ISIT 2009, Seoul, Korea, June 28 - July 3, 2009

On the Totally Asynchronous Interference Channel with Single-User Receivers

Eduard Calvo, Javier R. Fonollosa, and Josep Vidal
E-mail: {eduard,fono,pepe}@gps.tsc.upc.edu
SPCOM Group
Dpt. Signal Theory and Communications
Technical University of Catalonia (UPC), SPAIN

Abstract—The performance characterization of decentralized wireless networks with uncoordinated sender-destination pairs motivates the study of the totally asynchronous interference channel with single-user receivers. Since this channel is not information stable, its capacity region is determined resorting to information density, although more amenable single-letter inner and outer bounds are provided as well. Aiming at numerical evaluation of the achievable rates, we subsequently concentrate on the inner bound for the Gaussian case.

We show that taking Gaussian inputs is not the best choice in general and derive analytical conditions under which other input distributions may be optimal. Essentially, these conditions require the channel to be interference-limited. Finally, the existence of such non-Gaussian distributions with superior performance is validated numerically in different scenarios.

I. INTRODUCTION

The interference channel (IC) [1] is the network scenario that models the interactions between several disjoint non-cooperative (relayless) sender-receiver communication links sharing, generally in a non-orthogonal manner, the same physical medium. Interference couples the achievable rates, and the fact that each destination is interested in decoding only one among all the information-bearing codewords on which its channel output depends is what makes analysis difficult.

Finding a single-letter characterization of the capacity region of the IC remains an open problem which, however, has been solved in some particular cases: i) statistically equivalent channel outputs [2], ii) very strong [1] or strong interference [3], iii) a class of discrete degraded [4] and additive [5] ICs, and iv) a class of deterministic ICs [6].

Inherent to the definition of the channel is the perfect frame synchronization assumption, i.e. that the codewords sent by the transmitters are received at unison at each destination. This, however, is likely not to hold in decentralized wireless networks with autonomous sender-receiver pairs. Nonetheless, the lack of centralized signalling on who and when

E. Calvo was with the Dept. of Signal Theory and Communications, UPC. He is now with the Department of Production, Technology, and Operations Management, IESE Business School, Univ. of Navarra, Spain (e-mail: ecalvo@iese.edu).

This work has been partially funded by the European Commission, the Spanish Ministry of Education and Science, the Catalan Government and FEDER funds under contracts TEC2006-06481, CONSOLIDER INGENIO 2010 CSD2008-00010 COMONSENS, 215282 ROCKET, and 2005SGR-00639, and grant FPU-AP-2004-3549.

is transmitting together with fast time-varying changes on the network topology may render the destinations unaware of the potential interference hampering the transmission of their intended user. To mention but a few example scenarios, consider decentralized networks with simple receivers [8], networks with non-stationary interference [9], and the study of the throughput scaling law of multihop wireless networks [10].

Motivated by these operational and practical constraints, we restrict the receivers to treat interference as noise (single-user detection) and force total transmission asynchronism [11]. First, the channel model is defined in the discrete alphabet case, where the capacity region is characterized using an Information Spectrum approach [12]. Pursuing analytical results, we provide an inner and an outer bound to the capacity region using single-letter expressions. The single-letter inner bound is achieved by stationary inputs with i.i.d. letters and, given its simplicity, we subsequently focus on this achievable rate region throughout the rest of the paper.

Next, we study whether the natural appeal of Gaussian inputs in coding for the Gaussian IC (GIC) still holds when it comes to frame asynchronism and single-user detection. Despite [13] showed that Gaussian inputs fall short of achieving the capacity region when expressed as a limiting expression in the frame-synchronous setup, that tells us little about the their potential optimality when single-letter characterizations are used instead. Similarly, it does not preclude optimality in our setup. In fact, a finite expansion analysis of mutual information shows that Gaussian-distributed codes fall short of achievable rate under certain circumstances only. Additionally, analytic conditions for non-optimality of Gaussian-distributed codes are derived that only depend on the coupling coefficients of the channel and the transmit power constraints. For the symmetric GIC (equal coupling coefficients, equal transmit power constraints) they reduce to exceeding a transmit power threshold. Thus, in a nutshell, Gaussian codewords are not optimal when the channel is interference-limited (interference has to be stronger than moderate).

Indeed, the optimization of each input distribution is able to impact on mutual information and yield gains with respect to Gaussian-distributed codes only when the output distribution is dominated by interference. Numerical performance evaluation of some non-Gaussian-distributed codes shows excellent

agreement with the analytical conditions and constructively validates the assumptions considered in the finite series approximation of mutual information.

II. THE CAPACITY REGION

To start with, we clarify notation. X_k^n denotes an n-dimensional random vector taking on the value x_k^n over the finite set $\mathcal{X}_k^n$ with probability $P_{X_k^n}(x_k^n)$. Its i-th component is denoted by $X_{k,i}$, whereas $X_k^i = [X_{k,1} \ldots X_{k,i}]$ and $X_{k,i}^n = [X_{k,i} \ldots X_{k,n}]$.

Let $\mathcal{X}_1$ and $\mathcal{X}_2$ denote the input alphabets of the two senders, and $\mathcal{Y}_1$ and $\mathcal{Y}_2$ the output alphabets of the two destinations. The frame-asynchronous discrete memoryless IC with single-user receivers consists of two conditional distributions $\{P_{Y_k|X_1 X_2} : \mathcal{X}_1 \times \mathcal{X}_2 \to \mathcal{Y}_k\}_{k=1,2}$ that describe the underlying channels, and two collections of distributions $\{P_{D_1,n}, P_{D_2,n}\}_{n=1}^{\infty}$, defined on $\{0, 1, \ldots, n-1\}_{n=1}^{\infty}$, that describe the degree of asynchronism. We say the channel is totally asynchronous [11] when $P_{D_1,n}, P_{D_2,n}$ are uniform $\forall n$.

A $(2^{nR_1}, 2^{nR_2}, n)$ code for this channel consists of two encoding functions

$$X_k^n : \{1, \ldots, 2^{nR_k}\} \to \mathcal{X}_k^n, \quad k = 1, 2, \quad (1)$$

and two decoding functions

$$\hat{m}_k : \mathcal{Y}_k^n \to \{1, \ldots, 2^{nR_k}\}, \quad k = 1, 2. \quad (2)$$

Sender 1 draws a message M_1 uniformly from $\{1, \ldots, 2^{nR_1}\}$ and sends the corresponding codeword X_1^n, of length n, over the channel. We assume without loss of generality[1] that receiver 1 is frame-synchronized with sender 1, and thus Y_1^n is a sufficient statistic for the message M_1. Since receiver 1 is unaware of the presence of an interferer, the random delay D_1 (with distribution $P_{D_1,n}$) experienced by the codewords of sender 2 is unknown. The second link behaves analogously.

By treating interference as noise, the channel faced by the first link is determined by the value of D_1, drawn $\sim P_{D_1,n}$ at the beginning of transmission and held fixed thereafter, i.e.,

$$P_{Y_1^n|X_1^n,D_1}(y_1^n|x_1^n, d_1) = \sum_{x_2^n \in \mathcal{X}_2^n} P_{X_{2,n-d_1+1}^n}(x_{2,1}^{d_1})$$
$$\times P_{X_{2,1}^{n-d_1}}(x_{2,d_1+1}^n) \prod_{i=1}^n P_{Y_1|X_1 X_2}(y_{1,i}|x_{1,i} x_{2,i}). \quad (3)$$

Each received frame of receiver 1 depends on two independent codewords of user 2 and, as nothing is imposed on the distribution of X_2^n, the channel (3) may have memory in general. A related model to (3) is the composite channel [14], where each component channel (for each instance of D_1) is assumed stationary and ergodic. The composite channel is information unstable, so hence is (3). Essentially, this precludes single-letter characterizations of the capacity region. We thus treat each link as a general single-user channel adopt an Information Spectrum approach.

[1]Synchronization can be achieved via the use of periodic preamble sequences at negligible rate penalty [11, App. I].

Theorem 1: The capacity region of the totally-asynchronous DMIC with single-user receivers is

$$\mathcal{C} = \bigcup_{P_{\mathbf{X}_1}, P_{\mathbf{X}_2}} \{R_k \leq \sup\{\alpha_k : F_{\mathbf{X}_k}(\alpha_k) = 0\}, \ k = 1, 2\}, \quad (4)$$

where $P_{\mathbf{X}_k} = \{P_{X_k^n}\}_{n=1}^{\infty}$ is a sequence of finite-dimensional distributions, $F_{\mathbf{X}_k}$ is the limit of the cumulative distribution function of the normalized information density of the k-th link,

$$F_{\mathbf{X}_k}(\alpha_k) = \lim_{n \to \infty} P_{X_k^n W_k^n} \left\{ \frac{1}{n} i_{X_k^n W_k^n}(X_k^n; Y_k^n) \leq \alpha_k \right\}, \quad (5)$$

the information density amounts to

$$i_{X_k^n W_k^n}(x_k^n; y_k^n) = \log \frac{P_{W_k^n}(y_k^n|x_k^n)}{P_{Y_k^n}(y_k^n)}, \quad (6)$$

and W_k^n denotes the equivalent channel from sender k to receiver k,

$$P_{W_k^n}(y_k^n|x_k^n) = \frac{1}{n} \sum_{d_k=0}^{n-1} P_{Y_k^n|X_k^n,D_k}(y_k^n|x_k^n, d_k). \quad (7)$$

Proof: It follows from direct application to both links of the general capacity formula proved in [12]. ∎

Given the complex form of the expressions in Theorem 1, we derive simpler expressions upper and lower bounding $\mathcal{C}$.

Lemma 1: The achievable rate region

$$\mathcal{R} = \bigcup_{P_{X_1}, P_{X_2}} \{R_1 \leq I(X_1; Y_1), \ R_2 \leq I(X_2; Y_2)\} \quad (8)$$

is an inner bound of the capacity region; $\mathcal{R} \subseteq \mathcal{C}$.

Proof: The region $\mathcal{R}$ is the particularization of $\mathcal{C}$ when $P_{\mathbf{X}_1}$ and $P_{\mathbf{X}_2}$ are constrained to be i.i.d.. ∎

Lemma 2: The region

$$\mathcal{R}_o = \bigcup_{P_{X_1}, P_{X_2}} \{R_1 \leq I(X_1; Y_1|X_2), \ R_2 \leq I(X_2; Y_2|X_1)\} \quad (9)$$

is an outer bound to the capacity region; $\mathcal{C} \subseteq \mathcal{R}_o$.

Proof: The proof follows by noticing that

$$\sup\{\alpha_1 : F_{\mathbf{X}_1}(\alpha_1) = 0\} \overset{(a)}{\leq} \lim_{n \to \infty} \inf \frac{1}{n} I(X_1^n; Y_1^n) \quad (10)$$

$$\overset{(b)}{\leq} \lim_{n \to \infty} \inf \frac{1}{n} I(X_1^n; Y_1^n|X_2^n) \quad (11)$$

$$\overset{(c)}{\leq} I(X_1; Y_1|X_2), \quad (12)$$

where (a) follows from [12, Thm. 8.(h)], (b) is a consequence of the independence between X_1^n and X_2^n, and (c) holds for some pair of distributions P_{X_1}, P_{X_2} using [11, Thms. 3.3-3.4]. Analogous arguments hold for the second link, where [11, Thms. 3.3-3.4] guarantees that the *same* distributions P_{X_1}, P_{X_2} are used to evaluate $I(X_2; Y_2|X_1)$ in a bound similar to (12). ∎

Both $\mathcal{R}$ and $\mathcal{R}_o$ are characterized in terms of the union of regions, without any convex hull operation. Intuitively, the lack of frame synchronism precludes time-sharing between distributions, as happens in the discrete-multiple access channel [11]. Due to its simplicity and amenability for numerical computation, we subsequently focus on $\mathcal{R}$ throughout the rest of the paper.

III. THE GAUSSIAN IC

Consider the 2×2 standard-form Gaussian IC (GIC) [1],

$$Y_1 = X_1 + c_{21}X_2 + Z_1 \quad (13)$$
$$Y_2 = X_2 + c_{12}X_1 + Z_2, \quad (14)$$

where $Z_k \sim \mathcal{N}(0,1)$ $k = 1, 2$, and the codewords X_1 and X_2 are independent and independent of the noise samples Z_1 and Z_2. The input codewords satisfy the transmit power constraint $\mathbb{E}\{X_k^2\} \le P_k$, $k = 1, 2$. Whenever $c_{12} = c_{21} \triangleq c$ and $P_1 = P_2 \triangleq P$ we say the GIC is fully symmetric. The computation of $\mathcal{R}$ is inevitably bonded to the choice of the distribution of the codewords, described by the pdf's f_{X_1} and f_{X_2}.

A. Definition of optimality

Definition 1: The input distributions f_{X_1} and f_{X_2} are α-optimal, $\alpha \in [0, \pi/2]$, if they achieve the rate pair of the boundary of $\mathcal{R}$ that intersects the line $R_2 = \tan(\alpha)R_1$. Denote such pair by $(R_1^\star(\alpha), R_2^\star(\alpha))$.

It is not difficult to show that any pair of α-optimal distributions is a solution to the optimization problem

$$\underset{f_{X_1}, f_{X_2}}{\text{maximize}} \quad \min\left\{ \frac{I(X_1; Y_1)}{\cos(\alpha)}, \frac{I(X_2; Y_2)}{\sin(\alpha)} \right\} \quad (15)$$

$$\text{subject to} \quad f_{X_k}(x) \ge 0 \ \forall x \in \mathbb{R}, \ k = 1, 2 \quad (16)$$

$$\int f_{X_k}(x)dx = 1, \ k = 1, 2 \quad (17)$$

$$\int x^2 f_{X_k}(x)dx \le P_k, \ k = 1, 2. \quad (18)$$

The problem (15)-(18) is rather involved because of the intricate dependence of mutual information on f_{X_1} and f_{X_2}. Thus, instead of determining a pair of α-optimal distributions for each α and examining whether they are Gaussian or not, we shall study if the optimal value of (15)-(18) decreases when we consider only Gaussian inputs. In other words, we want to find out if the achievable rate region of Gaussian-distributed codes $\mathcal{R}^G$ and the capacity inner bound $\mathcal{R}$ coincide, where[2]

$$\mathcal{R}^G = \bigcup_{\substack{0 \le p_k \le P_k \\ k=1,2}} \left\{ R_k \le \frac{1}{2}\log\left(1 + \frac{p_k}{1 + c_{jk}^2 p_j}\right), k \ne j = 1, 2 \right\}. (19)$$

Intuitively, Gaussian-distributed codes behave as a (possibly local) extremum in the maximization of the achievable rates in that they are a greedy strategy. Although this input distribution maximizes mutual information if interference is absent, it also givers rise to the worst additive interference [15]. In other words, Gaussian-distributed codes maximize $h(Y_k)$ and $h(Y_k|X_k)$ simultaneously for $k = 1, 2$, but this does not necessarily imply that they maximize $I(X_k; Y_k)$ as well. Since direct construction of α-optimal distributions (15)-(18) seems overwhelming, we shall adopt a completely different approach for showing non-optimality of Gaussian-distributed codes. It is based on the relation between mutual information and the shape of the pdf of the codewords, as described by their cumulants.

[2]Unless the logarithm basis is indicated, it can be chosen arbitrarily as long as both sides of the equations have the same units.

B. Finite expansion analysis of mutual information

Let Ω_X be the support set of a zero-mean continuous r.v. X with pdf f_X and characteristic function $\varphi_X(\omega)$,

$$\varphi_X(\omega) = \mathbb{E}\{e^{j\omega X}\}. \quad (20)$$

The cumulants [16] $\{\kappa_i(X)\}_{i=1}^{+\infty}$ of X are the coefficients of the McLaurin series of the natural logarithm of $\varphi_X(\omega)$,

$$\log_e(\varphi_X(\omega)) = \sum_{i=1}^{+\infty} \kappa_i(X)\frac{(j\omega)^i}{i!}, \quad (21)$$

can be related to its (central) moments and have some interesting properties concerning the shape of f_X.

- Symmetry: $f_X(x) = f_X(-x) \Rightarrow \kappa_{2i-1}(X) = 0 \ \forall i \ge 1$.
- Independence: X_1, X_2 independent $\Rightarrow \kappa_i(X_1 + X_2) = \kappa_i(X_1) + \kappa_i(X_2) \ \forall i$.
- Scaling: $\kappa_i(aX) = a^i \kappa_i(X) \ \forall a \in \mathbb{R}$.
- Cumulants of the Gaussian distribution: $X \sim \mathcal{N}(0, P) \Rightarrow \kappa_2(X) = P, \ \kappa_i(X) = 0 \ \forall i \ne 2$.

The third-order cumulant, $\kappa_3(X)$ or skewness, measures the lack of symmetry of a distribution, whereas the fourth-order cumulant, $\kappa_4(X)$ or kurtosis, captures the non-Gaussianity (or peakedness) of X. Kurtosis is zero for a Gaussian r.v., it is typically positive for distributions with heavy tails and a peak at zero, and negative for flatter-than-Gaussian densities with lighter tails. Moreover, it is fundamentally lower bounded by

$$\kappa_4(X_k) = \mathbb{E}\{X_k^4\} - 3\sigma_X^4 \ge \mathbb{E}\{X_k^2\}^2 - 3\sigma_X^4 = -2\sigma_X^4, \quad (22)$$

due to Jensen's inequality and the fact that X is zero-mean.

Finally, the tool that will be used to analyze the optimality of Gaussian inputs is the Gram-Charlier expansion, which approximates $h(X)$ of the near-Gaussian r.v. X around the entropy of an equivalent Gaussian r.v. with the same variance as X. If $\frac{\kappa_i^2(X)}{\sigma_X^{2i}} \ll 1$ for $i > 2$ then

$$h(X) \approx \frac{1}{2}\log(2\pi e\sigma_X^2) - \left(\frac{1}{12}\frac{\kappa_3^2(X)}{\sigma_X^6} + \frac{1}{48}\frac{\kappa_4^2(X)}{\sigma_X^8}\right)\log(e), \quad (23)$$

is a fourth-order entropy approximation for X [16]. The gap between mutual information and cumulants is bridged in the next section. Before that, let us adopt w.l.o.g. the zero-mean assumption on X_1 and X_2.

IV. ON THE OPTIMALITY OF GAUSSIAN INPUTS

The Gram-Charlier expansion of entropy (23) can be used to relate mutual information to the cumulants of X_1 and X_2.

Lemma 3: Given the zero-mean r.v.'s X_1, X_2 with $\mathbb{E}\{X_k^2\} = p_k$, $k = 1, 2$, a fourth-order expansion approximation of mutual information is

$$I(X_1; Y_1) \approx \frac{1}{2}\log\left(1 + \frac{p_1}{1 + c_{21}^2 p_2}\right) + \boldsymbol{\kappa}^T \mathbf{A} \boldsymbol{\kappa} \quad (24)$$

$$I(X_2; Y_2) \approx \frac{1}{2}\log\left(1 + \frac{p_2}{1 + c_{12}^2 p_1}\right) + \boldsymbol{\kappa}^T \mathbf{B} \boldsymbol{\kappa}, \quad (25)$$

where $\boldsymbol{\kappa} \triangleq [\kappa_3(X_1) \; \kappa_4(X_1) \; \kappa_3(X_2) \; \kappa_4(X_2)]^T$,

$$[\mathbf{A}]_{1,1} = -\frac{\log(e)}{12\sigma_{Y_1}^6}, \quad [\mathbf{B}]_{1,1} = \frac{\log(e)c_{12}^6}{12}(\sigma_{Y_2|X_2}^{-6} - \sigma_{Y_2}^{-6}) \quad (26)$$

$$[\mathbf{A}]_{1,3} = -\frac{\log(e)c_{21}^3}{6\sigma_{Y_1}^6}, \qquad\qquad [\mathbf{B}]_{1,3} = -\frac{\log(e)c_{12}^3}{6\sigma_{Y_2}^6} \quad (27)$$

$$[\mathbf{A}]_{2,2} = -\frac{\log(e)}{48\sigma_{Y_1}^8}, \quad [\mathbf{B}]_{2,2} = \frac{\log(e)c_{12}^8}{48}(\sigma_{Y_2|X_2}^{-8} - \sigma_{Y_2}^{-8}) \quad (28)$$

$$[\mathbf{A}]_{2,4} = -\frac{\log(e)c_{21}^4}{24\sigma_{Y_1}^8}, \qquad\qquad [\mathbf{B}]_{2,4} = -\frac{\log(e)c_{12}^4}{24\sigma_{Y_2}^8} \quad (29)$$

$$[\mathbf{A}]_{3,3} = \frac{\log(e)c_{21}^6}{12}(\sigma_{Y_1|X_1}^{-6} - \sigma_{Y_1}^{-6}), \quad [\mathbf{B}]_{3,3} = -\frac{\log(e)}{12\sigma_{Y_2}^6} \quad (30)$$

$$[\mathbf{A}]_{4,4} = \frac{\log(e)c_{21}^8}{48}(\sigma_{Y_1|X_1}^{-8} - \sigma_{Y_1}^{-8}), \quad [\mathbf{B}]_{4,4} = -\frac{\log(e)}{48\sigma_{Y_2}^8}, \quad (31)$$

the rest of entries of $\mathbf{A}$ and $\mathbf{B}$ are zero, and

$$\sigma_{Y_k}^2 = 1 + c_{jk}^2 p_j + p_k, \; \sigma_{Y_k|X_k}^2 = 1 + c_{jk}^2 p_j \quad (32)$$

for $k, j = 1, 2, \; j \neq k$.

Proof: Expressions (24)-(31) follow from the application of the Gram-Charlier expansion (23) to each entropy term in

$$I(X_k; Y_k) = h(X_k + c_{jk}X_j + Z_k) - h(c_{jk}X_j + Z_k), \quad (33)$$

where $k \neq j = 1, 2$, together with the independence property, the scaling property, and the fact that Z_k is Gaussian. ∎

The matrices $\mathbf{A}$ and $\mathbf{B}$ are not negative definite and, thus, the possibility of finding some $\boldsymbol{\kappa}$ inducing a pair of non-Gaussian distributions outperforming Gaussian inputs is not precluded.

Lemma 4: Gaussian inputs are not optimal if, for some fixed $\alpha \in [0, \pi/2]$, the problem

$$\text{find} \qquad \boldsymbol{\kappa} \qquad\qquad\qquad\qquad (34)$$

$$\text{subject to} \quad \min\{\boldsymbol{\kappa}^T \mathbf{A}(\alpha)\boldsymbol{\kappa}, \boldsymbol{\kappa}^T \mathbf{B}(\alpha)\boldsymbol{\kappa}\} > 0 \quad (35)$$

$$[\boldsymbol{\kappa}]_2 \geq -2p_1^2(\alpha) \qquad\qquad (36)$$

$$[\boldsymbol{\kappa}]_4 \geq -2p_2^2(\alpha) \qquad\qquad (37)$$

is feasible, where $\mathbf{A}(\alpha), \mathbf{B}(\alpha)$ are equivalent to $\mathbf{A}, \mathbf{B}$ in Lemma 3 but with $\mathbb{E}\{X_k^2\} = p_k(\alpha)$ in (32), $k = 1, 2$, and $(p_1(\alpha), p_2(\alpha))$ is the power allocation that achieves the rate pair of the boundary of $\mathcal{R}^G$ that intersects the line $R_2 = \tan(\alpha)R_1$.

Proof: The Lemma follows from the fact that the feasibility of (34)-(37) implies that there exist a pair of distributions achieving a rate pair outside $\mathcal{R}^G$ in the direction given by the line $R_2 = \tan(\alpha)R_1$, in so implying $\mathcal{R} \neq \mathcal{R}^G$. ∎

In general, it is difficult to find a vector of cumulants satisfying Lemma 4 for a given GIC and α due to i) the lack of general closed-form expressions for $(p_1(\alpha), p_2(\alpha))$ (which are the solution to a non-convex problem), and ii) the fact that neither $\mathbf{A}(\alpha)$ nor $\mathbf{B}(\alpha)$ are positive/negative definite. Fortunately, we can bypass this by focusing in the symmetric setup.

Theorem 2: Gaussian-distributed codes are not optimal for the totally asynchronous GIC with single-user receivers. In the fully symmetric setup, it suffices to have

$$P > P_{\text{th}}(c) \triangleq \frac{\sqrt{1 + c^{-4}} - 1}{1 + c^2 - \sqrt{1 + c^4}}, \quad (38)$$

which implies that interference is at least moderate.

Proof: For $\alpha = \pi/4$ and the fully symmetric setup, $p_1(\pi/4) = p_2(\pi/4) = P$. Owing to the symmetry of the

channel and the choice of alpha, we impose that X_1 and X_2 have the same distribution, and that their pdf is symmetric around zero (this forces skewness to be zero). If κ_4 denotes their common kurtosis, (34)-(37) reduces to

$$\underset{\kappa_4 \geq -2P^2}{\text{find}} \qquad \kappa_4 \qquad\qquad\qquad\qquad (39)$$

$$\text{subject to} \quad \frac{\log(e)}{48}\left(\frac{c^8}{\sigma_{Y|X}^8} - \frac{(1 + c^4)^2}{\sigma_Y^8}\right)\kappa_4^2 > 0, \quad (40)$$

which is feasible when

$$\frac{\sigma_Y^2}{\sigma_{Y|X}^2} > \sqrt{1 + c^{-4}} \quad (41)$$

or, equivalently, $P > P_{\text{th}}(c)$ (38). Moderate interference[3] accounts for $\frac{\sqrt{1+2P}-1}{2P} < c^2$ or, equivalently for transmission with power above $P_{\text{mod}}(c) = \frac{1/2-c^2}{c^4}$. The inequality

$$P_{\text{th}}(c) - P_{\text{mod}}(c) = \frac{c^2 + \sqrt{1 + c^4} - 1}{2(1 + c^2 - \sqrt{1 + c^4})c^4} \geq 0, \quad (42)$$

concludes the proof. ∎

In essence, Theorem 2 shows that Gaussian-distributed codes are not optimal when interference is significant enough (at least moderate). The stronger the interference, the easier to outperform Gaussian-distributed codes at lower transmit powers ($P_{\text{th}}(c)$ is decreasing in c). Consistently with the fact that Gaussian codes are capacity-achieving in the AWGN, $P_{\text{th}}(c) \to +\infty$ when $c \to 0$.

Interestingly, [7] showed that Gaussian-distributed codes and single-user detection are sum-rate optimal in the frame-synchronous case provided that interference is low enough (noisy interference, as in the terminology of [7]). Our result is consistent with that of [7], which holds under weaker interference than weak interference [17], and rules out the optimality of Gaussian codes and single-user detection for frame-synchronous GICs with stronger than moderate interference.

V. NUMERICAL RESULTS

To show the existence of non-Gaussian distributions outperforming Gaussian inputs that support the fourth-order analysis of Section IV, the mutual information of other codes is numerically computed. In particular, and although these distributions are not near-Gaussian (in the sense $\frac{\kappa_i^2}{\sigma_X^{2i}} \ll 1, i > 2$), let us consider uniformly-distributed codes ($X_k \sim \mathcal{U}(-\sqrt{3p_k}, \sqrt{3p_k})$, $p_k \leq P_k, \; k = 1, 2$) and ternary-distributed codes ($\mathbb{P}\{X_k = A_k\} = \mathbb{P}\{X_k = -A_k\} = 0.5(1 - \mathbb{P}\{X_k = 0\})$, where A_k satisfies $\mathbb{E}\{X_k^2\} \leq P_k, \; k = 1, 2$). We shall see that the results derived under the Gram-Charlier expansion hold even in this situation.

In Figure 1, the achievable rate regions of Gaussian-, uniformly-, and ternary-distributed codes are computed for two different channels with $P = 15$. While for $c = 0.1$ transmission is clearly below the threshold $P_{\text{th}}(0.1) \approx 9950$ and none of the proposed non-Gaussian distributions can beat $\mathcal{R}^G$, when $c = 1/\sqrt{2}$ Theorem 2 holds ($P_{\text{th}}(1/\sqrt{2}) \approx 3.24$) and achievable rate gains over $\mathcal{R}^G$ are explicitly realized.

---

[3]Moderate interference [17] occurs when $c^2 < 1$ and time-sharing is better than Gaussian-distributed codes with single-user decoders.
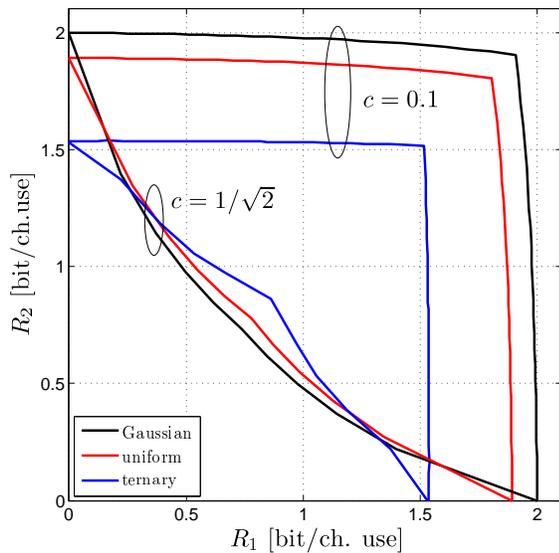
Fig. 1. Achievable rate regions of Gaussian-, uniformly-, and ternary-distributed codes: fully symmetric GIC with $P = 15$ and $c = \{0.1, 1/\sqrt{2}\}$.

To check the accuracy of the threshold power (38) of Theorem 2, Figure 2 (left) plots the achievable symmetric rate (maximum rate for $\alpha = \pi/4$), R_{sym}, of Gaussian-, uniformly-, and ternary-distributed codes as a function of P for $c = \{0.9245, 0.5436\}$, which yield the theoretical values of $P_{\text{th}} = \{1, 10\}$, respectively. Accuracy over a wider range of values of c is investigated in Figure 2 (right), where the theoretical value of the threshold power (38) and the actual threshold power of uniformly-distributed codes are compared. Agreement with Theorem 2 is excellent.

VI. CONCLUSIONS

Motivated by practical application scenarios, we studied the totally asynchronous IC with single-user receivers and derived a single-letter achievable rate region. In the Gaussian case, Gaussian inputs fell short of achieving the bound whenever the transmission powers exceeded a threshold that made inter-

ference to be at least moderate. This opens the door to optimal statistical signal design in realistic scenarios.

REFERENCES

[1] A. B. Carleial, "Interference channels", *IEEE Trans. Inform. Theory*, vol. IT-24, pp. 60-70, Jan. 1978.
[2] A. B. Carleial, "Outer bounds on the capacity of interference channels", *IEEE Trans. Inform. Theory*, vol. IT-29, pp. 602-606, July 1983.
[3] T. S. Han and K. Kobayashi, "A new achievable rate region for the interference channel", *IEEE Trans. Inform. Theory*, vol. IT-27, pp. 49-60, Jan. 1981.
[4] N. Liu and S. Ulukus, "The capacity region of a class of discrete degraded interference channels", *IEEE Trans. Inform. Theory*, vol. 54, pp. 4372-4378, Sep. 2008.
[5] R. Benzel, "The capacity region of a class of discrete additive degraded interference channels", *IEEE Trans. Inform. Theory*, vol. IT-25, pp. 228-231, March 1979.
[6] A. El Gamal and M. H. M. Costa, "The capacity region of a class of deterministic interference channels", *IEEE Trans. Inform. Theory*, vol. IT-28, pp. 343-346, Mar. 1982.
[7] X. Shang, G. Kramer, and B. Chen, "A new outer bound and the noisy-interference sum-rate capacity for Gaussian interference channels", *IEEE Trans. Inform. Theory*, vol. 55, pp. 689-699, Feb. 2009.
[8] G. Scutari, D. P. Palomar, and S. Barbarossa, "Asynchronous iterative water-filling for Gaussian frequency-selective interference channels", *IEEE Trans. Inform. Theory*, vol. 54, pp. 2868-2878, July 2008.
[9] E. Biglieri and M. Lops, "Multiuser detection in a dynamic environment. Part I: User identification and data detection", *IEEE Trans. Inform. Theory*, vol. 53, pp. 3158-3170, Sep. 2007.
[10] P. Gupta and P. R. Kumar, "The capacity of wireless networks", *IEEE Trans. Inform. Theory*, vol. 46, pp. 388-404, March 2000.
[11] J. Y. N. Hui and P. A. Humblet, "The capacity region of the totally asynchronous multiple-access channel", *IEEE Trans. Inform. Theory*, vol. IT-31, pp. 207-216, Mar. 1985.
[12] S. Verdú and T. S. Han, "A general formula for channel capacity", *IEEE Trans. on Inform. Theory*, vol. 40, pp. 1147-1157, July 1994.
[13] R. S. Cheng and S. Verdú, "On limiting characterizations of memoryless multiuser capacity regions", *IEEE Trans. on Inform. Theory*, vol. 39, pp. 609-612, Mar. 1993.
[14] M. Effros, A. Goldsmith, and Y. Liang, "Capacity definitions for general channels with receiver side information", submitted to *IEEE Trans. Inform. Theory*, Apr. 2008. arXiv:0804.4239v1.
[15] S. N. Diggavi and T. M. Cover, "The worst additive noise under a covariance constraint", *IEEE Trans. Inform. Theory*, vol. 47, pp. 3072-3081, Nov. 2001.
[16] C. W. Helstrom, *Elements of signal detection and estimation*, Prentice Hall, 1995.
[17] M. H. M. Costa, "On the Gaussian interference channel", *IEEE Trans. Inform. Theory*, vol. IT-31, pp. 607-615, Sep. 1985.
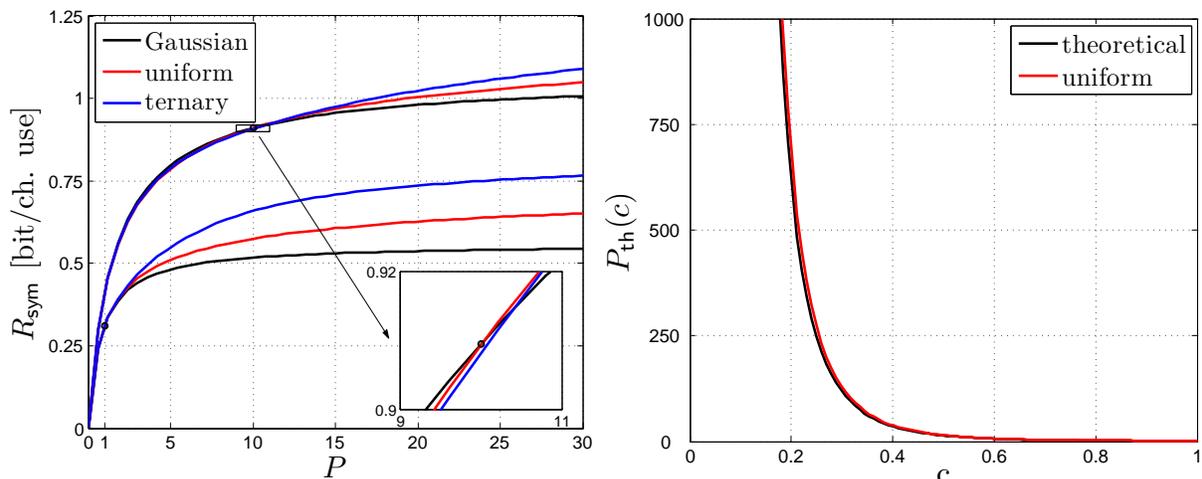
Fig. 2. Achievable symmetric rate of Gaussian-, uniformly-, and ternary-distributed codes for two different values of c yielding theoretical threshold powers of 1 and 10 (left). Comparison between the theoretical value of the threshold and the actual value of achieved by uniformly-distributed codes (right).