the achievable rate in (Eq. 4.30); the proof can be found in Appendix H.0.1, eq.(Eq. H.3).

Gap for $\mathcal{R}_{R_1+R_2}$

The gap between the outer bound region in (Eq. 4.42) and the achievable rate in (Eq. 4.30) with the parameters in (Eq. I.1) is

$$\begin{aligned}\Delta_{R_1} &= 2\mathsf{I_g}\left(\mathsf{snr}_{3,a,t}\right) - \log\left(\mathsf{N_d}\left(\mathsf{snr}_{3,a,t}\right)\right) - \mathsf{I_g}(\mathsf{snr}_{3,a,t}) + \Delta_{\text{(Eq. 4.30)}}\\ &\leq \log(2) + \Delta_{\text{(Eq. 4.30)}},\end{aligned}$$

and similarly

$$\Delta_{R_2} \leq \log(2) + \Delta_{\text{(Eq. 4.30)}}.$$

We are then left with bounding $\Delta_{\text{(Eq. 4.30)}}$, which depends on minimum distances of the received sum-set constellations. In Appendix H.0.1 we show

$$\min_{i\in[1:2]} \frac{d^2_{\min(S_i)}}{12} \geq \kappa^2_{\gamma,N_1,N_2} \cdot \frac{1}{24}, \tag{4.44}$$

where κ_{γ,N_1,N_2} is given in (Eq. 4.31e), and $\max(N_1^2, N_2^2) - 1 \leq \mathsf{inr} = \min(\mathsf{snr}, \mathsf{inr})$. With this, we finally get that the gap for this face is bounded by

$$\begin{aligned}\mathsf{G}_{\mathsf{d}\text{(Eq. 4.45)}} &\leq \max(\Delta_{R_1}, \Delta_{R_2}) = \log(2) + \Delta_{\text{(Eq. 4.30)}}\\ &\leq \frac{1}{2}\log\left(\frac{4\pi \mathrm{e}}{3}\right) + \frac{1}{2}\log\left(1 + 96 \cdot \frac{(1 + 1/2\ln(1+\min(\mathsf{snr},\mathsf{inr})))^2}{\gamma^2}\right) \text{ bits.}\end{aligned} \tag{4.45}$$

Gap for $\mathcal{R}_{2R_1+R_2}$

The derivation of the gap and other results for $\mathcal{R}_{2R_1+R_2}$ are delegated to Appendix H.0.2. The resulting gap is

$$\mathsf{G}_{\mathsf{d}(\text{Eq. 4.46})} \leq \frac{1}{2}\log\left(\frac{16\pi\mathrm{e}}{3}\right) + \frac{1}{2}\log\left(1 + 32 \cdot \frac{(1 + 1/2\ln(1 + \min(\mathsf{snr}, \mathsf{inr})))^2}{\gamma^2}\right) \text{ bits.} \tag{4.46}$$

Overall Gap for Weak2

To conclude the proof for this sub-regime, the gap is the maximum between the gaps of the different faces and is given by

$$\mathsf{G}_{\mathsf{d}(\text{Eq. 4.47})} = \max(\mathsf{G}_{\mathsf{d}(\text{Eq. 4.45})}, \mathsf{G}_{\mathsf{d}(\text{Eq. 4.46})}) = \mathsf{G}_{\mathsf{d}(\text{Eq. 4.46})}. \tag{4.47}$$

Another Overall Gap for Weak2

The choice of the mixed input parameters according to the *discrete→common map* in (Eq. 4.43) and in (Eq. 4.45) led to the $O\left(\log\left(\frac{\ln(\min(\mathsf{snr},\mathsf{inr}))}{\gamma}\right)\right)$ gap in (Eq. 4.47). This is so because we used Proposition 2.1.3 to bound the minimum distance. A interesting question is whether Proposition 2.1.2 could be used, possibly with a different choice of mixed input parameters.

With a gDoF-type analysis, one can show that it is possible to verify the condition in Proposition 2.1.2 with the proposed choice of parameters in (Eq. 4.43) but not with the input parameters in (Eq. 4.45). So, in this regime we are motivated to look at the *discrete→private map* as we hope to get a constant gap result for the whole region. In Appendix I we show that in this regime it is possible to use Proposition 2.1.2 and the *discrete→private map* to get a constant gap, namely,

$$\mathsf{G}_{\mathsf{d}(\text{Eq. 4.48})} \leq \frac{1}{2}\log\left(\frac{608\ \pi \mathrm{e}}{27}\right) \approx 3.79 \text{ bits}. \tag{4.48}$$

4.3.6 Very Weak Interference, i.e., $\mathsf{inr}(1+\mathsf{inr}) \leq \mathsf{snr}$

In this regime the capacity of the classical G-IC is achieved to within a constant gap by Gaussian inputs, treating interference as noise and power control. This strategy is compatible without the TINnoTS scheme (i.e., set $N_1 = N_2 = 1$ and vary δ_1 and δ_2), so the gap of

$$\mathsf{G}_{\mathsf{d}(\text{Eq. 4.49})} \leq 1/2 \text{ bit}, \tag{4.49}$$

as shown in (26) holds.

This concludes the proof of Theorem 4.3.1. □

4.4 Gap for Some Asymmetric Channels

In this Section we generalize the gap result of Theorem 4.3.1 to some general asymmetric settings.

Theorem 4.4.1. *For the general G-IC, except for the regime*

$$\frac{|h_{22}|^2}{1+|h_{21}|^2} < |h_{12}|^2 < \frac{|h_{22}|^2}{1+|h_{21}|^2}(1+|h_{11}|^2), \tag{4.50a}$$

$$\frac{|h_{11}|^2}{1+|h_{12}|^2} < |h_{21}|^2 < \frac{|h_{11}|^2}{1+|h_{12}|^2}(1+|h_{22}|^2), \tag{4.50b}$$

akin to the moderately weak interference regime for the symmetric setting, the TINnoTS achievable region and the outer bound in Proposition 4.2.2 are to within an additive gap that is either constant or of the order $O\left(\log \frac{\ln\left(\max(|h_{11}|^2,|h_{22}|^2)\right)}{\gamma}\right)$.

Remark 11 (Why is the regime in (Eq. 4.50) excluded?)**.** *The regime identified in* (Eq. 4.50) *involves numerous special cases, whose analysis gets very tedious. We do however strongly believe that our gap result generalizes to this regime as well, by using similar arguments to those developed so far. We note that the analysis in the rest of this section for the general asymmetric setting (which is characterized by four channel parameters) is restricted to those cases where it suffices to consider at most one rate split (thus reducing the number of parameters to be optimize for the mixed inputs) and for which the approximately optimal rate region does not require bounds on* $2R_1 + R_2$ *or* $R_1 + 2R_2$ *(thus reducing the achievability to the sum-capacity dominant face only).*

Proof. We shall treat different regimes separately in the rest of the section.

4.4.1 Very Strong Interference

In the general asymmetric case, the very strong interference regime is the regime in which a receiver can decode the interfering message while treating its intended signal as noise at a higher rate than the intended receiver in the absence of interference; this is the case when the channel gains satisfy (53)

$$|h_{11}|^2(1 + |h_{22}|^2) \le |h_{21}|^2, \tag{4.51a}$$

$$|h_{22}|^2(1 + |h_{11}|^2) \le |h_{12}|^2. \tag{4.51b}$$

Outer Bound

The capacity region of the classical G-IC in very strong interference coincides with that of two interference-free point-to-point links given by

$$\mathcal{R}_{\text{out}}^{(4.4.1)} = \left\{ \begin{array}{l} 0 \le R_1 \le \mathsf{I_g}\left(|h_{11}|^2\right) \\ 0 \le R_2 \le \mathsf{I_g}\left(|h_{22}|^2\right) \end{array} \right\}. \tag{4.52}$$

Inner Bound

The outer bound in (Eq. 4.52) can be matched to within a constant gap by our TINnoTS scheme by choosing, similarly to the symmetric case discussed in Section 4.3.1, the mixed inputs in (Eq. 4.6) with

$$N_1 = \mathsf{N_d}\left(\beta|h_{11}|^2\right) : N_1^2 - 1 \le \beta|h_{11}|^2, \tag{4.53a}$$

$$N_2 = \mathsf{N_d}\left(\beta|h_{22}|^2\right) : N_2^2 - 1 \le \beta|h_{22}|^2, \tag{4.53b}$$

$$\delta_1 = 0, \tag{4.53c}$$

$$\delta_2 = 0, \tag{4.53d}$$

for some $\beta \le 1$. The reason for the factor β in (Eq. 4.53) will be clear shortly (in Appendix I we use $\beta = 3/4$ for similar reasons; we could have used here $\beta = 3/4$ as well, but we will find next a value that gives a smaller gap).

We next show that Proposition 2.1.2 is applicable for the choice of mixed input parameters as in (Eq. 4.53). In particular, we aim to show that

$$N_1|h_{11}|d_{\min(X_1)} \le |h_{12}|d_{\min(X_2)} \text{ (for the received sum-set constellation at receiver 1)} \tag{4.54a}$$

$$N_2|h_{22}|d_{\min(X_2)} \le |h_{21}|d_{\min(X_1)} \text{ (for the received sum-set constellation at receiver 2),} \tag{4.54b}$$

or equivalently that

$$\frac{N_1^2}{N_1^2-1} \cdot \frac{|h_{11}|^2}{1+|h_{11}|^2} \cdot \frac{N_2^2-1}{|h_{22}|^2} \le \frac{|h_{12}|^2}{|h_{22}|^2(1+|h_{11}|^2)}, \tag{4.55a}$$

$$\frac{N_2^2}{N_2^2-1} \cdot \frac{|h_{22}|^2}{1+|h_{22}|^2} \cdot \frac{N_1^2-1}{|h_{11}|^2} \le \frac{|h_{21}|^2}{|h_{11}|^2(1+|h_{22}|^2)}. \tag{4.55b}$$

The condition in (Eq. 4.55) is verified, given the channel gain relationship in (Eq. 4.51), if

$$\frac{N_1^2}{N_1^2-1} \cdot \frac{|h_{11}|^2}{1+|h_{11}|^2} \cdot \frac{N_2^2-1}{|h_{22}|^2} \le 1, \tag{4.56a}$$

$$\frac{N_2^2}{N_2^2-1} \cdot \frac{|h_{22}|^2}{1+|h_{22}|^2} \cdot \frac{N_1^2-1}{|h_{11}|^2} \le 1. \tag{4.56b}$$

It can be easily seen that $\beta = 0.8277$ satisfies (Eq. 4.56) whenever $2 \le \min(N_1, N_2)$. For the found β we therefore have that the received constellations have $|S_1| = |S_2| = N_1N_2$ equally likely points and minimum distance

$$\min_{i\in[1:2]} \frac{d^2_{\min(S_i)}}{12} = \min\left(\frac{|h_{11}|^2}{N_1^2-1}, \frac{|h_{22}|^2}{N_2^2-1}\right) \ge \frac{1}{\beta}. \tag{4.57}$$

Thus, by following similar steps as in Section 4.3.1, the achievable region becomes

$$\mathcal{R}_{\text{in}}^{(4.4.1)} = \left\{ \begin{array}{c} 0 \le R_1 \le r_1 \\ \\ 0 \le R_2 \le r_2 \end{array} \right\} \text{ such that} \tag{4.58a}$$

$$r_1 \ge \mathsf{I_d}\left(S_1\right) - \min\left(\log(N_2), \mathsf{I_g}\left(|h_{12}|^2\right)\right) \ge \log(N_1) - \Delta_{\text{(Eq. 4.58)}}, \tag{4.58b}$$

$$r_2 \ge \mathsf{I_d}\left(S_2\right) - \min\left(\log(N_1), \mathsf{I_g}\left(|h_{21}|^2\right)\right) \ge \log(N_2) - \Delta_{\text{(Eq. 4.58)}}, \tag{4.58c}$$

$$\Delta_{\text{(Eq. 4.58)}} \le \frac{1}{2}\log\left(\frac{2\pi\mathrm{e}}{12}\right) + \frac{1}{2}\log\left(1 + \frac{12}{\min_{i\in[1:2]} d^2_{\min(S_i)}}\right) \le \frac{1}{2}\log\left(\frac{2\pi\mathrm{e}}{12}(\beta+1)\right). \tag{4.58d}$$

Gap

We can easily see, by comparing the inner bound in (Eq. 4.58) with the outer bound in (Eq. 4.52), that for the general asymmetric G-IC in very strong interference the TINnoTS region is optimal to within

$$\begin{aligned} \mathsf{G}_{\mathsf{d}\text{(Eq. 4.59)}} &\le \Delta_{\text{(Eq. 4.58)}} + \log(2) + \frac{1}{2}\log\left(\frac{1}{\beta}\right) \\ &\le \frac{1}{2}\log\left(\frac{2\pi\mathrm{e}}{3}\,\frac{1+\beta}{\beta}\right) \overset{\beta=0.8277}{\approx} 1.8260 \text{ bits}, \end{aligned} \tag{4.59}$$

where the term $\log(2)$ is the integrality gap and the term $\frac{1}{2}\log\left(\frac{1}{\beta}\right)$ because of the reduced number of points in (Eq. 4.53).

4.4.2 Strong (but not Very Strong) Interference

For the general asymmetric case, the strong interference regime is defined as

$$|h_{21}|^2 \geq |h_{11}|^2, \tag{4.60a}$$

$$|h_{12}|^2 \geq |h_{22}|^2. \tag{4.60b}$$

The strong (but not very strong) interference regime is the set of channel gains that satisfy the condition in (Eq. 4.60) but not the condition in (Eq. 4.51).

Outer Bound

The capacity region of the general G-IC in the strong interference regime is given by the 'compound MAC' region

$$\mathcal{R}_{\text{out}}^{(4.4.2)} = \left\{ \begin{array}{l} 0 \leq R_1 \leq \mathsf{I_g}\left(|h_{11}|^2\right) \\ 0 \leq R_2 \leq \mathsf{I_g}\left(|h_{22}|^2\right) \\ R_1 + R_2 \leq \mathsf{I_g}\left(\min\left(|h_{11}|^2 + |h_{12}|^2, |h_{22}|^2 + |h_{21}|^2\right)\right) \end{array} \right\}. \tag{4.61}$$

Inner Bound

The outer bound in (Eq. 4.61) can be matched to within a gap by our TINnoTS scheme by choosing, similarly to the symmetric case discussed in detail in Section 4.3.2, the parameters of the mixed inputs as

$$N_1 = \mathsf{N_d}\left(\mathsf{snr}_{5,a,t}\right),$$
$$\mathsf{snr}_{5,a,t} = (1+|h_{11}|^2)^{1-t}\left(\frac{1+\min\left(|h_{11}|^2+|h_{12}|^2, |h_{22}|^2+|h_{21}|^2\right)}{1+|h_{22}|^2}\right)^t - 1 \le |h_{11}|^2, \tag{4.62a}$$
$$N_2 = \mathsf{N_d}\left(\mathsf{snr}_{5,b,t}\right),$$
$$\mathsf{snr}_{5,b,t} = (1+|h_{22}|^2)^{t}\left(\frac{1+\min\left(|h_{11}|^2+|h_{12}|^2, |h_{22}|^2+|h_{21}|^2\right)}{1+|h_{11}|^2}\right)^{1-t} - 1 \le |h_{22}|^2, \tag{4.62b}$$
$$\delta_1 = 0, \tag{4.62c}$$
$$\delta_2 = 0, \tag{4.62d}$$

where the upper bounds on $\mathsf{snr}_{5,a,t}$ and $\mathsf{snr}_{5,b,t}$ are a consequence of not being in very strong interference, i.e.,

$$\min\left(1+|h_{11}|^2+|h_{12}|^2, 1+|h_{22}|^2+|h_{21}|^2\right) \le (1+|h_{11}|^2)(1+|h_{22}|^2).$$

Next, by using Proposition 2.1.3 we have

$$\frac{d^2_{\min(S_1)}}{12} \ge \kappa^2_{\gamma,N_1,N_2}\min\left(\frac{|h_{11}|^2}{N_1^2-1}, \frac{|h_{12}|^2}{N_2^2-1}, \max\left(\frac{|h_{12}|^2}{N_1^2(N_2^2-1)}, \frac{|h_{11}|^2}{N_2^2(N_1^2-1)}\right)\right), \tag{4.63a}$$
$$\frac{d^2_{\min(S_2)}}{12} \ge \kappa^2_{\gamma,N_1,N_2}\min\left(\frac{|h_{21}|^2}{N_1^2-1}, \frac{|h_{22}|^2}{N_2^2-1}, \max\left(\frac{|h_{21}|^2}{N_1^2(N_2^2-1)}, \frac{|h_{22}|^2}{N_2^2(N_1^2-1)}\right)\right), \tag{4.63b}$$

where the bounds in (Eq. 4.63) hold up to a set of measure γ and where κ_{γ,N_1,N_2} is defined in (Eq. 4.31e). By recalling the channel gain relationship, by noting that

$$N_1^2 N_2^2 - 1 \leq \min\left(|h_{11}|^2 + |h_{12}|^2, |h_{22}|^2 + |h_{21}|^2\right)$$

and by combining the two bounds in (Eq. 4.63) we get

$$\min_{i\in[1:2]} \frac{d^2_{\min(S_i)}}{12} \geq \min\left(1, \frac{\max(|h_{11}|^2, |h_{12}|^2)}{|h_{11}|^2 + |h_{12}|^2}, \frac{\max(|h_{21}|^2, |h_{22}|^2)}{|h_{22}|^2 + |h_{21}|^2}\right) \geq \frac{1}{2}.$$

Gap

By following the same reasoning and bounding steps as we did for the symmetric case in Section 4.3.2, we get that the proposed achievable scheme is optimal to within a gap of

$$\mathsf{G}_{\mathsf{d}(\text{Eq. 4.64})} \leq \frac{1}{2}\log\left(\frac{2\pi \mathrm{e}}{3}\right) + \frac{1}{2}\log\left(1 + 8 \cdot \frac{\left(1 + 1/2\ln\left(1 + \max\left(|h_{11}|^2, |h_{22}|^2\right)\right)\right)^2}{\gamma^2}\right) \text{ bits.} \tag{4.64}$$

4.4.3 Mixed Interference

The mixed interference regime occurs when one receiver experiences strong interference while the other experiences weak interference. This regime does not appear in the symmetric case, where both receiver are either in strong interference or in weak interference. The mixed interference is defined as (26)

$$\text{either } \left\{|h_{21}|^2 \geq |h_{11}|^2,\ |h_{12}|^2 \leq |h_{22}|^2\right\}, \tag{4.65a}$$

$$\text{or } \left\{|h_{21}|^2 \leq |h_{11}|^2,\ |h_{12}|^2 \geq |h_{22}|^2\right\}. \tag{4.65b}$$

In this Section we shall only focus on the sub-regime

$$|h_{21}|^2 \geq \frac{|h_{11}|^2}{1+|h_{12}|^2}(1+|h_{22}|^2),\ |h_{12}|^2 \leq |h_{22}|^2, \tag{4.66}$$

for which the rate region, as we shall see, does not require bounds on $2R_1 + R_2$ or $R_1 + 2R_2$. The regime $|h_{12}|^2 \geq \frac{|h_{22}|^2}{1+|h_{21}|^2}(1+|h_{11}|^2),\ |h_{21}|^2 \leq |h_{11}|^2$ can be analyzed similarly by swapping the role of the users.

Outer Bound

An outer bound to the capacity region of the general G-IC when (Eq. 4.66) holds is given by the 'Z-channel' outer bound (61)

$$\mathcal{R}_{\text{out}}^{(4.4.3)} = \left\{ \begin{array}{l} 0 \le R_1 \le \mathsf{I_g}\left(|h_{11}|^2\right) \\ 0 \le R_2 \le \mathsf{I_g}\left(|h_{22}|^2\right) \\ R_1 + R_2 \le \mathsf{I_g}\left(|h_{22}|^2\right) + \mathsf{I_g}\left(\frac{|h_{11}|^2}{1+|h_{12}|^2}\right) \end{array} \right\}$$
$$= \bigcup_{t\in[0,1]} \left\{ \begin{array}{l} 0 \le R_1 \le (1-t)\mathsf{I_g}\left(|h_{11}|^2\right) + t\mathsf{I_g}\left(\frac{|h_{11}|^2}{1+|h_{12}|^2}\right) \\ \quad = \mathsf{I_g}(\mathsf{snr}_{6,a,t}) \\ 0 \le R_2 \le (1-t)\left(\mathsf{I_g}\left(|h_{22}|^2\right) - \mathsf{I_g}\left(|h_{11}|^2\right) + \mathsf{I_g}\left(\frac{|h_{11}|^2}{1+|h_{12}|^2}\right)\right) + t\mathsf{I_g}(|h_{22}|^2) \\ \quad = \mathsf{I_g}(\mathsf{snr}_{6,b,t}) + \frac{1}{2}\log\left(\frac{1+|h_{22}|^2}{1+|h_{12}|^2}\right) \end{array} \right\}. \tag{4.67}$$

Inner Bound

The shape of the outer bound in (Eq. 4.67) suggests that a matching, to within a gap, inner region could be found by following steps similar to those used for the analysis of the strong interference regime (i.e., parameterize the points on the dominate sum-capacity face). The difference between this sub-regime and the strong interference regime is that here R_2 should be a combination of common and private rates because receiver 1 experiences weak interference (while receiver 2 experiences strong interference). Note that the interfering channel gain at receiver 2, h_{21}, does not appear in the outer bound in (Eq. 4.67). We therefore set

$$N_1 = \mathsf{N_d}\left(\mathsf{snr}_{6,a,t}\right) : \mathsf{snr}_{6,a,t} = \left(1 + \frac{|h_{11}|^2}{1+|h_{12}|^2}\right)^t \left(1+|h_{11}|^2\right)^{1-t} - 1 \le |h_{11}|^2, \tag{4.68a}$$

$$N_2 = \mathsf{N_d}\left(\mathsf{snr}_{6,b,t}\right) : \mathsf{snr}_{6,b,t} = \left(1+|h_{12}|^2\right)^t \left(1 + \frac{|h_{12}|^2}{1+|h_{11}|^2}\right)^{1-t} - 1 \le |h_{12}|^2, \tag{4.68b}$$

$$\delta_1 = 0, \tag{4.68c}$$

$$\delta_2 = \frac{1}{1+|h_{12}|^2}, \tag{4.68d}$$

in the achievable region in Proposition 4.2.1, which becomes

$$\mathcal{R}_{\text{in}}^{\text{(Eq. 4.69)}} = \left\{ \begin{array}{l} 0 \le R_1 \le \log(N_1) - \Delta_{\text{(Eq. 4.69)}} \\ 0 \le R_2 \le \log(N_2) + \mathsf{I_g}\left(\frac{|h_{22}|^2}{1+|h_{12}|^2}\right) - \Delta_{\text{(Eq. 4.69)}} \end{array} \right\}, \tag{4.69a}$$

$$\Delta_{\text{(Eq. 4.69)}} = \frac{1}{2}\log\left(\frac{2\pi e}{12}\right) + \frac{1}{2}\log\left(1 + \frac{12}{\min_{i\in[1:2]} d^2_{\min(S_i)}}\right), \tag{4.69b}$$

$$S_1 = \frac{1}{\sqrt{1+\frac{|h_{12}|^2}{1+|h_{12}|^2}}}\left(h_{11}X_{1D} + \sqrt{\frac{|h_{12}|^2}{1+|h_{12}|^2}}\, h_{12}X_{2D}\right), \tag{4.69c}$$

$$S_2 = \frac{1}{\sqrt{1+|h_{22}|^2\frac{1}{1+|h_{12}|^2}}}\left(h_{21}X_{1D} + \sqrt{\frac{|h_{12}|^2}{1+|h_{12}|^2}}\, h_{22}X_{2D}\right). \tag{4.69d}$$

Next, by using Proposition 2.1.3, we bound the minimum distance of the received constellations S_1 and S_2 as

$$\begin{aligned}
\frac{d^2_{\min(S_1)}}{12\,\kappa^2_{\gamma,N_1,N_2}} &\geq \frac{1}{1+|h_{12}|^2\delta_2}\min\left(\frac{|h_{11}|^2}{N_1^2-1},\frac{(1-\delta_2)|h_{12}|^2}{N_2^2-1},\max\left(\frac{(1-\delta_2)|h_{12}|^2}{N_1^2(N_2^2-1)},\frac{|h_{11}|^2}{N_2^2(N_1^2-1)}\right)\right)\\
&\overset{(a)}{\geq} \frac{1}{1+|h_{12}|^2\delta_2}\min\left(1,(1-\delta_2),\frac{\max\left((1-\delta_2)|h_{12}|^2,|h_{11}|^2\right)}{N_1^2N_2^2-1}\right)\\
&\geq \frac{1-\delta_2}{1+|h_{12}|^2\delta_2}\min\left(1,\frac{\max\left(|h_{12}|^2,|h_{11}|^2\right)}{N_1^2N_2^2-1}\right)\\
&\overset{(b)}{\geq} \frac{|h_{12}|^2}{2+|h_{12}|^2}\min\left(1,\frac{\max\left(|h_{12}|^2,|h_{11}|^2\right)}{|h_{12}|^2+|h_{11}|^2}\right)\\
&\overset{(c)}{\geq} \frac{1}{3}\min\left(1,\frac{1}{2}\right)=\frac{1}{6},
\end{aligned}$$

where the inequalities follow since: (a) by using the bounds in (Eq. 4.68a) and (Eq. 4.68b), (b) because $N_1^2N_2^2-1 \leq |h_{12}|^2+|h_{11}|^2$ from (Eq. 4.68a) and (Eq. 4.68b), and (c) by assuming $|h_{12}|^2 \geq 1$. Similarly we have that

$$\begin{aligned}
\frac{d^2_{\min(S_2)}}{12\,\kappa^2_{\gamma,N_1,N_2}} &\geq \frac{1}{1+|h_{22}|^2\delta_2}\min\left(\frac{|h_{21}|^2}{N_1^2-1},\frac{(1-\delta_2)|h_{22}|^2}{N_2^2-1},\max\left(\frac{|h_{21}|^2}{N_2^2(N_1^2-1)},\frac{(1-\delta_2)|h_{22}|^2}{N_1^2(N_2^2-1)}\right)\right)\\
&\overset{(a)}{\geq} \frac{1-\delta_2}{1+|h_{22}|^2\delta_2}\min\left(\frac{|h_{21}|^2}{|h_{11}|^2},\frac{|h_{22}|^2}{|h_{12}|^2},\frac{\max(|h_{21}|^2,|h_{22}|^2)}{|h_{12}|^2+|h_{11}|^2}\right)\\
&\overset{(b)}{\geq} \frac{1-\delta_2}{1+|h_{22}|^2\delta_2}\min\left(\frac{1+|h_{22}|^2}{1+|h_{12}|^2},\frac{|h_{22}|^2}{|h_{12}|^2},\frac{\max\left(|h_{11}|^2\frac{1+|h_{22}|^2}{1+|h_{12}|^2},|h_{12}|^2\frac{|h_{22}|^2}{|h_{12}|^2}\right)}{|h_{12}|^2+|h_{11}|^2}\right)\\
&\overset{(b)}{\geq} \frac{1-\delta_2}{1+|h_{22}|^2\delta_2}\ \frac{1+|h_{22}|^2}{1+|h_{12}|^2}\ \frac{1}{2}\\
&= \frac{1+|h_{22}|^2}{1+|h_{12}|^2+|h_{22}|^2}\ \frac{|h_{12}|^2}{1+|h_{12}|^2}\ \frac{1}{2}\\
&\overset{(c)}{\geq} \frac{2}{3}\ \frac{1}{2}\ \frac{1}{2}=\frac{1}{6},
\end{aligned}$$

where the inequalities follow since: (a) by using the bounds in (Eq. 4.68a) and (Eq. 4.68b) and because $N_1^2 N_2^2 - 1 \le |h_{12}|^2 + |h_{11}|^2$, (b) by the channel gain relationship in (Eq. 4.66), and (c) by assuming $1 \le |h_{12}|^2$ and since by assumption of this regime $|h_{12}|^2 \le |h_{22}|^2$. Note that the assumption $1 \le |h_{12}|^2$ is without loss of generality since if $|h_{12}|^2 < 1$ (i.e., interference below the noise floor of the receiver) then TIN with Gaussian inputs achieves the capacity outer bound (in this case essentially two interference-free point-to-point links) to within 1/2 bit.

This shows that

$$\min_{i \in [1:2]} \frac{d^2_{\min(S_i)}}{12} \ge \kappa^2_{\gamma,N_1,N_2} \cdot \frac{1}{6}, \tag{4.70}$$

up to an outage set of measure no more than γ, where γ affects κ_{γ,N_1,N_2}.

Gap

By following the same reasoning and bounding steps as we did for the symmetric case, we get that the proposed achievable scheme is optimal to within a gap of

$$\begin{aligned}
\Delta_{R_1} &\le \mathsf{I_g}(\mathsf{snr}_{6,a,t}) - \log(\mathsf{N_d}(\mathsf{snr}_{6,a,t})) + \Delta_{\text{(Eq. 4.69)}} \\
&\le \Delta_{\text{(Eq. 4.69)}} + \log(2), \\
\Delta_{R_2} &\le \mathsf{I_g}(\mathsf{snr}_{6,b,t}) + \frac{1}{2}\log\left(\frac{1+|h_{22}|^2}{1+|h_{12}|^2}\right) - \log(\mathsf{N_d}(\mathsf{snr}_{6,b,t})) - \mathsf{I_g}\left(\frac{|h_{22}|^2}{1+|h_{12}|^2}\right) + \Delta_{\text{(Eq. 4.69)}} \\
&\le \Delta_{\text{(Eq. 4.69)}} + \log(2),
\end{aligned}$$

where we used the fact that $\log(\mathsf{N_d}(x)) \geq \mathsf{I_g}(x) - \log(2)$. By including the minimum distance bound in (Eq. 4.70) into the expression for $\Delta_{\text{(Eq. 4.69)}}$ in (Eq. 4.69b), and by noticing that $\max(N_1^2, N_2^2) - 1 \leq \max(|h_{11}|^2, |h_{12}|^2) \leq \max(|h_{11}|^2, |h_{22}|^2)$ by the channel gain relationship in (Eq. 4.66), we finally get

$$\begin{aligned} \mathsf{G}_{\mathsf{d}\text{(Eq. 4.71)}} &\leq \frac{1}{2}\log\left(\frac{2\pi \mathrm{e}}{3}\right) + \frac{1}{2}\log\left(1 + \frac{6}{\kappa^2_{\gamma,N_1,N_2}}\right), \\ &\leq \frac{1}{2}\log\left(\frac{2\pi \mathrm{e}}{3}\right) + \frac{1}{2}\log\left(1 + 24 \cdot \frac{\left(1 + 1/2\ln\left(1 + \max\left(|h_{11}|^2, |h_{22}|^2\right)\right)\right)^2}{\gamma^2}\right) \text{ bits.} \end{aligned} \tag{4.71}$$

4.4.4 Weak Interference

For the general asymmetric G-IC, the weak interference is defined as

$$|h_{21}|^2 \leq |h_{11}|^2, \tag{4.72a}$$

$$|h_{12}|^2 \leq |h_{22}|^2, \tag{4.72b}$$

which involves numerous special cases whose analysis gets very tedious and is outside of the scope of this thesis – see also Remark 11.

4.4.5 Very Weak Interference

The very weak interference regime characterized in as (24) is defined as

$$|h_{12}|^2 \leq \frac{|h_{22}|^2}{1 + |h_{21}|^2}, \tag{4.73a}$$

$$|h_{21}|^2 \leq \frac{|h_{11}|^2}{1 + |h_{12}|^2}. \tag{4.73b}$$

In this regime, the outer bound to the capacity region of the classical G-IC is

$$\mathcal{R}_{\text{out}}^{(4.4.5)} = \left\{ \begin{array}{l} R_1 \leq \mathsf{I_g}\left(|h_{11}|^2\right) \\ R_2 \leq \mathsf{I_g}\left(|h_{22}|^2\right) \\ R_1 + R_2 \leq \mathsf{I_g}\left(|h_{12}|^2 + \frac{|h_{11}|^2}{1+|h_{21}|^2}\right) + \mathsf{I_g}\left(|h_{21}|^2 + \frac{|h_{22}|^2}{1+|h_{12}|^2}\right) \end{array} \right\} \tag{4.74}$$

and is achievable to within $1/2$ bit by Gaussian inputs with power control and TIN. Since the optimal strategy for the classical G-IC is compatible with our TINnoTS with mixed inputs, we conclude that a mixed-input is optimal to within $1/2$ bit in this regime.

This concludes the proof of Theorem 4.4.1. $\square$

4.5 TINnoTS is gDoF Optimal

In this section we show one of the consequences of Theorem 4.3.1, namely that TINnoTS is gDoF optimal almost surely. As mentioned in Section 3.6 the gDoF has become an important metric that sheds lights on the behavior of the capacity when exact capacity results are not available, and has been formally defined in (Eq. 3.18). The $O\left(\log\left(\frac{\ln(\min(\mathsf{snr},\mathsf{inr}))}{\gamma}\right)\right)$ additive gap result of Theorem 4.3.1 implies that:

Theorem 4.5.1. *For the symmetric G-IC the TINnoTS achievable scheme with mixed inputs is gDoF optimal for all channel gains up to a set of zero measure. The optimal inputs are given in Table II.*

Proof. We must show that as $\mathsf{snr} \to \infty$ the gap between the TINnoTS inner bound and outer bound in Proposition 4.2.2, normalized by $\mathsf{I_g}(\mathsf{snr})$, goes to zero almost everywhere.

In the proof of Theorem 4.3.1 we showed that for the very strong, the weak2 and the very weak interference regimes the gap between inner and outer bounds is $O(1)$ everywhere. Therefore, since $\lim_{\mathsf{snr}\to\infty} \frac{O(1)}{\mathsf{I_g}(\mathsf{snr})} = 0$, the result follows.

For the strong and weak1 interference regimes the gap is of the form $O\left(\log\left(\frac{\ln\min(\mathsf{snr},\mathsf{snr}^{\alpha})}{\gamma}\right)\right)$ for any $\gamma \in (0,1]$. Therefore, by choosing γ to be

$$\gamma(\mathsf{snr}) := \frac{1}{(\log\min(\mathsf{snr},\mathsf{snr}^{\alpha}))^{p}}, \text{ for some } p > 0 \text{ independent of } \mathsf{snr},$$

we have that $\lim_{\mathsf{snr}\to\infty} \frac{O\left(\log\left(\frac{\ln\min(\mathsf{snr},\mathsf{snr}^{\alpha})}{\gamma}\right)\right)}{\mathsf{I_g}(\mathsf{snr})} = 0$ and the measure of the outage set $\gamma(\mathsf{snr})$ vanishes as $\mathsf{snr} \to \infty$. This concludes the proof. $\square$

Remark 12. *We note that in (62), the authors showed that discrete-continuous mixtures are strictly sub-optimal in the DoF expression (62, Theorem 4) for $K > 2$ user. However, this does not mean that for an equivalent (but different) expression of the DoF continuous and discrete mixtures are not optimal. For example, in (63) it was shown that Gaussian inputs do not maximize the multi-letter capacity expression for Gaussian multiple access channels. However, in an equivalent single letter expression of the capacity Gaussian inputs are optimal. Something similar occurs in the context of K-user interference channel. On the one hand, in (62) it was shown that discrete-continuous mixtures are not optimal when used in a particular capacity expression (involving information dimension). On the other hand, from (36) we know that taking a discrete distribution that depends on the SNR (i.e., the number of points scales with SNR) and using this in another expression, one can achieve DoF= $\frac{K}{2}$ a.e.*

Moreover, the result of (62), does not imply that discrete-continuous mixtures can not be capacity achieving for some parameter regimes. In particular in (24) authors showed that Gaussian inputs are gDoF optimal (for $K \geq 2$) in the so called weak interference regime.

4.6 Totally Asynchronous and Codebook Oblivious G-IC

The only requirement for the implementation of the TINnoTS inner bound in (Eq. 4.5) is to have symbol synchronization and knowledge of the channel gains at all the terminals. Therefore, our TINnoTS achievable strategy applies to a large class of channels, besides the model considered thus far. Next, we outline two such examples for which very little was known in the past.

The first example is the block asynchronous G-IC, which is information unstable (14) and thus no single-letter capacity expression can be derived for it. Nonetheless, we are able to show that the capacity of this channel is to within a gap of the capacity of the fully synchronized channel. The second example is the G-IC with partial codebook knowledge at both receivers (20), which prevents using joint decoding or successive interference cancellation at the decoders. Still, we are able to show that the capacity of this channel is to within a gap of the capacity of the channel with full codebook knowledge.

The applications to oblivious and asynchronous ICs somewhat surprisingly implies that much less "global coordination" between nodes is needed than one might initially expect: synchronism and codebook knowledge might not be critical if one is happy with 'approximate' capacity results.

4.6.1 Block Asynchronous G-IC

Consider a G-IC with the following input-output relationship

$$Y_{1,t} = h_{11}X_{1,t} + h_{12}X_{2,t-D_1} + Z_{1,t}, \tag{4.75a}$$

$$Y_{2,t} = h_{21}X_{1,t-D_2} + h_{22}X_{2,t} + Z_{2,t}, \tag{4.75b}$$

for $t \in \mathbb{Z}^+$, and $X_{i,j}$ user i's input to the channel at channel use j, $X_{i,j} = 0$ for $j < 0$ (similarly for $Y_{i,j}$ and $Z_{i,j}$), where the delay D_i, $i \in [1:2]$, is chosen at the beginning of the transmission and held fixed thereafter. The channel is termed totally asynchronous if delay D_i is uniform on all n (14). Except for the introduction of random delay all definitions are identical to those given in Section 4.1. In (14) it has been shown that $\mathcal{R}_{\text{in}}^{\text{TINnoTS}}$ in (Eq. 4.5) is achievable for the channel in (Eq. 4.75). Moreover, because lack of synchronization can only harm communications, the outer bound in Proposition 4.2.2 is a valid outer bound for the asynchronous G-IC. Therefore, all of our previous results hold and we have:

Lemma 4.6.1. *For the block asynchronous G-IC the TINnoTS achievable region is to within an additive gap of the capacity of the fully synchronized G-IC, where the gap is given in Theorems 4.3.1 and 4.4.1.*

4.6.2 IC with No Codebook Knowledge

IC with partial codebook knowledge is practically relevant because it models the inability to use sophisticated decoding techniques such as joint decoding or successive inference cancellation. In Chapter 3 for the IC-OR with partial codebook knowledge at one receiver, it has been shown that using Gaussian input at the transmitter corresponding to the oblivious receiver and a mixed input at the transmitter corresponding to non-oblivious receiver is to within a constant gap from the capacity of the classical

G-IC with full codebook knowledge. In (20) it was shown that for IC-OR with both oblivious receivers the capacity is given by

$$\mathcal{C}^{\text{IC-OR}} = \bigcup_{P_Q P_{X_1|Q} P_{X_2|Q}} \left\{ \begin{array}{l} R_1 \le I(X_1;Y_1|Q) \\ \\ R_2 \le I(X_2;Y_2|Q) \end{array} \right\}. \tag{4.76}$$

Note that the region in (Eq. 4.76) is very similiar to TINoTS region in (Eq. 4.4) and $\mathcal{C}^{\text{IC-OR}}$ is upper bounded by the classical G-IC outer bound in Proposition 4.2.2. The set of optimizing distributions for (Eq. 4.76) and the cardinality bound for the alphabet of Q are not known (20, Section III.A). Based on our previous results, we have that:

Lemma 4.6.2. *For the G-IC with partial codebook knowledge the TINnoTS achievable region is to within an additive gap of the capacity of the G-IC with full codebook knowledge, where the gap is given in Theorems 4.3.1 and 4.4.1.*

4.7 TINnoTS with Mixed Inputs in Practice

4.7.1 A Simple TINnoTS Receiver in Very Strong Interference

In the introduction we mentioned that the optimal MAP decoder in an additive non-Gaussian noise channel, which one could implement for TIN when treating a non-Gaussian interference as noise, could be very complex. In the following we give an example of an approximate MAP decoder that is very simple to implement, thus making TINnoTS competitive in practical applications.

Let X_1, X_2 be from the PAM (N, d) with $N = 2Q + 1, Q \in \mathbb{N}$, and $d^2 = \frac{12}{N^2-1} = \frac{3}{Q(Q+1)}$. The restrictions to an odd number of points is just for simplicity of writing the constellation points. The received signal is

$$Y = \left(\sqrt{\mathsf{snr}}\, n_1 + \sqrt{\mathsf{inr}}\, n_2\right) d + Z_G,\ Z_G \sim \mathcal{N}(0, 1),$$

for some $(n_1, n_2) \in [-Q : Q]^2$ chosen independently with uniform probability. The condition in (Eq. 2.16b) is verified when

$$(2Q + 1)^2 \mathsf{snr} \leq \mathsf{inr}, \tag{4.77}$$

which corresponds to the very strong interference regime. In the regime identified by (Eq. 4.77), i.e., where the received points do not 'overlap' as in Fig Figure 2, the decoder could simply "modulo-out" the interference by "folding" the signal Y onto the interval $\mathcal{I} := [-\sqrt{\mathsf{inr}}d/2, +\sqrt{\mathsf{inr}}d/2]$. By doing so the resulting signal, given by

$$Y' = \left[\sqrt{\mathsf{snr}}\, n_1 d + Z'\right]_{\text{mod } \mathcal{I}},\ Z' := [Z_G]_{\text{mod } \mathcal{I}},$$

would be interference-free. Since

$$\begin{aligned} \Pr[Y' \neq \sqrt{\mathsf{snr}}\, n_1 d + Z_G] &\leq \Pr[|Z_G| \geq \sqrt{\mathsf{inr}}d/2 - \sqrt{\mathsf{snr}}Qd] \\ &\overset{\text{from (Eq. 4.77)}}{\leq} \Pr[|Z_G| \geq \sqrt{\mathsf{snr}}/2], \end{aligned}$$

and since $\Pr[|Z_G| \geq \sqrt{\mathsf{snr}}/2]$ is also an upper bound to the probability of error for PAM input on a Gaussian channel, we see that the simple modulo operation at the receiver results in a symbol-error rate that is at most double that of an interference-free Gaussian channel with the same PAM input.

4.7.2 Actual vs. Analytic Gap

Here we compare the gap derived in Theorems 4.3.1 and 4.4.1 to the actual gap evaluated numerically. The point is to show that our analytical closed-form (worst case scenario) bounds can be quite conservative and thus underestimate the actual achievable rates.

For example, we showed that in the very strong interference regime the TINnoTS achievable region with discrete inputs is at most $\frac{1}{2}\log\left(\frac{2\pi e}{3}\right)$ bits from capacity; the capacity in this case is the same as two parallel interference-free links. Consider the symmetric G-IC in very strong interference and the symmetric rate $R_1 = R_2 = R_{\text{sym}}(\mathsf{snr})$ with the same PAM input for each user, where the number of points is chosen as in (Eq. 4.11a). Fig. 8a shows $\mathsf{G_d}(\mathsf{snr}) := \mathsf{I_g}(\mathsf{snr}) - R_{\text{sym}}(\mathsf{snr})$ vs. snr expressed in dB, where

- the red line is the theoretical gap from Theorem 4.3.1, approximately $\frac{1}{2}\log\left(\frac{2\pi e}{3}\right) = 1.25$ bits;
- the green line is the gap by lower bounding $R_{\text{sym}}(\mathsf{snr})$ with the Ozarow-Wyner-B bound in Proposition 2.1.1, where the minimum distance of the received constellation was computed exactly (rather than lower bounded by Proposition 2.1.2); the gap in this case is approximately 0.75 bits;
- the magenta line is the gap by lower bounding $R_{\text{sym}}(\mathsf{snr})$ by the ‘full DTD-ITA’14 bound’ in (Eq. 2.13b), the gap in this case is approximately 0.37 bits;

- the cyan line is the gap when $R_{\text{sym}}(\mathsf{snr})$ is evaluated by Monte Carlo simulation; the gap in this case tends to the ultimate "shaping loss" $\frac{1}{2}\log\left(\frac{\pi \mathrm{e}}{6}\right) = 0.25$ bits at large snr; this shows that the actual gap is about 1 bit lower than the theoretical gap;

The figure also shows that the lower bound in (Eq. 2.13b) actually gives the tightest lower bound for the mutual information, but it is unfortunately not easy to deal with analytically.

We next consider the symmetric G-IC in strong interference. Theorem 4.3.1 upper bounds the gap in this regime by $\mathsf{G_d}(\mathsf{snr}) \le \frac{1}{2}\log\left(\frac{2\pi \mathrm{e}}{3}\right) + \frac{1}{2}\log\left(1 + 8\frac{(1+1/2\ln(1+\mathsf{snr}))^2}{\gamma^2}\right)$ where $\gamma \in (0, 1]$ is the measure of the outage set (i.e., those channel gains for which the gap lower bound is not valid). If we were to make the measure of the outage set very small, then we could end up finding that the gap is actually larger than capacity. Consider the case $\mathsf{snr} = 30$ dB and $\mathsf{inr} = \mathsf{snr}^{1.49} = 44.7$ dB; with $\gamma = 0.1$ it easy to see that $\frac{1}{2}\log\left(\frac{2\pi \mathrm{e}}{3}\right) + \frac{1}{2}\log\left(1 + 8\frac{(1+1/2\ln(1+\mathsf{snr}))^2}{\gamma^2}\right) = 6.977$ bits, which is larger than the interference-free capacity $\mathsf{I_g}(\mathsf{snr}) = 4.9836$ bits. This implies that our bounding steps, done for the sake of analytical tractability and especially meaningful at high SNR, are too crude for this specific example (where our result states the trivial fact that zero rate for each user is achievable to within $\mathsf{I_g}(\mathsf{snr})$ bits). We aim to convey next that, despite the fact that the closed-form gap result underestimates the achievable rates, it nonetheless provides valuable insights into the performance of practical systems, that is, that TINnoTS with discrete inputs performs quite well in the strong interference regime (where capacity is achieved by Gaussian codebooks and joint decoding of interfering and intended messages). To this end, Fig. 8b shows the achievable rate region for the symmetric G-IC with $\mathsf{snr} = 30$ dB and $\mathsf{inr} = \mathsf{snr}^{1.49} = 44.7$ dB and where the users employ a PAM input with the number of points given by (Eq. 4.22). We observe

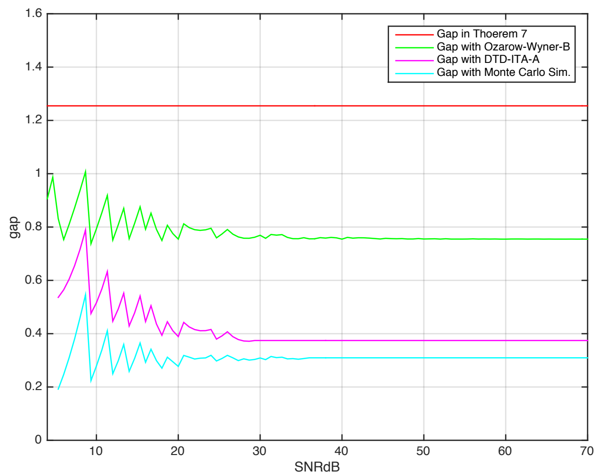
- The navy blue line shows the pentagon-shaped capacity region in (Eq. 4.16).

- The red point at the origin is the lower bound on the achievable rates from Theorem 4.3.1 with $\gamma = 0.1$.
- The green line is the achievable region when the rates are lower bounded by the Ozarow-Wyner-B bound in Proposition 2.1.1, where the minimum distances of the received constellations were computed exactly (rather than lower bounded by Proposition 2.1.3).
- For the magenta line we used the DTD-ITA'14-A lower bound in (Eq. 2.13b);
- For the cyan line we evaluated the rates by Monte Carlo simulation.

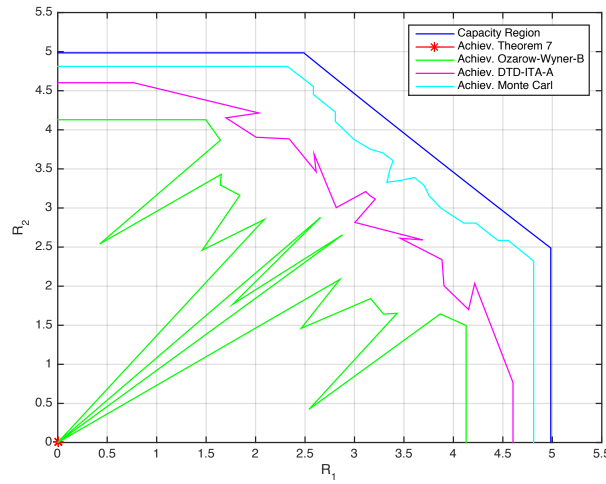
The reason why the green region has so many 'ups and downs' is because the Ozarow-Wyner-B bound in Proposition 2.1.1 depends on the constellation through its minimum distance; as we already saw in Fig. Figure 3, the minimum distance is very sensitive to the fractional values of the channel gains, which makes the corresponding bound looks very irregular. On the other hand, the magenta region is based on the lower bound in (Eq. 2.13b), which depends on the whole distance spectrum of the received constellation and as a consequence the corresponding bound looks smoother. The cyan region is the smoothest of all; its largest gap occurs at the symmetric rate point and is less than 0.7 bits – as opposed to the theoretical gap of 4.9836 bits. We thus conclude that, despite the large theoretical gap, a PAM input is quite competitive in this example.

4.7.3 Mixed (Gaussian+Discrete) vs. Discrete (Discrete+Discrete) Inputs

In the previous Sections we showed that TINnoTS with mixed (Gaussian+Discrete) inputs achieves the capacity to within a gap for several channels of interest. Practically, it may be interesting to understand what performance can be guaranteed when inputs are fully discrete, i.e., they do not contain a Gaussian component.



(a) Gap in the very strong interference regime vs. snr for $\mathsf{inr} = (\mathsf{snr}\,(1+\mathsf{snr}))^{1.2} \approx \mathsf{snr}^{2.4}$.



(b) Rate region in the strong interference regime for $\mathsf{snr} = 30$ dB and $\mathsf{inr} = \mathsf{snr}^{1.49} = 44.7$ dB.

Figure 8: Comparing Analytic with Numerical Gaps.

For the symmetric G-IC the following can be shown. Consider the TINnoTS region with $X_u \sim \mathsf{PAM}(N_u, d_u)$ such that the power constraints are met, that is, $\frac{N_u^2-1}{12}d_u^2 \leq 1$ for all $u \in [1:2]$, and lower bound the mutual informations with Proposition 2.1.1. Then, TINnoTS achieves the outer bound in Proposition 4.2.2 in very weak and in strong interference only, that is, for those regimes where 'rate splitting' was not used in Theorem 4.3.1. The proof of this result is omitted for sake of space. Thus it appears that in the moderately weak interference regime mixed inputs composed of 'two-layers' are necessary.

The next question we ask is thus whether we can show the same gap result of Theorem 4.3.1 for the moderately weak interference regime by using inputs that are the superposition of two PAM constellations, rather than a PAM and a Gaussian. The next proposition shows that the answer is in the affirmative, i.e., it is possible to 'switch' between Gaussian+Discrete and Discrete+Discrete inputs up to an additive gap.

Proposition 4.7.1. *Let*

$$\begin{aligned} X_D :=&\ X_c + X_p, \\ \textit{where } X_c \sim&\ \textit{discrete} : d_{\min(X_c)} > 0, \\ X_p \sim&\ \textit{discrete} : d_{\min(X_p)} > 0, \\ X_M :=&\ X_c + X_g, \\ \textit{where } X_g \sim&\, \mathcal{N}(0, \mathbb{E}[|X_g|^2]) \textit{ such that } \mathbb{E}[|X_p|^2] = \mathbb{E}[|X_g|^2], \end{aligned}$$

where X_c, X_g and X_p are mutually independent. Then, for $Z_G \sim \mathcal{N}(0,1)$ independent of everything else, we have

$$\begin{aligned} I(X_D; gX_D + Z_G) - I(X_M; gX_M + Z_G) &\le \frac{1}{2}\log(2), \\ I(X_M; gX_M + Z_G) - I(X_D; gX_D + Z_G) &\le \frac{1}{2}\log\left(\frac{\pi \mathrm{e}}{3}\right) + \frac{1}{2}\log\left(1 + \frac{12}{g^2\ d^2_{\min(X_D)}}\right). \end{aligned}$$

Proof. The first inequality follows since

$$
\begin{aligned}
I(X_D; gX_D + Z_G) &= I(X_\mathrm{c}, X_\mathrm{p}; gX_\mathrm{c} + gX_\mathrm{p} + Z_G) \\
&= I(X_\mathrm{p}; gX_\mathrm{c} + gX_\mathrm{p} + Z_G) + I(X_\mathrm{c}; gX_\mathrm{c} + gX_\mathrm{p} + Z_G | X_\mathrm{p}) \\
&= I\left(X_\mathrm{p}; gX_\mathrm{p} + N\right)|_{N:=gX_\mathrm{c}+Z_G} + I(X_\mathrm{c}; gX_\mathrm{c} + Z_G) \\
&\overset{(a)}{\leq} I\left(X_\mathrm{g}; gX_\mathrm{g} + N\right)|_{N:=gX_\mathrm{c}+Z_G} + \frac{1}{2}\log(2) + I(X_\mathrm{c}; gX_\mathrm{c} + Z_G) \\
&= I(X_M; gX_M + Z) + \frac{1}{2}\log(2),
\end{aligned}
$$

where in (a) we used (25, Theorem 1), which states that a Gaussian input for non-Gaussian additive noise channel results in at most $1/2$ bit loss.

The second inequality follows since

$$
\begin{aligned}
I(X_M; gX_M + Z_G) &\leq \mathsf{I}_\mathsf{g}(g^2 \mathrm{V}ar[X_M]) = \mathsf{I}_\mathsf{g}(g^2 \mathrm{V}ar[X_D]) \\
&\overset{(b)}{\leq} I(X_D; gX_D + Z_G) + \frac{1}{2}\log\left(\frac{\pi \mathrm{e}}{3}\right) + \frac{1}{2}\log\left(1 + \frac{12}{g^2\, d^2_{\min(X_D)}}\right),
\end{aligned}
$$

where in (b) we used the bound in Proposition 2.1.1. □

The question left is thus why 'two-layer' inputs, i.e., that comprise two random variables, are needed for approximate optimality in the moderately weak interference regime. Although at this point we do not have an answer for this question, the intuition for the moderately weak interference regime is as follows. With 'single-layer' PAM inputs and for the given power constraints, the number of points needed to attain a desired rate pair on the convex closure of the outer bound result in a minimum distance at the receivers that is too small. It may be that with ‘two-layer’ PAM inputs one effectively soft-estimates one of the layers whose effect can thus be removed from the received signal, thereby behaving as if there was an interfering common message jointly decoded at the non-intended receiver. Further investigation is needed to understand whether ‘multi-layer’ inputs are indeed necessary.

4.8 Conclusion

In this chapter, we evaluated a very simple, generally applicable lower bound, that neither requires joint decoding nor block synchronization, to the capacity of the Gaussian interference channel. This treating-interference-as-noise lower bound without time-sharing was evaluated for inputs that are a mixture of discrete and Gaussian random variables. We showed that, through careful choice of the mixed input parameters, namely the number of points of the discrete part and the amount of power assigned to the Gaussian part (that in general depends on the channel gains and on which point on the convex closure of the outer bound one wants to attain) the capacity of the classical Gaussian interference channel can be attained to within a gap. This result is of interest in several channels where this lower bound applies, such as block asynchronous channels and channels with partial codebook knowledge.

CHAPTER 5

ON COMMUNICATION THROUGH A GAUSSIAN CHANNEL WITH AN MMSE DISTURBANCE CONSTRAINT

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

$$\mathbf{Y} = \sqrt{\mathsf{snr}}\,\mathbf{X} + \mathbf{Z}, \tag{5.1a}$$

$$\mathbf{Y}_{\mathsf{snr}_0} = \sqrt{\mathsf{snr}_0}\,\mathbf{X} + \mathbf{Z}_0, \tag{5.1b}$$

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 $\mathbf{X}$ and $(\mathbf{Z}, \mathbf{Z}_0)$ are independent.[1] When it will be necessary to stress the SNR at $\mathbf{Y}$ in (Eq. 5.1a) we will denote it by $\mathbf{Y}_{\mathsf{snr}}$.

We also denote the mutual information[2] normalized by n as

$$I_n(\mathbf{X}, \mathsf{snr}) := \frac{1}{n} I(\mathbf{X}, \mathsf{snr}). \tag{5.2}$$

---

[1]Since there is no cooperation between receivers the capacity depends on $p_{\mathbf{Y}_1,\mathbf{Y}_2|\mathbf{X}}$ only thorough the marginals $p_{\mathbf{Y}_1|\mathbf{X}}$ and $p_{\mathbf{Y}_2|\mathbf{X}}$.

[2]In this Section, for convenience, the mutual information $I_n(\mathbf{X}, \mathsf{snr})$ is measured with respect to base e.

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 known only 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 (45).

2. *Partial interference cancellation*: by using the Han-Kobayashi (HK) achievable scheme (38), part of the interfering message is decoded and subtracted off the received signal, and the remaining part is treated as Gaussian noise. This approach has been shown to be capacity achieving in the strong interference regime (53) and optimal within 1/2 bit per channel per user otherwise (26).

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 Chapter 3 and 4 and shown to be optimal to within either a constant or a $O(\log\log(\mathsf{snr}))$ gap.

In this Chapter we look at a somewhat simplified scenario compared to the G-IC as shown in Fig. Figure 9. Formally, we aim to solve the following problem.

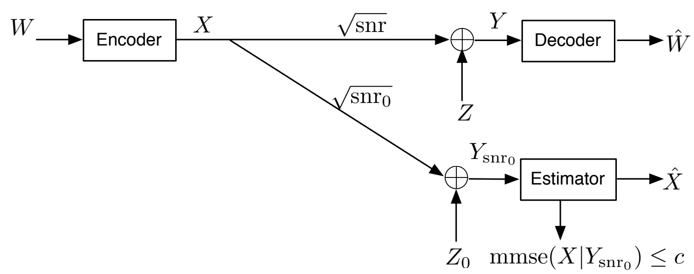


Figure 9: Channel Model.

Definition 13. *(*max-I problem.*) For some $\beta \in [0,1]$*

$$\mathcal{C}_n(\mathsf{snr}, \mathsf{snr}_0, \beta) := \sup_{\boldsymbol{X}} I_n(\boldsymbol{X}, \mathsf{snr}), \tag{5.3a}$$

$$\textit{s.t. } \frac{1}{n}\mathrm{Tr}\left(\mathbb{E}[\boldsymbol{X}\boldsymbol{X}^T]\right) \le 1, \textit{ power constraint}, \tag{5.3b}$$

$$\textit{and } \mathrm{mmse}(\boldsymbol{X}, \mathsf{snr}_0) \le \frac{\beta}{1+\beta\mathsf{snr}_0}, \textit{ MMSE constraint}. \tag{5.3c}$$

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 the G-IC (27) and (11).

In (27) the capacity of the channel in Fig. Figure 9 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 in (Eq. 5.3c) does not ‘single-letterize’. Moreover, in (11, Sec. VI.3) and (27, Sec. VIII) it was conjectured that the optimal input for $\mathcal{C}_1(\mathsf{snr}, \mathsf{snr}_0, \beta)$ is discrete.

Throughout this Chapter we will plot normalized quantities, where the normalization is with respect to the same quantity when the input is $\mathcal{N}(\mathbf{0}, \mathbf{I})$. For example, for mutual information $I_n(\mathbf{X}, \mathsf{snr})$ in (Eq. 5.2) we will plot

$$d(\mathbf{X}, \mathsf{snr}) := \frac{I_n(\mathbf{X}, \mathsf{snr})}{\frac{1}{2}\ln(1+\mathsf{snr})}, \tag{5.4}$$

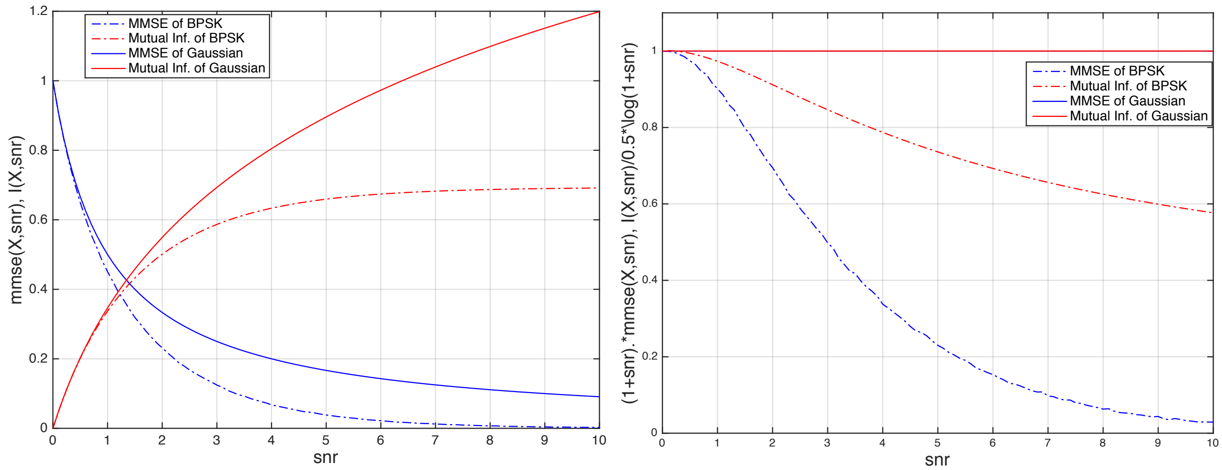
while for MMSE in (Eq. 1.6) we will plot

$$D(\mathbf{X}, \mathsf{snr}) := \frac{\mathrm{mmse}(\mathbf{X}, \mathsf{snr})}{\frac{1}{1+\mathsf{snr}}} = (1+\mathsf{snr}) \cdot \mathrm{mmse}(\mathbf{X}, \mathsf{snr}). \tag{5.5}$$

In particular, at high snr the quantity in (Eq. 5.4) is commonly referred to as the *degrees of freedom* (58) and the quantity in (Eq. 5.5) as the *MMSE dimension* (64). Moreover, it is well known that under the block-power constraint in (Eq. 5.3b), a Gaussian input maximizes both the mutual information and the MMSE (31), and thus the quantities $d(\mathbf{X}, \mathsf{snr})$, $D(\mathbf{X}, \mathsf{snr})$ have a natural meaning of multiplicative loss of the inputs $\mathbf{X}$ compared to the Gaussian input. Fig. Figure 10 compares normalized and unnormalized quantities.

5.1 Past Work and Contributions

The mutual information and the MMSE can be related, for any input $\mathbf{X}$, via the so called *I-MMSE relationship* (65, Theorem 1).



(a) Unnormalized plot.

(b) Normalized plot (solid red and blue lines overlap).

Figure 10: Comparing mutual informations and MMSE's for BPSK and Gaussian inputs. Fig. 10b clearly shows the multiplicative loss of BPSK, for both mutual information and MMSE, compared to a Gaussian input.

Proposition 5.1.1. *(*I-MMSE relationship *(65).) The I-MMSE relationship is given by the derivative relationship*

$$\frac{d}{d\mathsf{snr}} I_n(\boldsymbol{X}, \mathsf{snr}) = \frac{1}{2}\mathrm{mmse}(\boldsymbol{X}, \mathsf{snr}), \tag{5.6a}$$

or the integral relationship (65, Eq.(47))

$$I_n(\boldsymbol{X}, \mathsf{snr}) = \frac{1}{2}\int_0^{\mathsf{snr}} \mathrm{mmse}(\boldsymbol{X}, t)dt. \tag{5.6b}$$

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 5.1.2. *(*LMMSE bound (65).*) For any* $\boldsymbol{X}$ *and* $\mathsf{snr} > 0$ *it holds that*

$$\mathrm{mmse}(\boldsymbol{X}, \mathsf{snr}) \leq \frac{1}{\mathsf{snr}}. \tag{5.7a}$$

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

$$\mathrm{mmse}(\boldsymbol{X}, \mathsf{snr}) \leq \frac{\sigma^2}{1 + \sigma^2\mathsf{snr}}, \tag{5.7b}$$

where equality in (Eq. 5.7b) *is achieved iff* $\boldsymbol{X} \sim \mathcal{N}(0, \sigma^2\boldsymbol{I})$.

Another important bound for the MMSE is the *single-crossing point property* (SCPP) bound developed in (66) for $n = 1$ and extended in (67) to any $n \geq 1$.

Proposition 5.1.3. (SCPP (67).) *For any fixed* $\boldsymbol{X}$*, suppose that* $\mathrm{mmse}(\boldsymbol{X}, \mathsf{snr}_0) = \frac{\beta}{1+\beta\mathsf{snr}_0}$*, for some fixed* $\beta \geq 0$*. Then for all* $\mathsf{snr} \in [\mathsf{snr}_0, \infty)$ *we have that*

$$\mathrm{mmse}(\boldsymbol{X}, \mathsf{snr}) \leq \frac{\beta}{1 + \beta\mathsf{snr}}, \tag{5.8a}$$

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

$$\mathrm{mmse}(\boldsymbol{X}, \mathsf{snr}) \geq \frac{\beta}{1 + \beta\mathsf{snr}}. \tag{5.8b}$$

In words, Proposition 5.1.3 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 (Eq. 5.8a) and for all lower SNR values ($\mathsf{snr} \le \mathsf{snr}_0$) we have the lower bound in (Eq. 5.8b). Unfortunately, Proposition 5.1.3 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 chapter is to fill in this gap. Note that upper bounds on the MMSE are useful, thanks to the I-MMSE relationship, as tools to derive converse results, and have been used in (68), (66), (67), and (69) to name a few.

Motivated by the search for the complementary upper bound to the SCPP we define the following problem.

Definition 14. *(*max-MMSE problem.*) For some* $\beta \in [0,1]$

$$\mathrm{M}_n(\mathsf{snr}, \mathsf{snr}_0, \beta) := \sup_{\boldsymbol{X}} \mathrm{mmse}(\boldsymbol{X}, \mathsf{snr}), \tag{5.9a}$$

$$\textit{s.t. } \frac{1}{n}\mathrm{Tr}\left(\mathbb{E}[\boldsymbol{X}\boldsymbol{X}^T]\right) \le 1, \tag{5.9b}$$

$$\textit{and } \mathrm{mmse}(\boldsymbol{X}, \mathsf{snr}_0) \le \frac{\beta}{1+\beta\mathsf{snr}_0}. \tag{5.9c}$$

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 in (Eq. 5.9) and the max-I problem in (Eq. 5.3) have different objective functions but have the same constraints. This is also a good place to point out that neither of the max-MMSE and max-I problems falls under the category of convex optimization. This follows from the fact that the MMSE is a strictly concave function in the input distribution (70). Therefore, the set of input distributions, defined by (Eq. 5.3b) and (Eq. 5.3c), over which we are optimizing, might not be convex.

Note that Proposition 5.1.3 gives a solution to the max-MMSE problem in (Eq. 5.9) for $\mathsf{snr} \geq \mathsf{snr}_0$ and any $n \geq 1$ as follows:

$$\mathrm{M}_n(\mathsf{snr}, \mathsf{snr}_0, \beta) = \frac{\beta}{1+\beta\mathsf{snr}}, \text{ for } \mathsf{snr} \geq \mathsf{snr}_0, \tag{5.10}$$

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

The case $n = \infty$ of the max-MMSE problem in (Eq. 5.9) was solved in (71, Section V-C) and (27, Theorem 2) as follows:

$$\mathrm{M}_\infty(\mathsf{snr}, \mathsf{snr}_0, \beta) = \begin{cases} \frac{1}{1+\mathsf{snr}}, & \mathsf{snr} < \mathsf{snr}_0, \\ \frac{\beta}{1+\beta\mathsf{snr}}, & \mathsf{snr} \geq \mathsf{snr}_0, \end{cases}, \tag{5.11}$$

achieved by using superposition coding with Gaussian codebooks. Clearly there is a discontinuity in (Eq. 5.11) at $\mathsf{snr} = \mathsf{snr}_0$ for $\beta < 1$. This fact is a well known property of the MMSE, and it is referred to as a *phase transition* (71). It is also well known that, for any finite n, $\mathrm{mmse}(\mathbf{X}, \mathsf{snr})$ is a continuous function of snr (66). 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:

$$\mathrm{M}_n(\mathsf{snr}, \mathsf{snr}_0, \beta) = \begin{cases} \frac{1}{1+\mathsf{snr}}, & \mathsf{snr} \leq \mathsf{snr}_L, \\ T_n(\mathsf{snr}, \mathsf{snr}_0, \beta), & \mathsf{snr}_L \leq \mathsf{snr} \leq \mathsf{snr}_0, \\ \frac{\beta}{1+\beta\mathsf{snr}}, & \mathsf{snr}_0 \leq \mathsf{snr}, \end{cases} \tag{5.12}$$

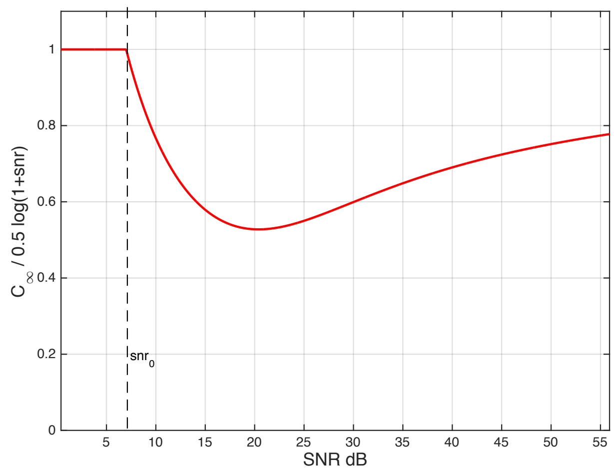


Figure 11: Plot of $\frac{\mathcal{C}_\infty(\mathsf{snr},\mathsf{snr}_0,\beta)}{\frac{1}{2}\ln(1+\mathsf{snr})}$ vs. snrdB, for $\beta = 0.01$, $\mathsf{snr}_0 = 5 = 6.989$ dB.

for some snr_L. In this Chapter we seek to characterize snr_L in (Eq. 5.12) and the continuous function $T_n(\mathsf{snr},\mathsf{snr}_0,\beta)$ such that

$$T_n(\mathsf{snr}_L,\mathsf{snr}_0,\beta) = \frac{1}{1+\mathsf{snr}_L}, \tag{5.13a}$$

$$T_n(\mathsf{snr}_0,\mathsf{snr}_0,\beta) = \frac{\beta}{1+\beta\mathsf{snr}_0}, \tag{5.13b}$$

and give scaling bounds on the width of the phase transition region defined as

$$W_n := \mathsf{snr}_0 - \mathsf{snr}_L. \tag{5.14}$$

Back to the max-I problem in (Eq. 5.3). Clearly $\mathcal{C}_n(\mathsf{snr}, \mathsf{snr}_0, \beta)$ in a non-decreasing function of n. In (27, Theorem. 3) it was shown that

$$\begin{aligned}
\mathcal{C}_\infty(\mathsf{snr}, \mathsf{snr}_0, \beta) &= \lim_{n\to\infty} \mathcal{C}_n(\mathsf{snr}, \mathsf{snr}_0, \beta), \\
&= \begin{cases} \frac{1}{2}\ln(1+\mathsf{snr}), & \mathsf{snr} \le \mathsf{snr}_0, \\ \frac{1}{2}\ln(1+\beta\mathsf{snr}) + \frac{1}{2}\ln\left(1+\frac{\mathsf{snr}_0(1-\beta)}{1+\beta\mathsf{snr}_0}\right), & \mathsf{snr} \ge \mathsf{snr}_0, \end{cases} \\
&= \frac{1}{2}\ln^+\left(\frac{1+\beta\mathsf{snr}}{1+\beta\mathsf{snr}_0}\right) + \frac{1}{2}\ln\left(1+\min(\mathsf{snr}, \mathsf{snr}_0)\right), \qquad (5.15)
\end{aligned}$$

which is achieved by using superposition coding with Gaussian codebooks. Fig. Figure 11 shows a plot of $\mathcal{C}_\infty(\mathsf{snr}, \mathsf{snr}_0, \beta)$ normalized by the capacity of the point-to-point channel $\frac{1}{2}\ln(1+\mathsf{snr})$. 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 the G-IC, but it is unknown whether it achieves capacity (26)).

The importance of studying models of communication systems with disturbance constraints has been recognized previously. For example, in (28) Bandemer *et al.* studied the following problem related to the max-I problem in (Eq. 5.3).

Definition 15. *(*Bandemer *et al.* problem.*) For some* $R \geq 0$

$$\mathcal{I}_n(\mathsf{snr}, \mathsf{snr}_0, R) := \max_{\boldsymbol{X}} I_n(\boldsymbol{X}, \mathsf{snr}), \tag{5.16a}$$

$$\textit{s.t. } \frac{1}{n}\mathrm{Tr}\left(\mathbb{E}[\boldsymbol{X}\boldsymbol{X}^T]\right) \leq 1, \tag{5.16b}$$

$$\textit{and } I_n(\boldsymbol{X}, \mathsf{snr}_0) \leq R. \tag{5.16c}$$

Observe that the max-I problems in (Eq. 5.3) and the one in (Eq. 5.16) have the same objective function but have different constraints. The relationship between the constraints in (Eq. 5.3c) and (Eq. 5.16c) can be explained as follows. The constraint in (Eq. 5.3c) imposes a maximum value on the function $\mathrm{mmse}(\mathbf{X}, \mathsf{snr})$ at $\mathsf{snr} = \mathsf{snr}_0$, while the constraint in (Eq. 5.16c), via the integral I-MMSE relationship in (Eq. 5.6), imposes a constraint on the area below the function $\mathrm{mmse}(\mathbf{X}, \mathsf{snr})$ in the range $\mathsf{snr} \in [0, \mathsf{snr}_0]$.

In (28) 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 (Eq. 5.16b) and (Eq. 5.16c) is satisfied with equality. In other words, the optimal input is i.i.d. Gaussian with power reduced such that the disturbance constraint in (Eq. 5.16c) is not violated.

Measuring the disturbance with the mutual information as in (Eq. 5.16), in contrast to the MMSE as in (Eq. 5.3), 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 (Eq. 5.16) 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 (27) and (11) that the max-I problem in (Eq. 5.3) with the MMSE disturbance constraint is a more suitable building block to study the G-IC and understand the key role of rate splitting.

5.2 Main Results

5.2.1 Max-MMSE Problem: Upper 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 (Eq. 5.12). The bound in Theorem 5.2.1 is referred to as the D-bound because it was derived through the technique of bounding the derivative of the MMSE.

Theorem 5.2.1. *(D-Bound.) For any* $\boldsymbol{X}$ *and* $0 < \mathsf{snr} \leq \mathsf{snr}_0$*, let* $\mathrm{mmse}(\boldsymbol{X}, \mathsf{snr}_0) = \frac{\beta}{1+\beta \mathsf{snr}_0}$ *for some* $\beta \in [0, 1]$*. Then*

$$\mathrm{mmse}(\boldsymbol{X}, \mathsf{snr}) \leq \mathrm{mmse}(\boldsymbol{X}, \mathsf{snr}_0) + k_n \left(\frac{1}{\mathsf{snr}} - \frac{1}{\mathsf{snr}_0} \right) - \Delta, \tag{5.17a}$$

$$k_n \leq n + 2, \ \Delta = 0. \tag{5.17b}$$

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

$$\begin{aligned} \Delta := \Delta_{\text{(Eq. 5.17c)}} &= \int_{\mathsf{snr}}^{\mathsf{snr}_0} \frac{1}{\gamma^2 (1+\gamma)^2} d\gamma. \\ &= 2 \ln \left(\frac{1+\mathsf{snr}_0}{1+\mathsf{snr}} \right) - 2 \ln \left(\frac{\mathsf{snr}_0}{\mathsf{snr}} \right) + \frac{1}{1+\mathsf{snr}} - \frac{1}{1+\mathsf{snr}_0} + \frac{1}{\mathsf{snr}} - \frac{1}{\mathsf{snr}_0}. \end{aligned} \tag{5.17c}$$

Proof. See Section 5.3.1. □

The bound on $\mathrm{M}_n(\mathsf{snr}, \mathsf{snr}_0, \beta)$ in (Eq. 5.17b) is depicted in Fig. 12a, 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 5.2.1.

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 5.2.2. *The bound in* (Eq. 5.17a) *from Theorem 5.2.1 intersects the LMMSE bound in* (Eq. 5.7a) *from Proposition 5.1.2 at*

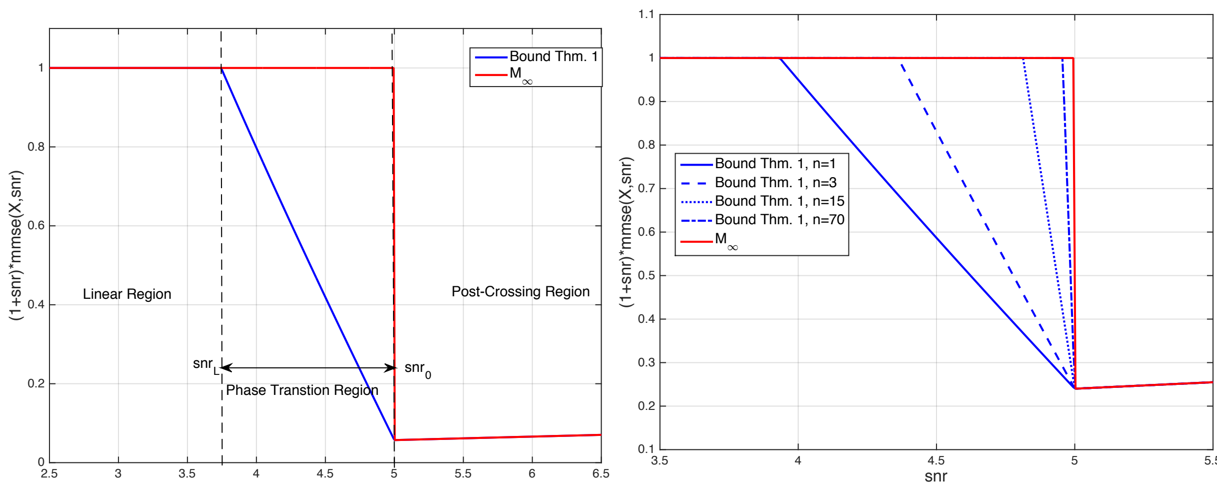
$$\mathsf{snr}_L = \mathsf{snr}_0 \frac{1 + \beta \mathsf{snr}_0}{\frac{k_n}{k_n - 1} + \beta \mathsf{snr}_0} = O\left(\left(1 - \frac{1}{n}\right) \mathsf{snr}_0\right). \tag{5.18a}$$

Thus, the width of the phase transition region is given, for k_n in (Eq. 5.17b)*, by*

$$W_n = \frac{1}{k_n - 1} \frac{\mathsf{snr}_0}{\frac{k_n}{k_n - 1} + \beta \mathsf{snr}_0} = O\left(\frac{1}{n}\right). \tag{5.18b}$$

Proof. See Appendix H.0.3. □

In Proposition 5.2.2 we found the intersection between the LMMSE bound $\frac{1}{\mathsf{snr}}$ in (Eq. 5.7a) and the bound in (Eq. 5.17a) from Theorem 5.2.1. Unfortunately, for the power constraint case, the intersection of the LMMSE bound $\frac{1}{1+\mathsf{snr}}$ in (Eq. 5.7b) and the bound in (Eq. 5.17c) cannot be found analytically. However, the solution can be computed efficiently by using numerical methods. Moreover, the asymp-



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

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

Figure 12: Bounds on $\mathrm{M}_n(\mathsf{snr}, \mathsf{snr}_0, \beta)$ vs. snr.

totic behavior of the phase transition region is still given by $O\left(\frac{1}{n}\right)$. The bound in Theorem 5.2.1 for several values of n is shown in Fig. 12b, 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 5.2.1 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 (Eq. 5.11) as $n \to \infty$.

5.2.2 Max-I Problem: Upper 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 5.2.1 we can find new upper bounds on $\mathcal{C}_n(\mathsf{snr}, \mathsf{snr}_0, \beta)$ by integration as follows:

$$\begin{aligned}
\mathcal{C}_n(\mathsf{snr}, \mathsf{snr}_0, \beta) &\leq \frac{1}{2}\int_0^{\mathsf{snr}} \mathrm{M}_n(t, \mathsf{snr}_0, \beta) dt \\
&= \frac{1}{2}\ln(1+\mathsf{snr}_L) + \frac{1}{2}\int_{\mathsf{snr}_L}^{\mathsf{snr}_0} T_n(t, \mathsf{snr}_0, \beta) dt + \frac{1}{2}\ln\left(\frac{1+\beta\mathsf{snr}}{1+\beta\mathsf{snr}_0}\right), \text{ for } \mathsf{snr}_0 \leq \mathsf{snr},
\end{aligned} \tag{5.19}$$

and

$$\begin{aligned}
\mathcal{C}_n(\mathsf{snr}, \mathsf{snr}_0, \beta) &\leq \frac{1}{2}\int_0^{\mathsf{snr}} \mathrm{M}_n(t, \mathsf{snr}_0, \beta) dt \\
&\leq \frac{1}{2}\ln(1+\min(\mathsf{snr}_L, \mathsf{snr})) + \frac{1}{2}\int_{\min(\mathsf{snr}_L, \mathsf{snr})}^{\mathsf{snr}} T_n(t, \mathsf{snr}_0, \beta) dt, \text{ for } \mathsf{snr}_0 \geq \mathsf{snr}.
\end{aligned} \tag{5.20}$$

By using Theorem 5.2.1 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 5.2.3. *For any* $0 \leq \mathsf{snr}_0$, $\beta \in [0,1]$, *and* snr_L *given in Proposition 5.2.2, we have that for* $\mathsf{snr}_0 \leq \mathsf{snr}$

$$\mathcal{C}_n(\mathsf{snr}, \mathsf{snr}_0, \beta) \leq \mathcal{C}_\infty(\mathsf{snr}, \mathsf{snr}_0, \beta) - \Delta_{\text{(Eq. 5.23)}}, \tag{5.21}$$

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

$$\mathcal{C}_n(\mathsf{snr}, \mathsf{snr}_0, \beta) \leq \mathcal{C}_\infty(\mathsf{snr}, \mathsf{snr}_0, \beta) - \Delta_{\text{(Eq. 5.24)}}, \tag{5.22}$$

where

$$0 \le \Delta_{\text{(Eq. 5.23)}} = \frac{1}{2}\ln\left(\frac{1+\mathsf{snr}_0}{1+\mathsf{snr}_L}\right) - \frac{1}{2}\frac{\beta(\mathsf{snr}_0-\mathsf{snr}_L)}{1+\beta\mathsf{snr}_0} - \frac{(n+2)}{2}\ln\left(\frac{\mathsf{snr}_0}{\mathsf{snr}_L}\right) + \frac{(n+2)(\mathsf{snr}_0-\mathsf{snr}_L)}{2\mathsf{snr}_0}$$
$$+\frac{1}{2}\left((2\mathsf{snr}_L+1)\ln\left(\frac{\mathsf{snr}_0(1+\mathsf{snr}_L)}{\mathsf{snr}_L(1+\mathsf{snr}_0)}\right) - \frac{\mathsf{snr}_0-\mathsf{snr}_L}{1+\mathsf{snr}_0} - \frac{\mathsf{snr}_0-\mathsf{snr}_L}{\mathsf{snr}_0}\right) = O\left(\frac{1}{n}\right), \tag{5.23}$$

and

$$\begin{aligned} 0 \le \Delta_{\text{(Eq. 5.24)}} &= \frac{1}{2}\ln\left(\frac{1+\mathsf{snr}}{1+\min(\mathsf{snr}_L,\mathsf{snr})}\right) - \frac{\beta(\mathsf{snr}-\min(\mathsf{snr}_L,\mathsf{snr}))}{2(1+\beta\mathsf{snr}_0)} \\ &- \frac{(n+2)}{2}\ln\left(\frac{\mathsf{snr}}{\min(\mathsf{snr}_L,\mathsf{snr})}\right) + \frac{(n+2)(\mathsf{snr}-\min(\mathsf{snr}_L,\mathsf{snr}))}{2\mathsf{snr}_0} \\ &+ \frac{1}{2}\left((2\min(\mathsf{snr}_L,\mathsf{snr})+1)\ln\left(\frac{1+\min(\mathsf{snr}_L,\mathsf{snr})}{\min(\mathsf{snr}_L,\mathsf{snr})}\right) - (2\mathsf{snr}+1)\ln\left(\frac{1+\mathsf{snr}}{\mathsf{snr}}\right)\right. \\ &\left.+ 2(\mathsf{snr}-\min(\mathsf{snr}_L,\mathsf{snr}))\ln\left(\frac{1+\mathsf{snr}_0}{\mathsf{snr}_0}\right) - \frac{\mathsf{snr}-\min(\mathsf{snr}_L,\mathsf{snr})}{\mathsf{snr}_0} - \frac{\mathsf{snr}-\min(\mathsf{snr}_L,\mathsf{snr})}{1+\mathsf{snr}_0}\right) \\ &= O\left(\frac{1}{n}\right). \end{aligned} \tag{5.24}$$

Fig. Figure 13 compares the bounds on $\mathcal{C}_n(\mathsf{snr},\mathsf{snr}_0,\beta)$ in (Eq. 5.15) from Proposition 5.2.3 with $\mathcal{C}_\infty(\mathsf{snr},\mathsf{snr}_0,\beta)$ for several values of n. The figure shows how the new bounds in Proposition 5.2.3 improve on the trivial $\mathcal{C}_\infty(\mathsf{snr},\mathsf{snr}_0,\beta)$ bound for finite n.

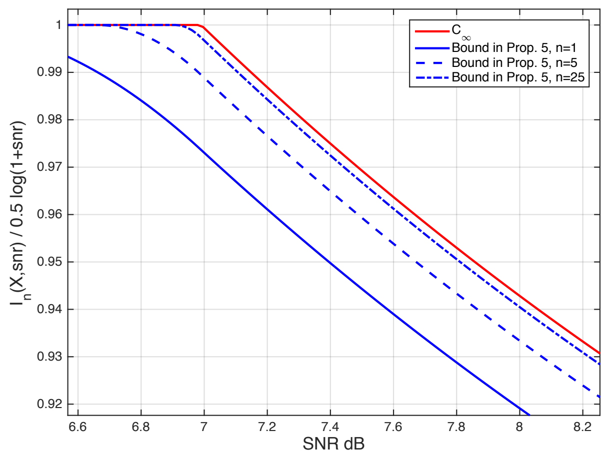


Figure 13: Bounds on $\mathcal{C}_n(\mathsf{snr}, \mathsf{snr}_0, \beta)$ vs. snr, for $\beta = 0.1$ and $\mathsf{snr}_0 = 5 = 6.9897$ dB.

5.2.3 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

$$X_{\text{mix}} := \sqrt{1-\delta} X_D + \sqrt{\delta} X_G, \ \delta \in [0,1], \tag{5.25}$$

where X_G and X_D are independent, $X_G \sim \mathcal{N}(0,1)$, $\mathbb{E}[X_D^2] \leq 1$, and where the distribution of X_D and the parameter δ are to be optimized over. The input X_{mix} exhibits a decomposition property via which the MMSE and the mutual information can be written as the sum of the MMSE and the mutual information of the X_D and X_G components, albeit at different SNR values.

Proposition 5.2.4. *For $X_{\rm mix}$ defined in* (Eq. 5.25) *we have that*

$$I(X_{\rm mix}, \mathsf{snr}) = I\left(X_D, \frac{\mathsf{snr}(1-\delta)}{1+\delta\mathsf{snr}}\right) + I(X_G, \mathsf{snr}\ \delta), \tag{5.26a}$$

$$\mathrm{mmse}(X_{\rm mix}, \mathsf{snr}) = \frac{1-\delta}{(1+\mathsf{snr}\delta)^2}\mathrm{mmse}\left(X_D, \frac{\mathsf{snr}(1-\delta)}{1+\delta\mathsf{snr}}\right) + \delta\ \mathrm{mmse}(X_G, \mathsf{snr}\ \delta). \tag{5.26b}$$

Proof. See Appendix K. □

Observe that Proposition 5.2.4 implies that, in order for mixed inputs (with $\delta < 1$) to comply with the MMSE constraint in (Eq. 5.3c) and (Eq. 5.9c), the MMSE of X_D must satisfy

$$\mathrm{mmse}\left(X_D, \frac{\mathsf{snr}_0(1-\delta)}{1+\delta\mathsf{snr}_0}\right) \le \frac{(\beta-\delta)(1+\delta\mathsf{snr}_0)}{(1-\delta)(1+\beta\mathsf{snr}_0)}. \tag{5.27}$$

The bound in (Eq. 5.27) will be helpful in choosing the parameter δ later on.

When X_D is a discrete random variable with $\mathsf{supp}(X_D) = N$ we use the following bounds from (72, App. C) and Chapter 2 in Remark 2.

Proposition 5.2.5. *For a discrete random variable X_D such that $p_i = \Pr(X_D = x_i)$, for $i \in [1:N]$, we have that*

$$\mathrm{mmse}(X_D, \mathsf{snr}) \le d_{\max}^2 \sum_{i=1}^{N} p_i \mathrm{e}^{-\frac{\mathsf{snr}}{8}d_i^2}, \tag{5.28a}$$

$$I(X_D, \mathsf{snr}) \ge H(X_D) - \frac{1}{2}\ln\left(\frac{\pi}{6}\right) - \frac{1}{2}\ln\left(1 + \frac{12}{d_{\min}^2}\mathrm{mmse}(X_D, \mathsf{snr})\right), \tag{5.28b}$$

where

$$d_\ell := \min_{x_i \in \mathsf{supp}(X_D): i \neq \ell} |x_\ell - x_i|, \tag{5.28c}$$

$$d_{\min} := \min_{\ell \in [1:N]} d_\ell, \tag{5.28d}$$

$$d_{\max} := \max_{x_k, x_i \in \mathsf{supp}(X_D)} |x_k - x_i|. \tag{5.28e}$$

Proposition 5.2.4 and Proposition 5.2.5 are particularly useful because they will allow us to design Gaussian and discrete components of the mixed input independently.

Fig. Figure 14 shows upper and lower bounds on $\mathrm{M}_1(\mathsf{snr}, \mathsf{snr}_0, \beta)$ where we show the following:

- The upper bound in (Eq. 5.11) (solid red line);
- The upper bound from Theorem 5.2.1 (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 (Eq. 5.25). We used Proposition 5.2.4 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 the 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 (Eq. 5.27). That is, we chose δ such that the power of X_G is approximately β while the MMSE constraint on X_D in (Eq. 5.27) is not equal to zero. The input X_D used in Fig. Figure 14 was found by a local search algorithm on the space of distributions with $N = 3$, and resulted in $X_D = [-1.8412, -1.7386, 0.5594]$ with $P_X = [0.1111, 0.1274, 0.7615]$, which we do not claim to be optimal;

- The discrete-only input lower bound (Discrete 1 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, and to compare the performance of the MMSE of the discrete component of the mixed input with and without the Gaussian component; and
- The discrete-only input lower bound (Discrete 2 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.

The choice of $N = 3$ is motivated by the fact that it requires roughly $N = \lfloor\sqrt{1 + \mathsf{snr}_0}\rfloor$ points for the PAM input to approximately achieve capacity of the point to point channel with SNR value snr_0.

On the one hand, Fig. Figure 14 shows that, for $\mathsf{snr} \geq \mathsf{snr}_0$, a Gaussian-only input with power reduced to β maximizes $\mathrm{M}_1(\mathsf{snr}, \mathsf{snr}_0, \beta)$ in agreement with the SCPP bound (green line). On the other hand, for $\mathsf{snr} \leq \mathsf{snr}_0$, 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 (Eq. 5.28a)) to comply with the MMSE constraint. However, for $\mathsf{snr} \geq \mathsf{snr}_0$, the behavior of the MMSE of discrete-only inputs, as opposed to mixed inputs, prevents it from being optimal; this is due to their exponential tail behavior in (Eq. 5.28a). This further motivates determining whether the MMSE constraint can imply a power constraint, which we shall investigate in Section 5.4.

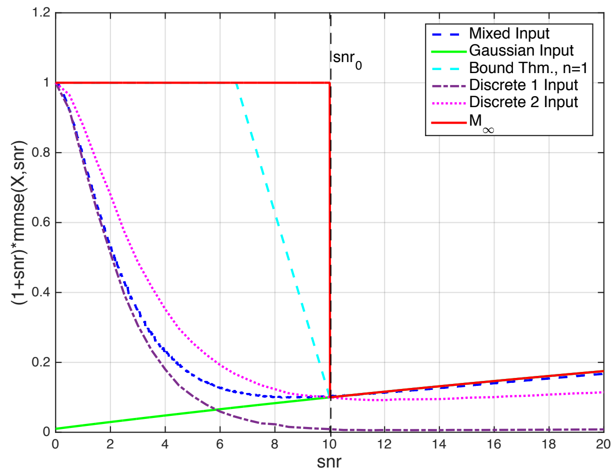


Figure 14: Upper and lower bounds on $\mathrm{M}_1(\mathsf{snr}, \mathsf{snr}_0, \beta)$ vs. snr, for $\beta = 0.01$, $\mathsf{snr}_0 = 10$.

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 $\mathsf{snr} \geq \mathsf{snr}_0$ (without any reduction in power) and the behavior of discrete-only inputs for $\mathsf{snr} \leq \mathsf{snr}_0$. 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. Figure 14 shows the achievable MMSE with another discrete-only input (dotted magenta line) that achieves higher MMSE than the mixed input for $\mathsf{snr} \leq \mathsf{snr}_0$ but lower than the mixed input for $\mathsf{snr} \geq \mathsf{snr}_0$. This is again due to the tail behavior of the MMSE of discrete inputs. The reason this second discrete input is not used as a component of the mixed inputs, is because this choice would violate the MMSE constraint on X_D in (Eq. 5.27).

TABLE III: PARAMETERS OF THE MIXED INPUT IN (**??**) USED IN THE PROOF OF PROPOSITION **??**.

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

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.

5.2.4 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 (Eq. 5.25) is to within an additive gap of the outer bound in Proposition 5.2.3.

Proposition 5.2.6. *A lower bound on $\mathcal{C}_1(\mathsf{snr}, \mathsf{snr}_0, \beta)$ with the mixed input in* (Eq. 5.25)*, with $X_D \sim \mathsf{PAM}(N)$ and with input parameters as specified in Table III, is to within $O\left(\log\log(\frac{1}{\mathrm{mmse}(X,\mathsf{snr}_0)}\right)$ of the outer bound in Proposition 5.2.3 with the exact gap value given by*

$$\mathsf{snr} \geq \mathsf{snr}_0 \geq 1 : C_1(\mathsf{snr}, \mathsf{snr}_0, \beta) - I_1(X_{\mathrm{mix}}, \mathsf{snr}) \leq \mathsf{G}_{\mathsf{d}1}, \tag{5.29a}$$

$$\mathsf{snr}_0 \geq \mathsf{snr} \geq 1 : C_1(\mathsf{snr}, \mathsf{snr}_0, \beta) - I_1(X_{\mathrm{mix}}, \mathsf{snr}) \leq \mathsf{G}_{\mathsf{d}2}, \tag{5.29b}$$

$$\mathsf{snr} \leq 1 : C_1(\mathsf{snr}, \mathsf{snr}_0, \beta) - I_1(X_{\mathrm{mix}}, \mathsf{snr}) \leq \frac{1}{2}\ln(2), \tag{5.29c}$$

where

$$\mathsf{G}_{\mathsf{d}1} := \frac{1}{2}\ln\left(\frac{2}{3}\ln\left(\frac{24(1+(1-\beta)\mathsf{snr}_0)}{\beta}\right)+\frac{6\beta}{1+\beta\mathsf{snr}_0}\right)+\frac{1}{2}\ln\left(\frac{4\pi}{3}\right)-\Delta_{\text{(Eq. 5.23)}}, \tag{5.29d}$$

$$\mathsf{G}_{\mathsf{d}2} := \frac{1}{2}\ln\left(1+\frac{2}{3}\ln\left(\frac{12(1+\beta\mathsf{snr}_0)}{\beta}\right)\right)+\frac{1}{2}\ln\left(\frac{4\pi}{6}\right)-\Delta_{\text{(Eq. 5.24)}}. \tag{5.29e}$$

and $\Delta_{\text{(Eq. 5.23)}}$ *and* $\Delta_{\text{(Eq. 5.24)}}$ *are given in* (Eq. 5.23) *and* (Eq. 5.24)*, respectively.*

Proof. See Appendix L. □

Please note that the gap result in Proposition 5.2.6 is constant in snr (i.e., independent of snr) but not in snr_0.

Fig. Figure 15 compares the inner bounds on $\mathcal{C}_1(\mathsf{snr}, \mathsf{snr}_0, \beta)$, normalized by the point-to-point capacity $1/2\ln(1+\mathsf{snr})$, with mixed inputs (dashed magenta line) in Proposition 5.2.6 to:

- The upper bound in (Eq. 5.15), (solid red line);
- The upper bound from Proposition 5.2.3 (dashed blue line); and
- 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. Figure 15 shows that Gaussian inputs are sub-optimal and that mixed inputs achieve large degrees of freedom 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 mixed input is used. This in particular supports the conjecture in (27) that discrete inputs may be optimal for $n = 1$ and $\mathsf{snr} \leq \mathsf{snr}_0$.

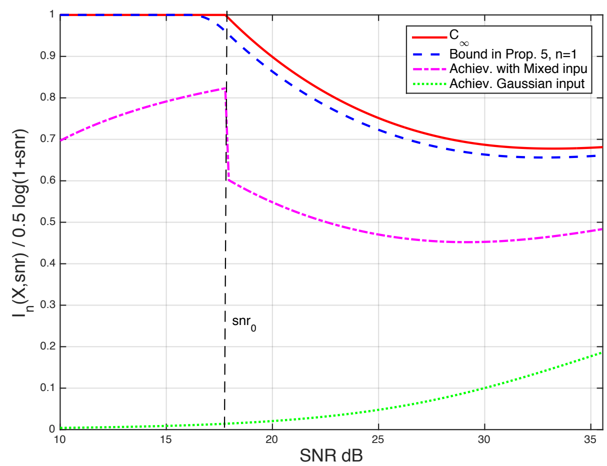


Figure 15: Upper and lower bounds on $\mathcal{C}_{n=1}(\mathsf{snr}, \mathsf{snr}_0, \beta)$ vs. snr, for $\beta = 0.001$ and $\mathsf{snr}_0 = 60 = 17.6815$ dB.

The above discussion completes the presentation of our bounds on max-I and max-MMSE problems. The remainder of the Chapter contains the proof of Theorem 5.2.1 and a discussion of when the MMSE constraint necessarily implies a power constraint.

5.3 Properties of the First Derivative of the MMSE

A key element in the proof of the SCPP in Proposition 5.1.3 was the characterization of the first derivative of the MMSE as

$$-\frac{d\mathrm{mmse}(\mathbf{X}, \mathsf{snr})}{d\mathsf{snr}} = \frac{1}{n}\mathrm{Tr}\left(\mathbb{E}\left[\mathbf{Cov}^2(\mathbf{X}|\mathbf{Y})\right)\right) := \frac{1}{n}\mathrm{Tr}\left(\mathbb{E}\left[\mathbf{Cov}^2(\mathbf{X}, \mathsf{snr})\right]\right), \tag{5.30}$$

which was given in (66, Proposition 9) for $n = 1$ and in (67, Lemma 3) for $n \geq 1$. The first derivative in (Eq. 5.30) turns out to be instrumental in proving Theorem 5.2.1 as well.

For ease of presentation, in the rest of the section, instead of focusing on the derivative 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 (66, Proposition 5) that

$$\mathbb{E}\left[\mathrm{Cov}^2(X|Y)\right] \leq \frac{k_1}{\mathsf{snr}^2}, \text{ where } k_1 \leq 3 \cdot 2^4. \tag{5.31}$$

Before using (Eq. 5.30) in the proof of Theorem 5.2.1, we will need to sharpen the existing constant for $n = 1$ in (Eq. 5.31) (given by $k_1 \leq 3 \cdot 2^4$) and generalize the bound to any $n \geq 1$, which to the best of our knowledge has not been considered before.

Proposition 5.3.1. *For any* $\boldsymbol{X}$ *and* $\mathsf{snr} > 0$ *we have*

$$\frac{1}{n}\mathrm{Tr}\left(\mathbb{E}[\mathbf{Cov}^2(\boldsymbol{X}|\boldsymbol{Y})]\right) \leq \frac{k_n}{\mathsf{snr}^2}, \tag{5.32a}$$

where

$$k_n \leq \frac{n(n+2) - n\,\mathrm{mmse}(\boldsymbol{Z}\boldsymbol{Z}^T|\boldsymbol{Y}) - \mathrm{Tr}\left(\boldsymbol{J}^2(\boldsymbol{Y})\right)}{n} \leq n+2. \tag{5.32b}$$

Proof. See Appendix M. □

In Proposition 5.3.1 the bound on k_1 in (Eq. 5.31) has been tightened from $k_1 \leq 3 \cdot 2^4$ in (Eq. 5.31) 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 5.3.2. *If $\boldsymbol{X}$ is such that $\frac{1}{n}\mathrm{Tr}\left(\mathbb{E}\left[\boldsymbol{X}\boldsymbol{X}^T\right]\right) \leq 1$, then*

$$\mathrm{Tr}(\boldsymbol{J}^2(\boldsymbol{Y})) \geq \frac{n}{(1+\mathsf{snr})^2}. \tag{5.33}$$

The equality in (Eq. 5.33) *is achieved if $\boldsymbol{X} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I})$.*

Proof. See Appendix N. □

Observe that, by using the bound in (Eq. 5.32) from Proposition 5.3.1 together with the lower bound on the Fisher information in Proposition 5.3.2, the bound on the constant k_n in (Eq. 5.32b) can be tightened to

$$k_n \leq \frac{n(n+2) - \frac{n}{(1+\mathsf{snr})^2}}{n} = n + 2 - \frac{1}{(1+\mathsf{snr})^2}. \tag{5.34}$$

By further assuming that $\mathbf{X}$ has a finite fourth moment we can arrive at the following bound that does not blow up around $\mathsf{snr} = 0^+$, as opposed to the bound in (Eq. 5.32a).

Proposition 5.3.3. *If $\boldsymbol{X}$ such that $\frac{1}{n}\mathrm{Tr}\left(\mathbb{E}\left[\left(\boldsymbol{X}\boldsymbol{X}^T\right)^2\right]\right) < \infty$ then*

$$\begin{aligned}&\mathrm{Tr}\left(\mathbb{E}[\mathbf{Cov}^2(\boldsymbol{X}|\boldsymbol{Y})]\right)\\ &\leq \min\left(\frac{\mathrm{Tr}\left(\mathbb{E}\left[\left(\left(\boldsymbol{X}-\sqrt{\mathsf{snr}}\boldsymbol{Z}\right)\left(\boldsymbol{X}-\sqrt{\mathsf{snr}}\boldsymbol{Z}\right)^T\right)^2\right]\right)}{(1+\mathsf{snr})^4}, \mathrm{Tr}\left(\mathbb{E}\left[\mathbb{E}^2\left[\boldsymbol{X}\boldsymbol{X}^T|\boldsymbol{Y}\right]\right]\right)\right),\end{aligned} \tag{5.35a}$$

where we can further bound

$$\mathrm{Tr}\left(\mathbb{E}\left[\mathbb{E}^2\left[\boldsymbol{X}\boldsymbol{X}^T|\boldsymbol{Y}\right]\right]\right) \leq \mathrm{Tr}\left(\mathbb{E}\left[\left(\boldsymbol{X}\boldsymbol{X}^T\right)^2\right]\right). \tag{5.35b}$$

Proof. See Appendix O. □

Note that evaluation of the first term of the minimum in (Eq. 5.35a) requires only the knowledge of second and fourth moments of **X**.

We are now ready to prove our main result.

5.3.1 Proof of Theorem 5.2.1

The proof of Theorem 5.2.1 relies on the fact that the MMSE is an infinitely differentiable function of snr (66, Proposition 7) and therefore can be written as the difference of two MMSE functions using the fundamental theorem of calculus

$$\begin{aligned}
&\mathrm{mmse}(\mathbf{X}, \mathsf{snr}) - \mathrm{mmse}(\mathbf{X}, \mathsf{snr}_0) \\
&= -\int_{\mathsf{snr}}^{\mathsf{snr}_0} \mathrm{mmse}'(\mathbf{X}, \gamma) d\gamma \\
&\overset{a)}{=} \int_{\mathsf{snr}}^{\mathsf{snr}_0} \frac{1}{n} \mathrm{Tr}\left(\mathbb{E}[\mathbf{Cov}^2(\mathbf{X}, \gamma)]\right) d\gamma \\
&\overset{b)}{\leq} \int_{\mathsf{snr}}^{\mathsf{snr}_0} \frac{(n+2)}{\gamma^2} d\gamma = (n+2)\left(\frac{1}{\mathsf{snr}} - \frac{1}{\mathsf{snr}_0}\right) - \Delta,\ \Delta = 0,
\end{aligned}$$

where the (in)-equalities follow by using: a) (Eq. 5.30), and b) the bound in Proposition 5.3.1 with $k_n \leq n+2$. If we further assume that **X** has finite power, instead of bounding $k_n \leq n+2$, we can use (Eq. 5.34), to obtain

$$0 \leq \Delta = \Delta_{\text{(Eq. 5.17c)}} = \int_{\mathsf{snr}}^{\mathsf{snr}_0} \frac{1}{\gamma^2(1+\gamma)^2} d\gamma.$$

This concludes the proof of Theorem 5.2.1.

5.4 When Does an MMSE Constraint Imply a Power Constraint

In this section we try to determine whether the MMSE constraint may imply a power constraint. For simplicity we focus on the case of $n = 1$. This question is motivated by the following limit, which exists iff $\mathbb{E}[X^2] < \infty$:

$$\lim_{\mathsf{snr} \to 0^+} \mathrm{mmse}(X, \mathsf{snr}) = \mathbb{E}[X^2]. \tag{5.36}$$

The limit in (Eq. 5.36) raises the question of whether the MMSE constraint at snr_0 around zero would imply a power constraint. In other words, are we required to reduce power to meet the MMSE constraint for very small snr_0? Surprisingly, the answer to this question is no.

Proposition 5.4.1. *There exists an input distribution X with maximum power as in* (Eq. 5.3b) *that satisfies the MMSE constraint in* (Eq. 5.3c) *for any $\mathsf{snr}_0 > 0$ and any $\beta > 0$.*

Proof. Consider an input distribution given by

$$X_a = [-a, 0, a], \quad P_{X_a} = \left[\frac{1}{2a^2}, 1 - \frac{1}{a^2}, \frac{1}{2a^2}\right], \tag{5.37}$$

for any $a \geq 1$. Note that for the input distribution in (Eq. 5.37) $\mathbb{E}[X_a^2] = 1$ for any a. The MMSE of X_a can be upper bounded by

$$\mathrm{mmse}(X_a, \mathsf{snr}) \leq \min\left(1, 4(a^2+1)e^{-\frac{a^2\mathsf{snr}}{8}}\right), \tag{5.38}$$

where the upper bound in (Eq. 5.38) follows by applying the upper bound in Proposition 5.2.5 together with the bound $\mathrm{mmse}(X_a, \mathsf{snr}) \leq \mathbb{E}[X_a^2] = 1$. Therefore, by choosing a large enough, any MMSE constraint can be met while transmitting at full power. This concludes the proof. □

The MMSE of X_a is shown on Fig. Figure 16. Here are some other properties of X_a that are easy to verify.

Proposition 5.4.2. *The random variable X_a has the following properties*

- $\lim_{a\to\infty} X_a = 0$ *almost surely (a.s.),*
- $\mathbb{E}[|X_a - 0|^n] = a^n p = E[X_a^2]a^{n-2} = a^{n-2}$.

The random variable X_a serves as a counterexample that shows that a.s. convergence does not imply L^p convergence.

An interesting question is whether we can characterize a family of input distributions for which the MMSE constraint implies a power constraint under some non-trivial condition. In other words, we

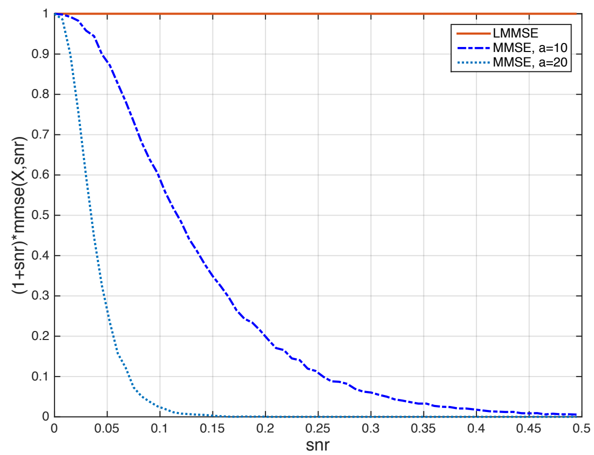


Figure 16: $\mathrm{mmse}(X_a, \mathsf{snr})$ vs. snr, for $a = 10$ and $a = 20$.

want to find a family of input distributions such that the power constraint can be related to the MMSE constraint at some snr_0, that is

$$\mathbb{E}[X^2] = f(\mathrm{mmse}(X, \mathsf{snr}_0)) \leq 1. \tag{5.39}$$

Towards this end we have the following:

Proposition 5.4.3. *For any X and any $\mathsf{snr}_0 \geq \mathsf{snr} > 0$, we have that*

$$\mathrm{mmse}(X, \mathsf{snr}) = \mathrm{mmse}(X, \mathsf{snr}_0) + k \cdot (\mathsf{snr}_0 - \mathsf{snr}), \tag{5.40}$$

where for some $\mathsf{snr}_c \in (\mathsf{snr}, \mathsf{snr}_0]$

$$k = \mathbb{E}\left[\mathrm{Cov}^2(X, \mathsf{snr}_c)\right] \leq \sup_{\gamma \in (\mathsf{snr}, \mathsf{snr}_0)} \mathbb{E}[\mathrm{Cov}^2(X, \gamma)] \leq \mathbb{E}[X^4]. \tag{5.41}$$

Moreover, for $\mathsf{snr} = 0^+$ *the inequality in* (Eq. 5.40) *is valid iff*

$$\lim_{\mathsf{snr} \to 0^+} \mathbb{E}[\mathrm{Cov}^2(X, \mathsf{snr})] < \infty. \tag{5.42}$$

Proof. The result easily follows by applying the mean value theorem

$$\begin{aligned} \mathrm{mmse}(X, \mathsf{snr}_0) - \mathrm{mmse}(X, \mathsf{snr}) &= \int_{\mathsf{snr}}^{\mathsf{snr}_0} \mathbb{E}[\mathrm{Cov}^2(X, \gamma)] d\gamma \\ &= \mathbb{E}[\mathrm{Cov}^2(X, \mathsf{snr}_c)](\mathsf{snr}_0 - \mathsf{snr}). \end{aligned} \tag{5.43}$$

for some $\mathsf{snr}_c \in (\mathsf{snr}, \mathsf{snr}_0)$. Note that for $\mathsf{snr} > 0$ the quantity $\mathbb{E}[\mathrm{Cov}^2(X, \gamma)]$ is finite due to Proposition 5.3.1. Therefore, we focus on the case when $\mathsf{snr} = 0^+$.

Observe that since the $\mathrm{mmse}(X, \mathsf{snr})$ is an analytic function of SNR for $\mathsf{snr} > 0$, its derivative or $\mathbb{E}[\mathrm{Cov}^2(X, \gamma)]$ is also an analytic function of $\mathsf{snr} > 0$. Therefore, if $\lim_{\mathsf{snr} \to 0^+} \mathbb{E}[\mathrm{Cov}^2(X, \mathsf{snr})] = K < \infty$ for some $K > 0$, by Jensen's inequality we have that

$$K = \lim_{\mathsf{snr} \to 0^+} \mathbb{E}[\mathrm{Cov}^2(X, \mathsf{snr})] \geq (\mathbb{E}[X^2])^2 = (\mathrm{mmse}(X, 0))^2. \tag{5.44}$$

So, in other words the existence of the derivative at $\mathsf{snr} = 0^+$ implies the existence of the power constraint and the integration in (Eq. 5.43) holds for $\mathsf{snr} = 0^+$.

Conversely, if the integration in (Eq. 5.43) is finite for $\mathsf{snr} = 0^+$ we have that

$\lim_{\mathsf{snr} \to 0^+} \mathbb{E}[\mathrm{Cov}^2(X, \mathsf{snr})] < \infty$.

Therefore, the bound in (Eq. 5.40) holds iff $\lim_{\mathsf{snr} \to 0^+} \mathbb{E}[\mathrm{Cov}^2(X, \mathsf{snr})] < \infty$. This concludes the proof. $\square$

From Proposition 5.4.3 we see that necessary and sufficient conditions for the MMSE at snr_0 to imply a reduction in power (i.e., $\mathbb{E}[X^2] < 1$) are

$$1)\ \mathrm{mmse}(X, \mathsf{snr}_0) + \mathsf{snr}_0 \cdot \mathbb{E}[\mathrm{Cov}^2(X, \mathsf{snr}_c)] < 1,$$

$$\Leftrightarrow \mathbb{E}[\mathrm{Cov}^2(X, \mathsf{snr}_c)] < \frac{1 - \mathrm{mmse}(X, \mathsf{snr}_0)}{\mathsf{snr}_0}, \tag{5.45a}$$

$$2)\ \lim_{\mathsf{snr} \to 0^+} \mathbb{E}[\mathrm{Cov}^2(X, \mathsf{snr})] < \infty, \tag{5.45b}$$

where snr_c is defined in Proposition 5.4.3.

Since snr_c might be difficult to compute, the following slightly stronger (i.e., sufficient condition) can be useful:

$$\sup_{\gamma \in (0, \mathsf{snr}_0)} \mathbb{E}[\mathrm{Cov}^2(X, \gamma)] < \frac{1 - \mathrm{mmse}(X, \mathsf{snr}_0)}{\mathsf{snr}_0}. \tag{5.46}$$

Finally, observe that $\lim_{a\to\infty} X_a$ does not satisfy this moment condition since

$$\lim_{a\to\infty} \mathbb{E}[\mathrm{Cov}^2(X_a|Y)] = \begin{cases} \infty & \mathsf{snr} = 0, \\ 0 & \mathsf{snr} > 0. \end{cases} \tag{5.47}$$

5.5 Conclusion

In this Chapter 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 demonstrated a matching lower bound that is to 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 SCPP of the MMSE and may be of independent interest.

CHAPTER 6

CONCLUSION

In this work we have develop several techniques for evaluating the approximate performance of discrete or more generally mixed inputs. The developed framework was then applied to several multi-user channels that are of interest in several information theoretic applications.

The first application of discrete and mixed inputs was in the context of the G-IC-OR. At first glance it appears that a system with oblivious nodes should suffer a considerable degradation since the oblivious receiver is incapable of doing sophisticated decoding techniques such as joint decoding or successive interference cancelation. The results of our work demonstrate that this intuition is wrong and almost no performance degradation takes place.

The intuition behind this result is given in Chapter 2 with an example of the channel with mixed noise (i.e. discrete+Gaussian). It was shown, provided that the minimum distance of the discrete component of the noise does not go to zero, that the discrete component of the noise has essentially no effect on the performance, and only the Gaussian part of the noise has an effect.

This observation was then used to show a quite surprising result – that a simple communication scheme of treating interference as noise is approximately optimal in the classical G-IC. This result has potential applications in wireless communications where treating interference as noise, albeit as Gaussian noise, has been the industry's standard. Our result suggests that accounting for the correct distribution of the noise can bring considerable gains.

In the last part of the thesis we gave an estimation theoretic explanation for the performance of the discrete inputs. In the past it has been recognized that soft-decoding strategies, in which the interference is estimated rather than decodes, have be advantageous. For example, soft-decoding strategies are easy to implement and have low computational complexity. However, no performance guarantees have been provided for such strategies. Interestingly, our technique of using discrete inputs falls into the category of soft-decoding, and this work gives a first performance guarantee result. That is, we have shown that a class of soft-decoding strategies is approximately optimal.

The interplay between estimation theoretic and information theoretic measures offers many potential advantages. For example, recently in (30), using an estimation framework under a non-quadric cost function, we have sharpened the gap term in the Ozarow-Wyner bound. This consequently improves all of the gap results given in this thesis.

Finally, we would like to mention that many of our results have potential applications in other fields. For example, the new MMSE bound in Chapter 5 has potential applications in the field of *statistical physics*. In particular, using recent connections (65) and (71) between information theory, estimation theory, and statistical physics, the bound in Chapter 5 can be related to the intrinsic free energy of thermodynamic systems with n particles. In many applications in statistical physics one is interested in computing thermodynamic limits as the number of particles goes to infinity (i.e. $n \to \infty$). Our new bound can potentially be useful for determining the thermodynamic limits, which are notoriously difficult to compute. In the past, this obstacle has led statistical physicists to resort to techniques such as the *replica method* or the *cavity method*, which lack mathematical rigor in general (73).

APPENDICES

Appendix A

PROOF OF EQ. 2.13

To prove the lower bound in (Eq. 2.13) we first find a lower bound on the differential entropy of output $Y = X_D + Z_G$. To that end let $p_i := \mathbb{P}[X_D = s_i], i \in [1 : N]$ then Y has the following Gaussian mixture density

$$Y \sim P_Y(y) := \sum_{i \in [1:N]} p_i \mathcal{N}(y; s_i, 1). \tag{A.1}$$

Appendix A (Continued)

where $\mathcal{N}(y; s_i, 1)$. We have

$$
\begin{aligned}
-h(Y) &= \int P_Y(y)\log(P_Y(y))dy \\
&\overset{(a)}{\leq} \log \int P_Y(y)P_Y(y)dy \\
&= \log \int \left(\sum_{i\in[1:N]} p_i \mathcal{N}(y; s_i, 1) \right)^2 dy \\
&= \log \left(\sum_{(i,j)\in[1:N]^2} p_i p_j \int \mathcal{N}(y; s_i, 1)\mathcal{N}(y; s_j, 1)dy \right) \\
&= \log \left(\sum_{(i,j)\in[1:N]^2} p_i p_j \frac{1}{\sqrt{4\pi}} \mathrm{e}^{\frac{-(s_i-s_j)^2}{4}} \int \mathcal{N}\left(y; \frac{s_i+s_j}{2}, \frac{1}{2}\right) dy \right) \\
&\overset{(b)}{=} \log \left(\sum_{(i,j)\in[1:N]^2} p_i p_j \frac{1}{\sqrt{4\pi}} \mathrm{e}^{-\frac{(s_i-s_j)^2}{4}} \right) \\
&\overset{(c)}{\leq} \log \left(\sum_{i\in[1:N]} p_i^2 \frac{1}{\sqrt{4\pi}} + \sum_{i\in[1:N]} p_i(1-p_i) \frac{1}{\sqrt{4\pi}} \mathrm{e}^{-\frac{d^2_{\min(X_D)}}{4}} \right) \\
&\overset{(d)}{\leq} -\log(N\sqrt{4\pi}) + \log\left(1 + (N-1)\mathrm{e}^{-\frac{d^2_{\min(X_D)}}{4}}\right),
\end{aligned}
$$

which implies

$$
\begin{aligned}
I(X_D; X_D + Z_G) = h(Y) - h(Z_G) \geq \log(N) - \mathsf{gap}_{\mathrm{AD}}, \\
\mathsf{gap}_{\mathrm{AD}} := \frac{1}{2}\log\left(\frac{\mathrm{e}}{2}\right) + \log\left(1 + (N-1)\mathrm{e}^{-\frac{d^2_{\min(X_D)}}{4}}\right),
\end{aligned}
\tag{A.2}
$$

Appendix A (Continued)

where the (in)-equalities follow from: (a) Jensen's inequality, (b) $\int \mathcal{N}(y; \mu, \sigma^2)dy = 1$, (c) $d_{\min(X_D)} \leq |s_i - s_j|, \forall i \neq j$, (d) by maximizing over the $\{p_i, i \in [1 : N]\}$. Combining this bound with the fact that mutual information is non-negative proves the claimed lower bound.

Appendix B

PROOF OF PROPOSITION 2.1.3

For convenience let $\mathcal{S} := \mathsf{supp}(h_{xx}X + h_{yy}Y)$. To prove that $|S| = |X||Y|$ a.e. we look at the measure of the set such that $|S| \neq |X||Y|$, that is, a set for which there exists $s_i = h_{xx}x_i + h_{yy}y_i$ and $s_j = h_{xx}x_j + h_{yy}y_j$ such that $s_i = s_j$ for some $i \neq j$; hence, we are interested in characterizing the measure of the set

$$A := \left\{ (h_{xx}, h_{yy}) \in [0,1]^2 : \begin{array}{c} h_{xx}x_i + h_{yy}y_i = h_{xx}x_j + h_{yy}y_j \\ (x_i, y_i) \neq (x_j, y_j) \end{array}, \forall x_i, x_j \in X \text{ and } \forall y_i, y_j \in Y \right\}. \tag{B.1}$$

Define

$$A(i,j) = \left\{ (h_{xx}, h_{yy}) \in [0,1]^2 : h_{xx}x_i + h_{yy}y_i = h_{xx}x_j + h_{yy}y_j, \text{ s.t. } (x_i, y_i) \neq (x_j, y_j) \right\}. \tag{B.2}$$

By the sub-additivity of measure we have

$$m(A) = m\left(\bigcup_{i,j} A(i,j) \right) \leq \sum_{i,j} m(A(i,j)). \tag{B.3}$$

For fixed x_i, x_j, y_i, y_j the set $A(i,j)$ is a line in $(h_{xx}, h_{yy}) \in \mathbb{R}^2$ and hence

$$m\left(A(i,j)\right) = 0.$$

Appendix B (Continued)

Thus, in (Eq. B.3) we have a countable sum of sets of measure zero, which implies that $m(A) = 0$.

Next, we bound the minimum distance $d_{\min(\mathcal{S})} := \min_{i \neq j}\{|s_i - s_j| : s_i, s_j \in \mathcal{S}\}$ with $|s_i - s_j| = |h_{xx}x_i + h_{yy}y_i - h_{xx}x_j - h_{yy}y_j|$. We distinguish two cases:

Case 1) $x_i = x_j$ and $y_i \neq y_j$, or $x_i \neq x_j$ and $y_i = y_j$: then trivially

$$|s_i - s_j| \geq |h_{yy}| d_{\min(Y)}, \text{ or}$$

$$|s_i - s_j| \geq |h_{xx}| d_{\min(X)}.$$

Case 2) $x_i \neq x_j$ and $y_i \neq y_j$: Let $z_* \in \mathbb{Z}$, then

$$\begin{aligned}|s_i - s_j| &= |h_{xx}x_i + h_{yy}y_i - h_{xx}x_j - h_{yy}y_j| \\ &= |h_{xx}(x_i - x_j) - h_{yy}(y_j - y_i)| \\ &= |h_{xx}d_{min(X)}(z_{xi} - z_{xj}) - h_{yy}d_{min(Y)}(z_{yj} - z_{yi})| \\ &= |a_{xx}\bar{a}z_x - b_{yy}\bar{b}z_y|\end{aligned}$$

where $z_x = (z_{xi} - z_{xj})$ and $z_y = (z_{yj} - z_{yi})$ and $a_{xx}\bar{a} = h_{xx}d_{\min(X)}$, $b_{yy}\bar{b} = h_{yy}d_{\min(Y)}$. That is a_{xx}, b_{yy} are the fractional parts of and $\bar{a}$ and $\bar{b}$ are the integer parts $h_{xx}d_{min(X)}$ and $h_{yy}d_{min(Y)}$ respectively. Hence, by Lemma B.1.1 in Appendix B.1 we have that

$$\begin{aligned}|s_i - s_j| &\geq \gamma \max\left(\frac{|h_{xx}|d_{min(X)}}{2|Y|(1 + \log(|X|))}, \frac{|h_{yy}|d_{min(Y)}}{2|X|(1 + \log(|Y|))}\right) \\ &\geq \kappa_{\gamma,|X|,|Y|} \max\left(\frac{|h_{xx}|d_{min(X)}}{|Y|}, \frac{|h_{yy}|d_{min(Y)}}{|X|}\right)\end{aligned}$$

up to an outage set of measure γ where $\kappa_{\gamma,|X|,|Y|} := \frac{\gamma}{1+\ln(\max(|X|,|Y|)}$ and $\gamma \in (0,1]$. Next, by taking the minimum over both cases we arrive at the result in Proposition 2.1.3.

B.1 Minimum Distance Auxiliary Lemma

Lemma B.1.1. *Let* $(a_{xx}, b_{yy}) \in [0,1]^2$ *then for fixed integers* $\bar{a}, \bar{b} \in \mathbb{Z}$

Then the function

$$f(z_x, z_y) = \min_{z_x, z_y} |a_{xx}\bar{a}z_x - b_{yy}\bar{b}z_y|,$$

subject to the constrains

$$z_x \in [-N_x : N_x]/\{0\},$$

$$z_y \in [-N_y : N_y]/\{0\},$$

satisfies

$$f(z_x, z_y) \geq \gamma \max\left(\frac{b_{yy}\bar{b}}{2N_x(1+\ln(N_y))}, \frac{a_{xx}\bar{a}}{2N_y(1+\ln(N_x))}\right)$$

for all $(a_{xx}, b_{yy}) \in [0,1]^2$ *except for an outage set of Lebesgue measure* γ *for any* $\gamma \in (0,1]$.

Proof. First observe that w.l.o.g. we can assume that $a_{xx}\bar{a}, b_{yy}\bar{b} \in \mathbb{R}^+$ and $z_x \in [1 : N_x]$ and $z_y \in [1 : N_y]$. This is because if $sign(az_x) \neq sign(bz_y)$ then the function is minimized by $|z_x| = 1$ and $|z_y| = 1$ and attains a value of $f = |a_{xx}\bar{a}| + |b_{yy}\bar{b}|$. Furthermore, we let

$$\begin{aligned}
A_\epsilon &= \left\{(a_{xx}, b_{yy}) \in [0,1]^2 : \min_{1\leq z_x\leq N_x,\, 1\leq z_y\leq N_y} |a_{xx}\bar{a}z_x - b_{yy}\bar{b}z_y| > \epsilon\right\} \\
&= \bigcap_{1\leq z_x\leq N_x,\, 1\leq z_y\leq N_y} \left\{(a_{xx}, b_{yy}) \in [0,1]^2 : |a_{xx}\bar{a}z_x - b_{yy}\bar{b}z_y| > \epsilon\right\} \\
&= \bigcap_{1\leq z_x\leq N_x,\, 1\leq z_y\leq N_y} A(z_x, z_y),
\end{aligned}$$

where $A_\epsilon(z_x, z_y) = \left\{(a_{xx}, b_{yy}) \in [0,1]^2 : |a_{xx}\bar{a}z_x - b_{yy}\bar{b}z_y| > \epsilon\right\}$ and for some $\epsilon > 0$. The shape of $A_\epsilon(z_x, z_y)$ is shown on Fig. Figure 17. Let A_ϵ^c be the complement of A_ϵ where we have

$$A_\epsilon^c = \bigcup_{1\leq z_x\leq N_x,\, 1\leq z_y\leq N_y} A_\epsilon^c(z_x, z_y) \tag{B.4}$$

where $A_\epsilon^c(z_x, z_y) = \left\{(a_{xx}, b_{yy}) \in [0,1]^2 : |a_{xx}\bar{a}z_x - b_{yy}\bar{b}z_y| \leq \epsilon\right\}$.

Next, we find the measure of the set A_ϵ^c as follows:

$$\begin{aligned}
m(A_\epsilon^c) &= m\left(\bigcup_{1\leq z_x\leq N_x,\, 1\leq z_y\leq N_y} A_\epsilon^c(z_x, z_y)\right) \\
&\leq \sum_{1\leq z_x\leq N_x,\, 1\leq z_y\leq N_y} m(A_\epsilon^c(z_x, z_y))
\end{aligned}$$

where the last inequality is due to the sub-additive of measure.

Appendix B (Continued)

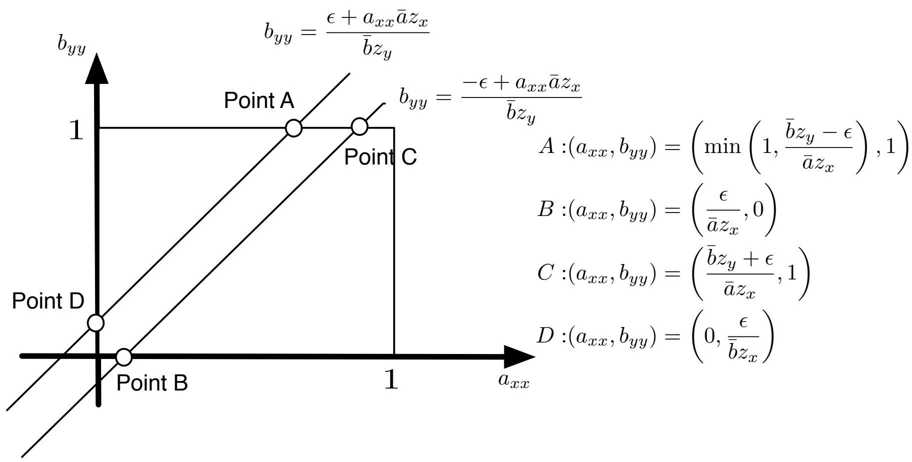


Figure 17: Shape of the outage strip.

Next, we compute $m(A_\epsilon^c(z_x, z_y))$ as follows

$$
\begin{aligned}
&m(A_\epsilon^c(z_x, z_y)) \\
&= \int_{a_{xx}=0}^{a_{xx}=\text{Point}A} \frac{\epsilon + a_{xx}\bar{a}z_x}{bz_y} da_{xx} - \int_{a_{xx}=\text{Point B}}^{a_{xx}=\text{Point A}} \frac{-\epsilon + a_{xx}\bar{a}z_x}{\bar{b}z_y} da_{xx} \\
&= \int_{a_{xx}=0}^{a_{xx}=\min\left(1, \frac{\bar{b}z_y - \epsilon}{\bar{a}z_x}\right)} \frac{\epsilon + a_{xx}\bar{a}z_x}{bz_y} da_{xx} - \int_{a_{xx}=\frac{\epsilon}{\bar{a}z_x}}^{a_{xx}=\min\left(1, \frac{\bar{b}z_y - \epsilon}{\bar{a}z_x}\right)} \frac{-\epsilon + a_{xx}\bar{a}z_x}{\bar{b}z_y} da_{xx} \\
&= \frac{a_{xx}\ (2\,\epsilon + \bar{a}\,a_{xx}\,z_x)}{2\,\bar{b}\,z_y}\Bigg|_{a_{xx}=0}^{a_{xx}=\min\left(1, \frac{\bar{b}z_y - \epsilon}{\bar{a}z_x}\right)} - \left(-\frac{a_{xx}\ (2\,\epsilon - \bar{a}\,a_{xx}\,z_x)}{2\,\bar{b}\,z_y}\Bigg|_{a_{xx}=\frac{\epsilon}{\bar{a}z_x}}^{a_{xx}=\min\left(1, \frac{\bar{b}z_y - \epsilon}{\bar{a}z_x}\right)}\right) \\
&= \frac{\min\left(1, -\frac{(\epsilon - \bar{b}\,z_y)}{\bar{a}\,z_x}\right)\ \left(2\,\epsilon + \bar{a}\,z_x\ \min\left(1, -\frac{(\epsilon - \bar{b}\,z_y)}{a\,z_x}\right)\right)}{2\,\bar{b}\,z_y} - \frac{\left(\epsilon - \bar{a}\,z_x\ \min\left(1, -\frac{(\epsilon - \bar{b}\,z_y)}{\bar{a}\,z_x}\right)\right)^2}{2\,\bar{a}\,\bar{b}\,z_x\,z_y} \\
&= -\frac{\epsilon\ \left(\epsilon - 4\,\bar{a}\,z_x\ \min\left(1, -\frac{(\epsilon - \bar{b}\,z_y)}{\bar{a}\,z_x}\right)\right)}{2\,\bar{a}\,\bar{b}\,z_x\,z_y}.
\end{aligned}
$$

Appendix B (Continued)

Next, compute $m(A_\epsilon^c)$ as follows

$$
\begin{aligned}
& m(A_\epsilon^c) \\
&= \sum_{1 \le z_x \le N_x,\, 1 \le z_y \le N_y} -\frac{\epsilon \left(\epsilon - 4\,\bar{a}\,z_x\, \min\left(1, -\frac{\left(\epsilon - \bar{b}\,z_y\right)}{\bar{a}\,z_x}\right)\right)}{2\,\bar{a}\,\bar{b}\,z_x\,z_y} \\
&= \sum_{1 \le z_x \le N_x,\, 1 \le z_y \le N_y} \frac{-\epsilon^2}{2\,\bar{a}\,\bar{b}\,z_x\,z_y} + \sum_{1 \le z_x \le N_x,\, 1 \le z_y \le N_y} \frac{4\epsilon\,\bar{a}\,z_x\, \min\left(1, -\frac{\left(\epsilon - \bar{b}\,z_y\right)}{\bar{a}\,z_x}\right)}{2\,\bar{a}\,\bar{b}\,z_x\,z_y} \\
&\le \sum_{1 \le z_x \le N_x,\, 1 \le z_y \le N_y} \frac{4\epsilon\,\bar{a}\,z_x\, \min\left(1, -\frac{\left(\epsilon - \bar{b}\,z_y\right)}{\bar{a}\,z_x}\right)}{2\,\bar{a}\,\bar{b}\,z_x\,z_y}.
\end{aligned}
$$

The term $\min\left(1, -\frac{\left(\epsilon - \bar{b}\,z_y\right)}{\bar{a}\,z_x}\right)$ can be upper bounded in two different ways

$$
\min\left(1, -\frac{\left(\epsilon - \bar{b}\,z_y\right)}{\bar{a}\,z_x}\right) \le 1, \tag{B.5}
$$

$$
\min\left(1, -\frac{\left(\epsilon - \bar{b}\,z_y\right)}{\bar{a}\,z_x}\right) \le -\frac{\left(\epsilon - \bar{b}\,z_y\right)}{\bar{a}\,z_x}. \tag{B.6}
$$

With the first upper bound in (Eq. B.5) we get

$$
\begin{aligned}
m(A_\epsilon^c) &\le \sum_{1 \le z_x \le N_x,\, 1 \le z_y \le N_y} \frac{4\epsilon\,\bar{a}\,z_x}{2\,\bar{a}\,\bar{b}\,z_x\,z_y} \\
&\le \frac{2\epsilon N_x (1 + \ln(N_y))}{\bar{b}}
\end{aligned} \tag{B.7}
$$

Appendix B (Continued)

where for the last inequality we have used $\sum_{zy=1}^{N_y} \frac{1}{z_y} \leq 1 + \ln(N_y)$. With the second upper bound in (Eq. B.6) we get

$$
\begin{aligned}
m(A_\epsilon^c) &\leq \sum_{1 \leq z_x \leq N_x,\, 1 \leq z_y \leq N_y} \frac{4\epsilon\,(\bar{b} z_y - \epsilon)}{2\,\bar{a}\,\bar{b}\,z_x\,z_y} \\
&\leq \sum_{1 \leq z_x \leq N_x,\, 1 \leq z_y \leq N_y} \frac{4\epsilon}{2\,\bar{a}\,z_x} \\
&\leq \frac{2\epsilon N_y (1 + \ln(N_x))}{\bar{a}}. \qquad \text{(B.8)}
\end{aligned}
$$

So by taking the tightest of the two bounds in (Eq. B.7) and in (Eq. B.8) we get

$$
m(A_\epsilon^c) \leq \min\left(\frac{2\epsilon N_x(1+\ln(N_y))}{\bar{b}}, \frac{2\epsilon N_y(1+\ln(N_x))}{\bar{a}}\right).
$$

Now let $m(A_\epsilon^c) = \gamma$ for some $\gamma \in [0,1]$ then we have that

$$
\gamma \leq \epsilon \min\left(\frac{2N_x(1+\ln(N_y))}{\bar{b}}, \frac{2N_y(1+\ln(N_x))}{\bar{a}}\right).
$$

Next, by solving for ϵ in terms of measure of the outage,

$$
\begin{aligned}
\epsilon &\geq \frac{\gamma}{\min\left(\frac{2N_x(1+\ln(N_y))}{\bar{b}}, \frac{2N_y(1+\ln(N_x))}{\bar{a}}\right)} = \gamma \max\left(\frac{\bar{b}}{2N_x(1+\ln(N_y))}, \frac{\bar{a}}{2N_y(1+\ln(N_x))}\right) \\
&\geq \gamma \max\left(\frac{b_{yy}\bar{b}}{2N_x(1+\ln(N_y))}, \frac{a_{xx}\bar{a}}{2N_y(1+\ln(N_x))}\right)
\end{aligned}
$$

This concludes the proof. $\square$

Appendix C

PROOF OF THEOREM 3.5.2

We proceed to evaluate the rate region in Proposition 3.4.1 with the inputs in (Eq. 3.16). With the chosen inputs, the outputs are

$$Y_1 = h_{11}\sqrt{1-\delta_1}X_{1D} + h_{11}\sqrt{\delta_1}X_{1G} + h_{12}\sqrt{1-\delta_2}X_{2Gc} + h_{12}\sqrt{\delta_2}X_{2Gp} + Z_1,$$
$$Y_2 = h_{21}\sqrt{1-\delta_1}X_{1D} + h_{21}\sqrt{\delta_1}X_{1G} + h_{22}\sqrt{1-\delta_2}X_{2Gc} + h_{22}\sqrt{\delta_2}X_{2Gp} + Z_2.$$

The achievable region in (Eq. 3.5) with $Q = \emptyset, U_2 = X_{2Gc}$ reduces to

$$\begin{aligned}
R_1 &\le I(X_1; Y_1|X_{2Gc}) \\
&= h(Y_1|X_{2Gc}) - h(Y_1|X_1, X_{2Gc}) \\
&= h(h_{11}\sqrt{1-\delta_1}X_{1D} + h_{11}\sqrt{\delta_1}X_{1G} + h_{12}\sqrt{\delta_2}X_{2Gp} + Z_1) \\
&- h(h_{12}\sqrt{\delta_2}X_{2Gp} + Z_1) \\
&= h\left(\sqrt{\frac{|h_{11}|^2(1-\delta_1)}{1+|h_{11}|^2\delta_1+|h_{12}|^2\delta_2}}X_{1D} + Z_1\right) - h(Z_1) \\
&+ \mathsf{l_g}\left(|h_{11}|^2\delta_1 + |h_{12}|^2\delta_2\right) - \mathsf{l_g}\left(|h_{12}|^2\delta_2\right);
\end{aligned}$$

Appendix C (Continued)

therefore, by (Eq. 2.10), we can further lower bound the rate of user 1 as

$$R_1 \leq \mathsf{l_g}\left(\min\left(N^2-1, \frac{|h_{11}|^2(1-\delta_1)}{1+|h_{11}|^2\delta_1+|h_{12}|^2\delta_2}\right)\right) - \frac{1}{2}\log\left(\frac{\pi \mathrm{e}}{3}\right) + \mathsf{l_g}\left(\frac{|h_{11}|^2\delta_1}{1+|h_{12}|^2\delta_2}\right),$$

thus proving (Eq. 3.17a).

For the rate of user 2 we have

$$\begin{aligned}
R_2 \leq I(X_2;Y_2) &= h\left(h_{21}\sqrt{1-\delta_1}X_{1D} + h_{21}\sqrt{\delta_1}X_{1G} + h_{22}\sqrt{1-\delta_2}X_{2Gc} + h_{22}\sqrt{\delta_2}X_{2Gp} + Z_2\right) \\
&- h\left(h_{21}\sqrt{1-\delta_1}X_{1D} + h_{21}\sqrt{\delta_1}X_{1G} + Z_2\right) \\
&= h\left(\sqrt{\frac{|h_{21}|^2(1-\delta_1)}{1+|h_{21}|^2\delta_1+|h_{22}|^2}}X_{1D} + Z_2\right) - h(Z_2) + \mathsf{l_g}\left(|h_{21}|^2\delta_1 + |h_{22}|^2\right) \\
&- h\left(\sqrt{\frac{|h_{21}|^2(1-\delta_1)}{1+|h_{21}|^2\delta_1}}X_{1D} + Z_2\right) + h(Z_2) - \mathsf{l_g}\left(|h_{21}|^2\delta_1\right)
\end{aligned}$$

therefore, by (Eq. 2.10), we can further lower bound the rate of user 2 as

$$\begin{aligned}
R_2 \leq \mathsf{l_g}&\left(\min\left(N^2-1, \frac{|h_{21}|^2(1-\delta_1)}{1+|h_{21}|^2\delta_1+|h_{22}|^2}\right)\right) - \frac{1}{2}\log\left(\frac{\pi \mathrm{e}}{3}\right) + \mathsf{l_g}\left(\frac{|h_{22}|^2}{1+|h_{21}|^2\delta_1}\right) \\
&- \mathsf{l_g}\left(\min\left(N^2-1, \frac{|h_{21}|^2(1-\delta_1)}{1+|h_{21}|^2\delta_1}\right)\right)
\end{aligned}$$

thus proving (Eq. 3.17b).

Appendix C (Continued)

Finally for the sum-rate we have

$$\begin{aligned}
R_1 + R_2 &\leq I(X_1, X_{2Gc}; Y_1) + I(X_2; Y_2|X_{2Gc}) \\
&= h(h_{11}\sqrt{1-\delta_1}X_{1D} + h_{11}\sqrt{\delta_1}X_{1G} + h_{12}\sqrt{1-\delta_2}X_{2Gc} + h_{12}\sqrt{\delta_2}X_{2Gp} + Z_1) \\
&- h(h_{12}\sqrt{\delta_2}X_{2Gp} + Z_1) + h(h_{21}\sqrt{1-\delta_1}X_{1D} + h_{21}\sqrt{\delta_1}X_{1G} + h_{22}\sqrt{\delta_2}X_{2Gp} + Z_2) \\
&- h(h_{21}\sqrt{1-\delta_1}X_{1D} + h_{21}\sqrt{\delta_1}X_{1G} + Z_2)'
\end{aligned}$$

therefore, by (Eq. 2.10), we can further lower bound the sum-rate as

$$\begin{aligned}
R_1 + R_2 \leq\ & \mathsf{lg}\left(\min\left(N^2-1, \frac{|h_{11}|^2(1-\delta_1)}{1+|h_{11}|^2\delta_1+|h_{12}|^2}\right)\right) - \frac{1}{2}\log\left(\frac{\pi \mathrm{e}}{3}\right) + \mathsf{lg}\left(|h_{11}|^2\delta_1 + |h_{12}|^2\right) \\
& - \mathsf{lg}\left(|h_{12}|^2\delta_2\right) \\
& + \mathsf{lg}\left(\min\left(N^2-1, \frac{|h_{21}|^2(1-\delta_1)}{1+|h_{21}|^2\delta_1+|h_{22}|^2\delta_2}\right)\right) - \frac{1}{2}\log\left(\frac{\pi \mathrm{e}}{3}\right) + \mathsf{lg}\left(|h_{21}|^2\delta_1 + |h_{22}|^2\delta_2\right) \\
& - \mathsf{lg}\left(\min\left(N^2-1, \frac{|h_{21}|^2(1-\delta_1)}{1+|h_{21}|^2\delta_1}\right)\right) - \mathsf{lg}\left(|h_{21}|^2\delta_1\right)
\end{aligned}$$

thus proving (Eq. 3.17c).

Appendix C (Continued)

Remark 16. *For future use, we specialized the derived achievable rate region for the power splits $\delta_1 = \frac{1}{1+|h_{21}|^2}$ and $\delta_2 = \frac{1}{1+|h_{12}|^2}$ inspired by (26); we thus have that the following region is achievable for any $N \in \mathbb{N}$*

$$R_1 \leq \mathsf{I_g}\left(\min\left(N^2-1, \frac{|h_{11}|^2 a}{1+\frac{|h_{11}|^2}{1+|h_{21}|^2}+b}\right)\right) - \frac{1}{2}\log\left(\frac{\pi \mathrm{e}}{3}\right) + \mathsf{I_g}\left(\frac{\frac{|h_{11}|^2}{1+|h_{21}|^2}}{1+b}\right), \tag{C.1a}$$

$$R_2 \leq \mathsf{I_g}\left(\min\left(N^2-1, \frac{|h_{21}|^2 a}{1+a+|h_{22}|^2}\right)\right) - \frac{1}{2}\log\left(\frac{\pi \mathrm{e}}{3}\right) + \mathsf{I_g}\left(\frac{|h_{22}|^2}{1+a}\right)$$
$$- \mathsf{I_g}\left(\min\left(N^2-1, |h_{21}|^2 a\right)\right), \tag{C.1b}$$

$$R_1 + R_2 \leq \mathsf{I_g}\left(\min\left(N^2-1, \frac{|h_{11}|^2 a}{1+\frac{|h_{11}|^2}{1+|h_{21}|^2}+|h_{12}|^2}\right)\right) + \mathsf{I_g}\left(\frac{|h_{11}|^2}{1+|h_{21}|^2}+|h_{12}|^2\right) - \mathsf{I_g}(b)$$
$$+ \mathsf{I_g}\left(\min\left(N^2-1, \frac{|h_{21}|^2 a}{1+a+\frac{|h_{22}|^2}{1+|h_{12}|^2}}\right)\right) + \mathsf{I_g}\left(\frac{\frac{|h_{22}|^2}{1+|h_{12}|^2}}{1+a}\right)$$
$$- \mathsf{I_g}\left(\min\left(N^2-1, |h_{21}|^2 a\right)\right) - \log\left(\frac{\pi \mathrm{e}}{3}\right). \tag{C.1c}$$

where $a := \frac{|h_{21}|^2}{1+|h_{21}|^2} \in [0,1]$ and $b := \frac{|h_{12}|^2}{1+|h_{12}|^2} \in [0,1]$.

Appendix C (Continued)

In the symmetric case the region in (Eq. C.1) *is further lower bounded by*

$$R_1 \leq \mathsf{I_g}\left(\min\left(N^2-1, \frac{\mathsf{snr}\cdot\mathsf{inr}}{1+\mathsf{snr}+2\mathsf{inr}}\right)\right) - \frac{1}{2}\log\left(\frac{\pi \mathrm{e}}{3}\right) + \mathsf{I_g}\left(\frac{\mathsf{snr}}{1+2\mathsf{inr}}\right), \tag{C.2a}$$

$$\begin{aligned} R_2 \leq\ & \mathsf{I_g}\left(\min\left(N^2-1, \frac{\mathsf{inr}^2}{(1+\mathsf{inr})(1+\mathsf{snr})+\mathsf{inr}}\right)\right) - \frac{1}{2}\log\left(\frac{\pi \mathrm{e}}{3}\right) \\ & + \mathsf{I_g}\left(\mathsf{snr}\frac{1}{2}\right) - \mathsf{I_g}\left(\min\left(N^2-1, \frac{\mathsf{inr}^2}{1+2\mathsf{inr}}\right)\right), \end{aligned} \tag{C.2b}$$

$$\begin{aligned} R_1 + R_2 \leq\ & \mathsf{I_g}\left(\min\left(N^2-1, \frac{\mathsf{snr}\cdot\mathsf{inr}}{(1+\mathsf{inr})^2+\mathsf{snr}}\right)\right) \\ & + \mathsf{I_g}\left(\mathsf{inr}+\frac{\mathsf{snr}}{1+\mathsf{inr}}\right) - \mathsf{I_g}\left(\frac{\mathsf{inr}}{1+\mathsf{inr}}\right) \\ & + \mathsf{I_g}\left(\min\left(N^2-1, \frac{\mathsf{inr}^2}{1+\mathsf{snr}+2\mathsf{inr}}\right)\right) \\ & + \mathsf{I_g}\left(\frac{\mathsf{snr}}{1+2\mathsf{inr}}\right) - \mathsf{I_g}\left(\min\left(N^2-1, \frac{\mathsf{inr}^2}{1+2\mathsf{inr}}\right)\right) - \log\left(\frac{\pi \mathrm{e}}{3}\right). \end{aligned} \tag{C.2c}$$

Appendix D

GAP DERIVATION FOR THE MODERATELY WEAK INTERFERENCE REGIME

In order to show achievability to within a constant gap of the outer bound in (Eq. 3.22) by means of the achievable region in (Eq. 3.32) (a further lower bound to the region in (Eq. C.2)), we distinguish two cases.

CASE 1 (regime corresponding to $\alpha \in [2/3, 1]$ in 7c)

Assume that the sum-rate in eq.(Eq. 3.32c) is redundant; under this condition we match the corner point of the rectangular achievable region, given by $(R_1, R_2) =$ (eq.(Eq. 3.32a), eq.(Eq. 3.32b)), to

$$\mathcal{R}_{out}^{\text{(G-IC mod P1)}}: \quad R_1 = \mathsf{I_g}\left(\mathsf{inr} + \frac{\mathsf{snr}}{1+\mathsf{inr}}\right), \tag{D.1a}$$

$$R_2 = \mathsf{I_g}\left(\mathsf{snr}\right) - \mathsf{I_g}\left(\mathsf{inr}\right) + \mathsf{I_g}(\mathsf{inr}+\mathsf{snr}) - \mathsf{I_g}\left(\mathsf{inr} + \frac{\mathsf{snr}}{1+\mathsf{inr}}\right), \tag{D.1b}$$

and

$$\mathcal{R}_{out}^{\text{(G-IC mod P2)}}: \quad R_1 = \mathsf{I_g}\left(\mathsf{snr}\right) - \mathsf{I_g}\left(\mathsf{inr}\right) + \mathsf{I_g}(\mathsf{inr}+\mathsf{snr}) - \mathsf{I_g}\left(\mathsf{inr} + \frac{\mathsf{snr}}{1+\mathsf{inr}}\right), \tag{D.2a}$$

$$R_2 = \mathsf{I_g}\left(\mathsf{inr} + \frac{\mathsf{snr}}{1+\mathsf{inr}}\right), \tag{D.2b}$$

which were obtained from the intersection of the sum-rate outer bound in (Eq. 3.22c) with either (Eq. 3.22e) or (Eq. 3.22f). In particular, for the corner point in (Eq. D.1) we use x in (Eq. 3.32d) (which corresponds to N in (Eq. 3.29)), and for the corner point in (Eq. D.2) we use x in (Eq. 3.32e) (which corresponds to N in (Eq. 3.30)).

The gap is readily computed as follows: for the corner point in (Eq. D.1) we have

$$\begin{aligned}
\Delta_1 &= \text{eq.(Eq. D.1a)-eq.(Eq. 3.32a)}|_{x \text{ in (Eq. 3.32d)}} \\
&\leq \mathsf{I_g}\left(\mathsf{inr} + \frac{\mathsf{snr}}{1+\mathsf{inr}}\right) - \mathsf{I_g}\left(\frac{\mathsf{snr}}{1+2\mathsf{inr}}\right) - \mathsf{I_g}\left(\frac{\mathsf{inr}^2}{1+\mathsf{snr}+2\mathsf{inr}}\right) \\
&+ \frac{1}{2}\log(4) + \frac{1}{2}\log\left(\frac{\pi \mathrm{e}}{3}\right) \\
&\leq \frac{1}{2}\log(2) + \frac{1}{2}\log(4) + \frac{1}{2}\log\left(\frac{\pi \mathrm{e}}{3}\right) = \frac{1}{2}\log\left(\frac{8\pi \mathrm{e}}{3}\right),
\end{aligned}$$

and

$$\begin{aligned}
\Delta_2 &= \text{eq.(Eq. D.1b)-eq.(Eq. 3.32b)}|_{x \text{ in (Eq. 3.32d)}} \\
&\leq \mathsf{I_g}(\mathsf{snr}) - \mathsf{I_g}(\mathsf{inr}) + \mathsf{I_g}(\mathsf{inr}+\mathsf{snr}) - \mathsf{I_g}\left(\mathsf{inr} + \frac{\mathsf{snr}}{1+\mathsf{inr}}\right) \\
&- \mathsf{I_g}\left(\frac{\mathsf{snr}}{2}\right) + \mathsf{I_g}\left(\frac{\mathsf{inr}^2}{1+\mathsf{snr}+2\mathsf{inr}}\right) - \mathsf{I_g}\left(\frac{\mathsf{inr}^2}{(1+\mathsf{inr})(1+\mathsf{snr})+\mathsf{inr}}\right) + \frac{1}{2}\log\left(\frac{\pi \mathrm{e}}{3}\right) \\
&\leq \frac{1}{2}\log(2) + \frac{1}{2}\log\left(\frac{\pi \mathrm{e}}{3}\right) = \frac{1}{2}\log\left(\frac{2\pi \mathrm{e}}{3}\right), \ \text{since } \mathsf{inr} \leq \mathsf{snr} \text{ in weak interfernce;}
\end{aligned}$$

Appendix D (Continued)

while for the corner point in (Eq. D.2) we have

$$\begin{aligned}\Delta_1 &= \text{eq.(Eq. D.2a)-eq.(Eq. 3.32a)}|_{x \text{ in (Eq. 3.32e)}}\\ &\le \mathsf{I_g}\left(\mathsf{snr}\right) - \mathsf{I_g}\left(\mathsf{inr}\right) + \mathsf{I_g}(\mathsf{inr}+\mathsf{snr}) - \mathsf{I_g}\left(\mathsf{inr} + \frac{\mathsf{snr}}{1+\mathsf{inr}}\right)\\ &- \mathsf{I_g}\left(\frac{\mathsf{snr}}{1+2\mathsf{inr}}\right) - \mathsf{I_g}\left(\frac{\mathsf{snr}\cdot\mathsf{inr}}{(1+\mathsf{inr})^2+\mathsf{snr}}\right) + \frac{1}{2}\log(4) + \frac{1}{2}\log\left(\frac{\pi e}{3}\right)\\ &\le \frac{1}{2}\log(2) + \frac{1}{2}\log(4) + \frac{1}{2}\log\left(\frac{\pi e}{3}\right) = \frac{1}{2}\log\left(\frac{8\pi e}{3}\right),\end{aligned}$$

and

$$\begin{aligned}\Delta_2 &= \text{eq.(Eq. D.2b)-eq.(Eq. 3.32b)}|_{x \text{ in (Eq. 3.32e)}}\\ &\le \mathsf{I_g}\left(\mathsf{inr} + \frac{\mathsf{snr}}{1+\mathsf{inr}}\right) - \mathsf{I_g}\left(\frac{\mathsf{snr}}{2}\right) + \mathsf{I_g}\left(\frac{\mathsf{snr}\cdot\mathsf{inr}}{(1+\mathsf{inr})^2+\mathsf{snr}}\right)\\ &-\mathsf{I_g}\left(\frac{\mathsf{inr}^2}{(1+\mathsf{inr})(1+\mathsf{snr})+\mathsf{inr}}\right) + \frac{1}{2}\log\left(\frac{\pi e}{3}\right)\\ &\le \frac{1}{2}\log(2) + \frac{1}{2}\log\left(\frac{\pi e}{3}\right) = \frac{1}{2}\log\left(\frac{2\pi e}{3}\right), \text{ since } \mathsf{inr} \le \mathsf{snr} \text{ in weak interference.}\end{aligned}$$

CASE 2 (regime corresponding to $\alpha \in [1/2, 2/3]$ in 7d)

Assume that the sum-rate in (Eq. 3.32) is not redundant, that is after simple algebraic manipulation,

$$\begin{aligned}&1+\min\left(x|_{x \text{ in (Eq. 3.32d)}}, x|_{x \text{ in (Eq. 3.32e)}}\right)\\ &< \underbrace{\frac{(1+2\mathsf{inr})(1+\frac{\mathsf{snr}}{2})}{(1+\mathsf{inr})(1+\mathsf{snr})+\mathsf{inr}}}_{\in[0.7358,1] \text{ for } \mathsf{inr}\le\mathsf{snr}\le\mathsf{inr}(1+\mathsf{inr}) \text{ see Appendix E}} \cdot \underbrace{\frac{(1+\mathsf{inr})(1+\mathsf{inr}+\mathsf{snr})}{(1+\mathsf{inr})^2+\mathsf{snr}}}_{=1+x|_{x \text{ in (Eq. 3.32e)}}},\end{aligned} \tag{D.3}$$

Appendix D (Continued)

which implies

$$x|_{x \text{ in (Eq. 3.32d)}} \le x|_{x \text{ in (Eq. 3.32e)}}. \tag{D.4}$$

Under the condition in (Eq. D.4) we match one of the corner point of the pentagon-shaped achievable region in (Eq. 3.32) to

$$\mathcal{R}_{out}^{\text{(G-IC weak P1)}} : R_1 = 3\mathsf{I_g}\left(\mathsf{inr} + \frac{\mathsf{snr}}{1+\mathsf{inr}}\right) - \mathsf{I_g}\left(\mathsf{snr}+\mathsf{inr}\right) - \mathsf{I_g}\left(\mathsf{snr}\right) + \mathsf{I_g}\left(\mathsf{inr}\right), \tag{D.5a}$$

$$R_2 = \mathsf{I_g}\left(\mathsf{snr}\right) - \mathsf{I_g}\left(\mathsf{inr}\right) + \mathsf{I_g}\left(\mathsf{snr}+\mathsf{inr}\right) - \mathsf{I_g}\left(\mathsf{inr} + \frac{\mathsf{snr}}{1+\mathsf{inr}}\right), \tag{D.5b}$$

and

$$\mathcal{R}_{out}^{\text{(G-IC weak P2)}} : R_1 = \mathsf{I_g}\left(\mathsf{snr}\right) - \mathsf{I_g}\left(\mathsf{inr}\right) + \mathsf{I_g}\left(\mathsf{snr}+\mathsf{inr}\right) - \mathsf{I_g}\left(\mathsf{inr} + \frac{\mathsf{snr}}{1+\mathsf{inr}}\right), \tag{D.6a}$$

$$R_2 = 3\mathsf{I_g}\left(\mathsf{inr} + \frac{\mathsf{snr}}{1+\mathsf{inr}}\right) - \mathsf{I_g}\left(\mathsf{snr}+\mathsf{inr}\right) - \mathsf{I_g}\left(\mathsf{snr}\right) + \mathsf{I_g}\left(\mathsf{inr}\right), \tag{D.6b}$$

which were obtained from the intersection of the sum-rate outer bound in (Eq. 3.22d) with either (Eq. 3.22e) or (Eq. 3.22f). In particular, for the corner point in (Eq. D.5) we use x in (Eq. 3.32d) (which corresponds to N in (Eq. 3.29)), and for the corner point in (Eq. D.6) we use x in (Eq. 3.32e) (which corresponds to N in (Eq. 3.30)).

Appendix D (Continued)

The gap is readily computed as follows: for the corner point in (Eq. D.5) we have

$$
\begin{aligned}
\Delta_1 &= \text{eq.(Eq. D.5a)-}\Big(\text{eq.(Eq. 3.32c)-eq.(Eq. 3.32b)}\Big)|_{x \text{ in (Eq. 3.32d)}}\\
&\leq 2\mathsf{I_g}\left(\mathsf{inr}+\frac{\mathsf{snr}}{1+\mathsf{inr}}\right)-\mathsf{I_g}\left(\mathsf{snr}+\mathsf{inr}\right)-\mathsf{I_g}\left(\mathsf{snr}\right)+\mathsf{I_g}\left(\mathsf{inr}\right)+\mathsf{I_g}\left(\frac{\mathsf{inr}}{1+\mathsf{inr}}\right)\\
&-\mathsf{I_g}\left(\frac{\mathsf{snr}}{1+2\mathsf{inr}}\right)+\mathsf{I_g}\left(\frac{\mathsf{snr}}{2}\right)+\mathsf{I_g}\left(\frac{\mathsf{inr}^2}{(1+\mathsf{inr})(1+\mathsf{snr})+\mathsf{inr}}\right)\\
&-2\mathsf{I_g}\left(\frac{\mathsf{inr}^2}{1+\mathsf{snr}+2\mathsf{inr}}\right)+\frac{1}{2}\log\left(4\right)+\frac{1}{2}\log\left(\frac{\pi \mathrm{e}}{3}\right)\\
&=\frac{1}{2}\log\left(\frac{\left(\frac{\mathsf{snr}}{2}+1\right)\left(\frac{\mathsf{inr}}{\mathsf{inr}+1}+1\right)\left(\frac{\mathsf{inr}^2}{\mathsf{inr}+(\mathsf{inr}+1)\,(\mathsf{snr}+1)}+1\right)\left(\mathsf{inr}+1\right)\left(\mathsf{inr}+\frac{\mathsf{snr}}{\mathsf{inr}+1}+1\right)^2}{\left(\frac{\mathsf{inr}^2}{2\,\mathsf{inr}+\mathsf{snr}+1}+1\right)^2\left(\frac{\mathsf{snr}}{2\,\mathsf{inr}+1}+1\right)\left(\mathsf{snr}+1\right)\left(\mathsf{inr}+\mathsf{snr}+1\right)}\right)\\
&+\frac{1}{2}\log\left(4\right)+\frac{1}{2}\log\left(\frac{\pi \mathrm{e}}{3}\right)\\
&=\frac{1}{2}\log\left(\frac{\left(2\,\mathsf{inr}+1\right)^2\left(\frac{\mathsf{snr}}{2}+1\right)\left(2\,\mathsf{inr}+\mathsf{snr}+1\right)}{\left(\mathsf{inr}+1\right)\left(\mathsf{snr}+1\right)\left(2\,\mathsf{inr}+\mathsf{snr}+\mathsf{inr}\,\mathsf{snr}+1\right)}\right)+\frac{1}{2}\log\left(\frac{\pi \mathrm{e}}{3}\right)\\
&\leq \frac{1}{2}\log(6)+\frac{1}{2}\log\left(4\right)+\frac{1}{2}\log\left(\frac{\pi \mathrm{e}}{3}\right)=\frac{1}{2}\log\left(8\pi \mathrm{e}\right)
\end{aligned}
$$

Appendix D (Continued)

and

$$
\begin{aligned}
\Delta_2 &= \text{eq.(Eq. D.5b)-eq.(Eq. 3.32b)}|_{x \text{ in (Eq. 3.32d)}} \\
&\leq \mathsf{I_g}\left(\mathsf{snr}\right) - \mathsf{I_g}\left(\mathsf{inr}\right) + \mathsf{I_g}\left(\mathsf{snr} + \mathsf{inr}\right) - \mathsf{I_g}\left(\mathsf{inr} + \frac{\mathsf{snr}}{1+\mathsf{inr}}\right) - \mathsf{I_g}\left(\frac{\mathsf{snr}}{2}\right) \\
&+ \mathsf{I_g}\left(\frac{\mathsf{inr}^2}{1+\mathsf{snr}+2\mathsf{inr}}\right) - \mathsf{I_g}\left(\frac{\mathsf{inr}^2}{(1+\mathsf{inr})(1+\mathsf{snr})+\mathsf{inr}}\right) + \frac{1}{2}\log\left(\frac{\pi \mathrm{e}}{3}\right) \\
&= \frac{1}{2}\log\left(\frac{\left(\frac{\mathsf{inr}^2}{2\,\mathsf{inr}+\mathsf{snr}+1}+1\right)\left(\mathsf{snr}+1\right)\left(\mathsf{inr}+\mathsf{snr}+1\right)}{\left(\frac{\mathsf{snr}}{2}+1\right)\left(\frac{\mathsf{inr}^2}{\mathsf{inr}+(\mathsf{inr}+1)\,(\mathsf{snr}+1)}+1\right)\left(\mathsf{inr}+1\right)\left(\mathsf{inr}+\frac{\mathsf{snr}}{\mathsf{inr}+1}+1\right)}\right) \\
&+ \frac{1}{2}\log\left(\frac{\pi \mathrm{e}}{3}\right) \\
&= \frac{1}{2}\log\left(\frac{2\,\left(\mathsf{snr}+1\right)\left(2\,\mathsf{inr}+\mathsf{snr}+\mathsf{inr}\,\mathsf{snr}+1\right)}{\left(\mathsf{inr}+1\right)\left(\mathsf{snr}+2\right)\left(2\,\mathsf{inr}+\mathsf{snr}+1\right)}\right) + \frac{1}{2}\log\left(\frac{\pi \mathrm{e}}{3}\right) \\
&\leq \frac{1}{2}\log\left(2\right) + \frac{1}{2}\log\left(\frac{\pi \mathrm{e}}{3}\right) = \frac{1}{2}\log\left(\frac{2\pi \mathrm{e}}{3}\right),
\end{aligned}
$$

Appendix D (Continued)

while for the corner point in (Eq. D.6) we have

$$
\begin{aligned}
\Delta_1 &= \text{eq.(Eq. D.6a)-eq.(Eq. 3.32a)}|_{x \text{ in (Eq. 3.32e)}} \\
&\leq \mathsf{I_g}(\mathsf{snr}) - \mathsf{I_g}(\mathsf{inr}) + \mathsf{I_g}(\mathsf{snr}+\mathsf{inr}) - \mathsf{I_g}\left(\mathsf{inr} + \frac{\mathsf{snr}}{1+\mathsf{inr}}\right) - \mathsf{I_g}\left(\frac{\mathsf{snr}}{1+2\mathsf{inr}}\right) \\
&- \mathsf{I_g}\left(\frac{\mathsf{snr}\cdot\mathsf{inr}}{(1+\mathsf{inr})^2+\mathsf{snr}}\right) + \frac{1}{2}\log\left(\frac{\pi e}{3}\right) + \frac{1}{2}\log(4) \\
&= \frac{1}{2}\log\left(\frac{(\mathsf{snr}+1)\ (\mathsf{inr}+\mathsf{snr}+1)}{\left(\frac{\mathsf{snr}}{2\,\mathsf{inr}+1}+1\right)\ (\mathsf{inr}+1)\ \left(\frac{\mathsf{inr}\,\mathsf{snr}}{\mathsf{snr}+(\mathsf{inr}+1)^2}+1\right)\ \left(\mathsf{inr}+\frac{\mathsf{snr}}{\mathsf{inr}+1}+1\right)}\right) \\
&+ \frac{1}{2}\log(4) + \frac{1}{2}\log\left(\frac{\pi e}{3}\right) \\
&= \frac{1}{2}\log\left(\frac{(2\,\mathsf{inr}+1)\ (\mathsf{snr}+1)}{(\mathsf{inr}+1)\ (2\,\mathsf{inr}+\mathsf{snr}+1)}\right) + \frac{1}{2}\log(4) + \frac{1}{2}\log\left(\frac{\pi e}{3}\right) \\
&\leq \frac{1}{2}\log(2) + \frac{1}{2}\log(4) + \frac{1}{2}\log\left(\frac{\pi e}{3}\right) = \frac{1}{2}\log\left(\frac{4\pi e}{3}\right),
\end{aligned}
$$

and

$$
\begin{aligned}
\Delta_2 &= \text{eq.(Eq. D.6b)-}\Big(\text{eq.(Eq. 3.32c)-eq.(Eq. 3.32a)}\Big)|_{x \text{ in (Eq. 3.32e)}} \\
&\leq 2\mathsf{I_g}\left(\mathsf{inr}+\frac{\mathsf{snr}}{1+\mathsf{inr}}\right) - \mathsf{I_g}(\mathsf{snr}+\mathsf{inr}) - \mathsf{I_g}(\mathsf{snr}) + \mathsf{I_g}(\mathsf{inr}) + \mathsf{I_g}\left(\frac{\mathsf{inr}}{1+\mathsf{inr}}\right) \\
&+ \frac{1}{2}\log\left(\frac{\pi e}{3}\right) \\
&= \frac{1}{2}\log\left(\frac{(1+2\mathsf{inr})((1+\mathsf{inr})^2+\mathsf{snr})}{(1+\mathsf{inr})^2(1+\mathsf{snr})(1+\mathsf{inr}+\mathsf{snr})}\right) + \frac{1}{2}\log\left(\frac{\pi e}{3}\right) \\
&\leq 0 + \frac{1}{2}\log\left(\frac{\pi e}{3}\right) = \frac{1}{2}\log\left(\frac{\pi e}{3}\right).
\end{aligned}
$$

This concludes the proof.

Appendix E

MINIMUM OF A FUNCTION IN EQ. D.3

The minimum of the function

$$f(x,y) = \frac{(1+2y)(1+\frac{x}{2})}{(1+y)(1+x)+y}, \quad \text{for } (x,y) \in \mathbb{R}_+^2 \text{ such that } \quad 1 \leq y \leq x \leq y(1+y),$$

is found by first taking the partial derivative with respect to x, given my $\frac{\partial f}{\partial x} = -\frac{2y^2+7y+3}{2(2x+y+xy+1)^2}$ which is easily seen to be monotone decreasing in x therefore attaining the minimum

$$f(y(1+y),y) = \frac{2y^3+3y^2+5y+2}{2y^3+6y^2+6y+2}, \quad \text{for} \quad 1 \leq y.$$

Now by taking the partial derivative with respect to y, given by $\frac{\partial f}{\partial y} = \frac{\left(3y^2-4y-1\right)}{2(y+1)^4}$ and setting it equal to zero we see that the minimum occurs at $y = \frac{\sqrt{7}+2}{3}$. Hence, the minimum of the function occurs at $f\left(\frac{\sqrt{7}+2}{3}\left(1+\frac{\sqrt{7}+2}{3}\right), \frac{\sqrt{7}+2}{3}\right) = 0.7359$. Conditions on the second derivatives can be easily checked to verify that indeed the claim stationary point is a global minimum (even easier still, by plotting the function for example with Matlab).

Appendix F

GAP FOR $\mathsf{inr} \leq \mathsf{snr} \leq 1 + \mathsf{inr}$

Outer Bound for $\mathsf{inr} \leq \mathsf{snr} \leq 1 + \mathsf{inr}$

It is well know that when $\mathsf{snr} \approx \mathsf{inr}$ time-division is approximately optimal (26). In this regime we outer bound the capacity region by the sum-rate constraint in (Eq. 4.9c) only, which in the symmetric case is

$$\begin{aligned}
R_1 + R_2 &\leq \mathsf{I_g}\left(\mathsf{snr}\right) - \mathsf{I_g}\left(\mathsf{inr}\right) + \mathsf{I_g}\left(\mathsf{snr} + \mathsf{inr}\right) \\
&= \mathsf{I_g}\left(\mathsf{snr}\right) + \mathsf{I_g}\left(\frac{\mathsf{snr}}{1 + \mathsf{inr}}\right) \\
&\leq \mathsf{I_g}\left(\mathsf{snr}\right) + \frac{1}{2}\log(2),
\end{aligned}$$

that is

$$\mathcal{R}_{\mathrm{out}}^{(\mathrm{F})} = \bigcup_{t \in [0,1]} \left\{ \begin{array}{l} R_1 \leq t\left(\mathsf{I_g}\left(\mathsf{snr}\right) + \frac{1}{2}\log(2)\right) \\ \\ R_2 \leq +(1-t)\left(\mathsf{I_g}\left(\mathsf{snr}\right) + \frac{1}{2}\log(2)\right) \end{array} \right\}.$$

Appendix F (Continued)

Inner Bound for $\mathsf{inr} \leq \mathsf{snr} \leq 1 + \mathsf{inr}$

We only use the discrete part of the mixed inputs and set

$$N_1 = \mathsf{N_d}\left(\mathsf{snr}_{1,t}\right), \ \mathsf{snr}_{1,t} := \left(1+\mathsf{snr}\right)^{t} - 1 \leq \mathsf{snr}, \tag{F.1a}$$

$$N_2 = \mathsf{N_d}\left(\mathsf{snr}_{2,t}\right), \ \mathsf{snr}_{2,t} := \left(1+\mathsf{snr}\right)^{1-t} - 1 \leq \mathsf{snr}, \tag{F.1b}$$

$$\delta_1 = 0, \tag{F.1c}$$

$$\delta_2 = 0. \tag{F.1d}$$

Note that

$$N_1^2 N_2^2 - 1 \leq (1+\mathsf{snr}_{1,t})(1+\mathsf{snr}_{2,t}) - 1 = \mathsf{snr}. \tag{F.1e}$$

We lower bound the minimum distance of the sum-set constellations as in (Eq. 4.20) and we get

$$\begin{aligned}
\min_{i\in[1:2]} \frac{d^2_{\min(S_i)}}{12} &\geq \kappa^2_{\gamma,N_1,N_2} \min\left(\frac{\min(\mathsf{snr},\mathsf{inr})}{\max(N_1^2,N_2^2)-1}, \frac{\max(\mathsf{snr},\mathsf{inr})}{N_1^2N_2^2-1}\right) \\
&\overset{\text{for } \mathsf{inr} \leq \mathsf{snr} \text{ and (Eq. F.1)}}{\geq} \kappa^2_{\gamma,N_1,N_2} \min\left(\frac{\mathsf{inr}}{\mathsf{snr}}, \frac{\mathsf{snr}}{\mathsf{snr}}\right) \\
&\overset{\text{for } \mathsf{snr} \leq 1+\mathsf{inr}}{\geq} \kappa^2_{\gamma,N_1,N_2} \min\left(\frac{\mathsf{inr}}{1+\mathsf{inr}}, 1\right) \\
&= \kappa^2_{\gamma,N_1,N_2} \frac{\mathsf{inr}}{1+\mathsf{inr}} \\
&\overset{\text{for } 1 \leq \mathsf{inr}}{\geq} \kappa^2_{\gamma,N_1,N_2} \frac{1}{2}.
\end{aligned}$$

Appendix F (Continued)

Gap for $\mathsf{inr} \leq \mathsf{snr} \leq 1 + \mathsf{inr}$

Similarly to the strong interference regime, we can upper bound the difference between the upper and lower bounds as

$$\begin{aligned}
\mathsf{G_d} &\leq \max\left(\mathsf{I_g}\left(\mathsf{snr}_{1,t}\right) + \frac{t}{2}\log(2) - \log(\mathsf{N_d}\left(\mathsf{snr}_{1,t}\right)),\ \mathsf{I_g}\left(\mathsf{snr}_{2,t}\right) + \frac{1-t}{2}\log(2) - \log(\mathsf{N_d}\left(\mathsf{snr}_{2,t}\right))\right) \\
&\quad + \frac{1}{2}\log\left(\frac{2\pi\mathrm{e}}{12}\right) + \frac{1}{2}\log\left(1 + \frac{2}{\kappa^2_{\gamma,N_1,N_2}}\right) \\
&\leq \frac{1}{2}\log\left(\frac{4\pi\mathrm{e}}{3}\right) + \frac{1}{2}\log\left(1 + 8\cdot\frac{\left(1 + 1/2\ln(1+\mathsf{snr})\right)^2}{\gamma^2}\right).
\end{aligned}$$

Appendix G

AUXILIARY RESULTS FOR REGIME WEAK1

We derive here some auxiliary results for the regime in (Eq. 4.28), namely

$$\begin{aligned}
&(1+\mathsf{inr}) \leq \mathsf{snr} \leq \mathsf{inr}(1+\mathsf{inr}), \\
&\frac{1+\mathsf{snr}}{1+\mathsf{inr}+\frac{\mathsf{snr}}{1+\mathsf{inr}}} \leq \frac{1+\mathsf{inr}+\frac{\mathsf{snr}}{1+\mathsf{inr}}}{1+\frac{\mathsf{snr}}{1+\mathsf{inr}}}.
\end{aligned}$$

G.0.1 Derivation of (Eq. 4.37)

The parameters of the mixed inputs are given in (Eq. 4.36). We aim to derive bounds on $\max(N_1^2, N_2^2)$ and $N_1^2 N_2^2$ and used them to find the lower bound on minimum distance in (Eq. 4.37).

The mixed input parameters are given in (Eq. 4.36). We have

$$\begin{aligned}
\max(N_1^2, N_2^2) - 1 &\leq \max(\mathsf{snr}_{1,a,t}, \mathsf{snr}_{1,b,t}) \\
&\leq \max\left(\frac{1+\mathsf{inr}+\frac{\mathsf{snr}}{1+\mathsf{inr}}}{1+\frac{\mathsf{snr}}{1+\mathsf{inr}}}, \frac{1+\mathsf{snr}}{1+\mathsf{inr}+\frac{\mathsf{snr}}{1+\mathsf{inr}}}\right) - 1 \\
&\overset{\text{from (Eq. 4.28)}}{\leq} \frac{1+\mathsf{inr}+\frac{\mathsf{snr}}{1+\mathsf{inr}}}{1+\frac{\mathsf{snr}}{1+\mathsf{inr}}} - 1 \\
&= \frac{\mathsf{inr}}{1+\frac{\mathsf{snr}}{1+\mathsf{inr}}} \qquad \text{(G.1)} \\
&\leq \mathsf{inr} = \min(\mathsf{snr}, \mathsf{inr}),
\end{aligned}$$

and

$$
\begin{aligned}
N_1^2 N_2^2 - 1 &\leq (\mathsf{snr}_{1,a,t}+1)(\mathsf{snr}_{1,b,t}+1) - 1 \\
&\overset{\text{from (Eq. 4.36)}}{\leq} \left(\frac{1+\mathsf{inr}+\frac{\mathsf{snr}}{1+\mathsf{inr}}}{1+\frac{\mathsf{snr}}{1+\mathsf{inr}}}\right)\left(\frac{1+\mathsf{snr}}{1+\mathsf{inr}+\frac{\mathsf{snr}}{1+\mathsf{inr}}}\right) - 1 \\
&= \frac{1+\mathsf{snr}}{1+\frac{\mathsf{snr}}{1+\mathsf{inr}}} - 1 \\
&= \frac{\mathsf{inr}\frac{\mathsf{snr}}{1+\mathsf{inr}}}{1+\frac{\mathsf{snr}}{1+\mathsf{inr}}} \\
&\leq \mathsf{inr}.
\end{aligned} \tag{G.2}
$$

Recall the definition of κ_{γ,N_1,N_2} in (Eq. 4.31e). By plugging the bounds in (Eq. G.1)-(Eq. G.2) into (Eq. 4.31) we get

$$
\begin{aligned}
\min_{i\in[1:2]} \frac{d^2_{\min(S_i)}}{12\,\kappa^2_{\gamma,N_1,N_2}} &\geq \frac{1-\max(\delta_1,\delta_2)}{1+(\mathsf{snr}+\mathsf{inr})\max(\delta_1,\delta_2)} \min\left(\frac{\mathsf{inr}}{\max(N_1^2,N_2^2)-1}, \frac{\mathsf{snr}}{N_1^2N_2^2-1}\right) \\
&\geq \frac{\mathsf{inr}}{1+\mathsf{snr}+2\mathsf{inr}} \min\left(\frac{\mathsf{inr}(1+\mathsf{snr}+\mathsf{inr})}{\mathsf{inr}(1+\mathsf{inr})}, \frac{\mathsf{snr}(1+\mathsf{inr}+\mathsf{snr})}{\mathsf{snr}\mathsf{inr}}\right) \\
&= \frac{1+\mathsf{snr}+\mathsf{inr}}{1+\mathsf{snr}+2\mathsf{inr}} \cdot \frac{\mathsf{inr}}{1+\mathsf{inr}} \\
&\overset{1\leq\mathsf{inr}\leq\mathsf{snr}}{\geq} \frac{1+2\mathsf{inr}}{1+3\mathsf{inr}} \cdot \frac{\mathsf{inr}}{1+\mathsf{inr}} \geq \frac{3}{8}.
\end{aligned} \tag{G.3}
$$

Note that the above derivation assumes $1 \leq \mathsf{inr}$; this restriction is without loss of generality since for $\mathsf{inr} \leq 1$ TIN with Gaussian codebooks is optimal to within $1/2$ bit (26). Note also that the minimum distance lower bound holds up to an outage set of measure less than γ, where γ is a tunable parameter; the reason why we need an outage set in this regime is the same as in Remark 10.

Appendix G (Continued)

G.0.2 Gap Derivation for $\mathcal{R}_{2R_1+R_2}$ for Regime Weak 1

Outer Bound $\mathcal{R}_{2R_1+R_2}$

With the corner point expressions in (Eq. 4.34) we write the outer bound in (Eq. 4.27b) as

$$\mathcal{R}^{(4.3.4)}_{2R_1+R_2} = \bigcup_{t\in[0,1]} \left\{ \begin{array}{l} R_1 \le \frac{t}{2}\log\left(\frac{1+\mathsf{inr}+\frac{\mathsf{snr}}{1+\mathsf{inr}}}{1+\frac{\mathsf{snr}}{1+\mathsf{inr}}}\right) + \frac{1-t}{2}\log\left(\frac{1+\mathsf{snr}}{1+\frac{\mathsf{snr}}{1+\mathsf{inr}}}\right) + \frac{1}{2}\log\left(1+\frac{\mathsf{snr}}{1+\mathsf{inr}}\right) \\ \qquad =: \mathsf{I_g}(\mathsf{snr}_{2,a,t}) + \mathsf{I_g}\left(\frac{\mathsf{snr}}{1+\mathsf{inr}}\right) \\ R_2 \le \frac{t}{2}\log\left(\frac{1+\mathsf{inr}+\mathsf{snr}}{1+\mathsf{inr}+\frac{\mathsf{snr}}{1+\mathsf{inr}}}\cdot\frac{1+\mathsf{snr}}{1+\mathsf{inr}}\right) + (1-t)c \\ \qquad =: \mathsf{I_g}(\mathsf{snr}_{2,b,t}) + \frac{t}{2}\log\left(\frac{1+\mathsf{snr}}{1+\mathsf{inr}}\right) + (1-t)c \end{array} \right\}, \quad \text{(G.4)}$$

where $(1-t)c \le c \le \log(2)$, where the parameter c is defined in (Eq. 4.26g).

Inner Bound for $\mathcal{R}_{2R_1+R_2}$

In order to approximately achieve the points in (Eq. G.4) we pick

$$N_1 = \mathsf{N_d}\left(\mathsf{snr}_{2,a,t}\right),\ \mathsf{snr}_{2,a,t} := \left(\frac{1+\mathsf{inr}+\frac{\mathsf{snr}}{1+\mathsf{inr}}}{1+\frac{\mathsf{snr}}{1+\mathsf{inr}}}\right)^t \left(\frac{1+\mathsf{snr}}{1+\frac{\mathsf{snr}}{1+\mathsf{inr}}}\right)^{1-t} - 1, \quad \text{(G.5a)}$$

$$N_2 = \mathsf{N_d}\left(\mathsf{snr}_{2,b,t}\right),\ \mathsf{snr}_{2,b,t} := \left(\frac{1+\mathsf{inr}+\mathsf{snr}}{1+\mathsf{inr}+\frac{\mathsf{snr}}{1+\mathsf{inr}}}\right)^t - 1, \quad \text{(G.5b)}$$

$$\delta_1 = \frac{1}{1+\mathsf{inr}}, \quad \text{(G.5c)}$$

$$\delta_2 :\ \mathsf{I_g}\left(\mathsf{snr}\delta_2\right) = \frac{t}{2}\log\left(\frac{1+\mathsf{snr}}{1+\mathsf{inr}}\right) \iff \delta_2 = \left(\left(\frac{1+\mathsf{snr}}{1+\mathsf{inr}}\right)^t - 1\right)\frac{1}{\mathsf{snr}}, \quad \text{(G.5d)}$$

where the power split δ_2 in (Eq. G.5d) satisfies

$$\delta_2 \le \frac{1-\mathsf{inr}/\mathsf{snr}}{1+\mathsf{inr}} \le \frac{1}{1+\mathsf{inr}},$$

as required for the achievable rate region in (Eq. 4.30).

Gap for $\mathcal{R}_{2R_1+R_2}$

The gap between the outer bound region in (Eq. G.4) and the achievable rate region in (Eq. 4.30) with the choice in (Eq. G.5) is

$$\begin{aligned}\Delta_{R_1} &= \mathsf{I_g}\left(\mathsf{snr}_{2,a,t}\right) + \mathsf{I_g}\left(\frac{\mathsf{snr}}{1+\mathsf{inr}}\right) - \log\left(\mathsf{N_d}\left(\mathsf{snr}_{2,a,t}\right)\right) - \mathsf{I_g}\left(\frac{\mathsf{snr}}{1+\mathsf{inr}}\right) + \Delta_{\text{(Eq. 4.30)}} \\ &\le \log(2) + \Delta_{\text{(Eq. 4.30)}},\end{aligned}$$

and similarly

$$\begin{aligned}\Delta_{R_2} &= \mathsf{I_g}\left(\mathsf{snr}_{2,b,t}\right) + \frac{t}{2}\log\left(\frac{1+\mathsf{snr}}{1+\mathsf{inr}}\right) + (1-t)c - \log\left(\mathsf{N_d}\left(\mathsf{snr}_{2,b,t}\right)\right) - \frac{t}{2}\log\left(\frac{1+\mathsf{snr}}{1+\mathsf{inr}}\right) + \Delta_{\text{(Eq. 4.30)}} \\ &\le \log(2) + \log(2) + \Delta_{\text{(Eq. 4.30)}},\end{aligned}$$

since $(1-t)c \le c \le \log(2)$, where the parameter c is defined in (Eq. 4.26g).

So we are left with bounding $\Delta_{\text{(Eq. 4.30)}}$ in (Eq. 4.30), which is related to the minimum distance of the received constellations S_1 and S_2 defined in (Eq. 4.8). In Appendix G.0.3 we show that

$$\min_{i\in[1:2]} \frac{d^2_{\min(S_i)}}{12} \ge \kappa^2_{\gamma,N_1,N_2} \cdot \frac{4}{45} \tag{G.6}$$

where κ_{γ,N_1,N_2} is given in (Eq. 4.31e), and $\max(N_1^2, N_2^2) - 1 \le \mathsf{inr} = \min(\mathsf{snr}, \mathsf{inr})$. With this, we finally get that the gap for this face is bounded by

$$\begin{aligned}
\mathsf{G}_{\mathsf{d}(\text{Eq. G.7})} &\le \max(\Delta_{R_1}, \Delta_{R_2}) \\
&\le \frac{1}{2}\log\left(\frac{16\pi\mathrm{e}}{3}\right) + \frac{1}{2}\log\left(1 + \frac{45}{4}\cdot\frac{1}{\kappa^2_{\gamma,N_1,N_2}}\right) \\
&\le \frac{1}{2}\log\left(\frac{16\pi\mathrm{e}}{3}\right) + \frac{1}{2}\log\left(1 + 45\cdot\frac{(1 + 1/2\ln(1+\min(\mathsf{snr},\mathsf{inr})))^2}{\gamma^2}\right) \text{ bits.} \qquad \text{(G.7)}
\end{aligned}$$

G.0.3 Derivation of (Eq. G.6)

We aim to derive different bounds involving N_1^2 and N_2^2 and used them in the minimum distance lower bound in (Eq. 4.31).

From (Eq. G.5a) we have

$$\begin{aligned}
N_1^2 - 1 &\le \mathsf{snr}_{2,a,t} \le \max(\mathsf{snr}_{2,a,0}, \mathsf{snr}_{2,a,1}) \\
&\le \frac{\max\left(1 + \mathsf{inr} + \frac{\mathsf{snr}}{1+\mathsf{inr}}, 1 + \mathsf{snr}\right)}{1 + \frac{\mathsf{snr}}{1+\mathsf{inr}}} - 1 \\
&\overset{\text{for } 1+\mathsf{inr}\le\mathsf{snr}}{\le} \frac{1+\mathsf{snr}}{1+\frac{\mathsf{snr}}{1+\mathsf{inr}}} - 1 \\
&= \mathsf{inr}\cdot\frac{\frac{\mathsf{snr}}{1+\mathsf{inr}}}{1+\frac{\mathsf{snr}}{1+\mathsf{inr}}} \le \mathsf{inr}; \qquad \text{(G.8)}
\end{aligned}$$

Appendix G (Continued)

from (Eq. G.5b) we have

$$\begin{aligned} N_2^2 - 1 \leq \mathsf{snr}_{2,b,t} &\leq \max(\mathsf{snr}_{2,b,0}, \mathsf{snr}_{2,b,1}) \\ &\leq \frac{1+\mathsf{inr}+\mathsf{snr}}{1+\mathsf{inr}+\frac{\mathsf{snr}}{1+\mathsf{inr}}} - 1 \\ &= \frac{\mathsf{inr}\cdot\frac{\mathsf{snr}}{1+\mathsf{inr}}}{1+\mathsf{inr}+\frac{\mathsf{snr}}{1+\mathsf{inr}}} \\ &\leq \min\left(\mathsf{inr}, \frac{\mathsf{snr}}{1+\mathsf{inr}}\right) = \frac{\mathsf{snr}}{1+\mathsf{inr}}, \end{aligned} \tag{G.9}$$

finally

$$\max(N_1^2, N_2^2) - 1 \leq \max\left(\mathsf{inr}, \frac{\mathsf{snr}}{1+\mathsf{inr}}\right) = \mathsf{inr} = \min(\mathsf{snr}, \mathsf{inr}). \tag{G.10}$$

We also have

$$\begin{aligned} N_1^2 N_2^2 - 1 &\leq (\mathsf{snr}_{2,a,t}+1)(\mathsf{snr}_{2,b,t}+1) - 1 \\ &= \frac{(1+\mathsf{snr})^{1-t}\,(1+\mathsf{inr}+\mathsf{snr})^t}{1+\frac{\mathsf{snr}}{1+\mathsf{inr}}} - 1 \\ &\leq \frac{1+\mathsf{inr}+\mathsf{snr}}{1+\frac{\mathsf{snr}}{1+\mathsf{inr}}} - 1 = \mathsf{inr}. \end{aligned} \tag{G.11}$$

In this regime, as we shall soon see, it also important to bound

$$\begin{aligned} (1+\mathsf{snr}\delta_2)(1+\mathsf{snr}_{2,a,t}) &= \frac{1+\mathsf{snr}}{1+\frac{\mathsf{snr}}{1+\mathsf{inr}}}\left(\frac{1+\mathsf{inr}+\frac{\mathsf{snr}}{1+\mathsf{inr}}}{1+\mathsf{inr}}\right)^t \\ &\leq \frac{1+\mathsf{snr}}{1+\mathsf{snr}+\mathsf{inr}}\left(1+\mathsf{inr}+\frac{\mathsf{snr}}{1+\mathsf{inr}}\right). \end{aligned} \tag{G.12}$$

Appendix G (Continued)

We next bound the minimum distances. Recall that κ_{γ,N_1,N_2} is given in (Eq. 4.31e).

With (Eq. 4.31c) we have

$$\begin{aligned}
\frac{d^2_{\min(S_1)}}{12\,\kappa^2_{\gamma,N_1,N_2}} &\geq \frac{1-\max(\delta_1,\delta_2)}{1+\mathsf{snr}\delta_1+\mathsf{inr}\delta_2}\min\left(\frac{\mathsf{inr}}{N_2^2-1},\frac{\mathsf{snr}}{N_1^2N_2^2-1}\right)\\
&\overset{(a)}{\geq} \frac{\mathsf{inr}}{1+\mathsf{snr}+2\mathsf{inr}}\min\left(\frac{\mathsf{inr}}{N_2^2-1},\frac{\mathsf{snr}}{N_1^2N_2^2-1}\right)\\
&\overset{(b)}{\geq} \frac{\mathsf{inr}}{1+\mathsf{snr}+2\mathsf{inr}}\min\left(\frac{\mathsf{inr}\,(1+\mathsf{inr})}{\mathsf{snr}},\frac{\mathsf{snr}}{\mathsf{inr}}\right)\\
&= \min\left(\frac{\mathsf{inr}^2\,(1+\mathsf{inr})}{\mathsf{snr}(1+\mathsf{snr}+2\mathsf{inr})},\frac{\mathsf{snr}}{1+\mathsf{snr}+2\mathsf{inr}}\right)\\
&\overset{(c)}{\geq} \min\left(\frac{\mathsf{inr}^2\,(1+\mathsf{inr})\left(1+\frac{\mathsf{snr}}{1+\mathsf{inr}}\right)}{(1+\mathsf{snr}+2\mathsf{inr})\left(1+\mathsf{inr}+\frac{\mathsf{snr}}{1+\mathsf{inr}}\right)^2},\frac{\mathsf{snr}}{1+\mathsf{snr}+2\mathsf{inr}}\right)\\
&\overset{(d)}{\geq} \min\left(\frac{\mathsf{inr}^2\,(1+\mathsf{inr}+\mathsf{snr})}{(1+\mathsf{snr}+2\mathsf{inr})(1+2\mathsf{inr})^2},\frac{1+\mathsf{inr}}{2+3\mathsf{inr}}\right)\\
&\overset{(e)}{\geq} \min\left(\frac{2\mathsf{inr}^2\,(1+\mathsf{inr})}{(2+3\mathsf{inr})\,(1+2\mathsf{inr})^2},\frac{1+\mathsf{inr}}{2+3\mathsf{inr}}\right)\\
&\overset{(f)}{\geq} \min\left(\frac{4}{45},\frac{1}{3}\right)=\frac{4}{45},
\end{aligned} \tag{G.13}$$

where the inequalities follow since: (a) $\delta_2 \leq \delta_1 = \frac{1}{1+\mathsf{inr}}$; (b) from (Eq. G.9) and (Eq. G.11); (c) from (Eq. 4.29) we have $\mathsf{snr} \leq 1+\mathsf{snr} \leq \frac{\left(1+\mathsf{inr}+\frac{\mathsf{snr}}{1+\mathsf{inr}}\right)^2}{1+\frac{\mathsf{snr}}{1+\mathsf{inr}}}$; (d) where we have used $1+\mathsf{inr} \leq \mathsf{snr}$, $\frac{\mathsf{snr}}{1+\mathsf{inr}} \leq \mathsf{inr}$ for 1st term must use largest $\frac{\mathsf{snr}}{1+\mathsf{inr}}$ while for 2nd smallest snr which $1+\mathsf{inr}$, (e) since $1+\mathsf{inr} \leq \mathsf{snr}$; and (f) comes from using $1 \leq \mathsf{inr}$.

Appendix G (Continued)

With (Eq. 4.31d), and recalling that $1+\mathsf{snr}\delta_2 = \left(\frac{1+\mathsf{snr}}{1+\mathsf{inr}}\right)^t$ from (Eq. G.5d) and $\mathsf{snr}_{2,a,t}$ in (Eq. G.5a), we have

$$
\begin{aligned}
\frac{d^2_{\min(S_2)}}{12\,\kappa^2_{\gamma,N_1,N_2}} &\geq \frac{1-\max(\delta_1,\delta_2)}{1+\mathsf{snr}\delta_2+\mathsf{inr}\delta_1}\min\left(\frac{\mathsf{inr}}{N_1^2-1},\frac{\mathsf{snr}}{N_1^2N_2^2-1}\right)\\
&\overset{(a)}{\geq} \frac{\frac{\mathsf{inr}}{1+\mathsf{inr}}}{1+\mathsf{snr}\delta_2+\frac{\mathsf{inr}}{1+\mathsf{inr}}}\min\left(\frac{\mathsf{inr}}{\mathsf{snr}_{2,a,t}},\frac{\mathsf{snr}}{\mathsf{inr}}\right)\\
&\overset{(b)}{\geq} \min\left(\frac{\frac{\mathsf{inr}^2}{1+\mathsf{inr}}}{2(1+\mathsf{snr}\delta_2)(1+\mathsf{snr}_{2,a,t})},\frac{\frac{\mathsf{snr}}{1+\mathsf{inr}}}{1+\frac{\mathsf{snr}}{1+\mathsf{inr}}+\frac{\mathsf{inr}}{1+\mathsf{inr}}}\right)\\
&\overset{(c)}{\geq} \min\left(\frac{\frac{\mathsf{inr}^2}{1+\mathsf{inr}}(1+\mathsf{snr}+\mathsf{inr})}{(1+\mathsf{snr})\left(1+\mathsf{inr}+\frac{\mathsf{snr}}{1+\mathsf{inr}}\right)},\frac{\mathsf{snr}}{1+\mathsf{snr}+2\mathsf{inr}}\right)\\
&\overset{(d)}{\geq} \min\left(\frac{\mathsf{inr}^2(1+\mathsf{snr}+\mathsf{inr})}{(1+\mathsf{inr})(1+\mathsf{snr})\,(1+2\mathsf{inr})},\frac{\mathsf{snr}}{1+\mathsf{snr}+2\mathsf{inr}}\right)\\
&\overset{(e)}{\geq} \min\left(\frac{\mathsf{inr}^2(1+\mathsf{inr})^2}{(1+\mathsf{inr})(1+\mathsf{inr}+\mathsf{inr}^2)\,(1+2\mathsf{inr})},\frac{1+\mathsf{inr}}{2+3\mathsf{inr}}\right)\\
&\overset{(f)}{\geq} \min\left(\frac{2}{9},\frac{1}{3}\right)=\frac{2}{9}
\end{aligned}
\tag{G.14}
$$

where the inequalities follow since: (a) $\delta_2 \leq \delta_1 = \frac{1}{1+\mathsf{inr}}$, $N_1^2-1 \leq \mathsf{snr}_{2,a,t}$ and (Eq. G.11); (b) $(1+\mathsf{snr}\delta_2)+\frac{\mathsf{inr}}{1+\mathsf{inr}} \leq 2+\mathsf{snr}\delta_2 \leq 2(1+\mathsf{snr}\delta_2)$ and $\delta_2 \leq \frac{1}{1+\mathsf{inr}}$, and the rest of the inequalities from the definition of weak interference $1 \leq \mathsf{inr},\ 1+\mathsf{inr} \leq \mathsf{snr} \leq \mathsf{inr}(1+\mathsf{inr})$; (c) from (Eq. G.12); (d) since $\frac{\mathsf{snr}}{1+\mathsf{inr}} \leq \mathsf{inr}$; (e) where we have used $1+\mathsf{inr} \leq \mathsf{snr}$, $\frac{\mathsf{snr}}{1+\mathsf{inr}} \leq \mathsf{inr}$ for 1st term must use largest $\frac{\mathsf{snr}}{1+\mathsf{inr}}$ while for 2nd smallest snr which $1+\mathsf{inr}$; and (f) comes from using $1 \leq \mathsf{inr}$.

By putting together (Eq. G.13) and (Eq. G.14), we obtain (Eq. G.6).

Appendix H

AUXILIARY RESULTS FOR REGIME WEAK2

We derive here some auxiliary results for the regime in (Eq. 4.29), namely

$$(1+\mathsf{inr}) \leq \mathsf{snr} \leq \mathsf{inr}(1+\mathsf{inr}),$$
$$\frac{1+\mathsf{snr}}{1+\mathsf{inr}+\frac{\mathsf{snr}}{1+\mathsf{inr}}} \geq \frac{1+\mathsf{inr}+\frac{\mathsf{snr}}{1+\mathsf{inr}}}{1+\frac{\mathsf{snr}}{1+\mathsf{inr}}}.$$

H.0.1 Derivation of (Eq. 4.44)

We aim to derive different bounds on N_1^2, N_2^2, δ_1 and δ_1 so as to obtaine in the minimum distance lower bound in (Eq. 4.44).

The mixed input parameters are in (Eq. 4.43). We have

$$\begin{aligned}
N_1^2 - 1 &\leq \mathsf{snr}_{3,a,t} \leq \max(\mathsf{snr}_{3,a,0}, \mathsf{snr}_{3,a,1}) \\
&= \max\left(\frac{\left(1+\frac{\mathsf{snr}}{1+\mathsf{inr}}\right)(1+\mathsf{snr})}{1+\mathsf{inr}+\frac{\mathsf{snr}}{1+\mathsf{inr}}}, \frac{\left(1+\mathsf{inr}+\frac{\mathsf{snr}}{1+\mathsf{inr}}\right)^3}{\left(1+\frac{\mathsf{snr}}{1+\mathsf{inr}}\right)(1+\mathsf{snr})}\right)^{\frac{1}{2}} - 1 \\
&\overset{\text{from (Eq. 4.29)}}{=} \left(1+\frac{\mathsf{snr}}{1+\mathsf{inr}}\right)\sqrt{\frac{1+\mathsf{snr}}{\left(1+\mathsf{inr}+\frac{\mathsf{snr}}{1+\mathsf{inr}}\right)\left(1+\frac{\mathsf{snr}}{1+\mathsf{inr}}\right)}} - 1 \\
&\leq \frac{\mathsf{snr}}{1+\mathsf{inr}}.
\end{aligned} \tag{H.1}$$

Appendix H (Continued)

Similarly we have

$$N_2^2 - 1 \leq \mathsf{snr}_{3,b,t} \leq \frac{\mathsf{snr}}{1+\mathsf{inr}}. \tag{H.2}$$

The bounds in (Eq. H.1)-(Eq. H.2) imply

$$\max(\delta_1, \delta_2) \leq \frac{\max(\mathsf{snr}_{3,a,t}, \mathsf{snr}_{3,b,t})}{\mathsf{snr}} \leq \frac{1}{1+\mathsf{inr}}. \tag{H.3}$$

Finally we have

$$\max(N_1^2, N_2^2) - 1 \leq \frac{\mathsf{snr}}{1+\mathsf{inr}} \leq \mathsf{inr} = \min(\mathsf{snr}, \mathsf{inr}), \tag{H.4}$$

by the definition of this regime.

We also have

$$\begin{aligned} N_1^2 N_2^2 - 1 &\leq (1+\mathsf{snr}_{3,a,t})(1+\mathsf{snr}_{3,b,t}) - 1 \\ &= (1+\mathsf{snr}\delta_1)(1+\mathsf{snr}_{3,b,t}) - 1 \\ &= (1+\mathsf{snr}_{3,a,t})(1+\mathsf{snr}\delta_2) - 1 \\ &= \mathsf{inr} + \frac{\mathsf{snr}}{1+\mathsf{inr}} \leq 2\mathsf{inr}. \end{aligned} \tag{H.5}$$

Appendix H (Continued)

With (Eq. 4.31c) we have

$$\begin{aligned}
\frac{d^2_{\min(S_1)}}{12\,\kappa^2_{\gamma,N_1,N_2}} &\geq \frac{1-\max(\delta_1,\delta_2)}{1+\mathsf{snr}\delta_1+\mathsf{inr}\delta_2}\min\left(\frac{\mathsf{inr}}{N_2^2-1},\frac{\mathsf{snr}}{N_1^2N_2^2-1}\right)\\
&\overset{(a)}{\geq} \frac{\frac{\mathsf{inr}}{1+\mathsf{inr}}}{1+\mathsf{snr}_{3,a,t}+\frac{\mathsf{inr}}{1+\mathsf{inr}}}\min\left(\frac{\mathsf{inr}}{\mathsf{snr}_{3,b,t}},\frac{\mathsf{snr}}{2\mathsf{inr}}\right)\\
&\overset{(b)}{\geq} \min\left(\frac{\frac{\mathsf{inr}^2}{1+\mathsf{inr}}}{2\left(1+\mathsf{inr}+\frac{\mathsf{snr}}{1+\mathsf{inr}}\right)},\frac{\mathsf{snr}\frac{\mathsf{inr}}{1+\mathsf{inr}}}{(1+\frac{\mathsf{snr}}{1+\mathsf{inr}}+\frac{\mathsf{inr}}{1+\mathsf{inr}})2\mathsf{inr}}\right)\\
&= \min\left(\frac{\mathsf{inr}^2}{2\left[(1+\mathsf{inr})^2+\mathsf{snr}\right]},\frac{\mathsf{snr}}{2\left[1+2\mathsf{inr}+\mathsf{snr}\right]}\right)\\
&\overset{(c)}{\geq} \min\left(\frac{\mathsf{inr}^2}{2\left[(1+\mathsf{inr})^2+\mathsf{inr}(1+\mathsf{inr})\right]},\frac{\mathsf{inr}(1+\mathsf{inr})}{2\left[1+2\mathsf{inr}+\mathsf{inr}(1+\mathsf{inr})\right]}\right)\\
&\overset{(c)}{\geq} \min\left(\frac{\mathsf{inr}^2}{2\left[(1+\mathsf{inr})^2+\mathsf{inr}(1+\mathsf{inr})\right]},\frac{1+\mathsf{inr}}{2\left[2+3\mathsf{inr}\right]}\right)\\
&= \min\left(\frac{1}{12},\frac{1}{6}\right)=\frac{1}{12},
\end{aligned} \tag{H.6}$$

where the inequalities follow from: (a) using (Eq. 4.43c), (Eq. H.3) and (Eq. H.5); (b) using (Eq. H.5); and (c) since $1 \leq \mathsf{inr}$ and $1+\mathsf{inr} \leq \mathsf{snr} \leq \mathsf{inr}(1+\mathsf{inr})$.

By symmetry an equivalent bound can be derived for $d^2_{\min(S_2)}$.

Hence minimum distance in (Eq. 4.44) is bounded by

$$\min_{i\in[1:2]}\frac{d^2_{\min(S_i)}}{12} \geq \kappa^2_{\gamma,N_1,N_2}\frac{1}{12}. \tag{H.7}$$

Appendix H (Continued)

H.0.2 Gap for $\mathcal{R}_{2R_1+R_2}$ using *common* $\to$ *discrete* map

Outer Bound $\mathcal{R}_{2R_1+R_2}$

With the corner point expressions in (Eq. 4.41) we write the outer bound in (Eq. 4.27b) as

$$\mathcal{R}^{(4.3.5)}_{2R_1+R_2} = \bigcup_{t\in[0,1]} \left\{ \begin{array}{l} R_1 \leq \frac{1-t}{2}\log\left(\frac{\left(1+\frac{\mathsf{snr}}{1+\mathsf{inr}}\right)(1+\mathsf{snr})}{1+\mathsf{inr}+\frac{\mathsf{snr}}{1+\mathsf{inr}}}\right) + \frac{t}{2}\log\left(1+\mathsf{snr}\right) \\ \qquad =: \mathsf{I_g}\left(\mathsf{snr}_{4,a,t}\right) + \mathsf{I_g}\left(\frac{\mathsf{snr}}{1+\mathsf{inr}}\right) \\ R_2 \leq \frac{1-t}{2}\log\left(\frac{\left(1+\mathsf{inr}+\frac{\mathsf{snr}}{1+\mathsf{inr}}\right)^3}{1+\mathsf{snr}}\frac{(1+\mathsf{inr})}{1+\mathsf{inr}+\mathsf{snr}}\right) + tc \\ \qquad =: \mathsf{I_g}(\mathsf{snr}_{4,b,t}) + \frac{1-t}{2}\log\left(\frac{\left(1+\mathsf{inr}+\frac{\mathsf{snr}}{1+\mathsf{inr}}\right)(1+\mathsf{inr})}{1+\mathsf{inr}+\mathsf{snr}}\right) + tc \end{array} \right\}, \tag{H.8}$$

where $tc \leq c \leq \log(2)$, where the parameter c is defined in (Eq. 4.26g), and $0 \leq t \leq 1$.

Inner Bound for $\mathcal{R}_{2R_1+R_2}$

In order to approximately achieve the points in $\mathcal{R}_{2R_1+R_2}$ in (Eq. H.8) we pick

$$N_1 = \mathsf{N_d}\left(\mathsf{snr}_{4,a,t}\right),\ \mathsf{snr}_{4,a,t} := \frac{1+\mathsf{snr}}{\left(1+\frac{\mathsf{snr}}{1+\mathsf{inr}}\right)^t\left(1+\mathsf{inr}+\frac{\mathsf{snr}}{1+\mathsf{inr}}\right)^{1-t}} - 1, \tag{H.9a}$$

$$N_2 = \mathsf{N_d}\left(\mathsf{snr}_{4,b,t}\right),\ \mathsf{snr}_{4,b,t} := \left(\frac{(1+\mathsf{inr}+\frac{\mathsf{snr}}{1+\mathsf{inr}})^2}{1+\mathsf{snr}}\right)^{1-t} - 1, \tag{H.9b}$$

$$\delta_1 := \frac{1}{1+\mathsf{inr}}, \tag{H.9c}$$

$$\delta_2 : \mathsf{I_g}\left(\mathsf{snr}\delta_2\right) = \frac{1-t}{2}\log\left(\frac{1+\mathsf{inr}+\frac{\mathsf{snr}}{1+\mathsf{inr}}}{1+\frac{\mathsf{snr}}{1+\mathsf{inr}}}\right) \Longleftrightarrow \delta_2 = \left(\left(\frac{1+\mathsf{inr}+\frac{\mathsf{snr}}{1+\mathsf{inr}}}{1+\frac{\mathsf{snr}}{1+\mathsf{inr}}}\right)^{1-t} - 1\right)\frac{1}{\mathsf{snr}}, \tag{H.9d}$$

Appendix H (Continued)

where in Appendix H.0.3, eq.(Eq. H.15), we show that

$$\delta_2 \leq \delta_1 = \frac{1}{1+\mathsf{inr}},$$

as required for the achievable region in (Eq. 4.30).

Gap for $\mathcal{R}_{2R_1+R_2}$

The gap between the outer bound in (Eq. 4.27b) and achievable region in (Eq. 4.30) with the parameters in (Eq. H.9) is

$$\begin{aligned}\Delta_{R_1} &= \mathsf{I_g}(\mathsf{snr}_{4,a,t}) + \mathsf{I_g}\left(\frac{\mathsf{snr}}{1+\mathsf{inr}}\right) - \log(\mathsf{N_d}(\mathsf{snr}_{4,a,t})) - \mathsf{I_g}(\mathsf{snr}\delta_1) + \Delta_{\text{(Eq. 4.30)}} \\ &\leq \log(2) + \Delta_{\text{(Eq. 4.30)}},\end{aligned}$$

and similarly we have that

$$\begin{aligned}\Delta_{R_2} &= \mathsf{I_g}(\mathsf{snr}_{4,b,t}) + \frac{1-t}{2}\log\left(\frac{\left(1+\mathsf{inr}+\frac{\mathsf{snr}}{1+\mathsf{inr}}\right)(1+\mathsf{inr})}{1+\mathsf{inr}+\mathsf{snr}}\right) + tc - \log(\mathsf{N_d}(\mathsf{snr}_{4,b,t})) - \mathsf{I_g}(\mathsf{snr}\delta_2) + \Delta_{\text{(Eq. 4.30)}} \\ &\leq \log(2) + \log(2) + \Delta_{\text{(Eq. 4.30)}},\end{aligned}$$

since $tc \leq c \leq \log(2)$, where the parameter c is defined in (Eq. 4.26g), and $0 \leq t \leq 1$.

So, we are left with bounding $\Delta_{\text{(Eq. 4.30)}}$ which is related to the minimum distances of the sum-set constellations. In Appendix H.0.3 we show that

$$\min_{i\in[1:2]} \frac{d^2_{\min(S_i)}}{12} \geq \kappa^2_{\gamma,N_1,N_2} \cdot \frac{1}{8}, \tag{H.10}$$

where κ_{γ,N_1,N_2} is given in (Eq. 4.31e), and that $\max(N_1^2, N_2^2) - 1 \leq \mathsf{inr}$; with this, we finally get that the gap for this face is bounded by

$$\begin{aligned}
\mathsf{G}_{\mathsf{d}\text{(Eq. H.11)}} &\leq \max(\Delta_{R_1}, \Delta_{R_2}) = 2\log(2) + \Delta_{\text{(Eq. 4.30)}} \\
&\leq \frac{1}{2}\log\left(\frac{16\pi\mathrm{e}}{3}\right) + \frac{1}{2}\log\left(1 + 32 \cdot \frac{(1 + 1/2\ln(1 + \min(\mathsf{snr}, \mathsf{inr})))^2}{\gamma^2}\right) \text{ bits.}
\end{aligned} \tag{H.11}$$

H.0.3 **<u>Proof of</u>** <u>(Eq. H.10)</u>

We first derive some bounds on N_1^2 and N_2^2 that will be useful in bounding minimum distance of the received constellations.

From (Eq. H.9a) we have

$$\begin{aligned}
N_1^2 - 1 \leq \mathsf{snr}_{4,a,t} &\leq \max(\mathsf{snr}_{4,a,0}, \mathsf{snr}_{4,a,1}) \\
&= \frac{1 + \mathsf{snr}}{\min\left(1 + \frac{\mathsf{snr}}{1+\mathsf{inr}}, 1 + \mathsf{inr} + \frac{\mathsf{snr}}{1+\mathsf{inr}}\right)} - 1 \\
&= \frac{1 + \mathsf{snr}}{1 + \frac{\mathsf{snr}}{1+\mathsf{inr}}} - 1 \\
&= \mathsf{inr} \cdot \frac{\frac{\mathsf{snr}}{1+\mathsf{inr}}}{1 + \frac{\mathsf{snr}}{1+\mathsf{inr}}} \leq \mathsf{inr}.
\end{aligned} \tag{H.12}$$

Appendix H (Continued)

Similarly form (Eq. H.9b)

$$
\begin{aligned}
N_2^2 - 1 \leq \mathsf{snr}_{4,b,t} &\leq \max(\mathsf{snr}_{4,b,0}, \mathsf{snr}_{4,b,1}) \\
&= \frac{(1 + \mathsf{inr} + \frac{\mathsf{snr}}{1+\mathsf{inr}})^2}{1 + \mathsf{snr}} - 1 \\
&\leq \frac{(1 + \mathsf{snr})(1 + \frac{\mathsf{snr}}{1+\mathsf{inr}})}{1 + \mathsf{snr}} - 1 = \frac{\mathsf{snr}}{1 + \mathsf{inr}},
\end{aligned} \tag{H.13}
$$

where inequality follow the definition of the regime in (Eq. 4.29). We also have

$$
\begin{aligned}
N_1^2 N_2^2 - 1 &\leq (1 + \mathsf{snr}_{4,a,t})(1 + \mathsf{snr}_{4,b,t}) - 1 \\
&= \left(1 + \mathsf{inr} + \frac{\mathsf{snr}}{1 + \mathsf{inr}}\right)^{1-t} \left(\frac{1 + \mathsf{snr}}{1 + \frac{\mathsf{snr}}{1+\mathsf{inr}}}\right)^{t} - 1 \\
&\leq \max\left(\mathsf{inr} + \frac{\mathsf{snr}}{1 + \mathsf{inr}}, \mathsf{inr}\frac{\frac{\mathsf{snr}}{1+\mathsf{inr}}}{1 + \frac{\mathsf{snr}}{1+\mathsf{inr}}}\right) \\
&= \mathsf{inr} + \frac{\mathsf{snr}}{1 + \mathsf{inr}} \leq 2\mathsf{inr},
\end{aligned} \tag{H.14}
$$

where the last inequality follows from $\frac{\mathsf{snr}}{1+\mathsf{inr}} \leq \mathsf{inr}$.

Appendix H (Continued)

From (Eq. H.9d) we have

$$\begin{aligned}
\mathsf{snr}\delta_2 &\leq \frac{1+\mathsf{inr}+\frac{\mathsf{snr}}{1+\mathsf{inr}}}{1+\frac{\mathsf{snr}}{1+\mathsf{inr}}} - 1 \\
&\overset{(a)}{\leq} \frac{1+\mathsf{snr}}{1+\mathsf{inr}+\frac{\mathsf{snr}}{1+\mathsf{inr}}} - 1 \\
&= \frac{\mathsf{inr}\frac{\mathsf{snr}}{1+\mathsf{inr}} - \mathsf{inr}}{1+\mathsf{inr}+\frac{\mathsf{snr}}{1+\mathsf{inr}}} \\
&\leq \frac{\mathsf{inr}\frac{\mathsf{snr}}{1+\mathsf{inr}}}{1+\mathsf{inr}+\frac{\mathsf{snr}}{1+\mathsf{inr}}} \\
&\leq \min\left(\mathsf{inr}, \frac{\mathsf{snr}}{1+\mathsf{inr}}\right) \\
&\overset{(b)}{=} \frac{\mathsf{snr}}{1+\mathsf{inr}}, \qquad \text{(H.15)}
\end{aligned}$$

where inequalities follow from: (a) using definition of the regime in (Eq. 4.29); and (b) using $\frac{\mathsf{snr}}{1+\mathsf{inr}} \leq \mathsf{inr}$.

As for the derivation in Section G.0.3, another key bound is

$$\begin{aligned}
(1+\mathsf{snr}\delta_1)(1+\mathsf{snr}_{4,b,t}) &\leq \left(1+\frac{\mathsf{snr}}{1+\mathsf{inr}}\right)\left(\frac{(1+\mathsf{inr}+\frac{\mathsf{snr}}{1+\mathsf{inr}})^2}{1+\mathsf{snr}}\right)^{1-t} \\
&\leq \left(1+\frac{\mathsf{snr}}{1+\mathsf{inr}}\right)\frac{(1+\mathsf{inr}+\frac{\mathsf{snr}}{1+\mathsf{inr}})^2}{1+\mathsf{snr}} \\
&\overset{(a)}{\leq} \frac{1+\mathsf{inr}+\mathsf{snr}}{1+\mathsf{snr}} \cdot \frac{(1+2\mathsf{inr})^2}{1+\mathsf{inr}} \qquad \text{(H.16)}
\end{aligned}$$

Appendix H (Continued)

where inequalities follow from: (a) using $\frac{\mathsf{snr}}{1+\mathsf{inr}} \le \mathsf{inr}$; and (b) using $\frac{1+\mathsf{inr}+\mathsf{snr}}{1+\mathsf{snr}} \le 2$ and $\frac{1+2\mathsf{inr}}{1+\mathsf{inr}} \le 2$. Similarly, we have

$$\begin{aligned}(1+\mathsf{snr}_{4,a,t})(1+\mathsf{snr}\delta_2) &\le \frac{1+\mathsf{snr}}{\left(1+\frac{\mathsf{snr}}{1+\mathsf{inr}}\right)^t\left(1+\mathsf{inr}+\frac{\mathsf{snr}}{1+\mathsf{inr}}\right)^{1-t}}\left(\frac{1+\mathsf{inr}+\frac{\mathsf{snr}}{1+\mathsf{inr}}}{1+\frac{\mathsf{snr}}{1+\mathsf{inr}}}\right)^{1-t} \\ &= \frac{1+\mathsf{snr}}{1+\frac{\mathsf{snr}}{1+\mathsf{inr}}} \le 1+\mathsf{inr}. \end{aligned} \tag{H.17}$$

By using (Eq. 4.31c), the minimum distance for S_1 can be bounded as

$$\begin{aligned}\frac{d^2_{\min(S_1)}}{12\,\kappa^2_{\gamma,N_1,N_2}} &\ge \frac{1-\max(\delta_1,\delta_2)}{1+\mathsf{snr}\delta_1+\mathsf{inr}\delta_2}\min\left(\frac{\mathsf{inr}}{N_2^2-1},\frac{\mathsf{snr}}{N_1^2N_2^2-1}\right) \\ &\overset{(a)}{\ge} \frac{\frac{\mathsf{inr}}{1+\mathsf{inr}}}{1+\mathsf{snr}\delta_1+\mathsf{inr}\delta_2}\min\left(\frac{\mathsf{inr}}{\mathsf{snr}_{4,b,t}},\frac{\mathsf{snr}}{2\mathsf{inr}}\right) \\ &\overset{(b)}{\ge} \min\left(\frac{\frac{\mathsf{inr}^2}{1+\mathsf{inr}}}{2(1+\mathsf{snr}\delta_1)(1+\mathsf{snr}_{4,b,t})},\frac{\mathsf{snr}}{2(1+\mathsf{snr}+2\mathsf{inr})}\right) \\ &\overset{(c)}{\ge} \min\left(\frac{\mathsf{inr}^2(1+\mathsf{snr})}{2(1+\mathsf{inr}+\mathsf{snr})(1+2\mathsf{inr})^2},\frac{1+\mathsf{inr}}{2(2+3\mathsf{inr})}\right) \\ &\overset{(d)}{\ge} \min\left(\frac{3}{8},\frac{1}{6}\right)=\frac{1}{6}\end{aligned}$$

where inequalities follow from: (a) $\max(\delta_1,\delta_2) \le \frac{1}{1+\mathsf{inr}}$ and from (Eq. H.13) and (Eq. H.14) we have that $N_2^2-1 \le \mathsf{snr}_{4,b,t}$ and $N_1^2N_2^2-1 \le 2\mathsf{inr}$; (b) from (Eq. H.15) $\delta_2 \le \frac{1}{1+\mathsf{inr}}$; (c) using (Eq. H.15) we have $\mathsf{snr}_{4,b,t}(2+\mathsf{snr}\delta_1) \le 2(1+\mathsf{snr}_{4,b,t})(1+\mathsf{snr}\delta_1)$ and (Eq. H.16); and (d) $\mathsf{snr} \ge (1+\mathsf{inr})$ and $\mathsf{inr} \ge 1$.

Appendix H (Continued)

Similarly,

$$\begin{aligned}
\frac{d^2_{\min(S_2)}}{12\ \kappa^2_{\gamma,N_1,N_2}} &\geq \frac{1-\max(\delta_1,\delta_2)}{1+\mathsf{snr}\delta_2+\mathsf{inr}\delta_1}\min\left(\frac{\mathsf{inr}}{N_1^2-1},\frac{\mathsf{snr}}{N_1^2N_2^2-1}\right)\\
&\overset{(a)}{\geq} \min\left(\frac{\frac{\mathsf{inr}^2}{1+\mathsf{inr}}}{2(1+\mathsf{snr}\delta_2)(1+\mathsf{snr}_{4,a,t})},\frac{\mathsf{snr}}{2(1+\mathsf{snr}+2\mathsf{inr})}\right)\\
&\overset{(b)}{\geq} \min\left(\frac{\mathsf{inr}^2}{2(1+\mathsf{inr})(1+\mathsf{inr})},\frac{1+\mathsf{inr}}{2(2+3\mathsf{inr})}\right)\\
&\overset{(c)}{\geq} \min\left(\frac{1}{8},\frac{1}{6}\right)=\frac{1}{8}
\end{aligned}$$

where inequalities follow from: (a) (Eq. H.9a) we have $N_1^2-1 \leq \mathsf{snr}_{4,a,t}$, from (Eq. H.15) and (Eq. H.9c) $\max(\delta_1,\delta_2) \leq \frac{1}{1+\mathsf{inr}}$ and from (Eq. H.14) $N_1^2N_2^2-1 \leq \mathsf{inr}$; (b) $\mathsf{snr}_{4,a,t}(2+\mathsf{snr}\delta_2) \leq 2(1+\mathsf{snr}_{4,a,t})(1+\mathsf{snr}\delta_2)$ and (Eq. H.17); and (c) from $\mathsf{snr} \geq (1+\mathsf{inr})$ and $\mathsf{inr} \geq 1$.

Hence, the minimum distance in (Eq. H.10) is bounded by

$$\min_{i\in[1:2]}\frac{d^2_{\min(S_i)}}{12\ \kappa^2_{\gamma,N_1,N_2}} \geq \frac{1}{8}.$$

Appendix I

CONSTANT GAP DERIVATION FOR REGIME WEAK2

I.0.1 Another Inner Bound for $\mathcal{R}_{R_1+R_2}$

In order to approximately achieve the points in $\mathcal{R}^{(4.3.5)}_{R_1+R_2}$ in (Eq. 4.42) we pick

$$N_1 = \mathsf{N_d}\left(\frac{1}{k}\mathsf{snr}_{3,a,t}\right),\ \mathsf{snr}_{3,a,t} \text{ in (Eq. 4.43a)} \tag{I.1a}$$

$$N_2 = \mathsf{N_d}\left(\frac{1}{k}\mathsf{snr}_{3,b,t}\right),\ \mathsf{snr}_{3,b,t} \text{ in (Eq. 4.43b)} \tag{I.1b}$$

$$\delta_1 : \mathsf{I_g}\left(\mathsf{snr}\delta_1\right) = \mathsf{I_g}\left(\mathsf{snr}_{3,a,t}\right) \Longleftrightarrow \delta_1 = \frac{\mathsf{snr}_{3,a,t}}{\mathsf{snr}}, \tag{I.1c}$$

$$\delta_2 : \mathsf{I_g}\left(\mathsf{snr}\delta_2\right) = \mathsf{I_g}\left(\mathsf{snr}_{3,b,t}\right) \Longleftrightarrow \delta_2 = \frac{\mathsf{snr}_{3,b,t}}{\mathsf{snr}}. \tag{I.1d}$$

where k is a parameter that we will tune in order to satisfy the non-overlap condition in Proposition 2.1.2. Indeed, in order to check whether we can use the bound in (Eq. 4.32a) we must check whether the condition in (Eq. 4.32b) holds. To simplify the analytical computations we choose to satisfy instead

$$\frac{(1-\delta_{i'})N_{i'}^2}{N_{i'}^2-1} \le k \le \frac{\mathsf{snr}}{\mathsf{inr}}\,\frac{(1-\delta_i)}{N_i^2-1} \quad \forall (i,i') \in \{(1,2),(2,1)\},$$

for some k; since $\frac{(1-\delta_{i'})N_{i'}^2}{N_{i'}^2-1} \le \frac{N_{i'}^2}{N_{i'}^2-1} \le \frac{4}{3}$ for all $N_{i'} \ge 2$, we set $\frac{4}{3} := k$. In other words, we accept an increase in gap of $\log(k) = \log(4/3)$, due to the reduction of the number of points of the discrete part of

Appendix I (Continued)

the mixed inputs from $\mathsf{N_d}(x)$ to $\mathsf{N_d}(3x/4)$ for some 'SNR' x, for ease of computations. Therefore, for the rest of this section instead of checking condition in (Eq. 4.32b) we will check the simpler condition

$$\frac{4}{3}\mathsf{inr} \leq \frac{\mathsf{snr}(1-\delta_i)}{N_i^2-1} \quad \forall i \in [1:2]. \tag{I.2}$$

The gap between the outer bound region in (Eq. 4.42) and the achievable rate in (Eq. 4.30) with the parameters in (Eq. I.1) is

$$\begin{aligned}\Delta_{R_1} &= 2\mathsf{I_g}\left(\mathsf{snr}_{3,a,t}\right) - \log\left(\mathsf{N_d}\left(3/4\ \mathsf{snr}_{3,a,t}\right)\right) - \mathsf{I_g}(\mathsf{snr}_{3,a,t}) + \Delta_{\text{(Eq. 4.30)}}\\ &\leq \log\left(\frac{8}{3}\right) + \Delta_{\text{(Eq. 4.30)}},\end{aligned}$$

and similarly

$$\Delta_{R_2} \leq \log\left(\frac{8}{3}\right) + \Delta_{\text{(Eq. 4.30)}}.$$

We are then left with bounding $\Delta_{\text{(Eq. 4.30)}}$, which depends on minimum distances of the received sum-set constellations. From (Eq. H.1)-(Eq. H.2) we have

$$\begin{aligned}N_1^2 - 1 &\leq \frac{3}{4}\mathsf{snr}_{3,a,t} \leq \frac{3}{4}\frac{\mathsf{snr}}{1+\mathsf{inr}}, \text{ from (Eq. H.1)},\\ N_2^2 - 1 &\leq \frac{3}{4}\mathsf{snr}_{3,b,t} \leq \frac{3}{4}\frac{\mathsf{snr}}{1+\mathsf{inr}}, \text{ from (Eq. H.2)},\end{aligned}$$

Appendix I (Continued)

and thus

$$\frac{\mathsf{snr}(1-\delta_i)}{N_i^2-1} \overset{\max(\delta_1,\delta_2)\le\frac{1}{1+\mathsf{inr}}}{\ge} \mathsf{inr}\frac{\frac{\mathsf{snr}}{1+\mathsf{inr}}}{N_i^2-1} \ge \frac{4}{3}\mathsf{inr}, \tag{I.3}$$

as needed in (Eq. I.2).

Therefore, by (Eq. 4.32a), for $d^2_{\min(S_1)}$ we have that

$$\begin{aligned}
\frac{d^2_{\min(S_1)}}{12} &= \frac{1}{1+\mathsf{snr}\delta_1+\mathsf{inr}\delta_2}\min\left(\frac{(1-\delta_1)\mathsf{snr}}{N_1^2-1},\frac{(1-\delta_2)\mathsf{inr}}{N_2^2-1}\right)\\
&\overset{(a)}{\ge} \frac{\frac{\mathsf{inr}}{1+\mathsf{inr}}}{1+\mathsf{snr}\delta_1+\frac{\mathsf{inr}}{1+\mathsf{inr}}}\min\left(\frac{\mathsf{snr}}{N_1^2-1},\frac{\mathsf{inr}}{N_2^2-1}\right)\\
&\overset{(b)}{\ge} \frac{\frac{\mathsf{inr}}{1+\mathsf{inr}}}{1+\mathsf{snr}\delta_1+\frac{\mathsf{inr}}{1+\mathsf{inr}}}\frac{4}{3}\min\left(1+\mathsf{inr},\frac{\mathsf{inr}}{\mathsf{snr}_{3,b,t}}\right)\\
&\overset{(c)}{\ge} \frac{4}{3}\min\left(\frac{\mathsf{inr}(1+\mathsf{inr})}{1+\mathsf{snr}+2\mathsf{inr}},\frac{\frac{\mathsf{inr}^2}{1+\mathsf{inr}}}{2(1+\mathsf{snr}\delta_1)(1+\mathsf{snr}_{3,b,t})}\right)\\
&\overset{(d)}{\ge} \frac{4}{3}\min\left(\frac{\mathsf{inr}(1+\mathsf{inr})}{1+\mathsf{snr}+2\mathsf{inr}},\frac{\mathsf{inr}^2}{2(1+\mathsf{inr})(1+\mathsf{inr}+\frac{\mathsf{snr}}{1+\mathsf{inr}})}\right)\\
&\overset{(e)}{\ge} \frac{4}{3}\min\left(\frac{\mathsf{inr}(1+\mathsf{inr})}{1+3\mathsf{inr}+\mathsf{inr}^2},\frac{\mathsf{inr}^2}{2(1+\mathsf{inr})(1+2\mathsf{inr})}\right)\\
&\ge \frac{4}{3}\min\left(\frac{2}{5},\frac{1}{12}\right)=\frac{1}{9},
\end{aligned} \tag{I.4}$$

where the inequalities follows from: (a) $\max(\delta_1,\delta_2)\le\frac{1}{1+\mathsf{inr}}$; (b) from (Eq. H.1) and (Eq. H.2); (c) $\max(\delta_1,\delta_2)\le\frac{1}{1+\mathsf{inr}}$; (d) from (Eq. H.5); and (e) from $1\le\mathsf{inr}\le\mathsf{snr}\le\mathsf{inr}(1+\mathsf{inr})$.

By symmetry, $\frac{d^2_{\min(S_2)}}{12}$ is bounded in the same way, thus

$$\min_{i\in[1:2]}\frac{d^2_{\min(S_i)}}{12}\ge\frac{1}{9}. \tag{I.5}$$

Appendix I (Continued)

Finally the gap for this face is

$$
\begin{aligned}
\mathsf{G}_{\mathsf{d}(\text{Eq. I.6})} &\le \max(\Delta_{R_1}, \Delta_{R_2}) = \log\left(\frac{8}{3}\right) + \Delta_{(\text{Eq. 4.30})} \\
&\le \log\left(\frac{8}{3}\right) + \frac{1}{2}\log\left(\frac{\pi \mathrm{e}}{3}\right) + \frac{1}{2}\log(1+9) \\
&= \frac{1}{2}\log\left(\frac{640\pi \mathrm{e}}{27}\right) \approx 3.83 \text{ bits.}
\end{aligned} \tag{I.6}
$$

I.0.2 Another Inner Bound for $\mathcal{R}_{2R_1+R_2}$

We choose the mixed input parameters as

$$N_1 = \mathsf{N_d}\left(\frac{3}{4}\,\frac{\mathsf{snr}-\mathsf{inr}}{1+\mathsf{inr}}\right), \tag{I.7a}$$

$$N_2 = \mathsf{N_d}\left(\frac{3}{4}\mathsf{snr}_{4,b,t}\right),\ \mathsf{snr}_{4,b,t} := \left(\frac{(1+\mathsf{inr}+\frac{\mathsf{snr}}{1+\mathsf{inr}})^2}{1+\mathsf{snr}}\right)^{1-t} - 1 \overset{\text{by eq.(Eq. H.13)}}{\le} \frac{\mathsf{snr}}{1+\mathsf{inr}}, \tag{I.7b}$$

$$\delta_1 = \frac{\mathsf{snr}_{4,a,t}}{\mathsf{snr}} \overset{\text{by eq.(Eq. H.12)}}{\le} \frac{\mathsf{inr}}{\mathsf{snr}}, \tag{I.7c}$$

$$\delta_2 = \frac{1+\mathsf{inr}+\frac{\mathsf{snr}}{1+\mathsf{inr}}}{\left(1+\frac{\mathsf{snr}}{1+\mathsf{inr}}\right)(1+\mathsf{snr})} \overset{\text{by eq.(Eq. 4.29)}}{\le} \frac{1}{1+\mathsf{inr}+\frac{\mathsf{snr}}{1+\mathsf{inr}}} \le \frac{1}{1+\mathsf{inr}}, \tag{I.7d}$$

where the factor $\frac{3}{4}$ in the number of points appears for the same reason as in Section I.0.1.

Appendix I (Continued)

An inequality we will need is

$$\begin{aligned}
\frac{\mathsf{snr}\delta_2}{1+\mathsf{inr}\delta_1} &\overset{(a)}{\geq} \frac{\mathsf{snr}\frac{\left(1+\mathsf{inr}+\frac{\mathsf{snr}}{1+\mathsf{inr}}\right)(1+\mathsf{inr})}{(1+\mathsf{inr}+\mathsf{snr})(1+\mathsf{snr})}}{\frac{\mathsf{inr}}{\mathsf{snr}}\frac{(1+\mathsf{snr})}{\left(1+\frac{\mathsf{snr}}{1+\mathsf{inr}}\right)^t\left(1+\mathsf{inr}+\frac{\mathsf{snr}}{1+\mathsf{inr}}\right)^{1-t}}} \\
&= \frac{\mathsf{snr}^2}{(1+\mathsf{snr})^2}\frac{\left(1+\mathsf{inr}+\frac{\mathsf{snr}}{1+\mathsf{inr}}\right)\left(1+\frac{\mathsf{snr}}{1+\mathsf{inr}}\right)^t\left(1+\mathsf{inr}+\frac{\mathsf{snr}}{1+\mathsf{inr}}\right)^{1-t}}{\mathsf{inr}\left(1+\frac{\mathsf{snr}}{1+\mathsf{inr}}\right)} \\
&= \frac{\mathsf{snr}^2}{(1+\mathsf{snr})^2}\frac{\left(1+\mathsf{inr}+\frac{\mathsf{snr}}{1+\mathsf{inr}}\right)(1+\mathsf{inr})^{1-t}\left(1+\mathsf{inr}+\frac{\mathsf{snr}}{1+\mathsf{inr}}\right)^{1-t}}{\mathsf{inr}\,(1+\mathsf{inr}+\mathsf{snr})^{1-t}} \\
&\overset{(b)}{\geq} \frac{3}{4}\left(\frac{1+\mathsf{inr}+\frac{\mathsf{snr}}{1+\mathsf{inr}}}{1+\frac{\mathsf{snr}}{1+\mathsf{inr}}}\right)^{1-t} \qquad \text{(I.8)}
\end{aligned}$$

where the inequalities follow from: (a) plugin in values of δ_1 and δ_2 and lower bounding the denominator; and (b) using $\mathsf{snr} \geq 1$ we have that $\frac{\mathsf{snr}^2}{(1+\mathsf{snr})^2} \geq \frac{1}{4}$ and using $\mathsf{snr} \geq (1+\mathsf{inr})$ we have $\frac{1+\mathsf{inr}+\frac{\mathsf{snr}}{1+\mathsf{inr}}}{\mathsf{inr}} \geq \frac{2+\mathsf{inr}}{\mathsf{inr}} \geq 3$.

Another inequality we will need is

$$\begin{aligned}
\mathsf{snr}\delta_2 &= \mathsf{snr}\frac{\left(1+\mathsf{inr}+\frac{\mathsf{snr}}{1+\mathsf{inr}}\right)(1+\mathsf{inr})}{(1+\mathsf{inr}+\mathsf{snr})(1+\mathsf{snr})} \\
&\leq \frac{\left(1+\mathsf{inr}+\frac{\mathsf{snr}}{1+\mathsf{inr}}\right)(1+\mathsf{inr})}{(1+\mathsf{inr}+\mathsf{snr})} \\
&= \frac{\left(1+\mathsf{inr}+\frac{\mathsf{snr}}{1+\mathsf{inr}}\right)}{(1+\frac{\mathsf{snr}}{1+\mathsf{inr}})} \\
&\leq \frac{(1+2\mathsf{inr})(1+\mathsf{inr})}{\mathsf{snr}} \qquad \text{(I.9)}
\end{aligned}$$

where the last inequality comes from using $\frac{\mathsf{snr}}{1+\mathsf{inr}} \leq \mathsf{inr}$ and dropping one in the denominator.

Appendix I (Continued)

Gap for $\mathcal{R}_{2R_1+R_2}$

The gap between the outer bound in (Eq. H.8) and the achievable rate in Proposition 4.2.1 with the choice of parameters in (Eq. I.7) is

$$\begin{aligned}\Delta_{R_1} &= \mathsf{I_g}\left(\mathsf{snr}_{4,a,t}\right) + \mathsf{I_g}\left(\frac{\mathsf{snr}}{1+\mathsf{inr}}\right) - \log\left(\mathsf{N_d}\left(\frac{3}{4}\,\frac{\mathsf{snr}-\mathsf{inr}}{1+\mathsf{inr}}\right)\right) - \mathsf{I_g}\left(\mathsf{snr}\delta_1\right) + \Delta_{\text{(Eq. 4.30)}}\\ &\le \log(2) + \frac{1}{2}\log(2) + \Delta_{\text{(Eq. 4.30)}},\end{aligned}$$

$$\mathsf{I_g}\left(\frac{\mathsf{snr}}{1+\mathsf{inr}}\right) - \mathsf{I_g}\left(\frac{3}{4}\,\frac{\mathsf{snr}-\mathsf{inr}}{1+\mathsf{inr}}\right) = \frac{1}{2}\log\frac{1+\mathsf{inr}+\mathsf{snr}}{1+\mathsf{inr}}\,\frac{1+\mathsf{inr}}{1+\mathsf{inr}/4+3\mathsf{snr}/4} \le \frac{1}{2}\log\frac{1+2\mathsf{snr}}{1+\mathsf{snr}} \le \frac{1}{2}\log(2),$$

and similarly

$$\begin{aligned}\Delta_{R_2} = \mathsf{I_g}(\mathsf{snr}_{4,b,t}) &+ \frac{1-t}{2}\log\left(\frac{\left(1+\mathsf{inr}+\frac{\mathsf{snr}}{1+\mathsf{inr}}\right)(1+\mathsf{inr})}{1+\mathsf{inr}+\mathsf{snr}}\right) + tc\\ &- \log\left(\mathsf{N_d}(\mathsf{snr}_{4,b,t})\right) - \mathsf{I_g}\log\left(\frac{\mathsf{snr}\delta_2}{1+\mathsf{inr}\delta_1}\right) - \frac{1}{2}\log(2) + \Delta_{\text{(Eq. 4.30)}}\\ &\le \log(2) + \frac{1}{2}\log\left(\frac{4}{3}\right) + \frac{1}{2}\log\left(\frac{4}{3}\right) + \log(2) - \frac{1}{2}\log(2) + \Delta_{\text{(Eq. 4.30)}} = \frac{1}{2}\log\left(\frac{2^7}{3^2}\right) + \Delta_{\text{(Eq. 4.30)}}\end{aligned}$$

where we have used $tc \le \log(2)$ and the bound in (Eq. I.8); the term '$-\frac{1}{2}\log(2)$' is because of the definition of $\Delta_{\text{(Eq. 4.30)}}$ that assumed $\max(\delta_1, \delta_2) \le \frac{1}{1+\mathsf{inr}}$, which is not the case here.

So, we are left with bounding $\Delta_{\text{(Eq. 4.30)}}$, which depends on the minimum distances of the received constellations. We must verify the condition in (Eq. I.2) at each receiver.

Appendix I (Continued)

For receiver 1 we have

$$\begin{aligned}\frac{\mathsf{snr}(1-\delta_1)}{N_1^2-1} &\overset{\text{from eq.(Eq. I.7c)}}{\geq} \frac{\mathsf{snr}-\mathsf{inr}}{N_1^2-1} \\ &\overset{\text{from eq.(Eq. I.7a)}}{\geq} \frac{\mathsf{snr}-\mathsf{inr}}{\frac{3}{4}\frac{\mathsf{snr}-\mathsf{inr}}{1+\mathsf{inr}}} \\ &= \frac{4}{3}(1+\mathsf{inr}) \geq \frac{4}{3}\mathsf{inr},\end{aligned} \tag{I.10}$$

and therefore

$$\begin{aligned}\frac{d^2_{\min(S_1)}}{12} &= \frac{1}{1+\mathsf{snr}\delta_1+\mathsf{inr}\delta_2}\min\left(\frac{(1-\delta_1)\mathsf{snr}}{N_1^2-1}, \frac{(1-\delta_2)\mathsf{inr}}{N_2^2-1}\right) \\ &\overset{(a)}{\geq} \frac{1}{1+\mathsf{snr}\delta_1+\mathsf{inr}\frac{1}{1+\mathsf{inr}}}\min\left(\frac{4}{3}(1+\mathsf{inr}), \frac{\frac{\mathsf{inr}^2}{1+\mathsf{inr}}}{\frac{3}{4}\mathsf{snr}_{4,b,t}}\right) \\ &= \frac{4}{3}\min\left(\frac{(1+\mathsf{inr})}{1+\mathsf{snr}\delta_1+\mathsf{inr}\frac{1}{1+\mathsf{inr}}}, \frac{\frac{\mathsf{inr}^2}{1+\mathsf{inr}}}{\left(1+\mathsf{snr}\delta_1+\mathsf{inr}\frac{1}{1+\mathsf{inr}}\right)\mathsf{snr}_{4,b,t}}\right) \\ &\overset{(b)}{\geq} \frac{4}{3}\min\left(\frac{1+\mathsf{inr}}{2+\mathsf{snr}\frac{\mathsf{inr}}{\mathsf{snr}}}, \frac{\frac{\mathsf{inr}^2}{1+\mathsf{inr}}}{2(1+\mathsf{snr}\delta_1)(1+\mathsf{snr}_{4,b,t})}\right) \\ &\overset{(c)}{\geq} \frac{4}{3}\min\left(\frac{1+\mathsf{inr}}{2+\mathsf{inr}}, \frac{\frac{\mathsf{inr}^2}{1+\mathsf{inr}}}{4(1+2\mathsf{inr})}\right) \\ &\overset{(d)}{\geq} \frac{4}{3}\min\left(\frac{2}{3}, \frac{1}{24}\right) = \frac{1}{18},\end{aligned}$$

where the bounds are obtained by: (a) using (Eq. I.10) and $\delta_2 \leq \frac{1}{1+\mathsf{inr}}$ and (Eq. H.13); (b) using $\delta_2 \leq \frac{\mathsf{inr}}{\mathsf{snr}}$ form (Eq. H.12) and $(1+\mathsf{snr}\delta_2+\mathsf{inr}\frac{1}{1+\mathsf{inr}}) \leq 2(1+\mathsf{snr}\delta_2)(1+\mathsf{snr}_{4,b,t})$; (c) using (Eq. I.7c) we have $(1+\mathsf{snr}\delta_2)(1+\mathsf{snr}_{4,b,t}) = (1+\mathsf{snr}_{4,a,t})(1+\mathsf{snr}_{4,b,t})$ and then using (Eq. H.14); and (d) come from minimizing over $\mathsf{inr} \geq 1$.

Appendix I (Continued)

For receiver 2 we have

$$\frac{\mathsf{snr}(1-\delta_2)}{N_2^2-1} \geq \frac{\mathsf{snr}\frac{\mathsf{inr}}{1+\mathsf{inr}}}{N_2^2-1} \geq \frac{\mathsf{snr}\frac{\mathsf{inr}}{1+\mathsf{inr}}}{\frac{3}{4}\frac{\mathsf{snr}}{1+\mathsf{inr}}} = \frac{4}{3}\mathsf{inr},$$

and therefore

$$\begin{aligned}
\frac{d^2_{\min(S_2)}}{12} &= \frac{1}{1+\mathsf{snr}\delta_2+\mathsf{inr}\delta_1}\min\left(\frac{(1-\delta_2)\mathsf{snr}}{N_2^2-1}, \frac{(1-\delta_1)\mathsf{inr}}{N_1^2-1}\right) \\
&\overset{(a)}{\geq} \frac{1}{1+\mathsf{snr}\delta_2+\mathsf{inr}\frac{\mathsf{inr}}{\mathsf{snr}}}\min\left(\frac{(1-\frac{1}{1+\mathsf{inr}})\mathsf{snr}}{N_2^2-1}, \frac{(1-\frac{\mathsf{inr}}{\mathsf{snr}})\mathsf{inr}}{N_1^2-1}\right) \\
&\overset{(b)}{\geq} \frac{4}{3}\frac{1}{1+\mathsf{snr}\delta_2+\mathsf{inr}\frac{\mathsf{inr}}{\mathsf{snr}}}\min\left(\frac{(1-\frac{1}{1+\mathsf{inr}})\mathsf{snr}}{\frac{\mathsf{snr}}{1+\mathsf{inr}}}, \frac{(1-\frac{\mathsf{inr}}{\mathsf{snr}})\mathsf{inr}(1+\mathsf{inr})}{\mathsf{snr}-\mathsf{inr}}\right) \\
&= \frac{4}{3}\frac{1}{1+\mathsf{snr}\delta_2+\mathsf{inr}\frac{\mathsf{inr}}{\mathsf{snr}}}\min\left(\mathsf{inr}, \frac{\mathsf{inr}(1+\mathsf{inr})}{\mathsf{snr}}\right) \\
&= \frac{4}{3}\min\left(\frac{\mathsf{inr}}{1+\mathsf{snr}\delta_2+\mathsf{inr}\frac{\mathsf{inr}}{\mathsf{snr}}}, \frac{\mathsf{inr}(1+\mathsf{inr})}{(1+\mathsf{snr}\delta_2+\mathsf{inr}\frac{\mathsf{inr}}{\mathsf{snr}})\mathsf{snr}}\right) \\
&\overset{(c)}{\geq} \frac{4}{3}\min\left(\frac{\mathsf{inr}}{1+\mathsf{snr}\frac{1}{1+\mathsf{inr}}+\mathsf{inr}\frac{\mathsf{inr}}{\mathsf{snr}}}, \frac{\mathsf{inr}(1+\mathsf{inr})}{(1+\mathsf{snr}\delta_2+\mathsf{inr}\frac{\mathsf{inr}}{\mathsf{snr}})\mathsf{snr}}\right) \\
&\overset{(d)}{\geq} \frac{4}{3}\min\left(\frac{\mathsf{inr}}{1+2\mathsf{inr}}, \frac{\mathsf{inr}(1+\mathsf{inr})}{(1+\frac{(1+2\mathsf{inr})(1+\mathsf{inr})}{\mathsf{snr}}+\mathsf{inr}\frac{\mathsf{inr}}{\mathsf{snr}})\mathsf{snr}}\right) \\
&= \frac{4}{3}\min\left(\frac{\mathsf{inr}}{1+2\mathsf{inr}}, \frac{\mathsf{inr}(1+\mathsf{inr})}{(\mathsf{snr}+(1+2\mathsf{inr})(1+\mathsf{inr})+\mathsf{inr}^2)}\right) \\
&\overset{(e)}{\geq} \frac{4}{3}\min\left(\frac{\mathsf{inr}}{1+2\mathsf{inr}}, \frac{\mathsf{inr}(1+\mathsf{inr})}{(\mathsf{inr}(1+\mathsf{inr})+(1+2\mathsf{inr})(1+\mathsf{inr})+\mathsf{inr}^2)}\right) \\
&\geq \frac{4}{3}\min\left(\frac{1}{3}, \frac{2}{10}\right) = \frac{4}{15}
\end{aligned}$$

Appendix I (Continued)

where the bounds are obtained by: (a) using $\delta_2 \leq \frac{1}{1+\mathsf{inr}}$ and $\delta_1 \leq \frac{\mathsf{inr}}{\mathsf{snr}}$; (b) from (Eq. I.7a) we have that $N_1^2 - 1 \leq \frac{3}{4}\frac{\mathsf{snr}-\mathsf{inr}}{1+\mathsf{inr}}$, (c) using $\delta_2 \leq \frac{1}{1+\mathsf{inr}}$; (d) used bound in (Eq. I.9); and (e) used bound $\mathsf{snr} \leq \mathsf{inr}(1+\mathsf{inr})$.

So, finally the gap is

$$\begin{aligned}
\mathsf{G}_{\mathsf{d}\text{(Eq. I.11)}} &\leq \max(\Delta_{R_1}, \Delta_{R_2}) \\
&= \frac{1}{2}\log\left(\frac{2^7}{3^2}\right) + \frac{1}{2}\log\left(\frac{\pi e}{3}\right) + \frac{1}{2}\log\left(1 + \frac{15}{4}\right) \\
&= \frac{1}{2}\log\left(\frac{608\ \pi e}{27}\right) \approx 3.79 \text{ bits} \qquad \text{(I.11)}
\end{aligned}$$

Overall Constant Gap for Weak 1

Therefore, the overall gap for Weak 1 is

$$\mathsf{G}_{\mathsf{d}} \leq \max(\mathsf{G}_{\mathsf{d}\text{(Eq. I.6)}}, \mathsf{G}_{\mathsf{d}\text{(Eq. I.11)}}) = \mathsf{G}_{\mathsf{d}\text{(Eq. I.6)}}.$$

Appendix J

PROOF OF PROPOSITION 5.2.2.

In order to find the point of intersection snr_L between (Eq. 5.7a) and (Eq. 5.17a) we must solve the following equation:

$$\frac{1}{\mathsf{snr}} - \frac{k_n}{\mathsf{snr}} + \frac{k_n}{\mathsf{snr}_0} - \frac{\beta}{1 + \beta\mathsf{snr}_0} = \frac{1}{\mathsf{snr}} - \frac{k_n}{\mathsf{snr}} + A = 0$$

where $A = \frac{k_n}{\mathsf{snr}_0} - \frac{\beta}{1+\beta\mathsf{snr}_0}$. By solving for snr we find that

$$\mathsf{snr}_L = \frac{k_n - 1}{A} = \frac{\mathsf{snr}_0(1 + \beta\mathsf{snr}_0)(k_n - 1)}{k_n + (k_n - 1)\beta\mathsf{snr}_0} = \mathsf{snr}_0 \frac{1 + \beta\mathsf{snr}_0}{\frac{k_n}{k_n - 1} + \beta\mathsf{snr}_0},$$

and the width of the phase transition is given by

$$\mathsf{snr}_0 - \mathsf{snr}_L = \mathsf{snr}_0 \left(1 - \frac{1 + \beta\mathsf{snr}_0}{\frac{k_n}{k_n - 1} + \beta\mathsf{snr}_0}\right) = \frac{1}{k_n - 1} \frac{\mathsf{snr}_0}{\frac{k_n}{k_n - 1} + \beta\mathsf{snr}_0},$$

as claimed in (Eq. 5.18b). This concludes the proof.

Appendix K

PROOF OF PROPOSITION 5.2.4.

We first show the decomposition for mutual information

$$\begin{aligned} I(X_{\text{mix}}, \mathsf{snr}) &= I(X_{\text{mix}}; Y) = I(X_G, X_D; Y) \\ &= I(X_D; Y) + I(X_G; Y|X_D) \\ &= I\left(X_D, \frac{\mathsf{snr}(1-\delta)}{1+\delta\mathsf{snr}}\right) + I(X_G, \mathsf{snr}\delta). \end{aligned} \tag{K.1}$$

Next we take the derivative of both sides of (Eq. K.1) with respect to snr. On the left side we get $\frac{d}{d\mathsf{snr}} I(X_{\text{mix}}, \mathsf{snr}) = \frac{1}{2}\text{mmse}(X_{\text{mix}}, \mathsf{snr})$ and on the right we get

$$\begin{aligned} &\text{mmse}(X_{\text{mix}}, \mathsf{snr}) \\ &= 2\frac{d}{d\mathsf{snr}} I\left(X_D, \frac{\mathsf{snr}(1-\delta)}{1+\delta\mathsf{snr}}\right) + 2\frac{d}{d\mathsf{snr}} I(X_G, \mathsf{snr}\delta) \\ &= \text{mmse}\left(X_D, \frac{\mathsf{snr}(1-\delta)}{1+\delta\mathsf{snr}}\right) \cdot \frac{d}{d\mathsf{snr}}\left(\frac{\mathsf{snr}(1-\delta)}{1+\delta\mathsf{snr}}\right) + \text{mmse}(X_G, \mathsf{snr}\delta) \cdot \frac{d}{d\mathsf{snr}}(\mathsf{snr}\delta) \\ &= \frac{1-\delta}{(1+\delta\mathsf{snr})^2}\text{mmse}\left(X_D, \frac{\mathsf{snr}(1-\delta)}{1+\delta\mathsf{snr}}\right) + \text{mmse}(X_G, \mathsf{snr}\delta)\delta \\ &= \frac{1-\delta}{(1+\delta\mathsf{snr})^2}\text{mmse}\left(X_D, \frac{\mathsf{snr}(1-\delta)}{1+\delta\mathsf{snr}}\right) + \frac{\delta}{1+\delta\mathsf{snr}}, \end{aligned}$$

as claimed in (Eq. 5.26). This concludes the proof.

Appendix L

PROOF OF PROPOSITION 5.2.6.

By letting $X_D \sim \mathsf{PAM}(N)$, and by using (Eq. 5.27) and the bound in Proposition 5.2.5, we further constrain the MMSE of X_D to satisfy

$$\mathrm{mmse}\left(X_D, \frac{\mathsf{snr}_0(1-\delta)}{1+\delta\mathsf{snr}_0}\right) \le d_{\max}^2 \mathrm{e}^{-\frac{\frac{\mathsf{snr}_0(1-\delta)}{1+\delta\mathsf{snr}_0}}{8} d_{\min}^2} \le \frac{(1+\mathsf{snr}_0\delta)(\beta-\delta)}{(1-\delta)(1+\beta\mathsf{snr}_0)}, \tag{L.1}$$

which ensures that the MMSE constraint in (Eq. 5.3c) is met. Since, the minimum distance of PAM is given by $d_{\min}^2 = \frac{12}{N^2-1}$, solving for N we have that

$$N \le \left\lfloor \sqrt{1 + c_1 \frac{(1-\delta)\mathsf{snr}_0}{1+\delta\mathsf{snr}_0}} \right\rfloor, \tag{L.2a}$$

$$c_1 = \frac{3}{2\log^+\left(\frac{d_{\max}^2(1-\delta)(1+\beta\mathsf{snr}_0)}{(1+\mathsf{snr}_0\delta)(\beta-\delta)}\right)} \le \frac{3}{2\log^+\left(\frac{12(1-\delta)(1+\beta\mathsf{snr}_0)}{(1+\mathsf{snr}_0\delta)(\beta-\delta)}\right)}, \tag{L.2b}$$

where the last inequality is due to the fact that for PAM

$$d_{\max}^2 = (N-1)^2 d_{\min}^2 = 12\frac{(N-1)^2}{N^2-1} = 12\frac{N-1}{N+1} \le 12. \tag{L.3}$$

For the case of $\mathsf{snr}_0 \le \mathsf{snr}$ we choose the number of points to satisfy (Eq. L.2) with equality and choose $\delta = \beta\frac{\mathsf{snr}_0}{1+\mathsf{snr}_0} := \beta c_2$.

Appendix L (Continued)

Next we compute the gap between the outer bound in Proposition 5.2.3 with the achievable mutual information of a mixed input in Proposition 5.2.4, where $I\left(X_D, \frac{\mathsf{snr}(1-\delta)}{1+\delta\mathsf{snr}}\right)$ is lower bounded by Proposition 5.2.5. We obtain

$$
\begin{aligned}
&\mathsf{G}_{\mathsf{d}1} + \Delta_{\text{(Eq. 5.23)}} \\
&= C_\infty - \left(\log(N) - \frac{1}{2}\log\left(\frac{\pi}{6}\right) - \frac{1}{2}\log\left(1 + \frac{12}{d_{\min}^2}\mathrm{mmse}\left(X_D, \frac{\mathsf{snr}(1-\delta)}{1+\delta\mathsf{snr}}\right)\right) + \frac{1}{2}\log(1+\delta\mathsf{snr})\right) \\
&\overset{a)}{\leq} C_\infty - \left(\frac{1}{2}\log\left(1 + c_1\frac{(1-\delta)\mathsf{snr}_0}{1+\delta\mathsf{snr}_0}\right) - \log(2) - \frac{1}{2}\log\left(\frac{\pi}{6}\right)\right. \\
&\left.-\frac{1}{2}\log\left(1 + \frac{12}{d_{\min}^2}\mathrm{mmse}\left(X_D, \frac{\mathsf{snr}(1-\delta)}{1+\delta\mathsf{snr}}\right)\right) + \frac{1}{2}\log(1+\delta\mathsf{snr})\right) \\
&= \frac{1}{2}\log\left(\frac{1 + \frac{\mathsf{snr}_0(1-\beta)}{1+\beta\mathsf{snr}_0}}{1 + c_1\frac{(1-\delta)\mathsf{snr}_0}{1+\delta\mathsf{snr}_0}}\right) + \frac{1}{2}\log\left(\frac{1+\beta\mathsf{snr}}{1+\delta\mathsf{snr}}\right) + \frac{1}{2}\log\left(1 + \frac{12}{d_{\min}^2}\mathrm{mmse}\left(X_D, \frac{\mathsf{snr}(1-\delta)}{1+\delta\mathsf{snr}}\right)\right) \\
&+ \frac{1}{2}\log\left(\frac{4\pi}{6}\right),
\end{aligned} \tag{L.4}
$$

where inequality in a) follows from getting an extra one bit gap from dropping the floor operation.

Appendix L (Continued)

We next bound each term in (Eq. L.4) individually. The first term in (Eq. L.4) can be bounded as follows:

$$
\begin{aligned}
\frac{1}{2}\log\left(\frac{1+\frac{\mathsf{snr}_0(1-\beta)}{1+\beta\mathsf{snr}_0}}{1+c_1\frac{(1-\delta)\mathsf{snr}_0}{1+\delta\mathsf{snr}_0}}\right) &= \frac{1}{2}\log\left(\frac{(1+\mathsf{snr}_0)(1+c_2\beta\mathsf{snr}_0)}{(1+\beta\mathsf{snr}_0)(1+c_1\mathsf{snr}+\beta c_2\mathsf{snr}_3-\beta c_1c_2\mathsf{snr})}\right)\\
&\overset{b)}{\leq} \frac{1}{2}\log\left(\frac{(1+\mathsf{snr}_0)(1+c_2\beta\mathsf{snr}_0)}{(1+\beta\mathsf{snr}_0)(1+c_1\mathsf{snr}+\beta c_2\mathsf{snr}_3-\beta c_1\mathsf{snr})}\right)\\
&= \frac{1}{2}\log\left(\frac{(1+\mathsf{snr}_0)(1+c_2\beta\mathsf{snr}_0)}{(1+\beta\mathsf{snr}_0)(1+(1-\beta)c_1\mathsf{snr}+\beta c_2\mathsf{snr}_0)}\right)\\
&\overset{c)}{\leq} \frac{1}{2}\log\left(\frac{(1+\mathsf{snr}_0)}{(1+(1-\beta)c_1\mathsf{snr}+\beta c_2\mathsf{snr}_0)}\right)\\
&\overset{d)}{\leq} \frac{1}{2}\log\left(\max\left(\frac{(1+\mathsf{snr}_0)}{(1+c_1\mathsf{snr})},\frac{(1+\mathsf{snr}_0)}{(1+c_2\mathsf{snr}_0)}\right)\right)\\
&\overset{e)}{\leq} \frac{1}{2}\log\left(\max\left(\frac{1}{c_1},2\right)\right),
\end{aligned}
\tag{L.5}
$$

where the inequalities follow from the facts: b) $c_2=\frac{\mathsf{snr}_0}{1+\mathsf{snr}_0}\leq 1$; c) used that $\frac{1+c_2\beta\mathsf{snr}_0}{1+\beta\mathsf{snr}_0}\leq 1$ since $c_2\leq 1$; d) the denominator term $1+(1-\beta)c_1\mathsf{snr}+\beta c_2\mathsf{snr}_0$ achieves its minimum at either $\beta=0$ or $\beta=1$; and e) $\frac{(1+\mathsf{snr}_0)}{(1+c_2\mathsf{snr}_0)}\leq\frac{1}{c_2}=\frac{1+\mathsf{snr}_0}{\mathsf{snr}_0}\leq 2$ for $\mathsf{snr}_0\geq 1$.

The second term in (Eq. L.4) can be bounded as follows:

$$
\frac{1}{2}\log\left(\frac{1+\beta\mathsf{snr}}{1+\delta\mathsf{snr}}\right)\leq\frac{1}{2}\log\left(\frac{1+\mathsf{snr}_0}{\mathsf{snr}_0}\right)\leq\frac{1}{2}\log(2), \tag{L.6}
$$

where the inequalities follow from using $\delta=\beta\frac{\mathsf{snr}_0}{1+\mathsf{snr}_0}$ and $\frac{1+\beta\mathsf{snr}_0}{1+\delta\mathsf{snr}_0}\leq\frac{1+\mathsf{snr}_0}{\mathsf{snr}_0}\leq 2$ for $\mathsf{snr}_0\geq 1$.

Appendix L (Continued)

The third term in (Eq. L.4) can be bounded as follows

$$\begin{aligned}
&\frac{1}{2}\log\left(1+\frac{12}{d_{\min}^2}\mathrm{mmse}\left(X_D,\frac{\mathsf{snr}(1-\delta)}{1+\delta\mathsf{snr}}\right)\right)\\
&\overset{f)}{\le}\frac{1}{2}\log\left(1+\frac{12}{d_{\min}^2}\mathrm{mmse}\left(X_D,\frac{\mathsf{snr}_0(1-\delta)}{1+\delta\mathsf{snr}_0}\right)\right)\\
&\overset{g)}{\le}\frac{1}{2}\log\left(1+c_1\frac{(1-\delta)\mathsf{snr}_0}{1+\delta\mathsf{snr}_0}\mathrm{mmse}\left(X_D,\frac{\mathsf{snr}_0(1-\delta)}{1+\delta\mathsf{snr}_0}\right)\right)\\
&\overset{h)}{\le}\frac{1}{2}\log\left(1+c_1\frac{(\beta-\delta)\mathsf{snr}_0}{1+\beta\mathsf{snr}_0}\right)\\
&\overset{i)}{\le}\frac{1}{2}\log\left(1+c_1\frac{\beta}{1+\beta\mathsf{snr}_0}\right),
\end{aligned}\tag{L.7}$$

where the (in)-equalities follow from: f) the fact that the MMSE is a decreasing function of SNR and $\frac{\mathsf{snr}(1-\delta)}{1+\delta\mathsf{snr}} \ge \frac{\mathsf{snr}_0(1-\delta)}{1+\delta\mathsf{snr}_0}$; g) using the bound on $d_{\min}^2 = \frac{12}{N^2-1}$ from (Eq. L.2); h) using the bound in (Eq. L.1); and i) using $\delta = \frac{\beta\mathsf{snr}_0}{1+\mathsf{snr}_0} \le \beta$.

By combining the bounds in (Eq. L.5), (Eq. L.6), and (Eq. L.7) we get

$$\begin{aligned}
\mathsf{G}_{\mathsf{d}1}+\Delta_{\text{(Eq. 5.23)}} &\le \frac{1}{2}\log\left(\max\left(\frac{1}{c_1},2\right)\right)+\frac{1}{2}\log\left(\frac{4\pi}{3}\right)+\frac{1}{2}\log\left(1+c_1\frac{\beta}{1+\beta\mathsf{snr}_0}\right)\\
&=\frac{1}{2}\log\left(\max\left(\frac{1}{c_1},2\right)+2\max\left(1,2c_1\right)\frac{\beta}{1+\beta\mathsf{snr}_0}\right)+\frac{1}{2}\log\left(\frac{4\pi}{3}\right)\\
&\overset{j)}{\le}\frac{1}{2}\log\left(\max\left(\frac{1}{c_1},2\right)+6\frac{\beta}{1+\beta\mathsf{snr}_0}\right)+\frac{1}{2}\log\left(\frac{4\pi}{3}\right)\\
&\overset{k)}{=}\frac{1}{2}\log\left(\max\left(\frac{2\log\left(\frac{12(1-\delta)(1+\beta\mathsf{snr}_0)}{(1+\mathsf{snr}_0\delta)(\beta-\delta)}\right)}{3},2\right)+6\frac{\beta}{1+\beta\mathsf{snr}_0}\right)+\frac{1}{2}\log\left(\frac{4\pi}{3}\right)\\
&\overset{l)}{\le}\frac{1}{2}\log\left(\max\left(\frac{2}{3}\log\left(\frac{24(1+(1-\beta)\mathsf{snr}_0}{\beta}\right),2\right)+6\frac{\beta}{1+\beta\mathsf{snr}_0}\right)+\frac{1}{2}\log\left(\frac{4\pi}{3}\right)\\
&\overset{m)}{=}\frac{1}{2}\log\left(\frac{2}{3}\log\left(\frac{24(1+(1-\beta)\mathsf{snr}_0}{\beta}\right)+6\frac{\beta}{1+\beta\mathsf{snr}_0}\right)+\frac{1}{2}\log\left(\frac{4\pi}{3}\right),
\end{aligned}$$

Appendix L (Continued)

where the inequalities follow from: j) the fact that $c_1 \leq \frac{3}{2}$; k) using the value of c_1 in (Eq. L.2); l) using $\delta = \beta \frac{\mathsf{snr}_0}{1+\mathsf{snr}_0}$ and $\frac{1+\beta\mathsf{snr}_0}{1+\delta\mathsf{snr}_0} \leq \frac{1+\mathsf{snr}_0}{\mathsf{snr}_0} \leq 2$ for $\mathsf{snr}_0 \geq 1$; and m) the fact that $\max\left(\frac{2\log\left(\frac{24(1+\beta\mathsf{snr}_0)}{\beta}\right)}{3}, 2\right) = \frac{2\log\left(\frac{24(1+\beta\mathsf{snr}_0)}{\beta}\right)}{3}$.

This concludes the proof of the gap result for the $\mathsf{snr} \geq \mathsf{snr}_0$ regime.

We next focus on the $\mathsf{snr} \leq \mathsf{snr}_0$ regime. We use only the discrete part of the mixed input and set $\delta = 0$. From (Eq. L.2) we have that the input parameters must satisfy

$$N \leq \left\lfloor \sqrt{1 + c_3 \mathsf{snr}_0} \right\rfloor, \tag{L.8a}$$

$$c_3 \leq \frac{3}{2\log\left(\frac{12(1+\beta\mathsf{snr}_0)}{\beta}\right)}, \tag{L.8b}$$

in order to comply with the MMSE constraint in (Eq. 5.3c). However, instead of choosing the number of points as in (Eq. L.8) we choose it to be

$$N = \left\lfloor \sqrt{1 + c_3 \mathsf{snr}} \right\rfloor \leq \left\lfloor \sqrt{1 + c_3 \mathsf{snr}_0} \right\rfloor. \tag{L.9}$$

The reason for this choice will be apparent from the gap derivation next.

Appendix L (Continued)

Similarly to the previous case, we compute the gap between the outer bound in Proposition 5.2.3 and the achievable mutual information of the mixed input in Proposition 5.2.4, where $I\left(X_D, \mathsf{snr}\right)$ is lower bounded using Proposition 5.2.5. We have,

$$\begin{aligned}
\mathsf{G}_{\mathsf{d}2} + \Delta_{\text{(Eq. 5.24)}} &\leq C_\infty - \log(N) + \frac{1}{2}\log\left(\frac{\pi \mathrm{e}}{6}\right) + \frac{1}{2}\log\left(1 + \frac{12}{d_{\min}^2}\mathrm{mmse}(X_D, \mathsf{snr})\right) \\
&\overset{n)}{\leq} \frac{1}{2}\log\left(\frac{1+\mathsf{snr}}{1+c_2\mathsf{snr}}\right) + \frac{1}{2}\log\left(\frac{4\pi \mathrm{e}}{6}\right) + \frac{1}{2}\log\left(1 + \frac{12}{d_{\min}^2}\mathrm{mmse}(X_D, \mathsf{snr})\right) \\
&\overset{o)}{\leq} \frac{1}{2}\log\left(\frac{1+\mathsf{snr}}{1+c_2\mathsf{snr}}\right) + \frac{1}{2}\log\left(\frac{4\pi \mathrm{e}}{6}\right) + \frac{1}{2}\log\left(1 + \frac{c_2\mathsf{snr}}{1+\mathsf{snr}}\right) \\
&= \frac{1}{2}\log\left(\frac{1+(1+c_2)\mathsf{snr}}{1+c_2\mathsf{snr}}\right) + \frac{1}{2}\log\left(\frac{4\pi \mathrm{e}}{6}\right) \\
&\overset{p)}{\leq} \frac{1}{2}\log\left(1 + \frac{1}{c_2}\right) + \frac{1}{2}\log\left(\frac{4\pi \mathrm{e}}{6}\right) \\
&\overset{r)}{=} \frac{1}{2}\log\left(1 + \frac{2}{3}\log\left(\frac{12(1+\beta\mathsf{snr}_0)}{\beta}\right)\right) + \frac{1}{2}\log\left(\frac{4\pi \mathrm{e}}{6}\right),
\end{aligned}$$

where the (in)-equalities follow from: n) getting an extra one bit gap by dropping the floor operation; o) using the bound on $d_{\min}^2 = \frac{12}{N^2-1}$ from (Eq. L.9) and bound $\mathrm{mmse}(X, \mathsf{snr}) \leq \frac{1}{1+\mathsf{snr}}$; p) using $\frac{1+(1+c_2)\mathsf{snr}}{1+c_2\mathsf{snr}} \leq \frac{1+c_2}{c_2} = 1 + \frac{1}{c_2}$; and r) using the value of c_2 from (Eq. L.8).

Note that had we chosen number of points to be $N = \left\lfloor \sqrt{1 + c_3\mathsf{snr}_0} \right\rfloor$, the inequality in p) would not hold.

This concludes the proof.

Appendix M

PROOF OF PROPOSITION 5.3.1.

Observe that

$$\mathbf{Cov}(\mathbf{Z}|\mathbf{Y}) = \mathbb{E}[\mathbf{Z}\mathbf{Z}^{\mathrm{T}}|\mathbf{Y}] - (\mathbb{E}[\mathbf{Z}|\mathbf{Y}])(\mathbb{E}[\mathbf{Z}|\mathbf{Y}])^{\mathrm{T}},$$

and so we have that

$$\begin{aligned}
\mathbf{Cov}^2(\mathbf{Z}|\mathbf{Y}) &= \left(\mathbb{E}[\mathbf{Z}\mathbf{Z}^{\mathrm{T}}|\mathbf{Y}] - \mathbb{E}[\mathbf{Z}|\mathbf{Y}]\mathbb{E}[\mathbf{Z}|\mathbf{Y}]^{\mathrm{T}}\right)^2 \\
&= (\mathbb{E}[\mathbf{Z}\mathbf{Z}^{\mathrm{T}}|\mathbf{Y}])^2 - \mathbb{E}[\mathbf{Z}|\mathbf{Y}]\mathbb{E}[\mathbf{Z}|\mathbf{Y}]^{\mathrm{T}}\mathbb{E}[\mathbf{Z}\mathbf{Z}^{\mathrm{T}}|\mathbf{Y}] \\
&- \mathbb{E}[\mathbf{Z}\mathbf{Z}^{\mathrm{T}}|\mathbf{Y}]\mathbb{E}[\mathbf{Z}|\mathbf{Y}]\mathbb{E}[\mathbf{Z}|\mathbf{Y}]^{\mathrm{T}} + (\mathbb{E}[\mathbf{Z}|\mathbf{Y}]\mathbb{E}[\mathbf{Z}|\mathbf{Y}]^{\mathrm{T}})^2 \\
&\stackrel{a)}{=} (\mathbb{E}[\mathbf{Z}\mathbf{Z}^{\mathrm{T}}|\mathbf{Y}])^2 - 2\mathbb{E}[\mathbf{Z}|\mathbf{Y}]\mathbb{E}[\mathbf{Z}|\mathbf{Y}]^{\mathrm{T}}\mathbb{E}[\mathbf{Z}\mathbf{Z}^{\mathrm{T}}|\mathbf{Y}] \\
&+ (\mathbb{E}[\mathbf{Z}|\mathbf{Y}]\mathbb{E}[\mathbf{Z}|\mathbf{Y}]^{\mathrm{T}})^2 \\
&\stackrel{b)}{\preceq} (\mathbb{E}[\mathbf{Z}\mathbf{Z}^{\mathrm{T}}|\mathbf{Y}])^2 - 2\mathbb{E}[\mathbf{Z}|\mathbf{Y}]\mathbb{E}[\mathbf{Z}|\mathbf{Y}]^{\mathrm{T}}\mathbb{E}[\mathbf{Z}|\mathbf{Y}]\mathbb{E}[\mathbf{Z}|\mathbf{Y}]^{\mathrm{T}} \\
&+ (\mathbb{E}[\mathbf{Z}|\mathbf{Y}]\mathbb{E}[\mathbf{Z}|\mathbf{Y}]^{\mathrm{T}})^2 \\
&= (\mathbb{E}[\mathbf{Z}\mathbf{Z}^{\mathrm{T}}|\mathbf{Y}])^2 - (\mathbb{E}[\mathbf{Z}|\mathbf{Y}]\mathbb{E}[\mathbf{Z}|\mathbf{Y}]^{\mathrm{T}})^2 \\
&\stackrel{c)}{=} \mathbb{E}[\mathbf{Z}\mathbf{Z}^{\mathrm{T}}(\mathbf{Z}\mathbf{Z}^{\mathrm{T}})^{\mathrm{T}}|\mathbf{Y}] - \mathbf{Cov}(\mathbf{Z}\mathbf{Z}^{\mathrm{T}}|\mathbf{Y}) - (\mathbb{E}[\mathbf{Z}|\mathbf{Y}]\mathbb{E}[\mathbf{Z}|\mathbf{Y}]^{\mathrm{T}})^2, \qquad \text{(M.1)}
\end{aligned}$$

where the order operations follow from: a) the fact that $\mathbb{E}[\mathbf{Z}|\mathbf{Y}]\mathbb{E}[\mathbf{Z}|\mathbf{Y}]^{\mathrm{T}}$ and $\mathbb{E}[\mathbf{Z}\mathbf{Z}^{\mathrm{T}}|\mathbf{Y}]$ are symmetric matrices; b) using $\mathbb{E}[\mathbf{Z}|\mathbf{Y}]\mathbb{E}[\mathbf{Z}|\mathbf{Y}]^{\mathrm{T}} \preceq \mathbb{E}[\mathbf{Z}\mathbf{Z}^{\mathrm{T}}|\mathbf{Y}]$; and c) the fact that, since $\mathbf{Cov}(\mathbf{Z}\mathbf{Z}^{\mathrm{T}}|\mathbf{Y}) = \mathbb{E}[\mathbf{Z}\mathbf{Z}^{\mathrm{T}}(\mathbf{Z}\mathbf{Z}^{\mathrm{T}})^{\mathrm{T}}|\mathbf{Y}] - \mathbb{E}[\mathbf{Z}\mathbf{Z}^{\mathrm{T}}|\mathbf{Y}](\mathbb{E}[\mathbf{Z}\mathbf{Z}^{\mathrm{T}}|\mathbf{Y}])^{\mathrm{T}}$ and by symmetry of $\mathbb{E}[\mathbf{Z}\mathbf{Z}^{\mathrm{T}}|\mathbf{Y}]$, we have that
$\mathbb{E}[\mathbf{Z}\mathbf{Z}^{\mathrm{T}}|\mathbf{Y}](\mathbb{E}[\mathbf{Z}\mathbf{Z}^{\mathrm{T}}|\mathbf{Y}])^{\mathrm{T}} = (\mathbb{E}[\mathbf{Z}\mathbf{Z}^{\mathrm{T}}|\mathbf{Y}])^2$. By using the monotonicity of the trace, properties of the expected value, and the inequality in (Eq. M.1), we have that

$$\begin{aligned}
\mathrm{Tr}\left(\mathbb{E}[\mathbf{Cov}^2(\mathbf{Z}|\mathbf{Y})]\right) &\leq \mathrm{Tr}\left(\mathbb{E}\left[\mathbb{E}[\mathbf{Z}\mathbf{Z}^{\mathrm{T}}(\mathbf{Z}\mathbf{Z}^{\mathrm{T}})^{\mathrm{T}}|\mathbf{Y}] - \mathbf{Cov}(\mathbf{Z}\mathbf{Z}^{\mathrm{T}}|\mathbf{Y}) - (\mathbb{E}[\mathbf{Z}|\mathbf{Y}]\mathbb{E}[\mathbf{Z}|\mathbf{Y}]^{\mathrm{T}})^2\right]\right) \\
&= \mathrm{Tr}\left(\mathbb{E}\left[\mathbb{E}[\mathbf{Z}\mathbf{Z}^{\mathrm{T}}(\mathbf{Z}\mathbf{Z}^{\mathrm{T}})^{\mathrm{T}}|\mathbf{Y}]\right]\right) - \mathrm{Tr}\left(\mathbb{E}\left[\mathbf{Cov}(\mathbf{Z}\mathbf{Z}^{\mathrm{T}}|\mathbf{Y})\right]\right) \\
&- \mathrm{Tr}\left(\mathbb{E}\left[(\mathbb{E}[\mathbf{Z}|\mathbf{Y}]\mathbb{E}[\mathbf{Z}|\mathbf{Y}]^{\mathrm{T}})^2\right]\right). \qquad \text{(M.2)}
\end{aligned}$$

We next focus on each term of the right hand side of (Eq. M.2) individually. The first term can be computed as follows:

$$\begin{aligned}
\mathrm{Tr}\left(\mathbb{E}\left[\mathbb{E}[\mathbf{Z}\mathbf{Z}^{\mathrm{T}}(\mathbf{Z}\mathbf{Z}^{\mathrm{T}})^{\mathrm{T}}|\mathbf{Y}]\right]\right) &\overset{d)}{=} \mathrm{Tr}\left(\mathbb{E}[\mathbf{Z}\mathbf{Z}^{\mathrm{T}}\mathbf{Z}\mathbf{Z}^{\mathrm{T}}]\right) \\
&\overset{e)}{=} \mathbb{E}\left[\mathrm{Tr}\left(\mathbf{Z}\mathbf{Z}^{\mathrm{T}}\mathbf{Z}\mathbf{Z}^{\mathrm{T}}\right)\right] \\
&= \mathbb{E}\left[\mathrm{Tr}\left(\mathbf{Z}^{\mathrm{T}}\mathbf{Z}\mathbf{Z}^{\mathrm{T}}\mathbf{Z}\right)\right] \\
&= \mathbb{E}\left[\left(\sum_{i=1}^{n} Z_i^2\right)^2\right] \\
&\overset{f)}{=} n(n+2), \qquad \text{(M.3)}
\end{aligned}$$

Appendix M (Continued)

where the (in)-equalities follow from: d) using the law of total expectation; e) since expectation is a linear operator and using fact that the trace can be exchanged with linear operators; and f) observing that $S = \sum_{i=1}^{n} Z_i^2$ is a chi-square distribution of degree n and hence $\mathbb{E}[S^2] = n(n+2)$.

For the second term in (Eq. M.2), by definition of the MMSE, we have

$$\mathrm{Tr}\left(\mathbb{E}\left[\mathbf{Cov}(\mathbf{Z}\mathbf{Z}^{\mathrm{T}}|\mathbf{Y})\right]\right) = n\mathrm{mmse}(\mathbf{Z}\mathbf{Z}^{\mathrm{T}}|\mathbf{Y}). \tag{M.4}$$

The third term in (Eq. M.2) satisfies

$$\begin{aligned}
\mathrm{Tr}\left(\mathbb{E}\left[(\mathbb{E}[\mathbf{Z}|\mathbf{Y}]\mathbb{E}[\mathbf{Z}|\mathbf{Y}]^{\mathrm{T}})^2\right]\right) &\overset{g)}{\geq} \mathrm{Tr}\left(\left(\mathbb{E}\left[\mathbb{E}[\mathbf{Z}|\mathbf{Y}]\mathbb{E}[\mathbf{Z}|\mathbf{Y}]^{\mathrm{T}}\right]\right)^2\right) \\
&= \mathrm{Tr}\left(\left(\mathbb{E}[\mathbf{Z}\mathbf{Z}^{\mathrm{T}}] - \mathbb{E}[\mathbf{Cov}(\mathbf{Z}|\mathbf{Y})]\right)^2\right) \\
&\overset{h)}{=} \mathrm{Tr}\left(\left(\mathbf{I} - \mathsf{snr}\,\mathbb{E}[\mathbf{Cov}(\mathbf{X}|\mathbf{Y})]\right)^2\right) \\
&\overset{i)}{=} \mathrm{Tr}\left(\mathbf{J}^2(\mathbf{Y})\right)
\end{aligned} \tag{M.5}$$

where the (in)-equalities follow from: g) using Jensen's inequality; h) using the property $\mathsf{snr}\,\mathbb{E}[\mathbf{Cov}(X|\mathbf{Y})] = \mathbb{E}[\mathbf{Cov}(\mathbf{Z}|\mathbf{Y})]$; and i) using identity (66)

$$\mathbf{I} - \mathsf{snr}\,\mathbb{E}[\mathbf{Cov}(\mathbf{X}|\mathbf{Y})] = \mathbf{J}(\mathbf{Y}).$$

By putting (Eq. M.3), (Eq. M.4), and (Eq. M.5) together, we have that

$$\mathbb{E}\left[\mathbf{Cov}^2(\mathbf{Z}|\mathbf{Y})\right] \leq k_n := \frac{n(n+2) - n\,\mathrm{mmse}(\mathbf{Z}\mathbf{Z}^{\mathrm{T}}|\mathbf{Y}) - \mathrm{Tr}\left(\mathbf{J}^2(\mathbf{Y})\right)}{n}.$$

Appendix M (Continued)

Finally, using the identity $\mathbb{E}\left[\mathbf{Cov}^2(\mathbf{Z}|\mathbf{Y})\right] = \mathsf{snr}^2\mathbb{E}\left[\mathbf{Cov}^2(\mathbf{X}|\mathbf{Y})\right]$ concludes the proof.

Appendix N

PROOF OF PROPOSITION 5.3.2.

Using the Cramer-Rao lower bound (74, Theorem 20) we have that

$$\mathbf{J}(\mathbf{Y}) \succeq \mathbf{Cov}^{-1}(\mathbf{Y}) \tag{N.1}$$

$$= \left(\mathsf{snr}\mathbb{E}[\mathbf{X}\mathbf{X}^T] + \mathbf{I}\right)^{-1} \tag{N.2}$$

$$= \mathbf{V}^{-1}\, \Lambda^{-1}\, \mathbf{V}, \tag{N.3}$$

where Λ is the eigen-matrix of $\mathsf{snr}\mathbb{E}[\mathbf{X}\mathbf{X}^T] + \mathbf{I}$, $\lambda_i = \mathsf{snr}\sigma_i + 1$, and σ_i is the i-th eigenvalue of matrix $\mathbb{E}[\mathbf{X}\mathbf{X}^T]$. Therefore,

$$\begin{aligned} \mathrm{Tr}\left(\mathbf{J}^2(\mathbf{Y})\right) &\geq \mathrm{Tr}\left(\mathbf{V}^{-1}\Lambda^{-1}\, \mathbf{V} \left(\mathbf{V}^{-1}\Lambda^{-1}\, \mathbf{V}\right)^T\right) \\ &= \mathrm{Tr}(\, \Lambda^{-2}) \\ &= \sum_{i=1}^{n} \frac{1}{(1+\mathsf{snr}\sigma_i)^2} \\ &\geq \frac{n}{(1+\mathsf{snr})^2}, \end{aligned}$$

where the last inequality comes from minimizing $\sum_{i=1}^n \frac{1}{(1+\mathsf{snr}\sigma_i)^2}$ subject to the constraint that $\mathrm{Tr}\left(\mathbb{E}[\mathbf{X}\mathbf{X}^T]\right) = \sum_{i=1}^n \sigma_i \leq n$ and where the minimum is attained with $\sigma_i = 1$.

Finally, note that all inequalities are equalities if $\mathbf{Y} \sim \mathcal{N}(\,\mathbf{0}, (1+\mathsf{snr})\mathbf{I})$ or equivalently if $\mathbf{X} \sim \mathcal{N}(\,\mathbf{0}, \mathbf{I})$. This concludes the proof.

Appendix O

PROOF OF PROPOSITION 5.3.3.

First observe that since the conditional expectation is the best estimator under a squared cost function

$$\begin{aligned}\mathbf{Cov}(\mathbf{X}|\mathbf{Y}=\mathbf{y}) &= \mathbb{E}\left[(\mathbf{X}-\mathbb{E}[\mathbf{X}|\mathbf{Y}])(\mathbf{X}-\mathbb{E}[\mathbf{X}|\mathbf{Y}])^{\mathrm{T}}|\mathbf{Y}=\mathbf{y}\right] \\ &\preceq \mathbb{E}\left[(\mathbf{X}-f(\mathbf{Y}))(\mathbf{X}-f(\mathbf{Y}))^{\mathrm{T}}|\mathbf{Y}=\mathbf{y}\right],\end{aligned} \tag{O.1}$$

for any deterministic function $f(\cdot)$. Therefore, the first bound in (Eq. 5.35a) follows by choosing $f(\mathbf{Y})=\frac{\sqrt{\mathsf{snr}}\mathbf{Y}}{1+\mathsf{snr}}$ in (Eq. O.1)

$$\begin{aligned}\mathrm{Tr}\left(\mathbb{E}\left[\mathbf{Cov}^2(\mathbf{X}|\mathbf{Y})\right]\right) &\leq \mathrm{Tr}\left(\mathbb{E}\left[\mathbb{E}^2\left[\left(\mathbf{X}-\frac{\sqrt{\mathsf{snr}}\mathbf{Y}}{1+\mathsf{snr}}\right)\left(\mathbf{X}-\frac{\sqrt{\mathsf{snr}}\mathbf{Y}}{1+\mathsf{snr}}\right)^{\mathrm{T}}|\mathbf{Y}\right]\right]\right) \\ &= \frac{1}{(1+\mathsf{snr})^4}\mathrm{Tr}\left(\mathbb{E}\left[\mathbb{E}^2\left[\left(\mathbf{X}-\sqrt{\mathsf{snr}}\mathbf{Z}\right)\left(\mathbf{X}-\sqrt{\mathsf{snr}}\mathbf{Z}\right)^{\mathrm{T}}|\mathbf{Y}\right]\right]\right) \\ &\leq \frac{1}{(1+\mathsf{snr})^4}\mathrm{Tr}\left(\mathbb{E}\left[\left(\left(\mathbf{X}-\sqrt{\mathsf{snr}}\mathbf{Z}\right)\left(\mathbf{X}-\sqrt{\mathsf{snr}}\mathbf{Z}\right)^{\mathrm{T}}\right)^2\right]\right),\end{aligned}$$

where the last inequality is due to Jensen's inequality.

The second bound in (Eq. 5.35a) follows by choosing $f(\mathbf{Y})=0$ in (Eq. O.1)

$$\mathrm{Tr}\left(\mathbb{E}\left[\mathbf{Cov}^2(\mathbf{X}|\mathbf{Y})\right]\right) \leq \mathrm{Tr}\left(\mathbb{E}\left[\mathbb{E}^2\left[(\mathbf{X}-\mathbf{0})(\mathbf{X}-\mathbf{0})^{\mathrm{T}}|\mathbf{Y}\right]\right]\right) = \mathrm{Tr}\left(\mathbb{E}\left[\mathbb{E}^2\left[\mathbf{X}\mathbf{X}^{\mathrm{T}}|\mathbf{Y}\right]\right]\right).$$

This concludes the proof.

Appendix P

COPYRIGHT PERMISSION

This Appendix includes the copyright permission granted from the IEEE to use published work in thesis. The following statement has been copied from the CopyRight Clearance Center (Rightslink).

The IEEE does not require individuals working on a thesis to obtain a formal reuse license, however, you may print out this statement to be used as a permission grant:

1. The following IEEE copyright/ credit notice should be placed prominently in the references: [year of original publication] IEEE. Reprinted, with permission, from [author names, paper title, IEEE publication title, and month/year of publication].

2. Only the accepted version of an IEEE copyrighted paper can be used when posting the paper or your thesis on-line.

3. In placing the thesis on the authors university website, please display the following message in a prominent place on the website: In reference to IEEE copyrighted material which is used with permission in this thesis, the IEEE does not endorse any of [university/educational entitys name goes here]s products or services. Internal or personal use of this material is permitted. If interested in reprinting/republishing IEEE copyrighted material for advertising or promotional purposes or for creating new collective works for resale or redistribution.

Thesis / Dissertation Reuse

The IEEE does not require individuals working on a thesis to obtain a formal reuse license, however, you may print out this statement to be used as a permission grant:

Requirements to be followed when using any portion (e.g., figure, graph, table, or textual material) of an IEEE copyrighted paper in a thesis:

1) In the case of textual material (e.g., using short quotes or referring to the work within these papers) users must give full credit to the original source (author, paper, publication) followed by the IEEE copyright line © 2011 IEEE.
2) In the case of illustrations or tabular material, we require that the copyright line © [Year of original publication] IEEE appear prominently with each reprinted figure and/or table.
3) If a substantial portion of the original paper is to be used, and if you are not the senior author, also obtain the senior author's approval.

Requirements to be followed when using an entire IEEE copyrighted paper in a thesis:

1) The following IEEE copyright/ credit notice should be placed prominently in the references: © [year of original publication] IEEE. Reprinted, with permission, from [author names, paper title, IEEE publication title, and month/year of publication]
2) Only the accepted version of an IEEE copyrighted paper can be used when posting the paper or your thesis on-line.
3) In placing the thesis on the author's university website, please display the following message in a prominent place on the website: In reference to IEEE copyrighted material which is used with permission in this thesis, the IEEE does not endorse any of [university/educational entity's name goes here]'s products or services. Internal or personal use of this material is permitted. If interested in reprinting/republishing IEEE copyrighted material for advertising or promotional purposes or for creating new collective works for resale or redistribution, please go to http://www.ieee.org/publications_standards/publications/rights/rights_link.html to learn how to obtain a License from RightsLink.

If applicable, University Microfilms and/or ProQuest Library, or the Archives of Canada may supply single copies of the dissertation.

CITED LITERATURE

1. Shannon, C.: A mathematical theory of communication. Bell Syst. Tech. J., 27(379-423, 623-656), Jul., Oct. 1948.

2. Dytso, A., Devroye, N., and Tuninetti, D.: On the capacity of interference channels with partial codebook knowledge. In Proc. IEEE Int. Symp. Inf. Theory, pages 2039–2043, 2013.

3. Dytso, A., Tuninetti, D., and Devroye, N.: On discrete alphabets for the two-user Gaussian interference channel with one receiver lacking knowledge of the interfering codebook. In Proc. Workshop on Info. Theory and Applications, pages 1–8, Feb 2014.

4. Dytso, A., Tuninetti, D., and Devroye, N.: On Gaussian interference channels with mixed Gaussian and discrete inputs. In Proc. IEEE Int. Symp. Inf. Theory, pages 261–265, June 2014.

5. Dytso, A., Tuninetti, D., and Devroye, N.: On the two-user interference channel with lack of knowledge of the interference codebook at one receiver. IEEE Trans. Inf. Theory, 61(3):1257–1276, March 2015.

6. Dytso, A., Tuninetti, D., and Devroye, N.: i.i.d. mixed inputs and treating interference as noise are gDoF optimal for the symmetric Gaussian two-user interference channel. In Proc. IEEE Int. Symp. Inf. Theory, pages 1716–1720, June 2015.

7. Dytso, A., Tuninetti, D., and Devroye, N.: Nearly optimal non-Gaussian codes for the Gaussian interference channel. pages 1287–1291, Nov 2015.

8. Dytso, A., Tuninetti, D., and Devroye, N.: Interference as noise: Friend or foe? IEEE Trans. Inf. Theory, PP(99):1–1, 2016.

9. Dytso, A., Bustin, R., Tuninetti, D., Devroye, N., Shamai, S., and Poor, H. V.: On communications through a Gaussian channel with a MMSE disturbance constraint. In Proc. Workshop on Info. Theory and Applications, 2016.

10. Ozarow, L. and Wyner, A.: On the capacity of the gaussian channel with a finite number of input levels. IEEE Trans. Inf. Theory, 36(6):1426–1428, Nov 1990.

11. Shamai, S.: From constrained signaling to network interference alignment via an information-estimation perspective. IEEE Information Theory Society Newsletter, 62(7):6–24, September 2012.

12. Shanonn, C.: Systems which approach the ideal as P/N $\rightarrow \infty$. Unpublished, 1948.

13. Wu, Y. and Verdu, S.: The impact of constellation cardinality on Gaussian channel capacity. In Proc. Allerton Conf. Commun., Control and Comp., September 2010.

14. Calvo, E., Fonollosa, J., and Vidal, J.: On the totally asynchronous interference channel with single-user receivers. In Proc. IEEE Int. Symp. Inf. Theory, pages 2587–2591, 2009.

15. Abbe, E. and Zheng, L.: A coordinate system for gaussian networks. IEEE Trans. Inf. Theory, 58(2):721–733, 2012.

16. Bennatan, A., Shamai, S., and Calderbank, A.: Soft-decoding-based strategies for relay and interference channels: Analysis and achievable rates using ldpc codes. IEEE Trans. Inf. Theory, 60(4):1977–2009, April 2014.

17. Smith, J. G.: The information capacity of amplitude-and variance-constrained sclar gaussian channels. Information and Control, 18(3):203–219, 1971.

18. Zhang, L., Li, H., and Guo, D.: Capacity of gaussian channels with duty cycle and power constraints. IEEE Trans. Inf. Theory, 60(3):1615–1629, 2014.

19. Sanderovich, A., Shamai, S., Steinberg, Y., and Kramer, G.: Communication via decentralized processing. IEEE Trans. Inf. Theory, 54(7):3008 –3023, July 2008.

20. Simeone, O., Erkip, E., and Shamai, S.: On codebook information for interference relay channels with out-of-band relaying. IEEE Trans. Inf. Theory, 57(5):2880 –2888, May 2011.

21. Tian, Y. and Yener, A.: Relaying for multiple sources in the absence of codebook information. pages 1845 –1849, November 2011.

22. Avestimehr, A., Diggavi, S., and Tse, D.: Wireless network information flow: A deterministic approach. IEEE Trans. Inf. Theory, 57(4):1872–1905, April 2011.

23. Bresler, G. and Tse, D.: The two-user gaussian interference channel: A deterministic view. 19:333–354, April 2008.

24. Geng, C., Naderializadeh, N., Avestimehr, A., and Jafar, S.: On the optimality of treating interference as noise. IEEE Trans. Inf. Theory, 61(4):1753–1767, April 2015.

25. Zamir, R. and Erez, U.: A Gaussian input is not too bad. IEEE Trans. Inf. Theory, 50(6):1362–1367, June 2004.

26. Etkin, R., Tse, D., and Wang, H.: Gaussian interference channel capacity to within one bit. IEEE Trans. Inf. Theory, 54(12):5534–5562, December 2008.

27. Bustin, R. and Shamai, S.: MMSE of 'bad' codes. IEEE Trans. Inf. Theory, 59(2):733–743, Feb 2013.

28. Bandemer, B. and El Gamal, A.: Communication with disturbance constraints. IEEE Trans. Inf. Theory, 60(8):4488–4502, Aug 2014.

29. Dytso, A., Bustin, R., Tuninetti, D., Devroye, N., Shamai, S., and Poor, H. V.: On communication through a Gaussian channel with an MMSE disturbance constraint. Submitted to IEEE Trans. Inf. Theory, 2016.

30. Dytso, A., Bustin, R., Tuninetti, D., Devroye, N., Shamai, S., and Poor, H. V.: On the applications of the minimum mean p-th error (MMPE) to information theoretic quantities. Submitted to IEEE Inf. Theory Workshop, Feb 2016.

31. Cover, T. and Thomas, J.: Elements of Information Theory: Second Edition. Wiley, 2006.

32. Moshksar, K., Ghasemi, A., and Khandani, A.: An alternative to decoding interference or treating int erference as Gaussian noise. IEEE Trans. Inf. Theory, 61(1):305–322, Jan 2015.

33. Tao, T. and Wu, V.: Additive Combinatorics. Cambridge University Press, 2006.

34. Niesen, U. and Maddah-Ali, M.: Interference alignment: From degrees of freedom to constant-gap capacity approximations. IEEE Trans. Inf. Theory, 59(8):4855–4888, Aug 2013.

35. Ordentlich, O., Erez, U., and Nazer, B.: The approximate sum capacity of the symmetric Gaussian K -user interference channel. IEEE Trans. Inf. Theory, 60(6):3450–3482, June 2014.

36. Motahari, A., Oveis-Gharan, S., Maddah-Ali, M.-A., and Khandani, A.: Real interference alignment: Exploiting the potential of single antenna systems. IEEE Trans. Inf. Theory, 60(8):4799–4810, Aug 2014.

37. A. El Gamal and Y.-H. Kim: Network Information Theory. Camrbidge University Press, 2012.

38. Han, T. and Kobayashi, K.: A new achievable rate region for the interference channel. IEEE Trans. Inf. Theory, IT-27(1):49–60, January 1981.

39. Telatar, E. and Tse, D.: Bounds on the capacity region of a class of interference channels. In Proc. IEEE Int. Symp. Inf. Theory, pages 2871 –2874, June 2007.

40. Csiszár, I. and Körner, J.: Information Theory: Coding Theorems for Discrete Memoryless Systems. New York, Academic Press, 1981.

41. Chong, H., Motani, M., Garg, H., and El Gamal, H.: On the han-kobayashi region for the interference channel. IEEE Trans. Inf. Theory, 54(7):3188–3195, July 2008.

42. El Gamal, A. and Costa, M.: The capacity region of a class of deterministic interference channels. IEEE Trans. Inf. Theory, 28(2):343–346, March 1982.

43. Ahlswede, R.: Multi-way communication channels. In Proc. IEEE Int. Symp. Inf. Theory, pages 23–52, Tsahkadsor, March 1973.

44. Dobrushin, R.: General formulation of Shannon's main theorem in information theory. American mathematical society translations: 9 papers on differential equations, 2 on information theory, page 323, 1963.

45. Shang, X., Kramer, G., and Chen, B.: A new outer bound and the noisy-interference sum x2013;rate capacity for gaussian interference channels. IEEE Trans. Inf. Theory, 55(2):689–699, 2009.

46. Cheng, R. and Verdu, S.: On limiting characterizations of memoryless multiuser capacity regions. IEEE Trans. Inf. Theory, 39(2):609–612, 1993.

47. Avestimehr, A., Diggavi, S., and Tse, D.: Wireless network information flow: a deterministic approach. IEEE Trans. Inf. Theory, 57(4):1872–1905, 2011.

48. Li, S., Huang, Y., Liu, T., and Pfizer, H.: On the limits of treating interference as noise for two-user symmetric Gassign interference channel. In Proc. IEEE Int. Symp. Inf. Theory, June 2015.

49. DoCoMo, N.: Requirements, candidate solutions & technology roadmap for lte rel-12 onward. 3GPP RWS-120010, June, 2012.

50. ETSI, L.: Evolved universal terrestrial radio access (e-utra). User Equipment (UE) radio access capabilities (3GPP TS 36.306 version 10.4. 0 Release 10), 2012.

51. MediaTek, R.: Broadcom,. RP-121995: Study on network-assisted interference cancellation and suppression for lte, in 3GPP RAN2, 2012.

52. Kramer, G.: Outer bounds on the capacity of Gaussian interference channels. IEEE Trans. Inf. Theory, 50(3), March 2004.

53. Sato, H.: The capacity of Gaussian interference channel under strong interference. IEEE Trans. Inf. Theory, IT-27(6):786–788, November 1981.

54. Tuninetti, D.: An outer bound region for interference channels with generalized feedback. In Proc. Annual Inf. Theory and Applications Workshop, La Jolla, January 2010.

55. Rini, S., Tuninetti, D., and Devroye, N.: Outer bounds for the interference channel with a cognitive relay. In Proc. IEEE Inf. Theory Workshop, Dublin, September 2010.

56. Tuninetti, D.: K-user interference channels: General outer bound and sum-capacity for certain gaussian channels. In Proc. IEEE Int. Symp. Inf. Theory, pages 1166–1170. IEEE, 2011.

57. Maamari, D., Tuninetti, D., and Devroye, N.: Approximate sum-capacity of k-user cognitive interference channels with cumulative message sharing. IEEE J. Select. Areas Commun., 32(3):654–666, March 2014.

58. Cadambe, V. and Jafar, S.: Interference alignment and the degrees of freedom for the K-user interference channel. IEEE Trans. Inf. Theory, 54(8):3425–3441, August 2008.

59. Etkin, R. and Ordentlich, E.: The degrees-of-freedom of the k -user Gaussian interference channel is discontinuous at rational channel coefficients. IEEE Trans. Inf. Theory, 55(11):4932–4946, Nov 2009.

60. Ordentlich, O., Erez, U., and Nazer, B.: The approximate sum capacity of the symmetric gaussian k-user interference channel. In Proc. IEEE Int. Symp. Inf. Theory, pages 2072–2076, 2012.

61. Sason, I.: On achievable rate regions for the Gaussian interference channel. IEEE Trans. Inf. Theory, June 2004.

62. Wu, Y., Shitz, S., and Verdu, S.: Information dimension and the degrees of freedom of the interference channel. IEEE Trans. Inf. Theory, 61(1):256–279, Jan 2015.

63. Cheng, R. and Verdu, S.: On limiting characterizations of memoryless multiuser capacity regions. IEEE Trans. Inf. Theory, 39(3):609–612, March 1993.

64. Wu, Y. and Verdu, S.: Mmse dimension. IEEE Trans. Inf. Theory, 57(8):4857–4879, 2011.

65. Guo, D., Shamai, S., and Verdú, S.: Mutual information and minimum mean-square error in Gaussian channels. IEEE Trans. Inf. Theory, 51(4):1261–1282, April 2005.

66. Guo, D., Wu, Y., Shamai, S., and Verdú, S.: Estimation in Gaussian noise: Properties of the minimum mean-square error. IEEE Trans. Inf. Theory, 57(4):2371–2385, April 2011.

67. Bustin, R., Payaró, M., Palomar, D. P., and Shamai, S.: On MMSE crossing properties and implications in parallel vector Gaussian channels. IEEE Trans. Inf. Theory, 59(2):818–844, Feb 2013.

68. Guo, D., Shamai, S., and Verdú, S.: The Interplay Between Information and Estimation Measures. now Publishers Incorporated, 2013.

69. Bustin, R., Schaefer, R. F., Poor, H. V., and Shamai, S.: On MMSE properties of optimal codes for the Gaussian wiretap channel. In Proc. IEEE Inf. Theory Workshop, pages 1–5, April 2015.

70. Wu, Y. and Verdú, S.: Functional properties of minimum mean-square error and mutual information. IEEE Trans. Inf. Theory, 58(3):1289–1301, March 2012.

71. Merhav, N., Guo, D., and Shamai, S.: Statistical physics of signal estimation in Gaussian noise: Theory and examples of phase transitions. IEEE Trans. Inf. Theory, 56(3):1400–1416, March 2010.

72. Lozano, A., Tulino, A. M., and Verdú, S.: Optimum power allocation for parallel Gaussian channels with arbitrary input distributions. IEEE Trans. Inf. Theory, 52(7):3033–3051, July 2006.

73. Talagrand, M.: Spin Glasses: A Challenge for Mathematicians: Cavity and Mean Field Models, volume 46. Springer Science & Business Media, 2003.

74. Dembo, A., Cover, T., and Thomas, J.: Information theoretic inequalities. IEEE Trans. Inf. Theory, 37(6):1501–1518, Nov 1991.

VITA

NAME: Alex Dytso

EDUCATION: B.S., Electrical Engineering, University of Illinois at Chicago, Chicago, Illinois, 2011

M.S., Electrical and Computer Engineering, University of Illinois at Chicago, Chicago, Illinois, 2014

Ph.D., Electrical and Computer Engineering, University of Illinois at Chicago, Chicago, Illinois, 2015

EXPERIENCE:

Research Assistant, University of Illinois at Chicago, Chicago; Fall 2011 -Spring 2016.

Intern, National Aeronautics and Space Administration (NASA) Goddard Space Flight Center, Maryland, Summer 2014

PUBLICATIONS: A. Dytso, R. Bustin, D. Tuninetti, N. Devroye, H. Poor, and S. Shamai(Shitz), “New bounds on MMSE and applications to communication with the disturbance constraint”, *Submitted to IEEE Trans. Inf. Theory'*, Mar. 2016.

A. Dytso, D. Tuninetti, and N. Devroye, “Interference as Noise: Friend or Foe?” To appear in to *IEEE Trans. Inf. Theory*, http://arxiv.org/abs/1506.02597, Jun. 2015

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), 12571276, Mar. 2015.

A. Dytso, S. Rini, N. Devroye, and D. Tuninetti, “On the capacity region of the two-user interference channel with a cognitive relay, *IEEE Trans. Wireless Comm.*, vol. 13, no.(12), 68246838, Dec. 2014.

VITA (Continued)

A. Dytso, R. Bustin, D. Tuninetti, N. Devroye, H. Poor, and S. Shamai(Shitz), "On the applications of the minimum mean p-th error (MMPE) to information theoretic quantities," Submitted to the *IEEE Information Theory Workshop*, 2016.

A. Dytso, R. Bustin, D. Tuninetti, N. Devroye, H. Poor, and S. Shamai(Shitz), "On communications through a Gaussian noise channel with an MMSE disturbance constraint", *Proc. Workshop on Info. Theory and Applications,* 18, Feb. 2016.

A. Dytso, D. Tuninetti, and N. Devroye, "Nearly optimal non-Gaussian codes for the Gaussian interference channel", 49th Asilomar Conference on Signals, Systems and Computers, Oct. 2015.

A. Dytso, N. Devroye, and D. Tuninetti, "The Gaussian interference channel with lack of codebook knowledge at one receiver symmetric capacity to within a gap with a PAM input", *Proc. IEEE Info. Theory Workshop*, 15, Apr. 2015.

A. Dytso, N. Devroye, and D. Tuninetti, "On Gaussian interference channels with mixed Gaussian and discrete inputs," *Proc. IEEE Int. Symp. Inf. Theory*, 261265, Jun. 2014.

A. Dytso, D. Tuninetti, and N. Devroye, "On discrete alphabets for the two -user Gaussian interference channel with one receiver lacking knowledge of the interfering codebook" ,*Proc. IEEE Int. Symp. Inf. Theory*, 18, Feb. 2014.

A. Dytso, D. Tuninetti, and N. Devroye, "On the capacity of interference channels with partial codebook knowledge, *Proc. IEEE Int. Symp. Inf. Theory*, 20392043, 2013.

A. Dytso, N. Devroye, and D. Tuninetti, "On the capacity of the symmetric interference channel with a cognitive relay at high snr," *Proc. IEEE Intern. Conf. Commun.*, 23502354, Jun. 2012.

A. Dytso, N. Devroye, and D. Tuninetti, "The sum-capacity of the symmetric linear deterministic complete k-user z-interference channel,"

VITA (Continued)

50th Annual Allerton Conference on Communication, Control, and Computing (Allerton), 12321237, Oct. 2012.