# Interference Decoding for Deterministic Channels

Bernd Bandemer, *Student Member, IEEE*, and Abbas El Gamal, *Fellow, IEEE*

*Abstract*—An inner bound to the capacity region of a class of deterministic interference channels with three user pairs is presented. The key idea is to simultaneously decode the combined interference signal and the intended message at each receiver. It is shown that this interference-decoding inner bound is tight under certain strong interference conditions. The inner bound is also shown to strictly contain the inner bound obtained by treating interference as noise, which includes interference alignment for deterministic channels. The gain comes from judicious analysis of the number of combined interference sequences in different regimes of input distributions and message rates. Finally, the inner bound is generalized to the case where each channel output is observed through a noisy channel.

*Index Terms*—Capacity region, deterministic model, interference alignment, interference channel, multiuser information theory, network information theory, simultaneous non-unique decoding.

## I. INTRODUCTION

INTERFERENCE channels with three or more user pairs exhibit the interesting property that decoding at each receiver is impaired by the *joint* effect of interference from all other senders rather than by each sender's signal separately. Consequently, dealing directly with the effect of the *combined interference signal* is expected to achieve higher rates.

One such coding scheme is *interference alignment*, e.g., [1] and [2], in which the code is designed so that the combined interference signal at each receiver is confined (*aligned*) to a subset of the receiver signal space. Depending on the specific channel, this alignment may be achieved via linear subspaces, signal scale levels, time delay slots, or number-theoretic bases of rationally independent real numbers [3], [4]. In some cases, e.g., the multiple-input multiple-output (MIMO) Gaussian interference channel [2], the decoder simply treats interference as noise. In general, however, decoding can be thought of as a two-step procedure. In the first step, the received signal is projected onto the desired signal subspace, e.g., by multiplying it by a matrix as for the MIMO case [1], [2] or by separating each received symbol into its constituent lattice points as for

the scalar Gaussian case [4]. In the second step, interference-unaware decoding is performed on the projection of the received signal. This decoding procedure often leads to an implicit decoding of some function of the undesired messages. Explicit decoding of the combined interference signals was first discussed in [5] for the many-to-one Gaussian interference channel. The authors argue that with Gaussian codes, decoding the combined interference is tantamount to decoding each interfering sender's codeword. On the other hand, with structured (lattice) codes, the combined interference can be made to appear essentially as a codeword from a single interferer.

In general, for channels with inherent linearity such as Gaussian interference channels, it is natural to consider decoding linear combinations of interfering codewords, instead of individual codewords. This idea is developed in [6] for Gaussian relay networks, leading to a *compute-forward* relaying scheme.

In this paper, we investigate interference decoding for the three receiver deterministic interference channel (3-DIC) depicted in Fig. 1. The channel consists of three sender-receiver alphabet pairs $(\mathcal{X}_k, \mathcal{Y}_k)$, loss functions $g_{lk}$, interference combining functions $h_k$, and receiver functions $f_k$ for $k, l \in \{1, 2, 3\}$. The outputs of the channel are

$$
\begin{aligned}
Y_k &= f_k(X_{kk}, S_k), \quad \text{where} \\
X_{lk} &= g_{lk}(X_l) \\
S_1 &= h_1(X_{21}, X_{31}) \\
S_2 &= h_2(X_{12}, X_{32}) \\
S_3 &= h_3(X_{13}, X_{23}).
\end{aligned}
\tag{1}
$$

We assume that $h_k$ and $f_k$ are one-to-one when either one of their arguments is fixed. For example, for $Y_1 = f_1(X_{11}, S_1)$, this assumption is equivalent to $H(X_{11}) = H(Y_1|S_1)$ and $H(S_1) = H(Y_1|X_{11})$ for every probability mass function (pmf) $p(x_{11}, s_1)$. Except for requiring the one-to-one property to hold for both arguments, this channel model is an extension of the Costa–El Gamal two-user-pair model [7]. Note that this model is more general than the class of deterministic interference channels studied in [8].

Each sender $k \in \{1, 2, 3\}$ wishes to convey an independent message $M_k$ at data rate $R_k$ to its corresponding receiver. We define a $(2^{nR_1}, 2^{nR_2}, 2^{nR_3}, n)$ code, probability of error, and achievability of a given rate triple $(R_1, R_2, R_3)$ in the standard way (see [9]).

We focus on this class of deterministic channels for several reasons. First, the capacity region for the two-user-pair version of this class [7] is known and is achieved by the Han–Kobayashi scheme [10]. This gives some hope that an appropriate extension of Han–Kobayashi where the combined interference is decoded partially or fully may be optimal for more than two user pairs. Second, this class includes the finite field deterministic model in
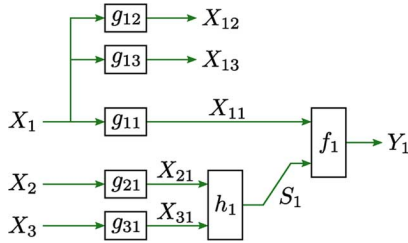
Fig. 1.   Block diagram of the 3-DIC for the first receiver.

[11], which approximates Gaussian interference channels in the high SNR regime [12]. For more than two user pairs, capacity results for the finite field deterministic model are known only in some special cases [13], [14], where interference is treated as noise. An interesting question is whether more sophisticated coding schemes can achieve higher rates for this class of channels. Finally, the combined interference signal in our channel takes values from a finite set, and therefore a certain type of alignment can be observed without resorting to complicated structured codes [15].

The main result of the paper is an inner bound on the capacity region of the 3-DIC, which is achieved via *interference decoding*. We assume point-to-point codes without rate splitting or superposition coding since such codes are widely deployed and it is interesting to investigate the benefit of using a more sophisticated receiver instead of treating interference as noise. Specifically, each receiver simultaneously decodes the intended message and the combined interference without penalizing incorrect decoding of the latter. Of course, one does not expect this scheme to be optimal in general, since even for the two-user-pair case, superposition coding is required for optimality. Note that for our class of deterministic channels, algebraic structures such as linear subspaces or lattices do not exist in general. Hence, our decoder does not use the two-step procedure as in the work on Gaussian channels and their corresponding high SNR deterministic models.

The key observation is that depending on the input pmfs and the message rates, the number of possible combined interference sequences can be equal to the number of interfering message pairs, the number of typical combined interference sequences, or some combination of the two. In our scheme, each sender does not need to know the other senders' codebooks. However, we use simultaneous decoding, which requires that the receivers know all codebooks. As in the recent characterization of the Han–Kobayashi region [16], we do not require the interference decoding to be correct with arbitrarily small probability of error.

In the following section, we summarize and discuss the main results in this paper. The proofs of these results are given in Sections III, IV and V, with some details deferred to the Appendix. In Section VI, we give final remarks on the optimality of interference decoding. Throughout the rest of the paper, notation and basic definitions follow [9].

## II. SUMMARY OF MAIN RESULTS

The main results in this paper are as follows.

### A. Interference-Decoding Inner Bound

Fix the random tuple $(Q, X_1, X_2, X_3) \sim p(q)p(x_1|q)p(x_2|q)p(x_3|q)$, where $Q$ is a time-sharing random variable from alphabet $\mathcal{Q}$. Define the region $\mathfrak{R}_1(Q, X_1, X_2, X_3)$ to consist of the rate triples $(R_1, R_2, R_3)$ such that

$$R_1 < H(X_{11}|Q) \tag{2}$$

$$R_1 + \min\{R_2, H(X_{21}|Q)\} < H(Y_1|X_{31}, Q) \tag{3}$$

$$R_1 + \min\{R_3, H(X_{31}|Q)\} < H(Y_1|X_{21}, Q) \tag{4}$$

$$\begin{aligned} R_1 + \min\{R_2 + R_3, \\ R_2 + H(X_{31}|Q), \\ H(X_{21}|Q) + R_3, \\ H(S_1|Q)\} < H(Y_1|Q). \end{aligned} \tag{5}$$

Similarly define the regions $\mathfrak{R}_2(Q, X_1, X_2, X_3)$ and $\mathfrak{R}_3(Q, X_1, X_2, X_3)$ by making the subscript replacements $1 \mapsto 2 \mapsto 3 \mapsto 1$ and $1 \mapsto 3 \mapsto 2 \mapsto 1$ in $\mathfrak{R}_1(Q, X_1, X_2, X_3)$, respectively.

*Theorem 1 (Interference-Decoding Inner Bound):* The region

$$\mathfrak{R}_{\mathrm{ID}} = \bigcup_{(Q, X_1, X_2, X_3)} \bigcap_{k=1}^{3} \mathfrak{R}_k(Q, X_1, X_2, X_3)$$

where $(Q, X_1, X_2, X_3) \sim p(q)p(x_1|q)p(x_2|q)p(x_3|q)$ and $|\mathcal{Q}| \leq 13$ is an inner bound to the capacity region of the 3-DIC.

The proof for this theorem is given in Section III.

Region $\mathfrak{R}_k(Q, X_1, X_2, X_3)$ ensures decodability at receiver $k$. The min terms on the left-hand side of the inequalities arise from counting the number of possible interfering sequences at various links of the channel. For example, consider the $\min\{R_2, H(X_{21}|Q)\}$ term in (3). If $R_2$ is small, the number of distinct sequences that can occur at $X_{21}$ is equal to the number of possible messages from sender 2. As $R_2$ increases beyond $H(X_{21}|Q)$, the number of possible sequences at $X_{21}$ "saturates" to the number of typical sequences, which is roughly $2^{nH(X_{21}|Q)}$. In this case, we can increase the rate of the second sender further without negatively impacting the first receiver. The min expressions in (4) and (5) likewise capture the saturation effects at $X_{31}$ and $S_1$, respectively.

An example of region $\mathfrak{R}_1(Q, X_1, X_2, X_3)$ is plotted in Fig. 2. The region is unbounded in the $R_2$ and $R_3$ directions, due to saturation. This is expected, since regardless of the values of $R_2$ and $R_3$, $S_1$ can always be treated as noise to achieve a nonzero rate. However, as $R_2$ and $R_3$ become smaller, the proposed scheme takes advantage of the structure in $S_1$ and can thereby increase $R_1$.

### B. Capacity Region Under Strong Interference

Consider the subclass of 3-DIC *with strong interference and invertible $h_k$* in which the following two conditions hold.

First, the loss functions $g_{lk}$ are such that

$$\min\{H(X_{12}), H(X_{13})\} \geq H(X_{11})$$
$$\min\{H(X_{21}), H(X_{23})\} \geq H(X_{22})$$
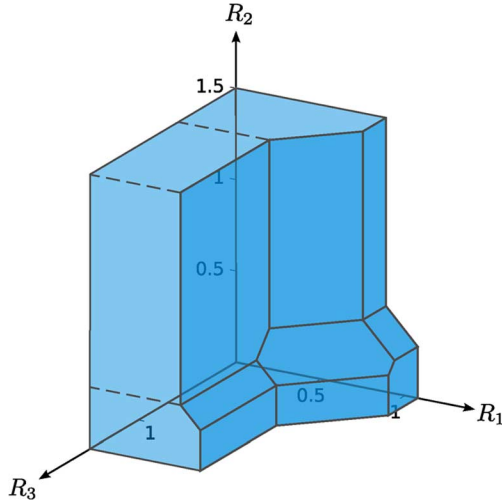$$\min\{H(X_{31}), H(X_{32})\} \geq H(X_{33})$$

Fig. 2. Region $\mathfrak{R}_1$ showing allowable rate triples for decodability at the first receiver.

for all product input pmfs $p(x_1)p(x_2)p(x_3)$. This condition implies that interference is strong.

Second, the functions $h_k$ are invertible, i.e.,

$$H(S_1) = H(X_{21}) + H(X_{31})$$
$$H(S_2) = H(X_{12}) + H(X_{32})$$
$$H(S_3) = H(X_{13}) + H(X_{23})$$

for all product input pmfs $p(x_1)p(x_2)p(x_3)$. With the conditional invertibility property of $f_k$, the channel becomes a nonsymmetric version of the deterministic model for the SIMO interference channel described in [8]. In both cases, a receiver can uniquely recover *both* interfering signals given the received sequence and the desired transmitted sequence. The capacity region under these conditions is achieved by interference decoding.

*Theorem 2:* The capacity region of the 3-DIC under strong interference and invertible $h_k$ functions is the set of rate triples $(R_1, R_2, R_3)$ such that

$$R_k < H(X_{kk}|Q), \quad k \in \{1, 2, 3\}$$
$$R_1 + R_2 < \min\{H(Y_1|X_{31}, Q), H(Y_2|X_{32}, Q)\}$$
$$R_1 + R_3 < \min\{H(Y_1|X_{21}, Q), H(Y_3|X_{23}, Q)\}$$
$$R_2 + R_3 < \min\{H(Y_2|X_{12}, Q), H(Y_3|X_{13}, Q)\}$$
$$R_1 + R_2 + R_3 < \min\{H(Y_1|Q), H(Y_2|Q), H(Y_3|Q)\}$$

for some $(Q, X_1, X_2, X_3) \sim p(q)p(x_1|q)p(x_2|q)p(x_3|q)$. The proof of this theorem is given in Section IV.

### C. Treating Interference as Noise

In the two-user-pair interference channel, decoding both messages at each receiver and treating interference as noise are considered as two extreme schemes. The extremes are bridged by the Han–Kobayashi scheme in which part of the interference is decoded and the rest is treated as noise [9]. While treating interference as noise is better for channels with weak interference, decoding both messages is optimal under strong interference.

We show surprisingly that for the 3-DIC under consideration, treating interference as noise is a *special case* of interference decoding!

By using randomly and independently generated codebooks as for the interference-decoding inner bound, but having each receiver decode only its message, we obtain the following inner bound.

*Lemma 1 (Treating Interference as Noise):* The set $\mathfrak{R}_{\text{TIN}}$ of rate triples $(R_1, R_2, R_3)$ such that

$$R_k < I(X_k; Y_k|Q), \quad k \in \{1, 2, 3\} \tag{6}$$

for some pmf $p(q)p(x_1|q)p(x_2|q)p(x_3|q)$ constitutes an inner bound to the capacity region of the general three-user-pair memoryless interference channel.

Note that in contrast to interference decoding, a user pair does not need to know the codebooks of other user pairs. Also note that this inner bound, with appropriate selections of the input pmfs, includes the interference alignment inner bounds in [2], [13], [14]. Maximum alignment is achieved when the number of combined interference sequences, e.g., $S_1^n$, is much smaller than the number of individual interference sequence pairs, e.g., $(X_{21}^n, X_{31}^n)$. Since $I(X_k; Y_k|Q) = H(Y_k|Q) - H(S_k|Q)$, this occurs when $H(S_k|Q)$ is small, causing the number of $S_k^n$ sequences to saturate.

In Section V, we establish the following result.

*Theorem 3:* The rate region achievable by treating interference as noise is included in the interference-decoding rate region, i.e., $\mathfrak{R}_{\text{TIN}} \subseteq \mathfrak{R}_{\text{ID}}$.

The difference between treating interference as noise and interference decoding is essentially that the former assumes that the combined interference signal $S_k$ is always saturated, while the latter distinguishes between saturated and non-saturated cases. Later in this section, we argue that the above inclusion result is tightly coupled to the definition of the 3-DIC.

The following example shows that the inclusion of Theorem 3 can be strict, i.e., the treating interference as noise region is strictly contained in the interference-decoding region.

*Additive 3-DIC Example:* Consider a cyclically symmetric 3-DIC with $\mathcal{X}_1 = \mathcal{X}_2 = \mathcal{X}_3 = \{0, 1, 2\}$, and $\mathcal{Y}_1 = \mathcal{Y}_2 = \mathcal{Y}_3 = \{0, 1, 2, 3, 4\}$, where $g_{11} = g_{22} = g_{33} = g$, $g_{12} = g_{23} = g_{31} = g_+$ and $g_{21} = g_{32} = g_{13} = g_-$ as well as $h_1 = h_2 = h_3 = h$ and $f_1 = f_2 = f_3 = f$. The direct path loss functions are the identity mapping, $g = \text{Id}$, while the cross path loss functions are given by $g_- = \{0 \mapsto 0, 1 \mapsto 1, 2 \mapsto 0\}$ and $g_+ = \{0 \mapsto 0, 1 \mapsto 1, 2 \mapsto 1\}$ (similar to the Blackwell broadcast channel [17]). Finally, the combining functions $h$ and receiver functions $f$ are taken to be addition. The resulting input-to-output mapping is shown in Fig. 3.

For this channel, the interference-decoding rate region strictly contains the region achievable by treating interference as noise. To demonstrate this, we computed the approximations of the inner bounds shown in Fig. 4(a) and (b). Since it is computationally infeasible to enumerate the 13 different conditional distributions of inputs given $Q$ as required by Theorem 1, we used the following procedure. We first assume $Q = \emptyset$ and consider a grid over all input distributions $p(x_1)p(x_2)p(x_3)$. For each grid
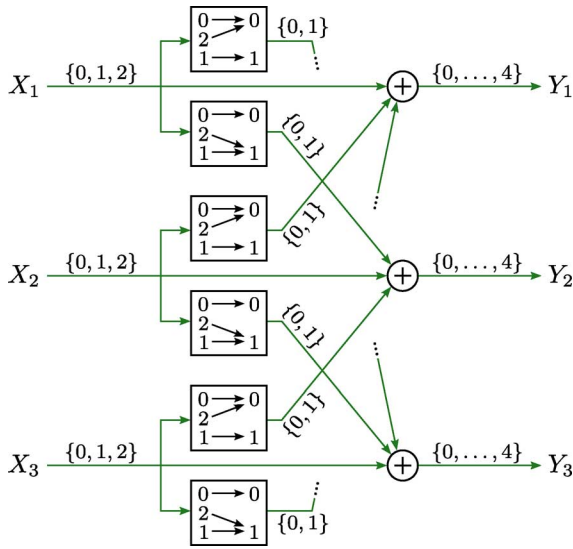
Fig. 3. Additive 3-DIC example.



(a)

(b)

(c)

Fig. 4. Inner bounds for the additive 3-DIC example. (a) Treating interference as noise. (b) Interference decoding. (c) Intersection with 45°-plane.
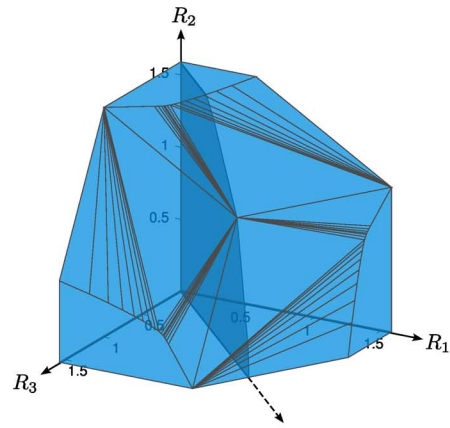
point, we compute the achievable rate regions as given by Theorem 1 and Lemma 1, respectively. We represent the regions as the convex hull of its corner points. The final approximation is obtained by taking the union of all such corner points over the grid. Due to the simple structure of $\mathfrak{R}_{\text{TIN}}$ in Lemma 1, which consists of a union of rectangular boxes, this method can compute $\mathfrak{R}_{\text{TIN}}$ to arbitrary precision provided the grid is sufficiently fine. On the other hand, our approximation method yields a possibly strictly smaller inner bound than $\mathfrak{R}_{\text{ID}}$.

Fig. 4(c) depicts the intersection of the three-dimensional regions with the plane defined by the $R_2$ axis and the 45°-line between the $R_1$ and $R_3$ axes. This plane is also shown in Fig. 4(a). Note that the same maximum sum rate $R_{\text{sum}} = 3$ is achieved by both schemes. However, while treating interference as noise does so at exactly one rate triple $(R_1 = R_2 = R_3 = 1)$, interference decoding achieves the maximal sum rate at many different asymmetric rate triples.
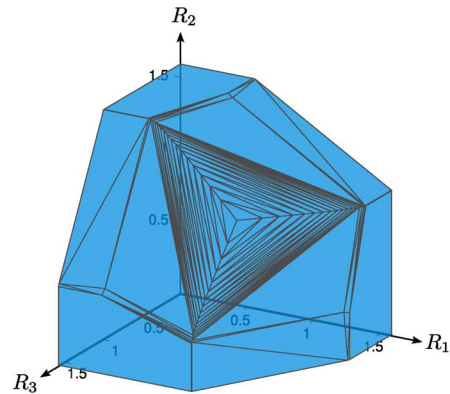
*Remark 1:* Treating interference as noise can be optimal in some cases. Consider the three-user-pair cyclically symmetric finite field deterministic model investigated in [14], which is a special case of the channel considered in this paper. The input and output alphabets for this channel are $\mathbb{F}_2^N$ and $\mathbb{F}_2^{2N}$, respectively, the loss functions $g_{lk}$ are vector shifting operations, where the amount of shift is parameterized by $(\alpha, \beta) \in [1, 2] \times [0, 1]$, and the interference combining functions $h_k$ and the receiver functions $f_k$ are componentwise additions over $\mathbb{F}_2$.

The sum capacity of this channel is computed in [14] for a large range of $(\alpha, \beta)$, and achievability is established by constructing linear encoding and decoding matrices for every $(\alpha, \beta)$. This scheme can be interpreted as treating interference as noise, and thus Lemma 1 subsumes the achievability results in [14]. In fact, the necessary input distributions are the ones implicitly stated there.
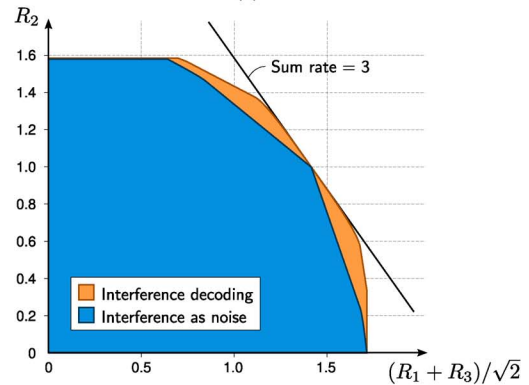
It would be interesting to investigate whether interference decoding can achieve higher sum rates than treating interference as noise in the $(\alpha, \beta)$ range where the sum capacity is not known.

Moreover, even in the range where we know the sum capacity, interference decoding may achieve higher asymmetric rates than treating interference as noise, as in the additive 3-DIC example. The main challenge in settling these questions is the prohibitively large space of possible input distributions in Theorem 1.

### D. Extension to 3-DIC With Noisy Observations

Finally, we consider the 3-DIC *with noisy observations*. In this generalization of 3-DIC, the channel outputs in (1) are observed through memoryless channels $Y_k \rightarrow Z_k$ for $k \in \{1, 2, 3\}$. Thus receiver $k$ now observes a noisy version $Z_k$ of $Y_k$, which may be from a discrete or a continuous alphabet.

The interference-decoding inner bound generalizes to the 3-DIC with noisy observations as follows. Let $(Q, X_1, X_2, X_3) \sim p(q)p(x_1|q)p(x_2|q)p(x_3|q)$. Define the region $\mathfrak{R}'_1(Q, X_1, X_2, X_3)$ as the set of rate triples $(R_1, R_2, R_3)$ such that

$$R_1 < I(X_1; Z_1|S_1, Q)$$
$$R_1 + \min\{R_2, H(X_{21}|Q)\} < I(X_1, X_{21}; Z_1|X_{31}, Q)$$
$$R_1 + \min\{R_3, H(X_{31}|Q)\} < I(X_1, X_{31}; Z_1|X_{21}, Q)$$
$$R_1 + \min\{R_2 + R_3,$$
$$R_2 + H(X_{31}|Q),$$
$$H(X_{21}|Q) + R_3,$$
$$H(S_1|Q)\} < I(X_1, S_1; Z_1|Q).$$

Similarly, define the regions $\mathfrak{R}'_2(Q, X_1, X_2, X_3)$ and $\mathfrak{R}'_3(Q, X_1, X_2, X_3)$ by making the subscript replacements $1 \mapsto 2 \mapsto 3 \mapsto 1$ and $1 \mapsto 3 \mapsto 2 \mapsto 1$ in $\mathfrak{R}'_1(Q, X_1, X_2, X_3)$, respectively.

*Theorem 4:* The region

$$\mathfrak{R}'_{\text{ID}} = \bigcup_{(Q,X_1,X_2,X_3)} \bigcap_{k=1}^{3} \mathfrak{R}'_k(Q, X_1, X_2, X_3)$$

where $(Q, X_1, X_2, X_3) \sim p(q)p(x_1|q)p(x_2|q)p(x_3|q)$ is an inner bound to the capacity region of the 3-DIC with noisy observations.

The proof of this theorem proceeds completely analogously to the proof of Theorem 1 as presented in Section III, and thus its details are omitted.

The following example demonstrates the inner bound for the 3-DIC with noisy observations. It also shows that the inclusion of Theorem 3 does not hold in general for this channel model.

*Gaussian Interference Channel Example:* Consider the Gaussian interference channel with finite input alphabets. The channel output at receiver $k$ is

$$Y_k = \sum_{l=1}^{3} g_{lk} X_l$$
$$Z_k = Y_k + N_k \qquad (7)$$

where $g_{lk} \in \mathbb{R}$ is the path gain from transmitter $l$ to receiver $k$, and $N_k$ is additive white Gaussian noise of average power $\sigma^2$. This is a realistic model for a wireless interference channel where the transmitter hardware is based on digital signal processing (DSP) and digital-to-analog conversion (DAC). For example, $\mathcal{X}_l = \{+1, -1\}$ represents a system with a binary constellation, e.g., binary phase-shift keying (BPSK). Equation (7) represents continuous-valued outputs (soft outputs), but our model would also apply if a quantizer is added (hard outputs), for example due to analog-digital conversion (ADC) at the receivers.

Fig. 5 shows approximations of the inner bounds for a cyclically symmetric Gaussian interference channel with BPSK inputs and continuous outputs. In contrast to the noiseless case, neither the interference-decoding region nor the region achieved by treating interference as noise contains the other, i.e., Theorem
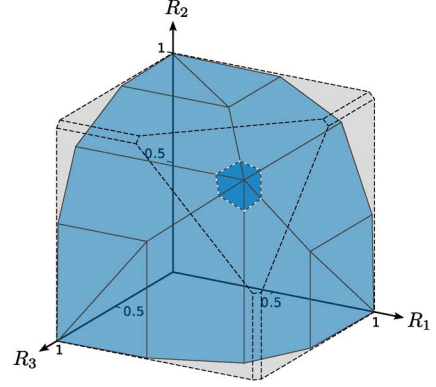


Fig. 5. Rate regions achieved by interference decoding (dashed outline) and treating interference as noise (shaded) for a cyclically symmetric Gaussian interference channel with $X_k \in \{+1, -1\}$, path gains $g_{11} = 1.8, g_{21} = 1.0, g_{31} = 1.1$, and noise power $\sigma^2 = 0.1$.

3 does not hold for 3-DIC with noisy observations. In particular, the sum rates achieved by treating interference as noise and interference decoding are 2.51 and 2.37, respectively. Intuitively, interference decoding attempts to separate the combined interference from the additive noise. As such, it may achieve lower rates than simply treating interference as noise for which this separation is not enforced. This discrepancy is more pronounced for low values of SNR, and it vanishes asymptotically as SNR grows.

## III. PROOF OF INTERFERENCE-DECODING INNER BOUND

We first present two key lemmas which formalize the notion of link saturation as discussed after Theorem 1. The proofs are deferred to the Appendix. The first lemma generalizes the packing lemma stated in [9].

*Lemma 2 (Packing Lemma for Pairs):* Let $(U, A, B, C) \sim p(u)p(a|u)p(b|u)p(c|a, b, u)$. Let $U^n \sim \prod_{i=1}^{n} p_U(u_i)$. For each $m \in [1 : 2^{nR_A}]$, let $A^n(m) \sim \prod_{i=1}^{n} p_{A|U}(a_i|u_i)$. For each $l \in [1 : 2^{nR_B}]$, let $B^n(l) \sim \prod_{i=1}^{n} p_{B|U}(b_i|u_i)$, conditionally independent of each $A^n(m)$ given $U^n$. Let $C^n \sim \prod_{i=1}^{n} p_{C|U}(c_i|u_i)$, conditionally independent of each $A^n(m)$ and $B^n(l)$ given $U^n$. There exists a $\delta(\varepsilon)$ with $\lim_{\varepsilon \to 0} \delta(\varepsilon) = 0$ such that if

$$\min\{R_A, H(A|U)\}$$
$$+ \min\{R_B, H(B|U)\} < I(A, B; C|U) - \delta(\varepsilon)$$

then $P\{(U^n, A^n(m), B^n(l), C^n) \in \mathcal{T}_\varepsilon^{(n)} \text{ for some } m, l\} \to 0$ as $n \to \infty$, where typicality, entropies and mutual information are with respect to $p(u, a, b, c)$.

The following lemma is a refined version of Lemma 2, where the sequences $B^n(l)$ are generated from two conditionally independent components $B_1^n(l_1)$ and $B_2^n(l_2)$.

*Lemma 3:* Let $(U, A, B_1, B_2, B, C) \sim p(u)p(a|u)p(b_1|u)p(b_2|u)p(b|b_1, b_2)p(c|a, b, u)$, where $p(b|b_1, b_2)$ corresponds to a deterministic mapping $h : (b_1, b_2) \mapsto b$. Let $U^n \sim \prod_{i=1}^{n} p_U(u_i)$. For each $m \in [1 : 2^{nR_A}]$, let $A^n(m) \sim \prod_{i=1}^{n} p_{A|U}(a_i|u_i)$. For each $l_1 \in [1 : 2^{nR_{B_1}}]$, let $B_1^n(l_1) \sim \prod_{i=1}^{n} p_{B_1|U}(b_{1i}|u_i)$, conditionally independent of each $A^n(m)$ given $U^n$. Likewise, for each $l_2 \in [1 : 2^{nR_{B_2}}]$,
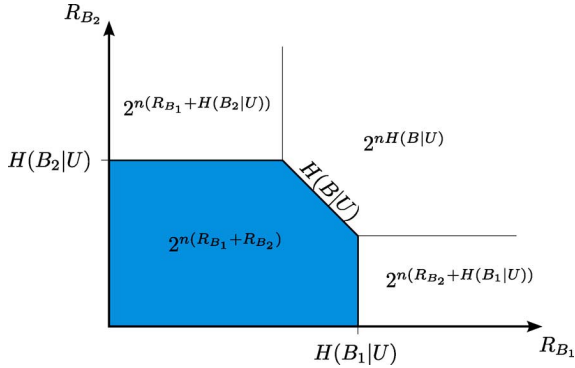
Fig. 6. Capacity region for a deterministic MAC. The number of output sequences as a function of the number of input sequences is stated in each region.

let $B_2^n(l_2) \sim \prod_{i=1}^{n} p_{B_2|U}(b_{2i}|u_i)$, conditionally independent of each $A^n(m)$ and $B_1^n(l_1)$ given $U^n$. For each $(l_1, l_2)$, let $B_i(l_1, l_2) = h(B_{1i}(l_1), B_{2i}(l_2))$ for $i \in [1:n]$. Finally, let $C^n \sim \prod_{i=1}^{n} p_{C|U}(c_i|u_i)$, conditionally independent of each $A^n(m)$, $B_1^n(l_1)$, and $B_2^n(l_2)$ given $U^n$.

There exists a function $\delta(\varepsilon)$ with $\lim_{\varepsilon \to 0} \delta(\varepsilon) = 0$ such that if

$$
\begin{aligned}
R_A + \min\{ & R_{B_1} + R_{B_2}, \\
& R_{B_1} + H(B_2|U), \\
& H(B_1|U) + R_{B_2}, \\
& H(B|U)\} < I(A, B; C|U) - \delta(\varepsilon)
\end{aligned}
$$

then $\mathrm{P}\{(U^n, A^n(m), B^n(l_1, l_2), C^n) \in \mathcal{T}_\varepsilon^{(n)}$ for some $m, l_1, l_2\} \to 0$ as $n \to \infty$, where typicality, entropies and mutual information are with respect to $p(u, a, b_1, b_2, b, c)$.

*Remark 2:* The intuition is that $B$ can be interpreted as the output of a deterministic multiple access channel with inputs $B_1$ and $B_2$ and input to output mapping $h$. Fig. 6 shows the number of output sequences for different ranges of $R_{B_1}$ and $R_{B_2}$ when $h$ is one-to-one in each argument. Note that when $(R_{B_1}, R_{B_2})$ is in the deterministic MAC capacity region, the number of output sequences is simply $2^{n(R_{B_1} + R_{B_2})}$. For $(R_{B_1}, R_{B_2})$ outside the capacity region, the number of output sequences saturates in one or both dimensions. The logarithm of the number of output sequences divided by $n$ appears in the $\min$ expression of the lemma.

We are now ready to prove Theorem 1. We begin by fixing a pmf $p(q)p(x_1|q)p(x_2|q)p(x_3|q)$.

*Codebook Generation:* Randomly generate a sequence $q^n$ according to $\prod_{i=1}^{n} p_Q(q_i)$. For each $k \in \{1, 2, 3\}$, randomly and conditionally independently generate sequences $x_k^n(m_k), m_k \in [1:2^{nR_k}]$, each according to $\prod_{i=1}^{n} p_{X_k|q}(x_{ki}|q_i)$. From the channel definition, this procedure induces intermediate sequences $x_{kl}^n(m_k)$ for $l \in \{1, 2, 3\}$, combined interference sequences $s_1^n(m_2, m_3), s_2^n(m_1, m_3), s_3^n(m_1, m_2)$, and output sequences $y_k^n(m_1, m_2, m_3)$.

*Encoding:* To send the message $m_k \in [1:2^{nR_k}], k \in \{1, 2, 3\}$, encoder $k$ transmits $x_k^n(m_k)$.

*Decoding:* Upon observing $y_1^n$, decoder 1 declares that $\hat{m}_1$ is sent if it is the unique message such that $(q^n, x_1^n(\hat{m}_1), s_1^n(\mathring{m}_2, \mathring{m}_3), x_{21}^n(\mathring{m}_2), x_{31}^n(\mathring{m}_3), y_1^n) \in \mathcal{T}_\varepsilon^{(n)}$ for some $\mathring{m}_2, \mathring{m}_3$, where $\mathcal{T}_\varepsilon^{(n)}$ is defined as in [9]. Decoding at the other receivers is performed similarly.

*Analysis of the Probability of Error:* Without loss of generality, assume that $m_k = 1$ for $k \in \{1, 2, 3\}$. Define $\mathcal{E}_{mlk} = (Q^n, X_1^n(m), S_1^n(l, k), X_{21}^n(l), X_{31}^n(k), Y_1^n(1, 1, 1)) \in \mathcal{T}_\varepsilon^{(n)}$, and the events

$$
\begin{aligned}
\mathcal{E}_0 &= \mathcal{E}_{111}^c \\
\mathcal{E}_1 &= \{\mathcal{E}_{m11} \text{ for some } m \neq 1\}, \\
\mathcal{E}_2 &= \{\mathcal{E}_{ml1} \text{ for some } m, l \neq 1\} \\
\mathcal{E}_3 &= \{\mathcal{E}_{m1k} \text{ for some } m, k \neq 1\}, \\
\mathcal{E}_4 &= \{\mathcal{E}_{mlk} \text{ for some } m, l, k \neq 1\}.
\end{aligned}
$$

Then the probability of decoding error at the first receiver averaged over codebooks is upper bounded as $\mathrm{P}(\mathcal{E}) = \mathrm{P}(\mathcal{E}_0 \cup \mathcal{E}_1 \cup \mathcal{E}_2 \cup \mathcal{E}_3 \cup \mathcal{E}_4) \leq \sum_{j=0}^{4} \mathrm{P}(\mathcal{E}_j)$. We bound each term. First, by the law of large numbers, $\mathrm{P}(\mathcal{E}_0) \to 0$ as $n \to \infty$.

Next consider

$$
\mathcal{E}_1 \subseteq \Big\{ (Q^n, X_1^n(m), S_1^n(1, 1), Y_1^n(1, 1, 1)) \in \mathcal{T}_\varepsilon^{(n)} \\
\text{for some } m \neq 1 \Big\}.
$$

By Lemma 2 with $U^n = (Q^n, S_1^n(1, 1))$, $A^n = X_1^n$, $B^n = \emptyset$, and $C^n = Y_1^n(1, 1, 1)$, the probability of this event tends to zero as $n \to \infty$ if

$$
R_1 < I(X_1; Y_1|S_1, Q)
$$

which simplifies to

$$
R_1 < H(X_{11}|Q). \tag{8}
$$

The event $\mathcal{E}_2$ can be treated similarly. Consider

$$
\mathcal{E}_2 \subseteq \Big\{ (Q^n, X_1^n(m), X_{21}^n(l), X_{31}^n(1), Y_1^n(1, 1, 1)) \in \mathcal{T}_\varepsilon^{(n)} \\
\text{for some } m, l \neq 1 \Big\}.
$$

Using Lemma 2 with $U^n = (Q^n, X_{31}^n(1))$, $A^n = X_1^n$, $B^n = X_{21}^n$, and $C^n = Y_1^n(1, 1, 1)$, we conclude that $\mathrm{P}(\mathcal{E}_2) \to 0$ if

$$
R_1 + \min\{R_2, H(X_{21}|X_{31}, Q)\} < I(X_1, X_{21}; Y_1|X_{31}, Q)
$$

or, equivalently

$$
R_1 + \min\{R_2, H(X_{21}|Q)\} < H(Y_1|X_{31}, Q). \tag{9}
$$

$\mathrm{P}(\mathcal{E}_3)$ can be analyzed in an identical fashion to $\mathrm{P}(\mathcal{E}_2)$, and we have $\mathrm{P}(\mathcal{E}_3) \to 0$ as $n \to \infty$ if

$$
R_1 + \min\{R_3, H(X_{31}|Q)\} < H(Y_1|X_{21}, Q). \tag{10}
$$

Finally, the event $\mathcal{E}_4$ is augmented as

$$
\mathcal{E}_4 \subseteq \Big\{ (Q^n, X_1^n(m), S_1^n(l, k), Y_1^n(1, 1, 1)) \in \mathcal{T}_\varepsilon^{(n)} \\
\text{for some } m, l, k \neq 1 \Big\}.
$$

Lemma 3 with $U^n = Q^n, A^n = X_1^n, B_1^n = X_{21}^n, B_2^n = X_{31}^n, B^n = S_1^n, h = h_1$, and $C^n = Y_1^n(1,1,1)$ shows that $P(\mathcal{E}_4) \to 0$ as $n \to \infty$ if

$$R_1 + \min\{R_2 + R_3, R_2 + H(X_{31}|Q),$$
$$H(X_{21}|Q) + R_3, H(S_1|Q)\} < H(Y_1|Q) \quad (11)$$

where we have used $I(X_1, S_1; Y_1|Q) = H(Y_1|Q)$. Collecting (8) to (11) yields the conditions of $\mathfrak{R}_1$. The probability of error at the second and third receiver can be bounded similarly, leading to the conditions of $\mathfrak{R}_2$ and $\mathfrak{R}_3$. Finally, the cardinality bound on $\mathcal{Q}$ can be established using the bounding technique described in [9]. ∎

## IV. PROOF OF THEOREM 2

*Proof of Achievability:* We prove achievability with interference decoding. Specifically, we show that under strong interference and invertible $h_k$, regions $\mathfrak{R}_k$ of Theorem 1 simplify to regions $\mathfrak{R}_k''$ below while maintaining $\mathfrak{R}_1 \cap \mathfrak{R}_2 \cap \mathfrak{R}_3 = \mathfrak{R}_1'' \cap \mathfrak{R}_2'' \cap \mathfrak{R}_3''$.

Recall the definition of $\mathfrak{R}_1(Q, X_1, X_2, X_3)$ as the set of rate triples $(R_1, R_2, R_3)$ that satisfy (2) to (5). Further recall the analogous definitions of $\mathfrak{R}_2$ and $\mathfrak{R}_3$, which include

$$R_2 < H(X_{22}|Q) \quad (12)$$
$$R_3 < H(X_{33}|Q). \quad (13)$$

When combined with the strong interference assumption, (12) and (13) imply that the $\min$ expressions in (3) and (4) simplify to $R_2$ and $R_3$, respectively. Furthermore, the sum of (12) and (13) implies that

$$R_2 + R_3 < H(X_{22}|Q) + H(X_{33}|Q)$$
$$\leq H(X_{21}|Q) + H(X_{31}|Q)$$
$$= H(S_1|Q)$$

where we have used the invertibility of $h_1$. Therefore, the $\min$ expression in (5) simplifies to $R_2 + R_3$.

Consequently, define $\mathfrak{R}_1''(Q, X_1, X_2, X_3)$ as the set of rate triples $(R_1, R_2, R_3)$ such that

$$R_1 < H(X_{11}|Q)$$
$$R_1 + R_2 < H(Y_1|X_{31}, Q)$$
$$R_1 + R_3 < H(Y_1|X_{21}, Q)$$
$$R_1 + R_2 + R_3 < H(Y_1|Q).$$

Likewise, define the regions $\mathfrak{R}_2''(Q, X_1, X_2, X_3)$ and $\mathfrak{R}_3''(Q, X_1, X_2, X_3)$ by replacing subscripts following $1 \mapsto 2 \mapsto 3 \mapsto 1$ and $1 \mapsto 3 \mapsto 2 \mapsto 1$ in $\mathfrak{R}_1''(Q, X_1, X_2, X_3)$, respectively. Then Theorem 1 implies that

$$\bigcup_{(Q, X_1, X_2, X_3)} \bigcap_{k=1}^{3} \mathfrak{R}_k''(Q, X_1, X_2, X_3)$$

is achievable, and the proposition follows by expanding the intersection operation. ∎

*Proof of Converse:* Consider a code with rates $(R_1, R_2, R_3)$, empirical pmf $p(x_1^n)p(x_2^n)p(x_3^n)$, and $P_e^{(n)}$ tending to 0 as $n \to \infty$. First, note that

$$nR_1 \leq I(X_1^n; Y_1^n) + n\varepsilon_n$$
$$= I(X_{11}^n; Y_1^n) + n\varepsilon_n$$
$$\leq H(X_{11}^n) + n\varepsilon_n$$
$$= nH(X_{11}|Q) + n\varepsilon_n$$

where $Q$ is a time-sharing random variable uniformly distributed over $[1 : n]$. Next, consider

$$n(R_1 + R_2)$$
$$\leq I(X_1^n; Y_1^n) + I(X_2^n; Y_2^n) + n\varepsilon_n$$
$$= H(Y_1^n) - H(Y_1^n|X_1^n) + H(Y_2^n)$$
$$\quad - H(Y_2^n|X_2^n) + n\varepsilon_n$$
$$= H(Y_1^n) - H(S_1^n) + H(Y_2^n) - H(S_2^n) + n\varepsilon_n$$
$$= H(Y_1^n) - H(X_{31}^n)$$
$$\quad + (H(Y_2^n) - H(X_{21}^n) - H(X_{12}^n) - H(X_{32}^n)) + n\varepsilon_n$$
$$\leq H(Y_1^n|X_{31}^n) + n\varepsilon_n$$
$$\leq nH(Y_1|X_{31}, Q) + n\varepsilon_n$$

where we have used $H(X_{22}^n) \leq H(X_{21}^n)$ and $H(Y_2^n) \leq H(X_{22}^n) + H(X_{12}^n) + H(X_{32}^n)$. In the same way, it can be shown that

$$n(R_1 + R_3) \leq nH(Y_1|X_{21}, Q) + n\varepsilon_n.$$

Finally

$$n(R_1 + R_2 + R_3)$$
$$\leq H(Y_1^n) - H(S_1^n) + H(Y_2^n) - H(S_2^n)$$
$$\quad + H(Y_3^n) - H(S_3^n) + n\varepsilon_n$$
$$= H(Y_1^n) + n\varepsilon_n$$
$$\quad + (H(Y_2^n) - H(X_{21}^n) - H(X_{12}^n) - H(X_{32}^n))$$
$$\quad + (H(Y_3^n) - H(X_{31}^n) - H(X_{13}^n) - H(X_{23}^n))$$
$$\leq nH(Y_1|Q) + n\varepsilon_n.$$

Thus, all four conditions related to the first receiver have been shown. Analogous steps yield the remaining bounds. ∎

## V. PROOF OF THEOREM 3

We show that the inner bound in Lemma 1 is included in the inner bound of Theorem 1. The conditions of region $\mathfrak{R}_1$ in Theorem 1 can be made more stringent by replacing the $\min$ expression with any one of its argument terms. For example, $(R_1, R_2, R_3) \in \mathfrak{R}_1$ is implied by

$$R_1 < H(X_{11}|Q)$$
$$R_1 + H(X_{21}|Q) < H(Y_1|X_{31}, Q)$$
$$R_1 + H(X_{31}|Q) < H(Y_1|X_{21}, Q)$$
$$R_1 + H(S_1|Q) < H(Y_1|Q)$$

or, equivalently

$$
\begin{aligned}
R_1 < \min\{ & H(X_{11}|Q), \\
& H(Y_1|X_{31}, Q) - H(X_{21}|Q), \\
& H(Y_1|X_{21}, Q) - H(X_{31}|Q), \\
& H(Y_1|Q) - H(S_1|Q)\}.
\end{aligned} \tag{14}
$$

To simplify this expression, consider

$$
\begin{aligned}
H(X_{11}|Q) &\geq I(X_{11}; Y_1|Q) \\
&= H(Y_1|Q) - H(Y_1|X_{11}, Q) \\
&= H(Y_1|Q) - H(S_1|Q)
\end{aligned}
$$

as well as

$$
\begin{aligned}
[H(Y_1 & |X_{21}, Q) - H(X_{31}|Q)] - [H(Y_1|Q) - H(S_1|Q)] \\
&= H(Y_1, X_{21}|Q) - H(X_{21}|Q) \\
&\quad - \underbrace{H(X_{31}|Q)}_{H(S_1|X_{21}, Q)} - H(Y_1|Q) + H(S_1|Q) \\
&= H(X_{21}|Y_1, Q) - H(X_{21}|S_1, Q) \\
&\geq H(X_{21}|Y_1, S_1, Q) - H(X_{21}|S_1, Q) \\
&= 0
\end{aligned}
$$

and, by symmetry

$$
\begin{aligned}
[H(Y_1|X_{31}, Q) - H(X_{21}|q)] & \\
& - [H(Y_1|Q) - H(S_1|Q)] \geq 0.
\end{aligned}
$$

Thus, the $\min$ in (14) is always achieved by the last term, and (14) simplifies to

$$
\begin{aligned}
R_1 &< H(Y_1|Q) - H(S_1|Q) \\
&= I(X_1; Y_1|Q).
\end{aligned}
$$

Using a similar argument, it follows that the conditions for $\mathfrak{R}_2$ and $\mathfrak{R}_3$ in Theorem 1 are implied by (6). ∎

*Remark 3:* In the case with noisy observations, this proof fails in the following manner. Interference decoding entails

$$
\begin{aligned}
R_1 &< I(X_1, S_1; Z_1|Q) - H(S_1|Q) \\
&= I(X_1; Z_1|Q) + I(S_1; Z_1|X_1, Q) - H(S_1|Q) \\
&= I(X_1; Z_1|Q) + H(S_1|X_1, Q) - H(S_1|X_1, Z_1, Q) \\
&\quad - H(S_1|Q) \\
&= I(X_1; Z_1|Q) - H(S_1|X_1, Z_1, Q).
\end{aligned}
$$

The first term is the achievable rate when treating interference as noise. The second term is zero when the channel is noiseless and acts as a penalty when noise is introduced.

## VI. FINAL REMARKS

This paper presented an interference-decoding inner bound to the capacity region of a class of three-user-pair deterministic interference channels. We showed that this inner bound strictly
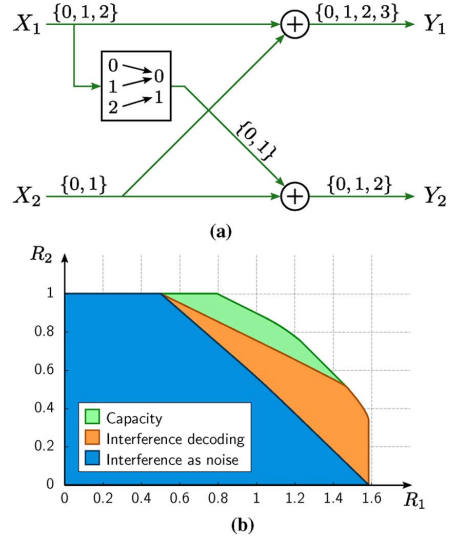


Fig. 7. 2-DIC example. (a) Block diagram of the channel. (b) Capacity region and inner bounds.

includes the interference-as-noise region. As in treating interference as noise, the interference-decoding scheme uses point-to-point codes. The decoder in interference decoding, however, is more sophisticated.

We showed that interference decoding is optimal under strong interference and function invertibility conditions. The scheme is not optimal in general, however. To exemplify this, we consider the following two-user-pair deterministic interference channel for which the capacity region is known.

Consider the 2-DIC in Fig. 7(a) with input alphabets $\mathcal{X}_1 = \{0, 1, 2\}, \mathcal{X}_2 = \{0, 1\}$, loss functions $g_{12} = \{0 \mapsto 0, 1 \mapsto 0, 2 \mapsto 1\}$ and $g_{11} = g_{22} = g_{21} = \mathrm{Id}$, and receiver functions $f_1 = f_2$ being addition. (The interference combining functions $h_1$ and $h_2$ are not relevant in this case.) The outputs of the channel are thus given by

$$
\begin{aligned}
Y_1 &= X_1 + X_2 \\
Y_2 &= g_{12}(X_1) + X_2.
\end{aligned}
$$

The interference-decoding inner bound in Theorem 1 reduces to the set of rate pairs $(R_1, R_2)$ such that

$$
\begin{aligned}
R_1 &< H(X_1|Q) \\
R_2 &< H(X_2|Q) \\
R_1 + \min\{R_2, H(S_1|Q)\} &< H(Y_1|Q) \\
R_2 + \min\{R_1, H(S_2|Q)\} &< H(Y_2|Q),
\end{aligned}
$$

for some $p(q)p(x_1|q)p(x_2|q)$.

Fig. 7(b) compares this inner bound to the capacity region given in [7] and to the region achievable by treating interference as noise (Lemma 1). Not surprisingly, interference decoding does not achieve the full capacity. To achieve capacity, Han–Kobayashi rate splitting and superposition coding are needed.

APPENDIX

### A. Proof of Lemma 2

Applying the packing lemma [9] with $U = U, X = (A, B)$, and $Y = C$ immediately establishes the convergence if $R_A + R_B < I(A, B; C|U) - \delta(\varepsilon)$. Next, we prove convergence when $R_B + H(A|U) < I(A, B; C|U) - \delta(\varepsilon)$. To this end, we bound the probability of the event in question as follows:

$$
\mathrm{P}\left\{(U^n, A^n(m), B^n(l), C^n) \in \mathcal{T}_\varepsilon^{(n)} \text{ for some } m, l\right\}
$$

$$
\leq \sum_{l=1}^{2^{nR_B}} \mathrm{P}\left\{(U^n, A^n(m), B^n(l), C^n) \in \mathcal{T}_\varepsilon^{(n)} \text{ for some } m\right\}
$$

$$
= \sum_{l=1}^{2^{nR_B}} \sum_{u^n \in \mathcal{T}_\varepsilon^{(n)}(U)} \mathrm{P}\{U^n = u^n\}
$$

$$
\cdot \sum_{b^n \in \mathcal{T}_\varepsilon^{(n)}(B|u^n)} \mathrm{P}\{B^n(l) = b^n | U^n = u^n\}
$$

$$
\cdot \sum_{\substack{a^n(m) \in \mathcal{T}_\varepsilon^{(n)}(A|u^n) \\ \text{for all } m \in [1:2^{nR_A}]}} \mathrm{P}\{A^n(m) = a^n(m)
$$

$$
\text{for all } m | U^n = u^n\}
$$

$$
\cdot \mathrm{P}\left\{\bigcup_{m=1}^{2^{nR_A}} \left\{(u^n, a^n(m), b^n, C^n) \in \mathcal{T}_\varepsilon^{(n)}(U, A, B, C)\right\}\right\}.
$$

Upon closer inspection, the union in the last probability term potentially contains duplicate events, for example if $a^n(1) = a^n(2)$. Those duplicates can be eliminated. By pessimistically assuming that the set $\{a^n(m) : m \in [1 : 2^{nR_A}]\}$ is equal to $\mathcal{T}_\varepsilon^{(n)}(A|u^n)$, we can write

$$
\mathrm{P}\left\{\bigcup_{m=1}^{2^{nR_A}} \left\{(u^n, a^n(m), b^n, C^n) \in \mathcal{T}_\varepsilon^{(n)}(U, A, B, C)\right\}\right\}
$$

$$
\leq \mathrm{P}\left\{\bigcup_{a^n \in \mathcal{T}_\varepsilon^{(n)}(A|u^n)} \left\{(u^n, a^n, b^n, C^n)\right.\right.
$$

$$
\left.\left. \in \mathcal{T}_\varepsilon^{(n)}(U, A, B, C)\right\}\right\}
$$

$$
\leq \sum_{a^n \in \mathcal{T}_\varepsilon^{(n)}(A|u^n)} \mathrm{P}\left\{(u^n, a^n, b^n, C^n) \in \mathcal{T}_\varepsilon^{(n)}(U, A, B, C)\right\}
$$

$$
\overset{(a)}{\leq} 2^{n(H(A|U)+\delta_1(\varepsilon))} \cdot 2^{-n(I(C;A,B|U)-\delta_2(\varepsilon))}
$$

$$
= 2^{n(H(A|U)-I(C;A,B|U)+\delta(\varepsilon))}. \tag{15}
$$

In step (a), we use the upper bound on the size of the conditional typical set $\mathcal{T}_\varepsilon^{(n)}(A|u^n)$ and the joint typicality lemma [9] with $X = U, Y = (A, B)$ and $Z = C$. Strictly speaking, the joint typicality lemma holds only if $(u^n, a^n, b^n) \in \mathcal{T}_\varepsilon^{(n)}(U, A, B)$, which does not necessarily follow from $(u^n, a^n) \in \mathcal{T}_\varepsilon^{(n)}(U, A)$ and $(u^n, b^n) \in \mathcal{T}_\varepsilon^{(n)}(U, B)$ as given in our sum expression. However, for the cases where $(u^n, a^n, b^n) \notin \mathcal{T}_\varepsilon^{(n)}(U, A, B)$, we have $\mathrm{P}\{(u^n, a^n, b^n, C^n) \in \mathcal{T}_\varepsilon^{(n)}(U, A, B, C)\} = 0$, and the bound from the joint typicality lemma still holds (though very

loosely). Substituting from (15) into the previous inequality, we obtain

$$
\mathrm{P}\left\{(U^n, A^n(m), B^n(l), C^n) \in \mathcal{T}_\varepsilon^{(n)} \text{ for some } m, l\right\}
$$

$$
\leq 2^{n(R_B+H(A|U)-I(C;A,B|U)+\delta(\varepsilon))}.
$$

Clearly, this probability converges to zero as $n \to \infty$ if $R_B + H(A|U) < I(A, B; C|U) - \delta(\varepsilon)$. Completely symmetrically, convergence follows from $R_A + H(B|U) < I(A, B; C|U) - \delta(\varepsilon)$. Thus convergence is implied by $\min\{R_A + R_B, R_A + H(B|U), H(A|U) + R_B\} < I(A, B; C|U) - \delta(\varepsilon)$, and the desired result follows by recalling that $H(A|U) + H(B|U) \geq I(A, B; C|U)$. ∎

### B. Proof of Lemma 3

The first and last term in the $\min$ expression follow immediately from Lemma 2 by disregarding the special structure of $B^n$. For the second term, we argue similarly

$$
\mathrm{P}\left\{(U^n, A^n(m), B^n(l_1, l_2), C^n) \in \mathcal{T}_\varepsilon^{(n)} \text{ for some } m, l_1, l_2\right\}
$$

$$
\leq \sum_{u^n \in \mathcal{T}_\varepsilon^{(n)}(U)} \mathrm{P}\{U^n = u^n\} \sum_{m=1}^{2^{nR_A}}
$$

$$
\cdot \sum_{a^n \in \mathcal{T}_\varepsilon^{(n)}(A|u^n)} \mathrm{P}\{A^n(m) = a^n | U^n = u^n\}
$$

$$
\cdot \sum_{l_1=1}^{2^{nR_{B_1}}} \sum_{b_1^n \in \mathcal{T}_\varepsilon^{(n)}(B_1|u^n)} \mathrm{P}\{B_1^n(l_1) = b_1^n | U^n = u^n\}
$$

$$
\cdot \sum_{\substack{b_2^n(l_2) \in \mathcal{T}_\varepsilon^{(n)}(B_2|u^n) \\ \text{for all } l_2 \in [1:2^{nR_{B_2}}]}} \mathrm{P}\{B_2^n(l_2) = b_2^n(l_2)
$$

$$
\text{for all } l_2 | U^n = u^n\}
$$

$$
\cdot \mathrm{P}\left\{\bigcup_{l_2=1}^{2^{nR_{B_2}}} \left\{(u^n, a^n, b^n(b_1^n, b_2^n(l_2)), C^n) \in \mathcal{T}_\varepsilon^{(n)}\right\}\right\}.
$$

There are at most $|\mathcal{T}_\varepsilon^{(n)}(B_2|u^n)|$ distinct events in the union expression. Using a similar line of reasoning as in the proof of Lemma 2, we can upper bound the last probability term by $2^{n(H(B_2|U)-I(C;A,B|U)+\delta(\varepsilon))}$. Substituting into the previous inequality, we obtain the upper bound $2^{n(R_A+R_{B_1}+H(B_2|U)-I(C;A,B|U)+\delta(\varepsilon))}$ on the probability of the event of interest. Clearly, this expression converges to zero as $n \to \infty$ if $R_A + R_{B_1} + H(B_2|U) < I(A, B; C|U) - \delta(\varepsilon)$. We have thus established the second term in the $\min$ expression. The third term follows in a symmetric manner. ∎

### REFERENCES

[1] M. A. Maddah-Ali, A. S. Motahari, and A. K. Khandani, "Communication over MIMO X channels: Interference alignment, decomposition, and performance analysis," *IEEE Trans. Inf. Theory*, vol. 54, no. 8, pp. 3457–3470, Aug. 2008.

[2] V. R. Cadambe and S. A. Jafar, "Interference alignment and degrees of freedom of the K-user interference channel," *IEEE Trans. Inf. Theory*, vol. 54, no. 8, pp. 3425–3441, Aug. 2008.

[3] R. H. Etkin and E. Ordentlich, "The degrees-of-freedom of the K-user Gaussian interference channel is discontinuous at rational channel coefficients," *IEEE Trans. Inf. Theory*, vol. 55, no. 11, pp. 4932–4946, Nov. 2009, arXiv:0901.1695.

[4] A. Motahari, S. O. Gharan, M. A. Maddah-Ali, and A. K. Khandani, "Real interference alignment: Exploiting the potential of single antenna systems," *IEEE Trans. Inf. Theory*, Nov. 2009, arXiv:0908.2282, submitted for publication.

[5] G. Bresler, A. Parekh, and D. Tse, "The approximate capacity of the many-to-one and one-to-many Gaussian interference channels," *IEEE Trans. Inf. Theory*, vol. 56, no. 9, pp. 4566–4592, Sep. 2010, arXiv:0804.4489.

[6] B. Nazer and M. Gastpar, "Compute-and-forward: Harnessing interference through structured codes," *IEEE Trans. Inf. Theory*, Aug. 2009, arXiv:0908.2119v2, submitted for publication.

[7] A. A. El Gamal and M. H. M. Costa, "The capacity region of a class of deterministic interference channels," *IEEE Trans. Inf. Theory*, vol. 28, no. 2, pp. 343–346, Mar. 1982.

[8] T. Gou and S. A. Jafar, "Sum capacity of a class of symmetric SIMO Gaussian interference channels within O (1)," *IEEE Trans. Inf. Theory*, vol. 57, no. 4, pp. 1932–1958, Apr. 2010, arXiv:0905.1745.

[9] A. El Gamal and Y.-H. Kim, Lectures on Network Inf. Theory 2010, arXiv:1001.3404.

[10] T. S. Han and K. Kobayashi, "A new achievable rate region for the interference channel," *IEEE Trans. Inf. Theory*, vol. 27, no. 1, pp. 49–60, Jan. 1981.

[11] A. S. Avestimehr, S. N. Diggavi, and D. N. C. Tse, "A deterministic approach to wireless relay networks," in *Proc. Allerton*, Sep. 2007.

[12] G. Bresler and D. Tse, "The two-user Gaussian interference channel: A deterministic view," *Eur. Trans. Telecomm.*, vol. 19, no. 4, pp. 333–354, Jun. 2008, arXiv:0807.3222.

[13] S. A. Jafar and S. Vishwanath, Generalized Degrees of Freedom of the Symmetric K User Gaussian Interference Channel arXiv:0804.4489.

[14] B. Bandemer, G. Vazquez-Vilar, and A. El Gamal, "On the sum capacity of a class of cyclically symmetric deterministic interference channels," in *Proc. ISIT 2009*, Seoul, Korea, Jun. 2009.

[15] B. Nazer and M. Gastpar, "The case for structured random codes in network capacity theorems," *Eur. Trans. Telecommun., Special Issue on New Directions in Inf. Theory*, vol. 19, no. 4, pp. 455–474, Jun. 2008.

[16] H.-F. Chong, M. Motani, H. K. Garg, and H. El Gamal, "On the Han–Kobayashi region for the interference channel," *IEEE Trans. Inf. Theory*, vol. 54, no. 7, pp. 3188–3195, Jul. 2008.

[17] E. van der Meulen, "A survey of multi-way channels in information theory: 1961–1976," *IEEE Trans. Inf. Theory*, vol. 23, no. 1, pp. 1–37, Jan. 1977.

**Bernd Bandemer** (S'06) received the Dipl.-Ing. degree in electrical and computer engineering in 2006 from Ilmenau University of Technology, Ilmenau, Germany.

During 2003–2004, he was awarded a Fulbright scholarship to study at Purdue University, West Lafayette, IN. He is currently working toward the Ph.D. degree in electrical engineering at Stanford University, Stanford, CA. His research interests are network information theory and wireless communications.

**Abbas El Gamal** (S'71–M'73–SM'83–F'00) received the B.Sc. (honors) degree in electrical engineering from Cairo University in 1972. He received the M.S. degree in statistics and the Ph.D. degree in electrical engineering, both from Stanford University, Stanford, CA, in 1977 and 1978, respectively.

From 1978 to 1980, he was an Assistant Professor of Electrical Engineering, University of Southern California (USC), Los Angeles. He has been on the Stanford University faculty since 1981, where he is currently the Hitachi America Professor in the School of Engineering. His research interest and contributions are in the areas of network information theory, digital imaging, and integrated circuit design. He has authored or coauthored more than 200 papers and 30 patents in these areas.

Dr. El Gamal is on the Board of Governors of the IEEE Information Theory Society.