Outage Analysis for Two-User Parallel Gaussian Interference Channels

Ehsan Ebrahimzadeh, Kamyar Moshksar and Amir K. Khandani

Dept. of Electrical and Computer Eng., University of Waterloo

e-mail: {eebrahim@ucla.edu, kmoshksa@uwaterloo.ca, khandani@uwaterloo.ca}

***Abstract*—We address outage analysis for a two-user parallel Gaussian interference channel consisting of two sub-channels. Each sub-channel is modelled by a two-user Gaussian interference channel with quasi-static and flat fading. Both users employ single-layer Gaussian codebooks and maintain a statistical correlation ρ between the signals transmitted over the underlying sub-channels. If the receivers treat interference as noise (TIN) or cancel interference (CI), the value of ρ minimizing the outage probability approaches 1 as the signal-to-noise ratio (SNR) approaches infinity, while $\rho = 0$ is optimum under joint decoding (JD) regardless of the value of SNR. Motivated by these observations, we let $\rho = 1$ under TIN and CI and $\rho = 0$ under JD and compute the outage probability in finite SNR assuming the direct and crossover channel coefficients are independent zero-mean complex Gaussian random variables with possibly different variances. In the asymptote of large SNR and assuming the transmission rate per user is $r \log \mathrm{snr}$, we show that the outage probability scales as $\mathrm{snr}^{-(1-r)}$ under both TIN and CI, while it vanishes at least as fast as $\mathrm{snr}^{-\min\{2-r,4(1-r)\}} \log \mathrm{snr}$ under JD.**

I. Introduction

A. Motivation

Due to the growing demand for higher data rates, modern wireless communication systems are required to serve a large number of users that simultaneously share resources such as time or frequency. This has motivated a remarkable body of research that explore fundamental limits of communication in frequency-selective interference networks.

Fading selectivity in frequency can be leveraged to create diversity through applying coding across different frequency bands. In an interference channel, the effect of fading selectivity is twofold, as the underlying diversity can be exploited in either decoding the desired data, or handling interference to facilitate the decoding of the desired data. This paper addresses the interactions and trade-offs between these two possibilities for interference management in frequency selective fading interference networks.

The channel model adopted in this article is based on assigning each group of such distant tones, for which the corresponding channel gains are assumed to be independent, among multiple users, resulting in multi-user interference. Namely, we cast the problem into the framework of parallel Gaussian interference channels (PGICs) where several transmitter-receiver pairs simultaneously share a number of independent Gaussian interference channels (GICs).

In such a setup, it is reasonable to assume that the transmitters do not have access to channel state information (CSI). Moreover, due to slow fading over each sub-channel, transmitted codewords are not spanned over all fading states, casting the problem into the realm of "outage analysis". In this non-ergodic setting, the probability that a target transmission rate falls out of the achievable rate region at the receiver is of particular interest, which is referred to as the outage probability.

In practice, the receivers rely on either decoding the interference, or treating the interference as additive noise. The main reason is the simplicity of the receiver structure and lower complexity of system design. Moreover, such schemes are well-suited for practical situations where CSI is not available at the transmitter side. It would be of interest to study if adding the capability of joint decoding to the PGIC model studied in the current article, for which the transmitters are unaware of the CSI, can be helpful in improving the decay rate of the outage probability.

B. Summary of Prior Art and Contributions

Characterizing the capacity region of GICs and hence, PGICs remains an open problem in general. Exploiting previously known capacity results for GICs [1]–[3], the authors in [4] derive the capacity region of a two-user PGIC in the strong interference regime. Sufficient conditions are derived in [5] that characterize the sum-capacity of a two-user PGIC in the so-called noisy interference regime, where separate encoding over different GICs at the transmitter side and treating interference as noise at the receiver side is optimal. It is shown in [6] that separate coding over the underlying GICs is not necessarily a capacity achieving scheme. The vector GIC, or multiple-input multiple-output (MIMO) GIC, is studied in [7]–[9] in both noisy interference and strong interference regimes. In a more recent research paper [10], the authors derive general sufficient conditions for a vector GIC to be in the noisy interference regime capturing the previously known results in [7], [8].

In fading GICs where transmitters are unaware of the realizations of channel coefficients outage probability turns out to be the right performance measure. As SNR grows to infinity, the so-called diversity-multiplexing gain tradeoff (DMT) is a standard approach to study the outage probability. The DMT of a two-user GIC with fading is investigated in [11]–[15] under different scenarios in terms of channel coefficients and transmitter cooperation.

In this paper we consider a two-user PGIC consisting of two independent GICs. The channel coefficients in each GIC are modelled by quasi-static and flat fading. We study the outage probability for a simple transmission scheme where both users utilize single-layer Gaussian codebooks and maintain a statistical correlation of ρ between the signals transmitted simultaneously over the two GICs. It is shown that whether the receivers treat interference as noise or cancel interference, the optimum ρ that minimizes the outage probability per user approaches 1 as SNR grows to infinity. In contrast, $\rho = 0$ is optimum under joint decoding regardless of the value of SNR. Motivated by these observations, we study the outage probability in finite SNR under TIN, CI and JD and for their corresponding optimum correlation coefficient in a scenario where the channel gains represent Rayleigh fading. We determine closed form expressions for the exact probability of outage under TIN and CI that decay like snr^{-d} in the asymptote of large SNR for some $d > 0$. Closed form expressions for the outage probability seem elusive under JD, however, we are able to derive an upper bound on the probability of outage in terms of the modified Bessel functions of second kind. In particular, it is shown that the leading term in the expansion of this upper bound scales like $\mathrm{snr}^{-d'} \log \mathrm{snr}$ where $0 < d < d'$. To the authors' best knowledge, this paper addresses outage analysis in a PGIC for the first time.

C. Notations

Vectors are shown by an arrow on top, e.g. $\vec{x}$. Random quantities are shown in bold, e.g. $\boldsymbol{x}$ and $\vec{\boldsymbol{y}}$ with realizations x and $\vec{y}$, respectively. The probability density function (PDF), expectation and covariance matrix of a random vector $\vec{\boldsymbol{x}}$ are shown by $p_{\vec{\boldsymbol{x}}}(\cdot)$, $\mathbb{E}[\vec{\boldsymbol{x}}]$ and $\mathrm{cov}(\vec{\boldsymbol{x}})$, respectively. The transpose and transpose conjugate of a matrix X are denoted by X^{t} and $X^{\dagger}$, respectively. The Frobenius norm of a vector $\vec{x}$ is shown by $\|\vec{x}\| = \sqrt{\vec{x}^{\dagger}\vec{x}}$. The probability and the indicator function of an event $\mathcal{E}$ are shown by $\mathbb{P}(\mathcal{E})$ and $\mathbb{1}_{\mathcal{E}}$, respectively. For two function f and g, we write $f = o(g)$ to mean $\lim_{x\to a} \frac{f(x)}{g(x)} = 0$ where a is clear from the context. A circularly symmetric complex Gaussian random vector $\vec{\boldsymbol{x}}$ with mean $\vec{m}$ and covariance matrix C is shown by $\mathcal{CN}(\vec{m}, C)$. A vector of length n whose all entries are equal to 0 or 1 is shown by $\vec{0}_n$ and $\vec{1}_n$, respectively. Finally, any equality, inequality or limit involving random quantities is understood to hold in the "almost surely" sense unless otherwise stated.

D. Organization

System model and the signalling scheme are discussed in Section II. Section III offers an overview of our main contributions. The proofs of all the Theorems can be found in [16], nevertheless, we include the proof of Theorem 1 in Section IV to give the reader a flavor of the proofs. Concluding remarks appear in Section V.

II. System Model and the Signalling Scheme

Consider the two-user PGIC in Fig. 1 which consists of two GICs. The channels are modelled by static and non frequency-selective coefficients. The channel coefficient of the direct link for user i in GIC k is shown by $a_{i,k}$ and the crossover channel coefficient from transmitter j to receiver i ($i \neq j$) in GIC k is shown by $b_{i,k}$. Denoting the signal at receiver i over GIC k during a transmission slot by $\boldsymbol{y}_{i,k}$, one can write

$$\underbrace{\begin{bmatrix}\boldsymbol{y}_{i,1}\\ \boldsymbol{y}_{i,2}\end{bmatrix}}_{\vec{\boldsymbol{y}}_i} = \underbrace{\begin{bmatrix}a_{i,1} & 0\\ 0 & a_{i,2}\end{bmatrix}}_{A_i} \underbrace{\begin{bmatrix}\boldsymbol{x}_{i,1}\\ \boldsymbol{x}_{i,2}\end{bmatrix}}_{\vec{\boldsymbol{x}}_i} + \underbrace{\begin{bmatrix}b_{i,1} & 0\\ 0 & b_{i,2}\end{bmatrix}}_{B_i} \underbrace{\begin{bmatrix}\boldsymbol{x}_{j,1}\\ \boldsymbol{x}_{j,2}\end{bmatrix}}_{\vec{\boldsymbol{x}}_j} + \underbrace{\begin{bmatrix}\boldsymbol{z}_{i,1}\\ \boldsymbol{z}_{i,2}\end{bmatrix}}_{\vec{\boldsymbol{z}}_i} \tag{1}$$

for $(i,j) \in \{(1,2),(2,1)\}$, where $\boldsymbol{x}_{i,k}$ is the signal sent by transmitter i over GIC k and $\boldsymbol{z}_{i,k} \sim \mathcal{CN}(0,1)$ is the additive ambient noise at receiver i over GIC k. The noise samples $\boldsymbol{z}_{i,1}$ and $\boldsymbol{z}_{i,2}$ are independent. Throughout the paper, we make the following assumptions:

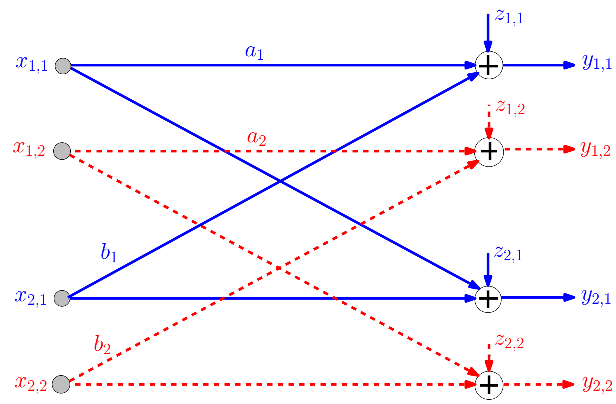

Fig. 1. A two-user PGIC consisting of two underlying GICs.

Assumption 1 *We focus on a "fair" scenario where the transmission rates and average transmission powers at both transmitters are identical.*

Denoting the average transmission power per transmitter by P, it is required that

$$\mathbb{E}\left[|\boldsymbol{x}_{i,1}|^2\right] = \mathbb{E}\left[|\boldsymbol{x}_{i,2}|^2\right] = P/2, \quad i = 1,2. \tag{2}$$

Each user utilizes a single-layer random Gaussian codebook. The signals transmitted over each GIC are independent from transmission slot to transmission slot, however, the signals transmitted over the two GICs at the same transmission slot have a correlation ρ, i.e.,

$$\vec{\boldsymbol{x}}_i \sim \mathcal{CN}\left(\begin{bmatrix}0\\0\end{bmatrix}, \frac{P}{2}\begin{bmatrix}1 & \rho\\ \rho^* & 1\end{bmatrix}\right), \quad i = 1,2, \quad |\rho| \leq 1. \tag{3}$$

III. Summary of Results

In this paper, we study three different decoding schemes at the receivers, i.e., treating interference as noise (TIN), cancelling interference (CI) and joint decoding (JD). The achievable rate region under decoding scheme S is shown by $\mathcal{R}^{(\mathsf{S})}(\rho, P, H)$ where S can be either TIN, CI or JD and

$$H = [H_1 \quad H_2], \quad H_i = \begin{bmatrix}a_{i,1} & b_{i,1}\\ a_{i,2} & b_{i,2}\end{bmatrix}, \quad i = 1,2. \tag{4}$$

In practice, the matrix H_i is a realization of a random matrix $\boldsymbol{H}_i$. For technical reasons, we make the following assumption:

Assumption 2 *$\boldsymbol{H}_1$ and $\boldsymbol{H}_2$ are independent random matrices, each having a probability density function.*

We consider a scenario where both transmitters are unaware of H_1 and H_2, while, receiver i has perfect knowledge of H_i. Denoting the transmission rate per user by $r \log P$ for $0 \leq r < 1$ and $P > 1$, the outage event under decoding scheme S is defined by

$$\mathcal{O}^{(\mathsf{S})}(\rho, r, P) = \{(r \log P, r \log P) \notin \mathcal{R}^{(\mathsf{S})}(\rho, P, \boldsymbol{H})\}. \tag{5}$$

Let us define $\rho^{(\mathsf{S})}(r,P)$ as the value of ρ that minimizes the outage probability under decoding scheme S, i.e.,

$$\rho^{(\mathsf{S})}(r,P) = \arg\min_{\rho:|\rho|\leq 1} \mathbb{P}(\mathcal{O}^{(\mathsf{S})}(\rho,r,P)). \tag{6}$$

Using "min" in (6) is meaningful, if $\mathbb{P}(\mathcal{O}^{(\mathsf{S})}(\rho,r,P))$ is a continuous function of ρ. In [16], we find a continuous function $R^{(\mathsf{S})}$ for $\mathsf{S} = \text{TIN, CI and JD}$, respectively, so that[1]

$$\begin{aligned}\mathbb{P}(\mathcal{O}^{(\mathsf{S})}(\rho,r,P)) = \\ \mathbb{P}\left(R^{(\mathsf{S})}(\rho,P,\boldsymbol{H}_1)\leq R'\right) + \mathbb{P}\left(R^{(\mathsf{S})}(\rho,P,\boldsymbol{H}_2)\leq R'\right) \\ -\mathbb{P}\left(R^{(\mathsf{S})}(\rho,P,\boldsymbol{H}_1)\leq R'\right)\mathbb{P}\left(R^{(\mathsf{S})}(\rho,P,\boldsymbol{H}_2)\leq R'\right),\end{aligned} \tag{7}$$

where $R' = r\log P$. Due to continuity of $R^{(\mathsf{S})}$, $\lim_{n\to\infty} R^{(\mathsf{S})}(\rho_n,P,\boldsymbol{H}_1) = R^{(\mathsf{S})}(\rho,P,\boldsymbol{H}_1)$ for any $0\leq\rho\leq 1$ and any sequence ρ_n converging to ρ. Therefore, the sequence of random variables $R^{(\mathsf{S})}(\rho_n,P,\boldsymbol{H}_1)$ converges weakly[2] to $R^{(\mathsf{S})}(\rho,P,\boldsymbol{H}_1)$ as n goes to infinity. Therefore, if

$$\mathbb{P}\left(R^{(\mathsf{S})}(\rho,P,\boldsymbol{H}_1) = R'\right) = 0, \tag{8}$$

then

$$\lim_{n\to\infty}\mathbb{P}\left(R^{(\mathsf{S})}(\rho_n,P,\boldsymbol{H}_1)\leq R'\right) = \mathbb{P}\left(R^{(\mathsf{S})}(\rho,P,\boldsymbol{H}_1)\leq R'\right). \tag{9}$$

Since ρ and the sequence ρ_n converging to ρ are arbitrary,[3] $\mathbb{P}\left(R^{(\mathsf{S})}(\rho,P,\boldsymbol{H}_1)\leq R'\right)$ is a continuous function of ρ by (9). Inspecting the explicit expressions for $R^{(\mathsf{S})}$ given in [16] and representing H_1 as a vector in $\mathbb{C}^4$, it is immediate to see that the level sets of the function $R^{(\mathsf{S})}(\rho,P,\cdot)$ have Lebesgue measure zero. Moreover, according to Assumption 2, $\boldsymbol{H}_1$ is a random matrix with density. As such,

$$\mathbb{P}\left(R^{(\mathsf{S})}(\rho,P,\boldsymbol{H}_1) = R'\right) = \int_{R^{(\mathsf{S})}(\rho,P,H_1)=R'} p_{\boldsymbol{H}_1}(H_1)\mathrm{d}H_1 = 0,$$

i.e., the sufficient condition in (8) holds. Similarly, one can show that $\mathbb{P}\left(R^{(\mathsf{S})}(\rho,P,\boldsymbol{H}_2)\leq R'\right)$ is continuous in terms of ρ. Therefore, the infimum of $\mathbb{P}(\mathcal{O}^{(\mathsf{S})}(\rho,r,P))$ is achieved over the compact region $|\rho|\leq 1$.

For arbitrary $P>1$, characterizing $\rho^{(\mathsf{S})}(r,P)$ in closed form turns out to be a difficult problem under TIN and CI. In this paper, we only study the effect of ρ on the outage probability in the asymptote of large P for these schemes. The first contribution of the paper is that under TIN and CI, transmitting the same signal over both GICs is "optimal" as P grows to infinity, while transmitting independent signals over the two GICs is optimal under JD regardless of r and P.

Theorem 1 *Let $0\leq r<1$. Under Assumption 2 in above and regardless of S being TIN or CI,*

(i) $\lim_{P\to\infty}\mathbb{P}(\mathcal{O}^{(\mathsf{S})}(\rho,r,P)) = \mathbb{1}_{0\leq\rho<1}$.

(ii) $\lim_{P\to\infty}\rho^{(\mathsf{S})}(r,P) = 1$.

Moreover $\rho^{(\mathrm{JD})}(r,P) = 0$ for any $P>1$.

Proof: We prove this for $\mathsf{S} = \text{TIN}$ in section IV. The other two cases are treated in [16]. ■

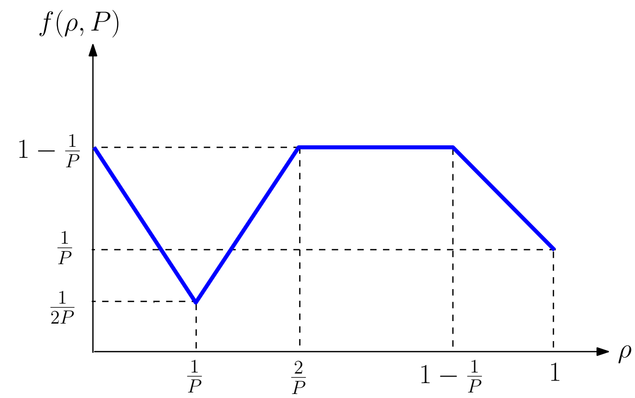

Fig. 2. An example of a function $f:[0,1]\times(1,\infty)\to[0,1]$ such that $\lim_{P\to\infty} f(\rho,P) = \mathbb{1}_{0\leq\rho<1}$, however, $\arg\min_{0\leq\rho\leq 1} f(\rho,P) = \frac{1}{P}$ which tends to 0 as P grows to infinity.

Remark 1- One may understand the importance of $\rho=1$ at high SNR as follows. Let $\boldsymbol{x}_i$ be the signal transmitted by user i over both underlying GICs. Then (1) can be written as

$$\vec{\boldsymbol{y}} = \boldsymbol{x}_1\vec{a} + \boldsymbol{x}_2\vec{b} + \vec{\boldsymbol{z}}, \tag{10}$$

where $\vec{a} = [a_{1,1}\quad a_{1,2}]^{\mathrm{t}}$ and $\vec{b} = [b_{1,1}\quad b_{1,2}]^{\mathrm{t}}$. By assumption, $\vec{a},\vec{b}\neq\vec{0}_2$ and $\det(H_1)\neq 0$ for almost all realizations H_1 of $\boldsymbol{H}_1$. Hence, receiver 1 can find a vector $\vec{b}_\perp$, say $\vec{b}_\perp = [b_{1,2}\quad -b_{1,1}]^{\mathrm{t}}$, such that $\vec{b}_\perp^{\mathrm{t}}\vec{b} = 0$ and $\vec{b}_\perp^{\mathrm{t}}\vec{a}\neq 0$. Multiplying both sides of (10) by $\vec{b}_\perp^{\mathrm{t}}$, we get $\vec{b}_\perp^{\mathrm{t}}\vec{\boldsymbol{y}} = \vec{b}_\perp^{\mathrm{t}}\vec{a}\boldsymbol{x}_1 + \vec{b}_\perp^{\mathrm{t}}\vec{\boldsymbol{z}}$ which represents a point-to-point channel with mutual information $\log\left(1+\frac{|\vec{b}_\perp^{\mathrm{t}}\vec{a}|^2}{\|\vec{b}_\perp\|^2}\frac{P}{2}\right)$. This quantity scales like $\log P$. As such, one expects the outage probability to vanish in the asymptote of large P for any $0\leq r<1$. □

Remark 2- In Theorem 1, part (ii) is not a direct consequence of part (i). For example, the function $f:[0,1]\times(1,\infty)\to[0,1]$ shown in Fig. 2 is such that $\lim_{P\to\infty} f(\rho,P) = \mathbb{1}_{0\leq\rho<1}$, however, $\arg\min_{0\leq\rho\leq 1} f(\rho,P) = \frac{1}{P}$ which tends to 0 as P grows to infinity. Nevertheless, we use part (i) to prove part (ii). □

Motivated by Theorem 1, we fix $\rho=1$ under TIN and CI and $\rho=0$ under JD in order to compute the outage probability under the assumption that the channel coefficients represent Rayleigh fading:

Theorem 2 *Let all the channel coefficients be independent and the direct and crossover channel coefficients be realizations of $\mathcal{CN}(0,1)$ and $\mathcal{CN}(0,\sigma^2)$, respectively. For any $0\leq r<1$,*

$$\mathbb{P}\left(\mathcal{O}^{(\mathrm{TIN})}(1,r,P)\right) = (4+o(1))P^{-(1-r)}$$

and

$$\mathbb{P}\left(\mathcal{O}^{(\mathrm{CI})}(1,r,P)\right) = \left(\frac{4}{\sigma^2}+o(1)\right)P^{-(1-r)}.$$

[1] The quantity $R^{(\mathsf{S})}(\rho,P,H_i)$ is an achievable rate for user i under $\mathsf{S} = \text{TIN}$, however, there is no such interpretation for $\mathsf{S} = \text{CI}$ and $\mathsf{S} = \text{JD}$.

[2] Almost sure convergence of a sequence of real-valued random variables implies weak convergence [18]. We say $\boldsymbol{X}_n$ converges weakly to a random variable $\boldsymbol{X}$, if $\lim_{n\to\infty}\mathbb{P}(\boldsymbol{X}_n\in C) = \mathbb{P}(\boldsymbol{X}\in C)$ for any Borel set C with topological boundary ∂C such that $\mathbb{P}(\boldsymbol{X}\in\partial C) = 0$.

[3] For a function $f:\mathbb{R}\to\mathbb{R}$, $\lim_{x\to x_0} f(x) = L$ if and only if $\lim_{n\to\infty} f(x_n) = L$ for any sequence x_n satisfying $\lim_{n\to\infty} x_n = a$.

Moreover,

$$\mathbb{P}(\mathcal{O}^{(\mathrm{JD})}(0,r,P)) \le (8(2-r)+o(1))P^{-(2-r)}\ln P,$$

for $0 \le r < \frac{2}{3}$ *and*

$$\mathbb{P}(\mathcal{O}^{(\mathrm{JD})}(0,r,P)) \le \left(\frac{32}{\sigma^4}(1-r)+o(1)\right)P^{-4(1-r)}\ln P,$$

for $\frac{2}{3} < r < 1$.

Proof: See [16]. ■

It is worth mentioning that if the two users are orthogonal, i.e., user 1 only transmits over GIC 1 and user 2 only transmits over GIC 2, the outage probability scales like $2P^{-(1-r)} + o(P^{-(1-r)})$ [17]. Comparing this with the performance under JD in Theorem 2 verifies the advantage of transmitting over both GICs compared to avoiding interference.

Remark 3- Characterizing the optimum correlation coefficient ρ for a two-user PGIC with an arbitrary number $N > 2$ of parallel GICs is a hard problem, even in the scenario where the receivers treat interference as noise. It is shown in [16] that the probability of the achievable rate per user having a local minimum at $\rho = 0$ approaches 1 in the asymptote of large P. Moreover, assuming the channel coefficients represent Rayleigh fading, the outage probability is computed for $\rho = 1$ in [16]. It is verified that the outage probability is given by $\frac{2N^{N-1}}{(N-1)!}P^{-(N-1)(1-r)} + o(P^{-(N-1)(1-r)})$ for any $0 \le r < 1$. This simplifies to (11) for $N = 2$.

Remark 4- For simplicity of presentation, we drop the index i and show $a_{i,j}$ and $b_{i,j}$ by a_j and b_j, respectively, throughout the rest of the paper. It is always clear from the context that the omitted index i is $i = 1$ or $i = 2$. □

IV. Proof of Theorem 1 for $\mathsf{S} = \mathrm{TIN}$

Assuming users treat each other as Gaussian noise, an achievable rate for user 1 is given by

$$R^{(\mathrm{TIN})}(\rho,P,H_1) = \log\frac{\det(\mathrm{cov}(\vec{\boldsymbol{y}}_1))}{\det(\mathrm{cov}(B_1\,\vec{\boldsymbol{x}}_2+\vec{\boldsymbol{z}}_1))}. \tag{11}$$

This can be expanded as

$$R^{(\mathrm{TIN})}(\rho,P,H_1) = \log\frac{\alpha(P,H_1)-\beta(P,H_1)|\rho|^2}{\gamma(P,H_1)-\delta(P,H_1)|\rho|^2}, \tag{12}$$

where

$$\alpha(P,H_1) = (1+\frac{P}{2}(|a_1|^2+|b_1|^2))(1+\frac{P}{2}(|a_2|^2+|b_2|^2)),$$
$$\beta(P,H_1) = \frac{P^2}{4}|a_1a_2^*+b_1b_2^*|^2,$$
$$\gamma(P,H_1) = (1+\frac{P}{2}|b_1|^2)(1+\frac{P}{2}|b_2|^2),$$
$$\delta(P,H_1) = \frac{P^2}{4}|b_1|^2|b_2|^2.$$

By (12), without loss of generality we may assume $\rho \in [0,1]$. By Assumption 2, $\boldsymbol{H}_1$ and $\boldsymbol{H}_2$ are independent and hence, $\mathbb{P}(\mathcal{O}^{(\mathrm{TIN})}(\rho,r,P)) = \mathbb{P}(R^{(\mathrm{TIN})}(\rho,P,\boldsymbol{H}_1) \le R'$ or $R^{(\mathrm{TIN})}(\rho,P,\boldsymbol{H}_2) \le R')$ can be expanded as in (7). Next, let us observe the following:

(a) If $\lim_{P\to\infty}\mathbb{P}(R^{(\mathrm{TIN})}(\rho,P,\boldsymbol{H}_i) \le R') = \mathbb{1}_{0\le\rho<1}$ for $i = 1,2$, then $\lim_{P\to\infty}\mathbb{P}(\mathcal{O}^{(\mathrm{TIN})}(\rho,r,P)) = \mathbb{1}_{0\le\rho<1}$ as well.

(b) The function $(x,y) \mapsto x+y-xy$ for $0 \le x,y \le 1$ is increasing in terms of x and y separately. Therefore, if we can show that the value of ρ minimizing $\mathsf{p}_i = \mathbb{P}(R^{(\mathrm{TIN})}(\rho,P,\boldsymbol{H}_i) \le R')$ approaches 1 as P grows to infinity regardless of $i = 1,2$, then the value of ρ minimizing $\mathsf{p}_1+\mathsf{p}_2-\mathsf{p}_1\mathsf{p}_2$ also approaches 1 as P increases.

Thus, it is enough to show that parts (i) and (ii) in Theorem 1 hold for $\mathbb{P}(R^{(\mathrm{TIN})}(\rho,P,\boldsymbol{H}_1) \le R')$ and $\mathbb{P}(R^{(\mathrm{TIN})}(\rho,P,\boldsymbol{H}_2) \le R')$ in place of $\mathbb{P}(\mathcal{O}^{(\mathrm{TIN})}(\rho,r,P))$. Here we only consider $i = 1$; the case $i = 2$ is treated similarly.

Proof of part (i) in Theorem 1 for $\mathsf{S} = \mathrm{TIN}$: We consider the cases $\rho = 1$ and $\rho < 1$ separately: First, suppose $\rho = 1$ in (12). Then $R^{(\mathrm{TIN})}(1,P,H_1)$ equals

$$\log\left(1+\frac{\frac{P}{2}\left(|a_1|^2+|a_2|^2\right)+\frac{P^2}{4}|\det(H_1)|^2}{1+\frac{P}{2}\left(|b_1|^2+|b_2|^2\right)}\right).$$

Since $\boldsymbol{b}_1,\boldsymbol{b}_2,\det(\boldsymbol{H}_1) \neq 0$, then $\lim_{P\to\infty}\frac{R^{(\mathrm{TIN})}(1,P,\boldsymbol{H}_1)}{\log P} = 1$. This shows that $R^{(\mathrm{TIN})}(1,P,\boldsymbol{H}_1)-R$ scales like $(1-r)\log P$ as P grows to infinity. Since $0 \le r < 1$, we get

$$\lim_{P\to\infty}\mathbb{1}_{R^{(\mathrm{TIN})}(1,P,\boldsymbol{H}_1)\le R'} = 0. \tag{13}$$

For $\rho < 1$, by (12), $R^{(\mathrm{TIN})}(\rho,P,H_1)$ equals

$$\log\frac{1+\frac{P}{2}c+\frac{P^2}{4}\left(c'-\rho^2|a_1a_2^*+b_1b_2^*|^2\right)}{1+\frac{P}{2}\left(|b_1|^2+|b_2|^2\right)+\frac{P^2}{4}|b_1|^2|b_2|^2(1-\rho^2)},$$

where $c = |a_1|^2+|a_2|^2+|b_1|^2+|b_2|^2$ and $c' = (|a_1|^2+|b_1|^2)(|a_2|^2+|b_2|^2)$. We have $(|\boldsymbol{a}_1|^2+|\boldsymbol{b}_1|^2)(|\boldsymbol{a}_2|^2+|\boldsymbol{b}_2|^2) - |\boldsymbol{a}_1\boldsymbol{b}_2^*+\boldsymbol{a}_2\boldsymbol{b}_1^*|^2 = |\boldsymbol{a}_1\boldsymbol{b}_2-\boldsymbol{a}_2\boldsymbol{b}_1|^2 = |\det(\boldsymbol{H}_1)|^2 > 0$. Since $\rho < 1$, we get $(|\boldsymbol{a}_1|^2+|\boldsymbol{b}_1|^2)(|\boldsymbol{a}_2|^2+|\boldsymbol{b}_2|^2) - \rho^2|\boldsymbol{a}_1\boldsymbol{b}_2^*+\boldsymbol{a}_2\boldsymbol{b}_1^*|^2 > 0$ as well. By assumption, $\boldsymbol{b}_1,\boldsymbol{b}_2 \neq 0$ and hence, $(1-\rho^2)|\boldsymbol{b}_1|^2|\boldsymbol{b}_2|^2 > 0$. Therefore, $R^{(\mathrm{TIN})}(\rho,P,\boldsymbol{H}_1)$ does not scale with $\log P$, i.e., $R^{(\mathrm{TIN})}(\rho,P,\boldsymbol{H}_1) - r\log P$ scales like $-r\log P$. So,

$$\lim_{P\to\infty}\mathbb{1}_{R^{(\mathrm{TIN})}(\rho,P,\boldsymbol{H}_1)\le r\log P} = 1. \tag{14}$$

Using (13) and (14) together with dominated convergence [18], the proof of (i) is complete.

Proof of part (ii) in Theorem 1 for $\mathsf{S} = \mathrm{TIN}$: Denote the value of ρ that minimizes $\mathbb{P}(R^{(\mathrm{TIN})}(\rho,P,\boldsymbol{H}_1) \le R')$ by $\rho_1^{(\mathrm{TIN})}(r,P_n)$. It is enough to prove

$$\lim_{n\to\infty}\rho_1^{(\mathrm{TIN})}(r,P_n) = 1, \tag{15}$$

where P_n is an arbitrary increasing and unbounded sequence of positive real numbers. Let us fix $0 < \epsilon < 1$. Since any probability measure on $\mathbb{C}^4$ is compact-regular [18], one can find a compact set $\mathcal{H}_\epsilon \subseteq \mathbb{C}^4$ such that $\mathbb{P}(\mathrm{vec}(\boldsymbol{H}_1) \in \mathcal{H}_\epsilon) \ge \epsilon$ where $\mathrm{vec}(\boldsymbol{H}_1)$ is a vector obtained by stacking the columns of $\boldsymbol{H}_1$ in a single column. Define the event $Q = \{\mathrm{vec}(\boldsymbol{H}_1) \in$

$\mathcal{H}_\epsilon\}$. Then

$$\begin{aligned}&\mathbb{P}\left(R^{(\mathrm{TIN})}(\rho, P_n, \boldsymbol{H}_1) \le r\log P_n\right)\\ &\ge \mathbb{P}\left(R^{(\mathrm{TIN})}(\rho, P_n, \boldsymbol{H}_1) \le r\log P_n, Q\right)\\ &= \mathbb{P}\left(R^{(\mathrm{TIN})}(\rho, P_n, \boldsymbol{H}_1) \le r\log P_n \,|\, Q\right)\mathbb{P}(Q)\\ &\ge \epsilon\,\mathbb{P}\left(R^{(\mathrm{TIN})}(\rho, P_n, \boldsymbol{H}_1) \le r\log P_n \,|\, Q\right). \end{aligned} \tag{16}$$

For $0 \le \rho \le 1$, we define

$$p_n(\rho) = \mathbb{P}\left(R^{(\mathrm{TIN})}(\rho, P_n, \boldsymbol{H}_1) \le r\log P_n \,|\, Q\right).$$

We note the following:

(a) $p_n(\rho)$ is continuous in terms of ρ for any n. This follows from similar lines of reasoning after (6) where we verified continuity of $\mathbb{P}(R^{(\mathrm{TIN})}(\rho, P, \boldsymbol{H}_1) \le R')$ in terms of ρ.

(b) There is $N_\epsilon \ge 1$ such that for $n \ge N_\epsilon$, $p_{n+1}(\rho) \ge p_n(\rho)$ for any $0 \le \rho \le \epsilon$. To see this, we observe that $\frac{\partial}{\partial P}\left(R^{(\mathrm{TIN})}(\rho, P, \boldsymbol{H}_1) - R'\right) < 0$ if and only if

$$\mu_4 P^4 + \mu_3 P^3 + \mu_2 P^2 + \mu_1 P - r < 0, \tag{17}$$

where $\mu_i = \mu_i(\rho, r, \boldsymbol{H}_1)$ for $i = 1, 2, 3, 4$ are polynomials in terms of r, ρ and real and imaginary parts of the channel coefficient. In particular, $\mu_4(\rho, r, \boldsymbol{H}_1) = -r\big((|\boldsymbol{a}_1|^2 + |\boldsymbol{b}_1|^2)(|\boldsymbol{a}_2|^2 + |\boldsymbol{b}_2|^2) - \rho^2|\boldsymbol{a}_1\boldsymbol{a}_2^* + \boldsymbol{b}_1\boldsymbol{b}_2^*|^2\big)|\boldsymbol{b}_1|^2|\boldsymbol{b}_2|^2(1-\rho^2)$. Since $\mu_4(\rho, r, \boldsymbol{H}_1)$ is the coefficient of P^4 which is the term with the largest exponent of P on the left side of (17) and $\mu_4(\rho, r, \boldsymbol{H}_1) < 0$, it follows that there is $P(\rho, r, \boldsymbol{H}_1) > 0$ such that for $P > P(\rho, r, \boldsymbol{H}_1)$, the inequality in (17) is valid. Since the roots of any polynomial are continuous functions of the coefficients of that polynomial, $P(\rho, r, H_1)$ is a continuous function of μ_is and hence, it is a continuous function of ρ and H_1. As $[0, \epsilon] \times \mathcal{H}_\epsilon$ is compact, $\sup_{0\le\rho\le\epsilon, \mathrm{vec}(H_1)\in\mathcal{H}_\epsilon} P(\rho, r, H_1)$ is finite and one can find $N_\epsilon \ge 1$ so that $P_n > \sup_{0\le\rho\le\epsilon, \mathrm{vec}(H_1)\in\mathcal{H}_\epsilon} P(\rho, r, H_1)$ holds for any $n \ge N_\epsilon$. Then it is guaranteed that the sequence $R^{(\mathrm{TIN})}(\rho, P_n, \boldsymbol{H}_1) - r\log P_n$ is decreasing in n as long as $0 \le \rho \le \epsilon$, $\boldsymbol{H}_1 \in \mathcal{H}_\epsilon$ and $n \ge N_\epsilon$. In turn, it follows that $p_n(\rho) \le p_{n+1}(\rho)$ for any $0 \le \rho \le \epsilon$ and $n \ge N_\epsilon$, as claimed.

(c) By Theorem 1(i), $\lim_{n\to\infty} p_n(\rho) = 1$ for any $0 \le \rho \le \epsilon$.

Putting these three observations together, $p_n(\cdot)$ for $n \ge N_\epsilon$ is an increasing sequence of continuous functions (in terms of ρ) that converges point-wise to the constant 1 over the compact interval $[0, \epsilon]$. Applying Dini's uniform convergence lemma [18], this point-wise convergence is indeed uniform, i.e.,

$$\lim_{n\to\infty} \inf_{0\le\rho\le\epsilon} p_n(\rho) = 1. \tag{18}$$

By (18), there exists $N'_\epsilon \ge 1$ such that if $n \ge N'_\epsilon$, then $\inf_{0\le\rho\le\epsilon} p_n(\rho) > \epsilon$. Combining this fact and (16), we obtain

$$\begin{aligned}\inf_{0\le\rho\le\epsilon} \mathbb{P}\left(R^{(\mathrm{TIN})}(\rho, P_n, \boldsymbol{H}_1) \le r\log P_n\right) &\ge \epsilon \inf_{0\le\rho\le\epsilon} p_n(\rho)\\ &\ge \epsilon^2, \end{aligned} \tag{19}$$

for any $n \ge N'_\epsilon$. Moreover, by part (i) in Theorem 1, there exists $N''_\epsilon \ge 1$ such that if $n \ge N''_\epsilon$,

$$\mathbb{P}\left(R^{(\mathrm{TIN})}(1, P_n, \boldsymbol{H}_1) \le r\log P_n\right) < \epsilon^2. \tag{20}$$

Combining (19) and (20), we conclude that

$$\rho_1^{(\mathrm{TIN})}(r, P_n) > \epsilon, \tag{21}$$

for any $n \ge \max\{N'_\epsilon, N''_\epsilon\}$. Since $0 < \epsilon < 1$ is arbitrary, (21) is equivalent to (15). This completes the proof of part (ii).

References

[1] A. Motahari and A. Khandani, "Capacity bounds for the Gaussian interference channel", *IEEE Trans. Inf. Theory*, vol. 55, no. 2, pp. 620–643, 2009.

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

[3] V. Annapureddy and V. Veeravalli, "Gaussian interference networks: Sum capacity in the low-interference regime and new outer bounds on the capacity region", *IEEE Trans. Inf. Theory*, vol. 55, no. 7, pp. 3032–3050, 2009.

[4] S. T. Chung and J. Cioffi, "The capacity region of frequency-selective Gaussian interference channels under strong interference", *IEEE Trans. Commun.*, vol. 55, no. 9, pp. 1812–1821, Sept. 2007.

[5] X. Shang, B. Chen, G. Kramer, and H. Poor, "Noisy-interference sum-rate capacity of parallel Gaussian interference channels", *IEEE Trans. Inf. Theory*, vol. 57, no. 1, pp. 210–226, 2011.

[6] V. Cadambe and S. Jafar, "Parallel Gaussian interference channels are not always separable", *IEEE Trans. Inf. Theory*, vol. 55, no. 9, pp. 3983–3990, Sept. 2009.

[7] X. Shang, B. Chen, G. Kramer and H. V. Poor, "Capacity regions and sum-rate capacities of vector Gaussian interference channels", *IEEE Trans. Inf. Theory*, vol. 56, no. 10, pp. 5030-5044, Oct. 2010.

[8] V. S. Annapureddy and V. V. Veeravalli, "Sum capacity of MIMO interference channels in the low interference regime", *IEEE Trans. Inf. Theory*, vol. 57, no. 5, pp. 2565-2581, May 2011.

[9] X. Shang, H. V. Poor, "Capacity region of vector Gaussian interference channels with generally strong interference", *IEEE Trans. Inf. Theory*, vol. 58, no. 6, pp. 3472-3496, June 2012.

[10] X. Shang, H. V. Poor, "Noisy-interference sum-rate capacity for vector Gaussian interference channels", *IEEE Trans. Inf. Theory*, vol. 59, no. 1, pp. 132-153, Jan. 2013.

[11] E. Akuiyibo and O. Leveque, "Diversity-multiplexing trade off for the slow fading interference channel", *IEEE Int. Zurich Seminar on Commun.* pp. 140–143, 2008.

[12] C. Akcaba and H. Bolcskei, "Diversity-multiplexing tradeoff in two-user fading interference channels", *IEEE Trans. Inf. Theory*, vol. 58, no. 7, pp. 4462–4480, 2012.

[13] A. Raja and P. Viswanath, "Diversity-multiplexing tradeoff of the two-user interference channel", *IEEE Trans. Inf. Theory*, vol. 57, no. 9, pp. 5782–5793, 2011.

[14] Y. Weng and D. Tuninetti, "Outage analysis of block-fading Gaussian interference channels", *IEEE Trans. Inf. Theory*, vol. 57, no. 10, pp. 6487–6501, 2011.

[15] H. Ebrahimzad and A. Khandani, "On the optimum diversity multiplexing tradeoff of the two-user Gaussian interference channel with Rayleigh fading", *IEEE Trans. Inf. Theory*, vol. 58, no. 7, pp. 4481–4492, 2012.

[16] E. Ebrahimzadeh, K. Moshksar and A. Khandani, "Signalling over two-user parallel Gaussian interference channels: outage analysis", submitted, 2015, available at `http://eebrahim.bol.ucla.edu/Files/PGIC.pdf`.

[17] D. Tse, P. Viswanath, "Fundamentals of Wireless Communication" *Cambridge university press*, 2005.

[18] R. Dudley, *Real analysis and probability*. Cambridge Uni. Press, 2005.