On the Two-User Interference Channel With Lack of Knowledge of the Interference Codebook at One Receiver

Alex Dytso, Daniela Tuninetti, and Natasha Devroye

Abstract—In multiuser information theory, it is often assumed that every node in the network possesses all codebooks used in the network. This assumption may be impractical in distributed ad hoc, cognitive, or heterogeneous networks. This paper considers the two-user interference channel with one oblivious receiver (IC-OR), i.e., one receiver lacks knowledge of the interfering cookbook, whereas the other receiver knows both codebooks. This paper asks whether, and if so how much, the channel capacity of the IC-OR is reduced compared with that of the classical IC where both receivers know all codebooks. A novel outer bound is derived and shown to be achievable to within a gap for the class of injective semideterministic IC-ORS; the gap is shown to be zero for injective fully deterministic IC-ORS. An exact capacity result is shown for the general memoryless IC-OR when the nonoblivious receiver experiences very strong interference. For the linear deterministic IC-OR that models the IC-OR when the nonoblivious receiver experiences very strong interference, Bernoulli(1/2) input bits are shown to achieve an achievable strategy is evaluated in which the input alphabets at the nonoblivious transmitter are a mixture of discrete and Gaussian random variables, where the cardinality of the discrete part is appropriately chosen as a function of the channel parameters. Surprisingly, as the oblivious receiver intuitively should not be able to jointly decode the intended and interfering messages whose codebook is unavailable, it is shown that with this choice of input, the capacity region of the symmetric Gaussian IC-OR is to within 1/2 log (12πe) ≈ 3.34 bits (per channel use per user) of an outer bound for the classical Gaussian IC with full codebook knowledge at both receivers.

Index Terms—Oblivious processing, codebooks, interference channel, discrete inputs.

I. 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.

To the best of our knowledge, systems with oblivious terminals were first introduced in [1]. In [1] 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 [2] and [3] this concept of partial codebook knowledge was extended to model oblivious relays and capacity results were derived. However, as pointed out in [2, Sec. III.A] and [3, 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.

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 [4] 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 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 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.

We are not the first to consider discrete inputs for Gaussian noise channels. In [5] the authors considered the point-to-point power-constrained Gaussian noise channel and derived lower
bounds on the achievable rate when the input is contained to be an equally spaced Pulse Amplitude Modulation (PAM) in which each equal point is used with equal probability; such an input was shown to be optimal to within 0.41 bits per channel use [5, eq. (9)]. As pointed out in [6], already in 1948 Claude Shannon in the unpublished work [7] argued the asymptotically optimality of a PAM input for the point-to-point power-constrained Gaussian noise channel.

In [8, Th. 6 and 7], the authors asymptotically characterized the optimal input distribution over $N$ masses at high and low 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 this work, 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 [8].

The sub-optimality of Gaussian inputs for Gaussian noise channels has been observed before. Past work on the asynchronous IC [9], [10] showed that non-Gaussian inputs may outperform i.i.d. Gaussian inputs by using local perturbations of an i.i.d. Gaussian input: [9, Lemma 3] considers a fourth order approximation of mutual information, while [10, Th. 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 [9, Fig. 1] numerically shows the performance of a ternary PAM input as well. For the cases considered in [9] and [10], 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 [9] and [10] 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 [11]: the main idea is to use sub-optimal point-to-point codes in which the reduction in achievable rates is more than compensated by the decrease in the interference created at the other users.

The rest of the paper is organized as follows. The general memoryless IC-OR channel model is introduced in Section II, together with the special class of injective semi-deterministic IC-ORs (ISD-IC-OR) of which the Gaussian noise channel is an example. Our main results are:

1) In Section III, in Theorem 2 we derive 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 ISD-IC-OR only.

2) In Section IV 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 we prove: (a) in Theorem 4 we obtain the capacity region for the general memoryless IC-OR in very strong interference at the non-oblivious receiver, (b) in Theorem 5 we demonstrate the capacity region to within a gap for the ISD-IC-OR, and (c) in Corollary 6 we show that for the injective fully deterministic IC-OR the gap is zero.

3) In Section IV-D, we look at the practically relevant G-IC-OR and its corresponding Linear Deterministic Approximation (LDA-IC-OR) in the spirit of [12], which models the G-IC-OR at high SNR, whose exact capacity is in principal given by Corollary 6 but whose convex closure we were not able to determine directly. Surprisingly, for the LDA-IC-OR we numerically demonstrate that for the proposed achievable scheme in Proposition 3, i.i.d. Bernoulli(1/2) input bits (known to be optimal for the LDA-IC with full codebook knowledge [13]) are outperformed by other (correlated and non-uniform) input distributions that achieve the same sum-rate performance as full codebook knowledge.

4) In Section IV-E, for the G-IC-OR, we show in Corollary 7 that our inner and outer bounds that 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 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 noise, 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 paper we consider the G-IC-OR, and demonstrate that even with partial codebook knowledge we are able to achieve to within $\frac{1}{2} \log(12\pi e) \approx 3.34$ bits per channel use of the symmetric capacity region of the G-IC with full codebook knowledge through the use of mixed inputs. The main tool, to derive the symmetric capacity to within a constant gap is Theorem 8, which is the lower bound from [5] on the mutual information achievable by a PAM input on a point-to-point Gaussian noise channel. With this tool, in Section VI in Theorems 9 and 10, we evaluate the achievable rate region presented in Proposition 3 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 11 and 12 provide the gDoF characterization of the achievable regions in Theorems 9 and 10.

6) In past work on networks with oblivious nodes no performance guarantees were provided as the capacity regions could not be evaluated. In Section VII we study the generalized degrees of freedom (gDoF) achievable with mixed inputs. In Theorem 13, we show that mixed inputs

1The restriction to the symmetric case, i.e., same direct links and same interference links, is just to reduce the number of parameters in our derivations. We strongly believe that an approximate capacity result (to within a constant gap) can be shown for the general asymmetric case, albeit through more tedious computations than those reported here.
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 VIII we turn our attention to the finite SNR regime and in Theorem 14 we show that the capacity of the symmetric G-IC-OR is within 1/3 log (12πe) ≈ 3.34 bits per channel use of the outer bound to the capacity region of the classical symmetric G-IC.

We conclude the paper with some final remarks and future directions in Section IX. Some proofs are reported in the Appendix.

II. CHANNEL MODEL

A. Notation

We adopt the following notation convention:

- Lower case variables are instances of upper case random variables which take on values in calligraphic alphabets.
- Throughout the paper log(·) denotes logarithms in base 2.
- \([n_1 : n_2]\) is the set of integers from \(n_1\) to \(n_2\). \(n\) is a real-valued random variable, \(x\) is the set of integers from 1 to \(n\).
- \(Y_j\) is a vector of length \(j\) with components \(Y_1, \ldots, Y_j\).
- \(\delta(\cdot)\) denotes the Dirac delta function.
- If \(A\) is a random variable (r.v.) we denote its support by \(\text{supp}(A)\).

The symbol \(|\cdot|\) may denote different things: \(|A|\) is the cardinality of the set \(A\), \(|X|\) is the cardinality of \(\text{supp}(X)\) of the r.v. \(X\), or \(|x|\) is the absolute value of the real-valued \(x\).

- For \(x \in \mathbb{R}\) we let \(|x|\) denote the largest integer smaller than \(x\).
- For \(x \in \mathbb{R}\) we let \(|x|^+ := \max(x, 0)\) and \(\log^+(x) := \log(\min(x, 0))\).

The functions \(I_g(x), I_d(n, x)\) and \(N_d(x)\), for \(n \in \mathbb{N}\) and \(x \in \mathbb{R}^+\), are defined as

\[
I_g(x) := \frac{1}{2} \log(1 + x),
\]

\[
I_d(n, x) := \left[ \frac{1}{2} \log(1 + \min(n^2 - 1, x)) \right]^+, \tag{1}
\]

\[
N_d(x) := \left[ \sqrt{1 + x} \right]^+, \tag{2}
\]

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

- \(\mathcal{N}(\mu, \sigma^2)\) denotes a real-valued Gaussian r.v. with mean \(\mu\) and variance \(\sigma^2\).
- \(\text{Unif}([n_1 : n_2])\) denotes the uniform distribution over the set \([n_1 : n_2]\).
- \(\text{Bernoulli}(p)\) denotes the Bernoulli distribution with parameter \(p \in [0, 1]\).
- \(X \sim \text{PAM}(N)\) denotes the uniform distribution over a zero-mean Pulse Amplitude Modulation (PAM) constellation with \(|X| = N\) points and unit-energy.
- \(\text{co}(\cdot)\) denotes the convex closure operator.

Fig. 1. The IC-OR, where \(F_1\) and \(F_2\) represent codebook indices known to one or both receivers.

B. General Memorless IC-OR

An IC-OR consists of the two-user memoryless interference channel \((X_1, X_2, Y_1, Y_2)\) 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 [1], where transmitters use randomized encoding functions, which are indexed by a message index and a “codebook index” \((F_1, F_2)\) in Fig. 1. An oblivious receiver is unaware of the “codebook index” \((F_1, F_2)\) is not given to decoder 2 in Fig. 1) 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 [2, Definition 2], a \((2^{n R_1}, 2^{n R_2}, n)\) code for the IC-OR with enabled time sharing is a six-tuple \((P_{F_1}(Q^n), \sigma_1^n, \sigma_2^n, \phi_1^n, \phi_2^n)\), where the distribution \(P_{F_1}(Q^n)\), \(i \in [1 : 2]\), 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 \(\sigma_i^n\) and the decoders \(\phi_i^n\), are mappings

\[
\sigma_1^n : [1 : 2^{n R_1}] \times [1 : |F_1|] \rightarrow \mathcal{X}_1^n, \\
\sigma_2^n : [1 : 2^{n R_2}] \times [1 : |F_2|] \rightarrow \mathcal{X}_2^n, \\
\phi_1^n : [1 : |F_1|] \times [1 : |F_2|] \times \mathcal{Y}_1^n \rightarrow [1 : 2^{n R_1}], \\
\phi_2^n : [1 : |F_2|] \times \mathcal{Y}_2^n \rightarrow [1 : 2^{n R_2}].
\]

Moreover, when user 1’s codebook index is unknown at decoder 2, the encoder \(\sigma_1^n\) and the distribution \(P_{F_1}(Q^n)\) must satisfy

\[
P[X_1^n = x_1^n | Q^n = q^n] = \prod_{i=1}^{2^{n R_1}} \mathbb{P}[F_1 | q^n = q_i] \mathbb{P}[F_2 | q^n = q_i] \delta(x_1^n - \sigma_1^n(w_1, f_1))
\]

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 [2, Remark 1] for further justifications of the condition in (4).

A non-negative rate pair \((R_1, R_2)\) is said to be achievable if there exist a sequence of encoding functions
The ISD property implies that entropies.

\[
\begin{align*}
H(Y_1^n | Y_2^n, F_2) &= \sum_{i=1}^{n} H(Y_i | X_i, Y_{i+1}^n, F_2) \\
&\leq H(Y_1^n | Y_2^n, F_2)
\end{align*}
\]

The capacity region is defined as the convex closure of all achievable rate pairs \((R_1, R_2)\) [14].

**Remark 1:** 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|^{2^{\frac{nR}{2}}}| [2]. 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 [1] and [2].

**C. Injective Semi-Deterministic IC-OR**

For a general memoryless IC-OR, no restrictions are imposed on the transition probability \(P_{Y_1 Y_2 | X_1 X_2}.\) The ISD-IC-OR is a special IC-OR with transition probability

\[
P_{Y_1 Y_2 | X_1 X_2}(y_1, y_2 | x_1, x_2) = \sum_{t_1, t_2} P_{T_1 | X_1}(t_1 | x_1) P_{T_2 | X_2}(t_2 | x_2) \delta(y_1 - g_1(x_1, t_2)) \delta(y_2 - g_2(x_2, t_1)),
\]

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 [15].

The ISD property implies that

\[
H(Y_1 | X_1) = H(T_2) \text{ and } H(Y_2 | X_2) = H(T_1)
\]

or in other words that the \(T_n\) is a deterministic function of the pair \((Y_n, X_n), u \in [1 : 2]\). For channels with continuous alphabets, the summation in (5) should be replaced with an integral and the discrete entropies in (6) with the differential entropies.

**III. 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 1:** The output of the oblivious decoder has a product distribution conditioned on the signal whose codebook is known, that is,

\[
P_{Y_2^n | X_2^n, F_2} = \prod_{i=1}^{n} P_{Y_i | X_i}(y_i | x_i).
\]

which implies

\[
H(Y_2^n | X_2^n, F_2) = \sum_{i=1}^{n} H(Y_i | X_i) \text{ for ISD-IC-OR } \equiv \sum_{i=1}^{n} H(T_i).
\]

**Proof of Proposition 1:** Starting from the joint distribution of \(Y_2^n, X_2^n\) conditioned on \(X_2^n, F_2\) we have that

\[
P_{Y_2^n, X_2^n | F_2} = \prod_{i=1}^{n} P_{Y_i | X_i}(y_i | x_i).
\]

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 (4), and c) the inputs are independent. By marginalizing with respect to \(X_1^n\) yields

\[
P_{Y_2^n, X_2^n | F_2} = \prod_{i=1}^{n} P_{Y_i | X_i}(y_i | x_i)
\]

as claimed.

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

**Theorem 2:** Any achievable rate pair \((R_1, R_2)\) for the IC-OR must satisfy

\[
R_1 \leq I(Y_1; X_1 | X_2, Q), \text{ (memoryless IC-OR)} \quad (7a)
\]

\[
R_2 \leq I(Y_2; X_2 | Q), \text{ (memoryless IC-OR)} \quad (7b)
\]

\[
R_1 + R_2 \leq H(Y_1 | Q) + H(Y_2 | U_2, Q) - H(T_2 | X_2, Q)
\]

\[
- H(T_1 | Q) \text{ (memoryless ISD-IC-OR)}
\]

\[
= I(Y_1; X_1, X_2 | Q) + I(Y_2; X_2 | U_2, Q), \quad (7c)
\]

for some input distribution that factors as

\[
P_{Q, X_1, X_2, U_2}(q, x_1, x_2, u_2) = P_{Q}(q) P_{X_1}(x_1 | q) P_{X_2}(x_2 | q) P_{U_2}(u_2 | x_2), \quad (7d)
\]

and with \(|Q| \leq 2\). We denote the region in (7d) as \(R_{\text{out}}\).

**Proof of Theorem 2:** By Fano’s inequality

\[
H(W_1 | Y_1^n, F_1, F_2) \leq n \epsilon_n \text{ and } H(W_2 | Y_2^n, F_2) \leq n \epsilon_n
\]

for some \(\epsilon_n \to 0\) as \(n \to \infty\). We begin with the \(R_1\)-bound (non-oblivious receiver) in (7a):

\[
n(R_1 - \epsilon_n) \leq I(W_1; Y_1^n, F_1, F_2)
\]

\[
\leq I(W_1; Y_1^n | F_1, F_2, W_2)
\]

\[
\leq I(X_1^n; Y_1^n | F_1, F_2, X_2^n)
\]

\[
= H(Y_1^n | F_1, F_2, X_2^n) - \sum_{i=1}^{n} H(Y_{1i} | X_{1i}, X_{2i})
\]

\[
\leq \sum_{i=1}^{n} H(Y_{1i} | X_{2i}) - \sum_{i=1}^{n} H(Y_{1i} | X_{1i}, X_{2i})
\]

\[
= \sum_{i=1}^{n} I(X_{1i}; Y_{1i} | X_{2i}),
\]
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 $(F_i, W_i) \rightarrow X_i^n \rightarrow Y_i^n$, for $i \in [1 : 2]$, and d) because the channel is memoryless, e) by chain rule of entropy and by “conditioning reduces entropy”. For the $R_2$-bound (oblivious receiver) in (7b) we have:

\[ n(R_2 - \epsilon_n) \leq I(W_2; Y_2^n, F_2) \]

\[ \leq I(W_2; Y_2^n | F_2) \]

\[ \leq I(X_2^n; Y_2^n | F_2) \]

\[ = H(Y_2^n | F_2) - \sum_{i=1}^{n} H(Y_2 | X_2) \]

\[ = \sum_{i=1}^{n} I(X_2; Y_2), \]

where the (in)equalities follow from: a) Fano’s inequality, b) the fact that $F_2$ and $W_2$ are independent, c) data processing $(F_i, W_i) \rightarrow X_i^n \rightarrow Y_i^n$, for $i \in [1 : 2]$, and d) by Proposition 1, 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 [15],\(^2\) where $U_2$ is jointly distributed with the inputs according to (7d), we have:

\[ n(R_1 + R_2 - 2\epsilon_n) \leq I(X_1^n; Y_1^n | F_1, F_2) + I(X_2^n; Y_2^n | U_2^n | F_2) \]

\[ = H(Y_1^n | F_1, F_2) - H(Y_1^n | F_1, F_2, X_1^n) \]

\[ + H(U_2^n | F_2) - H(U_2^n | F_2, X_2^n) \]

\[ + H(Y_2^n | F_2, U_2^n) - H(Y_2^n | F_2, X_2^n, U_2^n) \]

\[ \leq H(Y_1^n | F_1, F_2) - H(T_2^n | F_1, F_2) \]

\[ + H(U_2^n | F_2) - H(U_2^n | F_2, X_2^n) \]

\[ + H(Y_2^n | F_2, U_2^n) - H(T_2^n) \]

\[ \leq H(Y_1^n | F_1, F_2) - H(T_2^n | F_1, F_2) \]

\[ + H(T_2^n | F_2) - H(T_2^n | F_2, X_2^n) \]

\[ + H(Y_2^n | F_2, U_2^n) - H(T_2^n) \]

\[ \leq \sum_{i=1}^{n} H(Y_1 | F_1, F_2) + H(Y_2 | F_2, U_2) \]

\[ - H(T_2^n | X_2^n) - H(T_1^n) \]

\[ \leq \sum_{i=1}^{n} H(Y_1 | F_1, F_2) + H(Y_2 | F_2, U_2) \]

\[ - H(T_2 | X_2) - H(T_1), \]

\[ \leq \sum_{i=1}^{n} H(Y_1 | F_1, F_2) + H(Y_2 | F_2, U_2) \]

\[ - H(T_2 | X_2) - H(T_1), \]

where (in)equalities follow from: a) 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 (5) and the independence of $(X_1^n, T_1^n)$ and $X_2^n$, c) by definition of $U_2$ in (7d) 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 $H(T_2^n | F_2, X_2^n) = H(T_2^n | X_2^n)$ and since $H(T_1^n) = H(Y_1^n | X_2^n)$ can be single-letterized by using Proposition 1, and f) by conditioning reduces entropy.

The introduction of a time-sharing random variable $Q \sim \text{Unif}[1 : n]$ yields the bounds in (7). The Fenchel-Eggleston-Carathéodory theorem [16, Ch. 14] guarantees that we may restrict attention to $|Q| \leq 2$ without loss of optimality.

Finally, the equality in (7c) follows from the injective property in (5), the independence of the inputs and the memoryless property of the channel, i.e.,

\[ H(T_2 | X_2) = H(T_2 | X_1, X_2) = H(Y_1 | X_1, X_2, Q), \]

\[ H(T_1 | Q) = H(T_1 | U_2, Q, X_2). \]

This concludes the proof.

\[ \blacksquare \]

IV. CAPACITY RESULTS

In this section we prove that the outer bound in (7d) 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 [4] 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.

A. 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 [4]. It may then be shown that the following rates are achievable,

Proposition 3: The set of non-negative rate pairs $(R_1, R_2)$ satisfying

\[ R_1 \leq I(Y_1; X_1 U_2, Q), \]

\[ R_2 \leq I(Y_2; X_2 | Q), \]

\[ R_1 + R_2 \leq I(Y_1; X_1, U_2 | Q) + I(Y_2; X_2 | U_2, Q), \]

is achievable for every input distribution that factorizes as

\[ P_{Q,X_1,X_2,U_2} = P_Q P_{X_1} P_{X_2} P_{U_2} P_{U_2 | X_2}, \]

and where $|Q| \leq 8$ from [17]. We denote the region in (8) as $\mathcal{R}_{\text{in}}$, which is achievable for any memoryless IC-OR.

Proof of Proposition 3: The proof follows by setting the auxiliary r.v. $U_1$ in the Han and Kobayashi rate region in [14, Sec. 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

\[ \text{Random variable } U_2 \text{ is obtained by passing } X_2 \text{ through an auxiliary channel described by } P_{X_2 | X_2}. \text{ Intuitively, } U_2 \text{ represents interference caused by } X_2 \text{ plus noise at the output } Y_1. \text{ The idea is that providing a noisy version of } X_2 \text{ as side information will result in a tighter bound than for example giving just } X_2. \]
codebook of transmitter 1 is not needed at receiver 2 (the oblivious receiver).

Remark 2: By comparing the outer bound region $\mathcal{R}_{\text{out}}$ in Theorem 2 to the inner bound region $\mathcal{R}_{\text{in}}$ in Proposition 3 we notice the following differences: 1) in (7d) the side information random variable $U_2$ is distributed as $T_2$ conditioned on $X_2$, while in (8d) 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.

B. 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 2 is tight.

A general memoryless IC-OR for which

$$I(X_2;Y_2|X_1) \leq I(X_2;Y_1), \quad \forall P_{X_1}, X_2 = P_{X_1}P_{X_2}, \quad (9)$$

is said to have very strong interference at the non-oblivious receiver (receiver 1). Intuitively, when the condition in (9) 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 (7c). Next we formalize this intuition.

Theorem 4: When the condition in (9) holds the capacity region of the IC-OR is given by

$$R_1 \leq I(X_1; Y_1|X_2, Q), \quad (10a)$$

$$R_2 \leq I(X_2; Y_2|Q), \quad (10b)$$
taken over the union of all input distributions that factor as $P_{Q,X_1,X_2} = P_QP_{X_1|Q}P_{X_2|Q}$ and where $|Q| \leq 2$.

Proof of Theorem 4: By dropping the sum-rate outer bound in (7c) we see that the region in (10a) is an outer bound for a general memoryless IC-OR. By setting $U_2 = X_2$ in the achievable region in (8), the region

$$R_1 \leq I(X_1; Y_1|X_2, Q), \quad (11a)$$

$$R_2 \leq I(X_2; Y_2|Q), \quad (11b)$$

$$R_1 + R_2 \leq I(X_1; X_2; Y_1|Q), \quad (11c)$$
taken over the union of all $P_{Q,X_1,X_2} = P_QP_{X_1|Q}P_{X_2|Q}$, is achievable. We see that the single rate bounds in (11a) match the upper bounds in (10a). We next intend to show that when the condition in (9) holds, the sum-rate bound in (11c) is redundant. By summing (11a) and (11b)

$$R_1 + R_2 \leq I(X_1; Y_1|X_2, Q) + I(X_2; Y_2|Q) \leq I(X_1; Y_1|X_2, Q) + I(X_2; Y_2, X_1|Q) \leq I(X_1; Y_1|X_2, Q) + I(X_2; Y_1|Q) = I(X_1, X_2; Y_1|Q) = \text{eq.}(11c),$$

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 (9). Therefore, the sum-rate bound in (11c) can be dropped without affecting the achievable rate region. This shows that the outer bound in (10a) is achievable thereby proving the claimed capacity result.

Remark 3: For the classical IC, the very strong interference regime is defined as

$$I(X_1; Y_1|X_2) \leq I(X_1; Y_2),$$

$$I(X_2; Y_2|X_1) \leq I(X_2; Y_1),$$

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 [1] 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.

C. Capacity to Within a Constant Gap for the ISD-IC-OR

We now show that $\mathcal{R}_{\text{in}}$ in Proposition 3 lies to within a gap of the outer bound $\mathcal{R}_{\text{out}}$ in Theorem 2 for the general ISD-IC-OR. We have

Theorem 5: 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 5: The proof is as in [15]. First, we define a new outer bound region $\mathcal{R}_{\text{out}}$ 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 $\mathcal{R}_{\text{out}} \subseteq \mathcal{R}_{\text{out}}$. We next compare $\mathcal{R}_{\text{out}}$ and $\mathcal{R}_{\text{in}}$ term by term (we only need to compare the mutual informations involving $Y_1$ as those involving $Y_2$ are the same in both bounds, see Remark 2, thus implying a zero gap for rate $R_2$): the difference is that $\mathcal{R}_{\text{out}}$ 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.

Remark 4: Note that

$$I(X_2; T_2|U_2, Q) = H(T_2|U_2, Q) - H(T_2|X_2) \leq H(T_2) - H(T_2|X_2) \leq \max I(T_2; X_2),$$

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 IV-E after having discussed in Section IV-D a special class of ISD-IC-OR for which the gap to capacity is zero.
D. Exact Capacity for the Injective Fully Deterministic IC-OR

We now specialize Theorem 5 to the class of injective fully deterministic ICs [18]. For this class of channels the mappings $T_1$ and $T_2$ in (5) are deterministic functions of $X_1$ and $X_2$, respectively. We have

**Corollary 6**: For the injective fully deterministic IC-OR the outer bound in Theorem 2 is tight.

**Proof of Corollary 6**: The injective fully deterministic IC-OR has $T_2 = U_2$ and therefore $I(X_2; T_2|U_2, Q) = 0$ in Theorem 5.

As an application of Corollary 6 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 [12]. The LDA-IC-OR has input/output relationship

$$Y_1 = S^{q_{11}} X_1 + S^{q_{12}} X_2, \quad T_2 = S^{q_{21}} X_1, \quad T_1 = S^{q_{22}} X_2,$$

where inputs and outputs are binary-valued vectors of length $q$, $S$ is the $q \times q$ shift matrix [12], $(n_{11}, n_{12}, n_{21}, n_{22})$ 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 $n_{11} = n_{22} = n_S$ and $n_{12} = n_{21} = n_1 := n_S/\alpha$ for some non-negative $\alpha$. The maximum symmetric rate, or symmetric sum-capacity normalized by the sum-capacity of an interference-free channel, is defined as

$$d(\alpha) := \frac{\max(R_1 + R_2)}{2n_S},$$

where the maximization is over all achievable rate pairs $(R_1, R_2)$ satisfying Theorem 2, which is the capacity region by Corollary 6. 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

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

where $d^{(W)}(\alpha)$, the so-called W-curve [19], is the maximum symmetric rate of the classical LDA-IC. In [13] it was shown that i.i.d. Bernoulli$(1/2)$ input bits in the Han and Kobayashi region yield $d^{(W)}(\alpha)$.

Although Theorem 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 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. 2 and Table I. Although we lack a formal proof that we can achieve the whole W-curve with a non i.i.d. Bernoulli$(1/2)$ input we do, however, conjecture that it is indeed possible with the scheme in Proposition 3. 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. 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 $\alpha$ and $n_S, n_1$, the input distributions to be used in $R_{\text{out}}$ in Theorem 2. We notice that, at least when evaluating the region in Theorem 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. Bernoulli$(1/2)$ input bits in Theorem 2 for $|Q| = 1$ we would obtain a normalized sum-rate of $1/2 = 3/6$, the same as achieved by time division [13].

Also, i.i.d. 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. 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 paper tries to use intuition gained in this section to construct non-Gaussian inputs that will be capacity or constant gap capacity approaching.
E. 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

\[ Y_1 = h_{11}X_1 + h_{12}X_2 + Z_1 = h_{11}X_1 + T_2, \]
\[ T_2 = h_{12}X_2 + Z_1, \] (15a)
\[ Y_2 = h_{21}X_1 + h_{22}X_2 + Z_2 = h_{22}X_2 + T_1, \]
\[ T_1 = h_{21}X_1 + Z_2, \] (15b)

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 5 to the G-IC-OR we may show the following:

**Corollary 7:** For the G-IC-OR the gap is at most 1/2 bit per channel use.

**Proof of Corollary 7:** For the G-IC-OR \( T_2 = h_{12}X_2 + Z_1, \) and thus we set \( U_2 \) in Theorem 2 to \( U_2 = h_{12}X_2 + Z_1^*, \) where \( Z_1^* \) and mutually independent. We thus have

\[ I(X_2; U_2| U_2, Q) = h(T_2| U_2, Q) \leq h(T_2) - h(Z_1) \]
\[ = h(Z_1 - Z_1^*) - h(Z_1) = \frac{1}{2} \log(2), \]

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 [19]. From the rate expression in Theorem 2 it is not clear whether Gaussian inputs are optimal for \( R_{\text{out}}. \) 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 \( |h_{11}|^2 = |h_{22}|^2 = \text{SNR} \) and \( |h_{12}|^2 = |h_{21}|^2 = \text{INR}, \) with INR = SNR\(^{\alpha}\) for some non-negative \( \alpha. \) The sum-gDoF is defined as

\[
d(\alpha) := \lim_{\text{SNR} \to \infty} \frac{\max[R_1 + R_2]}{2 \cdot \frac{1}{2} \log(1 + \text{SNR})},
\]

where the maximization is over all possible achievable rate pairs. By using the classical G-IC as a trivial upper bound, we have \( d(\alpha) \leq d^{(\text{GG})}(\alpha) \) where \( d^{(\text{GG})}(\alpha) \) is given in (14).

By evaluating Theorem 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

\[
(R_1 + R_2)^{(\text{GG})} = \min \left\{ I_g(SNR) + I_g\left(\frac{\text{SNR}}{1 + \text{INR}}\right), \right. \]
\[ I_g\left(\frac{\text{SNR}}{1 + \text{INR}}\right) + I_g\left(\frac{\text{SNR}}{1 + \text{INR}}\right), \]
\[ \iff d^{(\text{GG})}(\alpha) = \frac{1}{2} + \left\lceil \frac{1}{2} - \alpha \right\rceil, \]

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

\[
(R_1 + R_2)^{(\text{TD})} = \frac{1}{2} \log(1 + 2 \cdot \text{SNR}),
\]

\[
\iff d^{(\text{TD})}(\alpha) = \frac{1}{2}\]

We plot the achievable gDoF vs. \( \alpha \) in Fig. 2, together with the gDoF of the classical G-IC given by \( d^{(\text{W})}(\alpha) \) [19]. We note that Gaussian inputs are indeed optimal for \( 0 \leq \alpha \leq 1/2, \) i.e., \( d^{(\text{GG})}(\alpha) = d^{(\text{W})}(\alpha), \) 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 \( d^{(\text{GG})}(\alpha) = d^{(\text{TD})}(\alpha), \) 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 \( d^{(\text{GG})}(4/3) = d^{(\text{TD})}(4/3) = 1/2. \) 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 [1, Sec. VI.A], which demonstrated that binary signaling

| \( \alpha, (n_1, n_1) \) | Probability mass function with \(|Q| = 1\) |
|---|---|
| \( \frac{1}{2}, (2,1) \) | \( P_{X_1} = [0.5, 0.5, 0.5, 0.0] \)
| \( P_{X_2} = [0.5, 0.5, 0.5] \) |
| \( \frac{2}{3}, (3,2) \) | \( P_{X_1} = [0.0, 0.25, 0.25, 0.0, 0.25, 0.25, 0.0, 0.25, 0.25] \)
| \( P_{X_2} = [0.0, 0.25, 0.25, 0.0, 0.25, 0.25, 0.0, 0.25, 0.25] \) |
| \( 1, (2,2) \) | \( P_{X_1} = [0.0, 0.5, 0.5, 0.0] \)
| \( P_{X_2} = [0.0, 0.5, 0.5] \) |
| \( \frac{2}{3}, (3,4) \) | \( P_{X_1} = [0.0, 0.0, 0.0, 0.25, 0.25, 0.0, 0.0, 0.0, 0.25, 0.25] \)
| \( P_{X_2} = [0.0, 0.0, 0.0, 0.25, 0.25, 0.0, 0.0, 0.0, 0.0, 0.25, 0.25] \) |
| \( 2, (2,1) \) | \( P_{X_1} = [0.0, 0.5, 0.5, 0.0] \)
| \( P_{X_2} = [0.0, 0.5, 0.5] \) |
consider a uniform PAM constellation with 
outperforms Gaussian signaling for a fixed finite SNR, we 
within the Gaussian noise channel) to within 
gap ≤ \log (SNR) − \log (N_{d}(SNR), SNR)

\begin{align*}
\text{shaping loss} & + \frac{1}{2} \log \left( \frac{1 + \text{SNR}}{\sqrt{1 + \text{SNR}}} \right)^2 \\
\text{integer penalty} & = \frac{1}{2} \log \left( \frac{\pi e}{3} \right) + \frac{1}{2} \log(4) = \frac{1}{2} \log \left( \frac{4\pi e}{3} \right),
\end{align*}

where the largest integer penalty is attained for 
2^{2 - \epsilon}, 0 < \epsilon \ll 1, \text{for which } \sqrt{1 + \text{SNR}}^2 = (2 - 1)^2 = 1.

VI. ACHIEVABLE REGIONS FOR THE G-IC-OR

We now analyze the G-IC-OR by using Theorem 8 
i.e., bounds on the mutual information achievable by 
A CHIEVABLE REGIONS FOR THE G-IC-OR
Gaussian noise channel) and the insight on the nature of the gap due to a PAM input from Remark 5. 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.

A. Achievable Scheme I

We first derive an achievable rate region from Proposition 3 with inputs

Scheme I: $X_{1D} \sim \text{PAM} (N), N \in \mathbb{N}$, independent of

\begin{align}
X_{2G} & \sim \mathcal{N}(0, 1), \\
X_1 & = X_{1D}, \quad X_2 = X_{2G}, \\
U_2 & = X_2, \quad Q = \emptyset.
\end{align}

(20a)

(20b)

(20c)

(20d)

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

Theorem 9: For the G-IC-OR the following rate region is achievable by the input in (20a)

\begin{align}
R_1 & \leq I_d \left( N, |h_{11}|^2 \right), \\
R_2 & \leq I_d \left( N, |h_{21}|^2 \right) + \log \left( |h_{22}|^2 \right) - \log \left( \min (N^2 - 1, |h_{21}|^2) \right), \\
R_1 + R_2 & \leq I_d \left( N, \frac{|h_{11}|^2}{1 + |h_{12}|^2} \right) + \log \left( |h_{12}|^2 \right).
\end{align}

(21a)

(21b)

(21c)

Proof of Theorem 9: We proceed to evaluate the rate region in Proposition 3 with the inputs in (20a), that is, the achievable region in (11a) with $|Q| = 1$.

The rate of the user 1 is bounded by $R_1 \leq I(X_1; Y_1|X_2) = I(X_{1D}; h_{11}X_{1D} + Z_1)$, where $I(X_{1D}; h_{11}X_{1D} + Z_1)$ can be further lower bounded by using (17) from Theorem 8 with $\text{SNR} = |h_{11}|^2$; by doing so we obtain the bound in (21a).

The rate of the user 2 is bounded by

\begin{align}
R_2 & \leq I(X_2; Y_2) \\
& = h(h_{21}X_{1D} + h_{22}X_{2G} + Z_2) - h(h_{21}X_{1D} + Z_2) \\
& \geq \log \left( \frac{h_{21}X_{1D} + Z_2}{1 + |h_{22}|^2} \right) \\
& \geq \log \left( \min (N^2 - 1, |h_{21}|^2) \right) \quad \text{from (17)}
\end{align}

(22a)

(22b)

(22c)

from which we conclude that the achievable rate for user 2 is lower bounded as in (21b).

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 Theorem 8 with $\text{SNR} = |h_{11}|^2$ and where $I(X_2; Y_1|X_1) = I(X_{2G}; h_{12}X_{2G} + Z_1) = \log (|h_{12}|^2)$; by combining the terms we obtain the bound in (21c).

B. Achievable Scheme II

The input in (20a) might not be optimal in general and may be generalized as follows. Consider the rate region in Proposition 3 with inputs

Scheme II:

\begin{align}
X_{1D}, X_{1G}, X_{2Gc}, X_{2Gp} & \text{ independent and distributed as} \\
X_{1D} & \sim \text{PAM} (N), \quad N \in \mathbb{N}, \\
\text{all the others} & \sim \mathcal{N}(0, 1), \\
X_1 & = \sqrt{1 - \delta_1}X_{1D} + \sqrt{\delta_1}X_{1G}, \quad \delta_1 \in [0, 1], \\
X_2 & = \sqrt{1 - \delta_2}X_{2Gc} + \sqrt{\delta_2}X_{2Gp}, \quad \delta_2 \in [0, 1], \\
U_2 & = X_{2Gc}, \quad Q = \emptyset.
\end{align}

(22a)

(22b)

(22c)

(22d)

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 [4]. 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 the paper 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 [1] 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 with the inputs in (22a) 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 (22a) reduces to Scheme I in (20a).

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 10: For the G-IC-OR the following rate region is achievable with inputs as in (22a)
\[
R_1 \leq \frac{1}{1} \left( N, \frac{|h_{11}|^2 (1 - \delta_1)}{1 + |h_{11}|^2 \delta_1 + |h_{12}|^2 \delta_2} \right) + \frac{1}{1} \left( N, \frac{|h_{21}|^2 (1 - \delta_1)}{1 + |h_{21}|^2 \delta_1 + |h_{22}|^2 \delta_2} \right),
\]
(23a)
\[
R_2 \leq \frac{1}{1} \left( N, \frac{|h_{21}|^2 (1 - \delta_1)}{1 + |h_{21}|^2 \delta_1 + |h_{22}|^2 \delta_2} \right) + \frac{1}{1} \left( N, \frac{|h_{21}|^2 (1 - \delta_1)}{1 + |h_{21}|^2 \delta_1 + |h_{22}|^2 \delta_2} \right) - \frac{1}{1} \left( N, \frac{|h_{11}|^2 (1 - \delta_1)}{1 + |h_{11}|^2 \delta_1 + |h_{12}|^2 \delta_2} \right),
\]
(23b)
\[
R_1 + R_2 \leq \frac{1}{1} \left( N, \frac{|h_{11}|^2 (1 - \delta_1)}{1 + |h_{11}|^2 \delta_1 + |h_{12}|^2 \delta_2} \right) + \frac{1}{1} \left( N, \frac{|h_{21}|^2 (1 - \delta_1)}{1 + |h_{21}|^2 \delta_1 + |h_{22}|^2 \delta_2} \right) + \frac{1}{1} \left( N, \frac{|h_{11}|^2 (1 - \delta_1)}{1 + |h_{11}|^2 \delta_1 + |h_{12}|^2 \delta_2} \right) - \frac{1}{1} \left( N, \frac{|h_{21}|^2 (1 - \delta_1)}{1 + |h_{21}|^2 \delta_1 + |h_{22}|^2 \delta_2} \right),
\]
(23c)

Proof of Theorem 10: The proof can be found in Appendix A and follows similarly to the proof of Theorem 9.

VII. HIGH SNR PERFORMANCE

We now analyze the performance of the schemes in Theorems 9 and 10 for the symmetric G-IC-OR at high-SNR by using the gDoF region as performance metric. The gDoF region is formally defined as follows. For an achievable pair \((R_1, R_2)\), let
\[
D(\alpha) := \left\{ (d_1, d_2) \in \mathbb{R}_+^2 : d_i := \lim_{\text{SNR} \to \infty} \frac{R_i}{\frac{1}{1} \log(1 + \text{SNR})} \right\},
\]
for any \(\alpha \geq 0\). By Theorem 9, the gDoF region \(D(\alpha, \beta)\) is achievable.

Proof of Corollary 11: We prove the bound in (26b) only as the other bounds follow similarly. With \(\text{INR} = \text{SNR}^\alpha\) and \(N = N_d(\text{SNR}^\beta)\) we have
\[
\lim_{\text{SNR} \to \infty} \frac{\log(N^2)}{\log(1 + \text{SNR})} = \beta, \quad \lim_{\text{SNR} \to \infty} \frac{\log(1 + \text{INR})}{\log(1 + \text{SNR})} = \alpha.
\]
Therefore \(d_2\) can be bounded as
\[
d_2 = \lim_{\text{SNR} \to \infty} \frac{\text{left hand side of eq.}(21b)}{\text{SNR} \to \infty} \left( \frac{1}{1} \log(1 + \text{SNR}) \right) = \min(\beta, [\alpha - 1]^+) + 1 - \min(\beta, \alpha),
\]
thus proving (26b).

Next, by using Theorem 10 with the power split as in [19] we show yet another achievable gDoF region.

Corollary 12 (gDoF Region From Achievable Scheme II): Let \(N = N_d(\text{SNR}^\beta)\) and
\[
D^{II}(\alpha, \beta) : d_1 \leq \min(\beta, 1 + \alpha - \max(1, \alpha)) + [1 - \alpha]^+,
\]
(27a)
\[
d_2 \leq \min(\beta, [\alpha - 1]^+) + 1 - \min(\beta, \alpha),
\]
(27b)
\[
d_1 + d_2 \leq \min(\beta, [1 + \alpha - \max(1, 2\alpha)]^+) + \max(\alpha, 1 - \alpha) + \min(\beta, [2\alpha - \max(1, \alpha)]^+) + [1 - \alpha]^+ - \min(\beta, \alpha).
\]
(27c)

Proof of Theorem 13: We consider several regimes:

A. Very Strong Interference Regime \(\alpha \geq 2\)

In this regime the gDoF region outer bound \(D^{G-IC}(\alpha)\) is characterized by (25a) and (25b). For achievability we consider Corollary 11 with \(\beta = 1\), that is,
\[
D^{I}(\alpha, 1) : d_1 \leq \min(1,1) = 1,
\]
(25a)
\[
d_2 \leq \min(1, [\alpha - 1]^+) + 1 - \min(1, \alpha)
\]
(25b)
\[
d_1 + d_2 \leq \min(1, [1 - \alpha]^+) + \alpha
\]
(25c)
\[
d_1 + d_2 \leq \alpha \text{ (redundant because } \alpha \geq 2).}
\]
Since the sum-gDoF is redundant, we get that
\[
D^{I}(\alpha, 1) = \{d_1 \in [0, 1], i \in [1, 2]\} = D^{G-IC-OR}(\alpha) = D^{G-IC}(\alpha).
\]

Fig. 4(a) illustrates the region \(D^{I}(\alpha, \beta = 1)\).
B. Strong Interference Regime 1 ≤ α < 2

In this regime the gDoF region outer bound $\mathcal{D}^{G\text{-IC}}(\alpha)$ is characterized by (25a)-(25c) 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

$$D^I(\alpha, 1) : d_1 \leq 1,$$
$$d_2 \leq \alpha - 1,$$
$$d_1 + d_2 \leq \alpha \text{ (redundant).}$$

and

$$D^I(\alpha, \alpha - 1) : d_1 \leq \alpha - 1,$$
$$d_2 \leq 1,$$
$$d_1 + d_2 \leq \alpha \text{ (redundant).}$$

Fig. 4(b) illustrates that

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

C. Moderately Weak Interference Regime $\frac{1}{3} < \alpha < 1$

In this regime the gDoF region outer bound $\mathcal{D}^{G\text{-IC}}(\alpha)$ is characterized by all the constraints in (25a) 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

$$D^{II}(\alpha, 2\alpha - 1) : d_1 \leq \min(2\alpha - 1, 1 + \alpha - 1) + 1 - \alpha = \alpha,$$
$$d_2 \leq \min(2\alpha - 1, 0) + 1 - \min(2\alpha - 1, \alpha) = 2 - 2\alpha,$$
$$d_1 + d_2 \leq \min(2\alpha - 1, [1 + \alpha - \max(1, 2\alpha)]^+ + \min(2\alpha - 1, [2\alpha - 1]^+) + \max(\alpha, 1 - \alpha) + 1 - \alpha - \min(2\alpha - 1, \alpha),$$
$$= \min(2\alpha, 2 - \alpha),$$
(redundant for $\alpha \in \left[\frac{2}{3}, 1\right]$).
and
\[
D^H(a, 1 - a) : d_1 \leq \min(1 - a, 1 + a - 1) \\
+1 - a = 2 - 2a, \\
d_2 \leq \min(1 - a, 0) + 1 \\
- \min(1 - a, a) = a, \\
d_1 + d_2 \leq \min(1 - a, \\
[1 + a - \max(1, 2a)]^+ \\
+ \max(a, 1 - a) \\
+ \min(1 - a, [2a - 1]^+) \\
+ 1 - a - \min(1 - a, a) \\
= \min(2a, 2 - a), \\
\text{(redundant for } a \in [2/3, 1]).
\]

Fig. 4(c) (for \( a \in [2/3, 1] \)) and Fig. 4(d) (for \( a \in [1/2, 2/3] \)) illustrate that
\[
\begin{align*}
\mathcal{C} \left( \{d_1, d_2\} = (1, 0) \cup \{d_1, d_2\} = (0, 1) \right) \\
\cup D^H(2a - 1) \cup D^H(a, 1 - a) = D^{G-IC}(a) \\
= D^{G-IC-OR}(a).
\end{align*}
\]

D. Noisy Interference \( 0 \leq a \leq \frac{1}{2} \)

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 13 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.

VIII. 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 show that this is indeed possible. For future use, the capacity region of the classical G-IC is outer bounded by [19]
\[
\begin{align*}
R_1^\text{G-IC} : R_1 &\leq I_g(\text{SNR}), \\
R_2 &\leq I_g(\text{SNR}),
\end{align*}
\]
which is tight for \( \text{SNR} \leq \text{INR} \) and optimal to with 1/2 bit (per channel use per user) otherwise.

The main result of this section is:

**Theorem 14:** For the G-IC-OR it is possible to achieve the outer bound region in (28f) to within \( \frac{1}{2} \log (12\pi e) \approx 3.34 \) bits per channel use per user.

**Proof of Theorem 14:** We consider different regimes separately.

A. Very Strong Interference \( \text{SNR}(1 + \text{SNR}) \leq \text{INR} \)

In the regime the capacity region of the classical G-IC is given by (28a) and (28b). For achievability we consider the achievable region in Theorem 9 with

\[
N = N_d(\text{SNR}) \quad \text{(equivalent of } \beta = 1) \Rightarrow N^2 - 1 \leq \text{SNR} \leq \frac{\text{INR}}{1 + \text{SNR}} \leq \text{INR}.
\]

Recall that the achievable region in Theorem 9 is the region in (11a) with the inputs as in (20a); the sum-rate in Theorem 9 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 (20a). With a Gaussian \( X_2 \) as in (20a):

\[
I(X_2; Y_2) \leq I(X_2; Y_2|X_1) = I(X_2; Y_2; \sqrt{\text{SNR}} X_2 + Z_2) \\
= I_g(\text{SNR}),
\]

and

\[
I(X_2; Y_1) = I(X_2; Y_1; \sqrt{\text{INR}} X_2 + \sqrt{\text{SNR}} X_1 + Z_2) \\
\geq I_g \left( \frac{\text{INR}}{1 + \text{SNR}} \right),
\]

because a Gaussian noise is the worst noise for a Gaussian input. Since in very strong interference we have \( I_g(\text{SNR}) \leq I_g \left( \frac{\text{INR}}{1 + \text{SNR}} \right) \), the condition \( I(X_2; Y_2) \leq I(X_2; Y_1) \) is verified for all inputs in (20a) and hence we can drop the sum-rate constraint in (21c) from Theorem 9. Therefore, in this regime the following rates are achievable:
\( R_{\text{in}}^{\text{(G-IC-OR very strong)}}: \quad R_1 \leq I_g(\text{SNR}) - \Delta_1, \quad (30a) \)
\( R_2 \leq I_g(\text{SNR}) - \Delta_2, \quad (30b) \)

where for \( N = N_d(\text{SNR}) \)
\( \Delta_1 := I_g(\text{SNR}) - I_d(N, \text{SNR}) \)
\( \leq \frac{1}{2} \log \left( \frac{4\pi e}{3} \right) \) for the reasoning leading to eq.(19),
\( \Delta_2 := I_g \left( \min \left( N^2 - 1, \text{INR} \right) \right) - I_d(N, \frac{\text{INR}}{1 + \text{SNR}}) \)
\( = \log(N) - \left[ \log(N) - \frac{1}{2} \log \left( \frac{\pi e}{3} \right) \right]^+ \)
\( \leq \frac{1}{2} \log \left( \frac{\pi e}{3} \right), \quad (30d) \)

where the equality in (30d) is a consequence of the relationships in (29).

It is immediate to see that (30c) is the gap for \( R_1 \) and that (30d) is the gap for \( R_2 \). Therefore in this regime the gap is at most \( \frac{1}{2} \log \left( \frac{4\pi e}{3} \right) \) per channel use per user, and it is due to shaping loss and integer penalty.

\( \text{B. Strong Interference } \text{SNR} \leq \text{INR} < \text{SNR}(1 + \text{SNR}) \)

In this regime the capacity region of the classical G-IC is given by (28a)-(28c), and has two dominant corner points
\[ R_{\text{out}}^{\text{(G-IC strong P1)}}: \quad (R_1, R_2) = \left( I_g(\text{SNR}), I_g \left( \frac{\text{INR}}{1 + \text{SNR}} \right) \right), \quad (31a) \]
and
\[ R_{\text{out}}^{\text{(G-IC strong P2)}}: \quad (R_1, R_2) = \left( I_g \left( \frac{\text{INR}}{1 + \text{SNR}} \right), I_g(\text{SNR}) \right). \quad (31b) \]

The other two corner points are \((R_1, R_2) = (I_g(\text{SNR}), 0)\) and \((R_1, R_2) = (0, I_g(\text{SNR}))\) 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. 4(b)), that is, we show the achievability to within a constant gap of the corner points in (31a) and (31b) by choosing two different values of \( N \) in Theorem 9. For the corner point in (31a) we consider the achievable region in Theorem 9 with
\( N = N_d(\text{SNR}) \) (equivalent of \( \beta = 1 \))
\( \implies N^2 - 1 \leq \text{SNR} \leq \text{INR} \leq \text{SNR}(1 + \text{SNR}), \quad (32a) \)
and for the corner point (31b) we consider the achievable region in Theorem 9 with
\( N = N_d \left( \frac{\text{INR}}{1 + \text{SNR}} \right) \) (equivalent of \( \beta = \alpha - 1 \))
\( \implies N^2 - 1 \leq \frac{\text{INR}}{1 + \text{SNR}} \leq \text{SNR} \leq \text{INR}. \quad (32b) \)

For the choice of \( N \) in (32a) the achievable region in Theorem 9 can be written as
\[ R_1 \leq I_d(N, \text{SNR}) \]
\[ = \left[ \log(N) - \frac{1}{2} \log \left( \frac{\pi e}{3} \right) \right]^+, \]
\[ R_2 \leq I_d \left( N, \frac{\text{INR}}{1 + \text{SNR}} \right) + I_g(\text{SNR}) \]
\[ - I_g \left( \min \left( N^2 - 1, \text{INR} \right) \right) \]
\[ = \left[ I_g \left( \frac{\text{INR}}{1 + \text{SNR}} \right) - \frac{1}{2} \log \left( \frac{\pi e}{3} \right) \right]^+ \]
\[ + I_g(\text{SNR}) - \log(N), \]
\[ R_1 + R_2 \leq I_d(N, \frac{\text{SNR}}{1 + \text{INR}}) + I_g(\text{INR}) \]
\[ = \left[ I_g \left( \frac{\text{SNR}}{1 + \text{INR}} \right) - \frac{1}{2} \log \left( \frac{\pi e}{3} \right) \right]^+ + I_g(\text{INR}), \]

which can further be lower bounded as
\[ R_{\text{in}}^{\text{(G-IC strong P1)}}: \quad R_1 \leq \log(N) - \frac{1}{2} \log \left( \frac{\pi e}{3} \right) \]
\[ = I_g(\text{SNR}) - \Delta_1, \quad (33a) \]
\[ R_2 \leq I_g(\text{SNR} + \text{INR}) - \log(N) \]
\[ - \frac{1}{2} \log \left( \frac{\pi e}{3} \right) \]
\[ = I_g \left( \frac{\text{INR}}{1 + \text{SNR}} \right) - \Delta_2, \quad (33b) \]
\[ R_1 + R_2 \leq I_g(\text{SNR} + \text{INR}) - \frac{1}{2} \log \left( \frac{\pi e}{3} \right) \]
\[ = (I_g(\text{SNR}) - \Delta_1) \]
\[ + \left( I_g \left( \frac{\text{INR}}{1 + \text{SNR}} \right) - \Delta_2 \right) \]
\[ + \frac{1}{2} \log \left( \frac{\pi e}{3} \right), \quad (33c) \]

where the sum-rate bound is clearly redundant and where
\[ \Delta_1 := I_g(\text{SNR}) - \log(N) + \frac{1}{2} \log \left( \frac{\pi e}{3} \right) \]
\[ \leq \frac{1}{2} \log \left( \frac{4\pi e}{3} \right), \quad (34a) \]
\[ \Delta_2 := \log(N) - I_g(\text{SNR}) + \frac{1}{2} \log \left( \frac{\pi e}{3} \right) \]
\[ \leq \frac{1}{2} \log \left( \frac{\pi e}{3} \right), \quad (34b) \]

Therefore, with \( N \) as in (32a) in Theorem 9, the gap to the corner point in (31a) is at most \( \frac{1}{2} \log \left( \frac{4\pi e}{3} \right) \) per channel use per user, as for the very strong interference regime.

By following similar steps, for the choice of \( N \) in (32b) in Theorem 9, the gap to the corner point in (31b) is still
given by (34a), that is, the gap is at most $\frac{1}{2}\log\left(\frac{4\pi e}{3}\right)$ per channel use per user, as for the very strong interference regime.

C. Moderately Weak Interference

**INR ≤ SNR ≤ INR(1 + INR)**

In this regime the capacity of the G-IC is outer bounded by (28f).

As we did for the gDoF region (see Figs. 4(c) and 4(d)), we show here that we can achieve, up to a constant gap, all dominant corner points of (28f). By silencing one of the users, we can achieve $(R_1, R_2) = (\log (\text{SNR}), 0)$ and $(R_1, R_2) = (0, \log (\text{SNR}))$; these rate points are to within 1 bit of the corner points of (28f) given by $(R_1, R_2) = (A, \log (\text{SNR}))$ and $(R_1, R_2) = (\log (\text{SNR}), A)$ where

\[
A := \log (\text{SNR} + \text{INR}) + \log \left(\frac{\text{INR} + \text{SNR}}{1 + \text{INR}}\right) - \log (\text{SNR}) - \log (\text{INR}) = \log \left(\frac{\text{INR}}{1 + \text{INR}}\right) + \log \left(\frac{\text{SNR}}{1 + \text{INR}}\right) \leq \log \left(\frac{\text{INR}}{1 + \text{INR}}\right) \leq 2 \cdot \frac{1}{2}\log(2) = 1.
\]

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.(28c), eq. (28d))) with either (28e) or (28f). For these corner points, the gDoF-optimal choices of $\beta$ were $2\alpha - 1$ and $1 - \alpha$, which we mimic here by choosing the following values of $N$ in the region in (40c) (a simplified achievable region from Theorem 10)

\[
N = N_d \left(\frac{\text{SNR} \cdot \text{INR}}{1 + \text{INR}^2 + \text{SNR}}\right) \text{ (equivalent of } \beta = 2\alpha - 1) \implies N^2 - 1 \leq \frac{\text{INR}^2}{1 + \text{SNR} + 2\text{INR}}
\]

\[
N = N_d \left(\frac{\text{SNR} \cdot \text{INR}}{1 + \text{INR}^2 + \text{SNR}}\right) \text{ (equivalent of } \beta = 1 - \alpha) \implies N^2 - 1 \leq \frac{\text{SNR} \cdot \text{INR}}{1 + \text{INR}^2 + \text{SNR}}
\]

\[
N = N_d \left(\frac{\text{SNR} \cdot \text{INR}}{1 + \text{INR}^2 + \text{SNR}}\right) \text{ (equivalent of } \beta = 1 - \alpha) \implies N^2 - 1 \leq \frac{\text{SNR} \cdot \text{INR}}{1 + \text{INR}^2 + \text{SNR}}
\]

because $\text{INR} \leq \text{SNR}$, and

\[
N = N_d \left(\frac{\text{SNR} \cdot \text{INR}}{1 + \text{INR}^2 + \text{SNR}}\right) \text{ (equivalent of } \beta = 1 - \alpha) \implies N^2 - 1 \leq \frac{\text{SNR} \cdot \text{INR}}{1 + \text{INR}^2 + \text{SNR}}
\]

because $\text{SNR} \leq \text{INR}(1 + \text{INR})$. In the regime $\text{INR} \leq \text{SNR} \leq \text{INR}(1 + \text{INR})$ we also have

\[
\frac{\text{INR}^2}{(1 + \text{INR})(1 + \text{SNR}) + \text{INR}} \leq \frac{\text{INR}^2}{(1 + \text{INR})^2 + \text{INR}} \leq 1 \leq N^2 - 1, \quad \forall N \geq 2.
\]

With (35)-(37), and by recalling that $\log(x) - \frac{1}{2}\log(4) \leq \log(N_d(x)) \leq \log(x)$, $x \geq 0$, the region in (40c) can be further lower bounded as follows\(^3\)

\[
R_{\text{in}}^{\text{(G-IC-OR weak)}},
\]

\[
R_1 \leq \log(x) - \frac{1}{2}\log(4)
\]

\[
R_2 = \log \left(\frac{\text{SNR} \cdot \text{INR}}{1 + \text{SNR} + 2\text{INR}}\right) \leq \frac{1}{2}\log(4)
\]

\[
R_1 + R_2 \leq \log \left(\frac{\text{SNR} \cdot \text{INR}}{1 + \text{SNR} + 2\text{INR}}\right) - \frac{1}{2}\log(4)
\]

where

\[
x := \frac{\text{SNR} \cdot \text{INR}}{1 + \text{SNR} + 2\text{INR}} \text{ if } N \text{ as in (35), or } \frac{\text{SNR} \cdot \text{INR}}{1 + \text{INR}^2 + \text{SNR}} \text{ if } N \text{ as in (36).}
\]

In Appendix A we show that region in (38) achieves the classical G-IC outer bound to within $\frac{1}{2}\log(12\pi e) \approx 3.34$ bits (per channel use per user).

D. Noisy Interference **INR ≤ 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.

IX. Conclusion

In this paper 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

\(3\)In order to get the sum-rate, let $n = N^2 - 1 \in \mathbb{N}$ and consider either $N = N_d(a) \cdot n_d := N_d(a)^2 \cdot 1 \leq a \in \mathbb{R}^+$ or $N = N_d(b) \cdot n_p := N_d(a)^2 \cdot 1 \leq b \in \mathbb{R}^+$ in the expression $y(a) := \log(\min(a,\cdot)) + \log(\min(b,\cdot)) - \log(\min)$. That appears in the sum-rate. It follows easily that for $N = N_d(a) : y = \log(\min(a,\cdot)) \geq \log(\min(a,\cdot)) \geq \log(\min(a,\cdot)) - \frac{1}{2}\log(4)$, and for $N = N_d(b) : y = \log(\min(a,\cdot)) \geq \log(\min(a,\cdot)) \geq \log(\min(a,\cdot)) - \frac{1}{2}\log(4)$, where the term $\frac{1}{2}\log(4)$ is due to the “integer penalty”.
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 \( \frac{1}{2} \log (12 \pi e) \approx 3.34 \) 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 future 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 [2]. 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). For progress on this problem see our conference work [21], [22].

We studied the performance of mixed inputs on the Gaussian 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”. Extension to other ICs and to more than two users are subject of current investigation.

**APPENDIX A**

**PROOF OF THEOREM 10**

We proceed to evaluate the rate region in Proposition 3 with the inputs in (22a). With the chosen inputs, the outputs are

\[
\begin{align*}
Y_1 &= h_{11} \sqrt{1 - \delta_1 X_{1D}} + h_{11} \sqrt{\delta_1 X_{1G}} \\
&\quad + h_{12} \sqrt{1 - \delta_2 X_{2G}c} + h_{12} \sqrt{\delta_2 X_{2Gp}} + Z_1, \\
Y_2 &= h_{21} \sqrt{1 - \delta_1 X_{1D}} + h_{21} \sqrt{\delta_1 X_{1G}} \\
&\quad + h_{22} \sqrt{1 - \delta_2 X_{2Gc}} + h_{22} \sqrt{\delta_2 X_{2Gp}} + Z_2.
\end{align*}
\]

The achievable region in (8) with \( Q = \emptyset, U_2 = X_{2Gc} \) reduces to

\[
\begin{align*}
R_1 &\leq I(X_1; Y_1|X_{2Gc}) \\
&= h(Y_1|X_{2Gc}) - h(Y_1|X_1, X_{2Gc}) \\
&= h(h_{11} \sqrt{1 - \delta_1 X_{1D}} + h_{11} \sqrt{\delta_1 X_{1G}} \\
&\quad + h_{12} \sqrt{\delta_2 X_{2Gp}} + Z_1) \\
&\quad - h(h_{12} \sqrt{\delta_2 X_{2Gp}} + Z_1) \\
&= h \left( \frac{|h_{11}|^2 (1 - \delta_1)}{1 + |h_{11}|^2 \delta_1} X_{1D} + Z_1 \right) \\
&\quad - h(Z_1) + \frac{1}{2} \log \left( \frac{|h_{11}|^2 (1 - \delta_1)}{1 + |h_{11}|^2 \delta_1} \right).
\end{align*}
\]

For the rate of user 2 we have

\[
\begin{align*}
R_2 &\leq I(X_2; Y_2) \\
&= h(h_{21} \sqrt{1 - \delta_1 X_{1D}} + h_{21} \sqrt{\delta_1 X_{1G}} \\
&\quad + h_{22} \sqrt{1 - \delta_2 X_{2Gc}} + h_{22} \sqrt{\delta_2 X_{2Gp}} + Z_2) \\
&\quad - h(h_{21} \sqrt{\delta_1 X_{1D}} + h_{21} \sqrt{\delta_1 X_{1G}} + Z_2) \\
&= h \left( \frac{|h_{21}|^2 (1 - \delta_1)}{1 + |h_{21}|^2 \delta_1} X_{1D} + Z_2 \right) - h(Z_2) \\
&\quad + \frac{1}{2} \log \left( \frac{|h_{21}|^2 (1 - \delta_1)}{1 + |h_{21}|^2 \delta_1} \right) \\
&\quad - h \left( \frac{|h_{22}|^2 \delta_1}{1 + |h_{22}|^2} \right) + h(Z_2) \\
&\quad - \frac{1}{2} \log \left( \frac{|h_{21}|^2 (1 - \delta_1)}{1 + |h_{21}|^2 \delta_1} \right) + \frac{1}{2} \log \left( \frac{|h_{22}|^2}{1 + |h_{22}|^2} \right) + \frac{1}{2} \log \left( \frac{|h_{22}|^2 (1 - \delta_1)}{1 + |h_{22}|^2 \delta_1} \right).
\end{align*}
\]

therefore, by Theorem 8, we can further lower bound the rate of user 2 as

\[
\begin{align*}
R_2 &\leq \left( N, \frac{|h_{21}|^2 (1 - \delta_1)}{1 + |h_{21}|^2 \delta_1} \right) + \frac{1}{2} \log \left( \frac{|h_{22}|^2}{1 + |h_{22}|^2} \right) \\
&\quad + \left( \min \{ N^2 - 1, \frac{|h_{21}|^2 (1 - \delta_1)}{1 + |h_{21}|^2 \delta_1} \} \right) \\
&\quad + \left( \min \{ N^2 - 1, \frac{|h_{22}|^2 (1 - \delta_1)}{1 + |h_{22}|^2 \delta_1} \} \right)
\end{align*}
\]

thus proving (23b).

Finally, for the sum-rate we have

\[
\begin{align*}
R_1 + R_2 &\leq I(X_1, X_{2Gc}; Y_1) + I(X_2; Y_2|X_{2Gc}) \\
&= h(h_{11} \sqrt{1 - \delta_1 X_{1D}} + h_{11} \sqrt{\delta_1 X_{1G}} \\
&\quad + h_{12} \sqrt{\delta_2 X_{2Gp}} + Z_1) \\
&\quad - h(h_{12} \sqrt{\delta_2 X_{2Gp}} + Z_1) \\
&\quad + h(h_{21} \sqrt{1 - \delta_1 X_{1D}} + h_{21} \sqrt{\delta_1 X_{1G}} \\
&\quad + h_{22} \sqrt{\delta_2 X_{2Gp}} + Z_2) \\
&\quad - h(h_{21} \sqrt{\delta_1 X_{1D}} + h_{21} \sqrt{\delta_1 X_{1G}} + Z_2)
\end{align*}
\]
therefore, by Theorem 8, we can further lower bound the sum-rate as

\[
R_1 + R_2 \leq I_d(N, \frac{|h_{11}|^2(1 - \delta_1)}{1 + |h_{11}|^2\delta_1 + |h_{12}|^2})
+ I_d(h_{11} \delta_1^2 + |h_{12}|^2)
- I_d(h_{21} \delta_1^2 + |h_{22}|^2\delta_2)
+ I_d(N, \frac{|h_{21}|^2(1 - \delta_1)}{1 + |h_{21}|^2\delta_1 + |h_{22}|^2\delta_2})
+ I_d(h_{21} \delta_1^2 + |h_{22}|^2\delta_2)
- I_d \left( \min \left( N^2 - 1, \frac{|h_{21}|^2(1 - \delta_1)}{1 + |h_{21}|^2\delta_1} \right) \right)
- I_d(h_{21} \delta_1^2),
\]

thus proving (23c).

Remark 6: For future use, we specialized the derived achievable rate region for the power splits \( \delta_1 = \frac{1}{1 + |h_{21}|^2} \) and \( \delta_2 = \frac{1}{1 + |h_{22}|^2} \), inspired by [19]; we thus have that the following region is achievable for any \( N \in \mathbb{N} \):

\[
R_1 \leq I_d(N, \frac{|h_{11}|^2a}{1 + |h_{11}|^2\alpha + b}) + I_d(h_{11} \delta_1^2 + |h_{12}|^2)
\]

\[
R_2 \leq I_d(N, \frac{|h_{21}|^2a}{1 + a + |h_{21}|^2}) + I_d(h_{21} \delta_1^2 + |h_{22}|^2\delta_2)
- I_d \left( \min \left( N^2 - 1, \frac{|h_{21}|^2}{1 + a} \right) \right),
\]

\[
R_1 + R_2 \leq I_d(N, \frac{|h_{11}|^2a}{1 + |h_{11}|^2\alpha + |h_{12}|^2})
+ I_d(h_{11} \delta_1^2 + |h_{12}|^2) - I_d \left( \min \left( N^2 - 1, \frac{|h_{21}|^2a}{1 + |h_{21}|^2\alpha + |h_{12}|^2} \right) \right)
+ I_d(N, \frac{|h_{21}|^2a}{1 + a + |h_{21}|^2\alpha + |h_{22}|^2\delta_2})
+ I_d(h_{21} \delta_1^2 + |h_{22}|^2\delta_2)
- I_d \left( \min \left( N^2 - 1, \frac{|h_{21}|^2a}{1 + a + |h_{21}|^2\alpha + |h_{22}|^2\delta_2} \right) \right),
\]

where \( a := \frac{|h_{21}|^2}{1 + |h_{21}|^2} \in [0, 1] \) and \( b := \frac{|h_{21}|^2}{1 + |h_{12}|^2} \in [0, 1] \).

In the symmetric case the region in (39c) is further lower bounded by

\[
R_1 \leq I_d \left( \min \left( N^2 - 1, \frac{\text{SNR} \cdot \text{INR}}{1 + \text{SNR} + 2\text{INR}} \right) \right)
- \frac{1}{2} \log \left( \frac{\pi e}{3} \right) + I_d \left( \frac{\text{SNR}}{1 + 2\text{INR}} \right),
\]

\[
R_2 \leq I_d \left( \min \left( N^2 - 1, \frac{\text{INR}^2}{(1 + \text{INR})(1 + \text{SNR} + \text{INR})} \right) \right)
- \frac{1}{2} \log \left( \frac{\pi e}{3} \right) + I_d \left( \frac{\text{SNR}}{1 + 2\text{INR}} \right)
- \frac{1}{2} \log \left( \frac{\pi e}{3} \right) + I_d \left( \frac{\text{INR} + \text{SNR}}{1 + \text{INR}} \right).
\]

APPENDIX B

GAP DERIVATION FOR THE MODERATELY WEAK INTERFERENCE REGIME

In order to show achievability to within a constant gap of the outer bound in (28f) by means of the achievable region in (38) (a further lower bound to the region in (40c)), we distinguish two cases.

Case 1 (Regime Corresponding to \( a \in [2/3, 1] \) in Fig. 4(c)): Assume that the sum-rate in eq.(38c) is redundant; under this condition we match the corner point of the rectangular achievable region, given by \( (R_1, R_2) = (\text{eq.(38a), eq.(38b)} \), to

\[
\mathcal{R}^{(\text{G-IC mod P1})}:
\]

\[
R_1 = I_d \left( \frac{\text{INR + SNR}}{1 + \text{INR}} \right),
\]

\[
R_2 = I_d \left( \text{SNR} - I_d \left( \text{INR} + \text{INR} + \text{SNR} \right) \right)
- I_d \left( \frac{\text{INR + SNR}}{1 + \text{INR}} \right).
\]

and

\[
\mathcal{R}^{(\text{G-IC mod P2})}:
\]

\[
R_1 = I_d \left( \text{SNR} - I_d \left( \text{INR} + \text{INR} + \text{SNR} \right) \right)
- I_d \left( \frac{\text{INR + SNR}}{1 + \text{INR}} \right).
\]
which were obtained from the intersection of the sum-rate outer bound in (28c) with either (28e) or (28f). In particular, for the corner point in (41b) we use $x$ in (38d) (which corresponds to $N$ in (35)), and for the corner point in (42b) we use $x$ in (38e) (which corresponds to $N$ in (36)).

The gap is readily computed as follows: for the corner point in (41b) we have
\[
\Delta_1 = \text{eq.}(41a) - \text{eq.}(38a), \text{in (38d)}
\]
\[
\leq I_g \left( \frac{\text{SNR}}{1 + \text{INR}} \right) - I_g \left( \frac{\text{SNR}}{1 + 2 \text{INR}} \right)
- I_g \left( \frac{\text{INR}}{1 + \text{SNR} + 2 \text{INR}} \right) + \frac{1}{2} \log(4) + \frac{1}{2} \log \left( \frac{\pi e}{3} \right)
\leq \frac{1}{2} \log(2) + \frac{1}{2} \log(4) + \frac{1}{2} \log \left( \frac{\pi e}{3} \right)
= \frac{1}{2} \log \left( \frac{8 \pi e}{3} \right),
\]
and
\[
\Delta_2 = \text{eq.}(41b) - \text{eq.}(38b), \text{in (38d)}
\]
\[
\leq I_g \left( \frac{\text{SNR}}{1 + \text{INR} + \text{SNR}} \right) - I_g \left( \frac{\text{SNR}}{2} \right)
+ I_g \left( \frac{\text{INR}}{1 + \text{SNR} + 2 \text{INR}} \right)
- I_g \left( \frac{\text{INR}^2}{1 + \text{SNR}(1 + \text{INR}) + \text{INR}} \right) + \frac{1}{2} \log \left( \frac{\pi e}{3} \right)
\leq \frac{1}{2} \log(2) + \frac{1}{2} \log \left( \frac{\pi e}{3} \right)
= \frac{1}{2} \log \left( \frac{2 \pi e}{3} \right), \text{since INR} \leq \text{SNR} \text{ in weak interf.,}
\]
while for the corner point in (42b) we have
\[
\Delta_1 = \text{eq.}(42a) - \text{eq.}(38a), \text{in (38e)}
\]
\[
\leq I_g \left( \frac{\text{SNR}}{1 + \text{INR} + \text{SNR}} \right) - I_g \left( \frac{\text{SNR}}{1 + 2 \text{INR}} \right)
- I_g \left( \frac{\text{SNR}}{1 + 2 \text{INR}} \right) - I_g \left( \frac{\text{SNR} \cdot \text{INR}}{(1 + \text{INR})^2 + \text{SNR}} \right)
+ \frac{1}{2} \log(4) + \frac{1}{2} \log \left( \frac{\pi e}{3} \right)
\leq \frac{1}{2} \log(2) + \frac{1}{2} \log(4) + \frac{1}{2} \log \left( \frac{\pi e}{3} \right)
= \frac{1}{2} \log \left( \frac{8 \pi e}{3} \right),
\]
and
\[
\Delta_2 = \text{eq.(42b)} - \text{eq.(38b)}, \text{in (38e)}
\]
\[
\leq I_g \left( \frac{\text{INR} + \frac{\text{SNR}}{1 + \text{INR}}}{2} \right) - I_g \left( \frac{\text{SNR}}{2} \right)
+ I_g \left( \frac{\text{SNR} \cdot \text{INR}}{(1 + \text{INR})^2 + \text{SNR}} \right)
- I_g \left( \frac{\text{INR}^2}{(1 + \text{INR})(1 + \text{SNR} + \text{INR})} \right) + \frac{1}{2} \log \left( \frac{\pi e}{3} \right)
\leq \frac{1}{2} \log(2) + \frac{1}{2} \log \left( \frac{\pi e}{3} \right)
= \frac{1}{2} \log \left( \frac{2 \pi e}{3} \right), \text{since INR} \leq \text{SNR} \text{ in weak interf.}
\]

Case 2 [Regime Corresponding to $\alpha \in \{1/2, 2/3\}$ in Fig. 4(d)] Assume that the sum-rate in (38) is not redundant, that is after simple algebraic manipulation,
\[
1 + \min \left( x \mid x \text{ in (38d)}, x \text{ in (38e)} \right)
< \frac{(1 + 2 \text{INR})(1 + \frac{\text{SNR}}{1 + \text{INR} + \text{SNR}})}{(1 + \text{INR})(1 + \text{SNR} + \text{INR})}
\in [0.7358, 1] \text{ for INR} \leq \text{SNR} \leq \text{INR}(1 + \text{INR}) \text{ see Appendix C}
\]
\[
= \frac{(1 + \text{INR})(1 + \text{INR} + \text{SNR})}{(1 + \text{INR})^2 + \text{SNR}}
\]
which implies
\[
\left. x \right|_{x \text{ in (38d)}} \leq \left. x \right|_{x \text{ in (38e)}}.
\]

Under the condition in (43) we match one of the corner point of the pentagon-shaped achievable region in (38) to $R_{\text{out}}^{(G-IC \text{ weak P1})} :$
\[
R_1 = 3 I_g \left( \frac{\text{INR} + \frac{\text{SNR}}{1 + \text{INR}}}{2} \right) - I_g \left( \frac{\text{SNR} \cdot \text{INR}}{(1 + \text{INR})^2 + \text{SNR}} \right),
\]
and
\[
R_{\text{out}}^{(G-IC \text{ weak P2})} :$
\[
R_1 = I_g \left( \frac{\text{SNR}}{1 + \text{INR} + \text{SNR}} \right) - I_g \left( \frac{\text{SNR}}{1 + \text{INR}} \right)
- I_g \left( \frac{\text{SNR} \cdot \text{INR}}{(1 + \text{INR})^2 + \text{SNR}} \right),
\]
\[
R_2 = 3 I_g \left( \frac{\text{SNR} \cdot \text{INR}}{(1 + \text{INR})^2 + \text{SNR}} \right),
\]
which were obtained from the intersection of the sum-rate outer bound in (28d) with either (28e) or (28f). In particular, for the corner point in (44b) we use $x$ in (38d)
\[ \Delta_1 = \text{eq.(44a)} - (\text{eq.(38c)} - \text{eq.(38b)})|_{x \text{ in (38d)}} \]
\[ \leq 2 \log \left( \frac{\text{SNR}}{1 + \text{INR}} \right) - \log \left( \frac{\text{SNR} + \text{INR}}{1 + \text{INR}} \right) + \log \left( \frac{\text{INR}}{1 + \text{INR}} \right) \]
\[ + \log \left( \frac{\text{SNR}}{2} \right) - \log \left( \frac{\text{SNR} + \text{INR}}{2 + \text{SNR}} \right) - \log \left( \frac{\text{SNR}}{1 + \text{INR}} \right) - \log \left( \frac{\text{SNR}}{1 + 2\text{INR}} \right) \]
\[ \leq 2 \log \left( \frac{\text{SNR}}{2} \right) + \log \left( \frac{\text{SNR}^2}{(1 + \text{INR})(1 + \text{SNR} + \text{INR})} \right) - \log \left( \frac{\text{SNR}}{1 + \text{INR}} \right) + \log \left( \frac{\text{SNR}}{1 + \text{SNR} + \text{INR}} \right) \]
\[ + \frac{1}{2} \log (4) + \frac{1}{2} \log \left( \frac{\pi e}{3} \right) \]
\[ = \frac{1}{2} \log \left( \left( \frac{\text{SNR}}{2} + 1 \right) \left( \frac{\text{SNR} + \text{INR}}{1 + \text{SNR} + \text{INR}} \right) \right) + \frac{1}{2} \log (4) + \frac{1}{2} \log \left( \frac{\pi e}{3} \right) \]
\[ \leq \frac{1}{2} \log (6) + \frac{1}{2} \log (4) + \frac{1}{2} \log \left( \frac{\pi e}{3} \right) = \frac{1}{2} \log (8\pi e) \]  
(46)

\[ \Delta_2 = \text{eq.(44b)} - \text{eq.(38b)}|_{x \text{ in (38d)}} \]
\[ \leq \log \left( \frac{\text{SNR}}{1 + \text{INR}} \right) + \log \left( \frac{\text{SNR} + \text{INR}}{1 + \text{INR}} \right) - \log \left( \frac{\text{SNR}}{2} \right) \]
\[ + \log \left( \frac{\text{SNR}^2}{1 + \text{SNR} + \text{INR} + \text{INR}^2} \right) + \frac{1}{2} \log \left( \frac{\pi e}{3} \right) \]
\[ = \frac{1}{2} \log \left( \left( \frac{\text{SNR}}{2} + 1 \right) \left( \frac{\text{SNR} + \text{INR}}{1 + \text{SNR} + \text{INR} + \text{INR}^2} \right) \right) + \frac{1}{2} \log \left( \frac{\pi e}{3} \right) \]
\[ = \frac{1}{2} \log \left( \frac{\text{SNR} + \text{INR}}{1 + \text{SNR} + \text{INR} + \text{INR}^2} + \frac{1}{2} \log \left( \frac{\pi e}{3} \right) \right) \]
\[ \leq \frac{1}{2} \log (2) + \frac{1}{2} \log \left( \frac{\pi e}{3} \right) = \frac{1}{2} \log \left( \frac{2\pi e}{3} \right) \]
(47)

\[ \Delta_1 = \text{eq.(45b)} - \text{eq.(38a)}|_{x \text{ in (38e)}} \]
\[ \leq \log \left( \frac{\text{SNR}}{1 + \text{INR}} \right) + \log \left( \frac{\text{SNR} + \text{INR}}{1 + \text{INR}} \right) - \log \left( \frac{\text{SNR}}{2} \right) \]
\[ + \log \left( \frac{\text{SNR} \cdot \text{INR}}{(1 + \text{INR})^2 + \text{SNR}} \right) + \frac{1}{2} \log \left( \frac{\pi e}{3} \right) + \frac{1}{2} \log (4) \]
\[ = \frac{1}{2} \log \left( \frac{\text{SNR} \cdot \text{INR}}{(1 + \text{INR})^2 + \text{SNR}} \right) + \frac{1}{2} \log \left( \frac{\pi e}{3} \right) \]
\[ + \frac{1}{2} \log (4) + \frac{1}{2} \log \left( \frac{\pi e}{3} \right) \]
\[ = \frac{1}{2} \log \left( \frac{\text{SNR} + \text{INR}}{1 + \text{SNR} + \text{INR} + \text{INR}^2} + \frac{1}{2} \log \left( \frac{\pi e}{3} \right) \right) \]
\[ \leq \frac{1}{2} \log (2) + \frac{1}{2} \log (4) + \frac{1}{2} \log \left( \frac{\pi e}{3} \right) = \frac{1}{2} \log \left( \frac{4\pi e}{3} \right) \]
(48)

\[ \Delta_2 = \text{eq.(45c)} - \text{eq.(38c)} - \text{eq.(38a)}|_{x \text{ in (38e)}} \]
\[ \leq 2 \log \left( \frac{\text{SNR} + \text{INR}}{1 + \text{INR}} \right) - \log \left( \frac{\text{SNR} + \text{INR}}{1 + \text{INR}} \right) + \log \left( \frac{\text{INR}}{1 + \text{INR}} \right) + \frac{1}{2} \log \left( \frac{\pi e}{3} \right) \]
\[ = \frac{1}{2} \log \left( \frac{\text{INR}}{1 + \text{INR}} \right) + \frac{1}{2} \log \left( \frac{\pi e}{3} \right) \]
\[ \leq 0 + \frac{1}{2} \log \left( \frac{\pi e}{3} \right) = \frac{1}{2} \log \left( \frac{\pi e}{3} \right) \]
(49)
(which corresponds to \( N \) in (35)), and for the corner point in (45c) we use \( x \) in (38e) (which corresponds to \( N \) in (36)). The gap is readily computed as follows: for the corner point in (44b) see (46) and (47) while for the corner point in (45c) see (48) and (49). This concludes the proof.

**APPENDIX C**

**MINIMUM OF A FUNCTION**

The minimum of the function

\[
f(x, y) = \frac{(1 + 2y)(1 + \frac{1}{x})}{(1 + y)(1 + x) + y}, \quad \text{for } (x, y) \in \mathbb{R}^2_+,
\]

such that \( 1 \leq y \leq x \leq (1 + y), \)

is found by first taking the partial derivative with respect to \( x \), given by

\[
\frac{\partial f}{\partial x} = -\frac{2y(1 + 2y + 3)}{2(2x + y + 1)(1 + y)},
\]

which is easily seen to be monotone decreasing in \( x \) therefore attaining the minimum

\[
f(y(1 + y), y) = \frac{2y^3 + 3y^2 + 2y}{2y^3 + 5y^2 + 6y + 2}, \quad \text{for } 1 \leq y.
\]

Now by taking the partial derivative with respect to \( y \), given by

\[
\frac{\partial f}{\partial y} = \frac{3y^2 + 4y - 1}{2(y + 1)^2},
\]

and setting it equal to zero we see that the minimum occurs at \( y = \sqrt{\frac{7}{2}} \). Hence, the minimum of the function occurs at

\[
f\left(\sqrt{\frac{7}{2}}, \left(1 + \sqrt{\frac{7}{2}}\right), \frac{\sqrt{7} + 1}{3}\right) = 0.7359.
\]

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).

**REFERENCES**


**Alex Dytso** is currently a Ph.D. student at University of Illinois at Chicago. He received his bachelor’s degrees in sciences in 2011 from University of Illinois at Chicago where he received International Engineering Consortium’s William L. Everitt Student Award of Excellence for outstanding seniors who have demonstrated an interest in the communications field. His current research topic focuses on multi-user information theory and its applications to wireless networks.

**Daniela Tuninetti** received her M.S. in Telecommunication Engineering from Politecnico di Torino (Turin, Italy) in 1998, and her Ph.D. in Electrical Engineering from ENST/Telecom ParisTech (Paris, France with work done at the Eurecom Institute in Sophia Antipolis, France) in 2002. From 2002 to 2004 she was a postdoctoral research associate at the School of Communication and Computer Science at the EPFL/Swiss Federal Institute of Technology in Lausanne (Lausanne, Switzerland). Since January 2005, she is with the Department of Electrical and Computer Engineering at the University of Illinois at Chicago, Chicago, IL, USA, where she currently is an Associate Professor. Dr. Tuninetti was the editor-in-chief of the IEEE INFORMATION THEORY SOCIETY NEWSLETTER from 2006 to 2008, an associate editor for the IEEE COMMUNICATION LETTERS from 2006 to 2009, and an editor for the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS from 2009 to 2014; she currently is an associate editor for the IEEE TRANSACTIONS ON INFORMATION THEORY. She regularly serves on the Technical Program Committee of IEEE workshops and conferences, and she was the Communication Theory symposium co-chair of the 2010 IEEE International Conference on Communications (ICC 2010). Dr. Tuninetti received the best student paper award at the European Wireless Conference in 2002, and was the recipient of an NSF CAREER award in 2007. Her research interests are in the ultimate performance limits of wireless interference networks, with special emphasis on cognition and user cooperation.

**Natasha Devroye** has been an Assistant Professor in the Department of Electrical and Computer Engineering at the University of Illinois at Chicago (UIC) since January 2009. From July 2007 until July 2008 she was a Lecturer at Harvard University. Dr. Devroye obtained her Ph.D in Engineering Sciences from the School of Engineering and Applied Sciences at Harvard University in 2007, a Sc from Harvard University in 2003 and a Honors B. Eng in Electrical Engineering from McGill University in 2001. Dr. Devroye was a recipient of an NSF CAREER award in 2011, is an Editor for the IEEE TRANSACTION ON WIRELESS COMMUNICATIONS and the IEEE JOURNAL OF SELECTED AREAS OF COMMUNICATIONS: COGNITIVE RADIO SERIES and was named UIC’s Researcher of the Year in the “Rising Star” category in 2012. Her research focuses on multi-user information theory and applications to cognitive and software-defined radio, radar, and relay and two-way communication networks.