Signalling Over Two-User Parallel Gaussian Interference Channels: Outage Analysis

Ehsan Ebrahimzadeh, Kamyar Moshksar and Amir K. Khandani

Coding & Signal Transmission Laboratory (www.cst.uwaterloo.ca)

Dept. of Elec. and Comp. Eng., University of Waterloo

Waterloo, ON, Canada, N2L 3G1

Tel: 519-725-7338, Fax: 519-888-4338

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

Abstract

This paper addresses 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 $r \log \mathrm{snr}$, it is shown that the outage probability scales like $\mathrm{snr}^{-(1-r)}$ 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. The paper is concluded by extending some of the results to an arbitrary number of sub-channels.

Index Terms

Interference Cancellation, Joint Decoding, Outage Probability, Parallel Gaussian Interference Channel, Treating Interference as Noise.

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 (or time) can be leveraged to create diversity through applying coding across different frequency bands (or across different time symbols). In an interference channel, the effect of fading selectivity is twofold, as the underlying diversity can be exploited in: (i) decoding the desired data, or (ii) handling the interference to facilitate the decoding of the desired data. The current article addresses the interactions and trade-offs between these two possibilities for interference management in frequency selective fading interference networks.

Orthogonal frequency division multiplexing (OFDM) converts a frequency selective static channel into a set of sub-channels, each with a flat and fixed channel gain. In orthogonal frequency division multiple access (OFDMA), sub-channels are distributed among transmitter-receiver pairs in such a way that distant enough sub-channels are allocated to a given transmitter-receiver pair. For example, in a system with $T = AB$ tones shared among A users, a given user can be assigned tones which are at least A units apart. Such an allocation helps enhancing frequency diversity and reducing the inter-career interference among sub-channels assigned to a given transmitter-receiver pair. 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 order to justify the assumption of independent channel gains over different sub-channels, we note that the channel impulse response can be manipulated to induce and/or enhance selectivity, which can be in turn exploited to increase diversity, thereby leading to a lower outage probability. A common method for enhancing frequency selectivity, widely used in practice [1]–[3], is based on the so-called delay diversity in OFDM/OFDMA. Practically, delay diversity results in large variations of the channel gains among different tones (albeit still dependent). This feature, in conjunction with allocation of distant tones to each user, is the motivation for considering interference channels with independent fading model. Another method to enhance selectivity in time is based on using multiple transmit antennas and applying a different phase shift to each antenna in subsequent time symbols [4]. Again, by multiplexing A users in AB units of time, a given user can be assigned time symbols which are at least A units apart, providing further grounds for the PGIC model considered in this article.

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 [5]–[7], the authors in [8] derive the capacity region of a two-user PGIC in the strong interference regime. Sufficient conditions are derived in [9] 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 [10] 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 [11]–[13] in both noisy interference and strong interference regimes. In a more recent research paper [14], the authors derive general sufficient conditions for a vector GIC to be in the noisy interference regime capturing the previously known results in [11], [12].

In fading GICs where transmitters are unaware of the realizations of channel coefficients outage probability turns out to be the right performance measure. Computing outage probability in finite SNR can be a challenging task. For example, a conjecture made in [15] regarding the outage probability in a single-user MIMO channel is only partially answered [16], [17]. As SNR grows to infinity, the so-called diversity-multiplexing gain tradeoff (DMT) coined in [18] is a standard approach to study the outage probability. Motivated by [19] where the capacity region of a two-user GIC is determined to within one bit, the DMT of a two-user GIC with fading is investigated in [20]–[25] 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

Here is a list of notations adopted throughout the paper. Vectors are shown by an arrow on top such as $\vec{x}$. Random quantities are shown in bold such as $\boldsymbol{x}$ 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. An $n \times n$ diagonal matrix with diagonal elements $x_1, \cdots, x_n$ is denoted by $\mathrm{diag}(x_1, \cdots, x_n)$. 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 function f is said to be $\Theta(1)$ if $c_1 \leq |f(x)| \leq c_2$ for all $x \geq x_0$ where x_0 and $c > 0$ are constants. 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

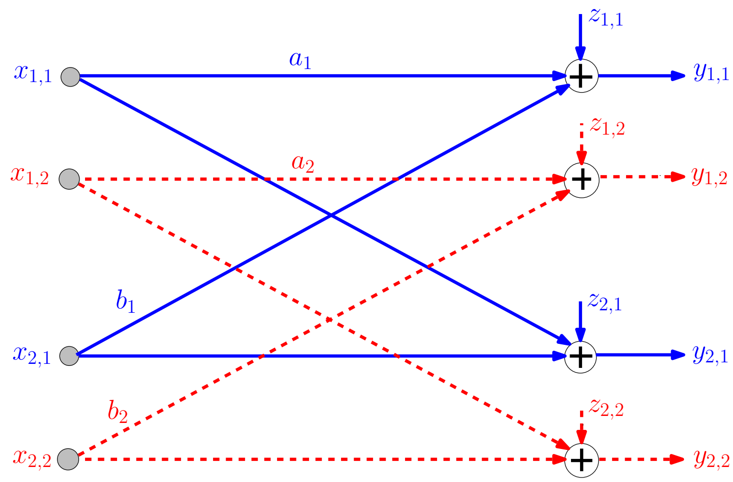


Fig. 1. A two-user PGIC consisting of two underlying GICs.

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*

The rest of the paper is organized as follows. System model and the signalling scheme are discussed in Section II. Section III offers an overview of contributions made in the paper. Sections IV, V and VI are devoted to prove Theorem 1 for TIN, CI and JD, respectively. In Section VII, we investigate the interactions between the variance of crossover channel coefficients and the correlation coefficient ρ among the transmitted signals in determining the outage behaviour of the system. Section VIII is an attempt to extend the results to a PGIC with more than two parallel GICs in a setup where both users apply TIN. Finally, Section IX provides a summary of the paper and offers some observations in a PGIC with more than two users.

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

$$\begin{bmatrix} \boldsymbol{y}_{i,1} \\ \boldsymbol{y}_{i,2} \end{bmatrix} = \begin{bmatrix} a_{i,1} & 0 \\ 0 & a_{i,2} \end{bmatrix} \begin{bmatrix} \boldsymbol{x}_{i,1} \\ \boldsymbol{x}_{i,2} \end{bmatrix} + \begin{bmatrix} b_{i,1} & 0 \\ 0 & b_{i,2} \end{bmatrix} \begin{bmatrix} \boldsymbol{x}_{j,1} \\ \boldsymbol{x}_{j,2} \end{bmatrix} + \begin{bmatrix} \boldsymbol{z}_{i,1} \\ \boldsymbol{z}_{i,2} \end{bmatrix}, \quad 1 \leq i,j \leq 2, \ i \neq j, \tag{1}$$

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.

One can alternatively write (1) as

$$\vec{\boldsymbol{y}}_i = A_i \, \vec{\boldsymbol{x}}_1 + B_i \, \vec{\boldsymbol{x}}_j + \vec{\boldsymbol{z}}_i, \quad 1 \leq i, j \leq 2, \; i \neq j, \tag{2}$$

where $A_i = \mathrm{diag}(a_{i,1}, a_{i,2})$ and $B_i = \mathrm{diag}(b_{i,1}, b_{i,2})$. The definitions for $\vec{\boldsymbol{x}}_i, \vec{\boldsymbol{x}}_j, \vec{\boldsymbol{y}}_i$ and $\vec{\boldsymbol{z}}_i$ are clear by comparing (1) and (2). Throughout the paper, we make the following assumption:

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

$$\mathbb{E}\left[|\boldsymbol{x}_{i,1}|^2\right] = \mathbb{E}\left[|\boldsymbol{x}_{i,2}|^2\right] = \frac{P}{2}, \;\; i = 1, 2. \tag{3}$$

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.,

$$\vec{\boldsymbol{x}}_i \sim \mathcal{CN}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \frac{P}{2}\begin{bmatrix} 1 & \rho \\ \rho^* & 1 \end{bmatrix}\right), \;\; i = 1, 2, \quad |\rho| \leq 1. \tag{4}$$

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

$$H = \begin{bmatrix} H_1 & H_2 \end{bmatrix}, \quad H_i = \begin{bmatrix} a_{i,1} & b_{i,1} \\ a_{i,2} & b_{i,2} \end{bmatrix}, \;\; i = 1, 2. \tag{5}$$

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 $0 \le r < 1$ and $P > 1$, the outage event under decoding scheme S is defined by

$$\mathcal{O}^{(\mathsf{S})}(\rho, r, P) = \{(r\log P, r\log P) \notin \mathcal{R}^{(\mathsf{S})}(\rho, P, \boldsymbol{H})\}. \tag{6}$$

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

$$\rho^{(\mathsf{S})}(r, P) = \arg\min_{\rho:|\rho|\le 1} \mathbb{P}(\mathcal{O}^{(\mathsf{S})}(\rho, r, P)). \tag{7}$$

Using "min" in (7) is meaningful, if $\mathbb{P}(\mathcal{O}^{(\mathsf{S})}(\rho, r, P))$ is a continuous function of ρ. In Sections IV, V and VI, we find a continuous function $R^{(\mathsf{S})}$ for $\mathsf{S} = \mathrm{TIN}, \mathrm{CI}$ and JD, respectively, so that[1]

$$\begin{aligned}\mathbb{P}(\mathcal{O}^{(\mathsf{S})}(\rho, r, P)) &= \mathbb{P}\left(R^{(\mathsf{S})}(\rho, P, \boldsymbol{H}_1) \le r\log P\right) + \mathbb{P}\left(R^{(\mathsf{S})}(\rho, P, \boldsymbol{H}_2) \le r\log P\right) \\ &\quad -\mathbb{P}\left(R^{(\mathsf{S})}(\rho, P, \boldsymbol{H}_1) \le r\log P\right)\mathbb{P}\left(R^{(\mathsf{S})}(\rho, P, \boldsymbol{H}_2) \le r\log P\right).\end{aligned} \tag{8}$$

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 $0 \le \rho \le 1$ 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 grows to infinity. Therefore, if

$$\mathbb{P}\left(R^{(\mathsf{S})}(\rho, P, \boldsymbol{H}_1) = r\log P\right) = 0, \tag{9}$$

then

$$\lim_{n\to\infty} \mathbb{P}\left(R^{(\mathsf{S})}(\rho_n, P, \boldsymbol{H}_1) \le r\log P\right) = \mathbb{P}\left(R^{(\mathsf{S})}(\rho, P, \boldsymbol{H}_1) \le r\log P\right). \tag{10}$$

Since ρ and the sequence ρ_n converging to ρ are arbitrary,[3] $\mathbb{P}\left(R^{(\mathsf{S})}(\rho, P, \boldsymbol{H}_1) \le r\log P\right)$ is a continuous function of ρ by (10). Inspecting the explicit expressions for $R^{(\mathsf{S})}$ given in Sections IV, V and VI 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.

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

[2]Almost sure convergence of a sequence of real-valued random variables implies weak convergence [28]. 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 $\lim_{n\to\infty} x_n = a$ [29].

As such,

$$\mathbb{P}\left(R^{(\mathsf{S})}(\rho,P,\boldsymbol{H}_1)=r\log P\right)=\int_{R^{(\mathsf{S})}(\rho,P,H_1)=r\log P}p_{\boldsymbol{H}_1}(H_1)\mathrm{d}H_1=0, \tag{11}$$

i.e., the sufficient condition in (9) holds. Similarly, one can show that $\mathbb{P}\left(R^{(\mathsf{S})}(\rho,P,\boldsymbol{H}_2)\le r\log P\right)$ 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 $|\rho|\le 1$.

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

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

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

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

Proof: See Sections IV, V and VI for $\mathsf{S}=\mathrm{TIN},\mathrm{CI}$ and JD, respectively. ■

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 (2) can be written as

$$\vec{\boldsymbol{y}}=\boldsymbol{x}_1\vec{a}+\boldsymbol{x}_2\vec{b}+\vec{\boldsymbol{z}}, \tag{12}$$

where $\vec{a}=\begin{bmatrix}a_{1,1} & a_{1,2}\end{bmatrix}^{\mathrm{t}}$ and $\vec{b}=\begin{bmatrix}b_{1,1} & b_{1,2}\end{bmatrix}^{\mathrm{t}}$. 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 $\vec{b}_\perp=\begin{bmatrix}b_{1,2} & -b_{1,1}\end{bmatrix}^{\mathrm{t}}$, such that $\vec{b}_\perp^{\mathrm{t}}\vec{b}=0$ and $\vec{b}_\perp^{\mathrm{t}}\vec{a}\neq 0$. Multiplying both sides of (12) 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 $0\le r<1$. □

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 $\lim_{P\to\infty}f(\rho,P)=\mathbb{1}_{0\le\rho<1}$, however, $\arg\min_{0\le\rho\le1}f(\rho,P)=\frac{1}{P}$ which tends to 0 as P grows to infinity. Nevertheless, we use part (i) to prove

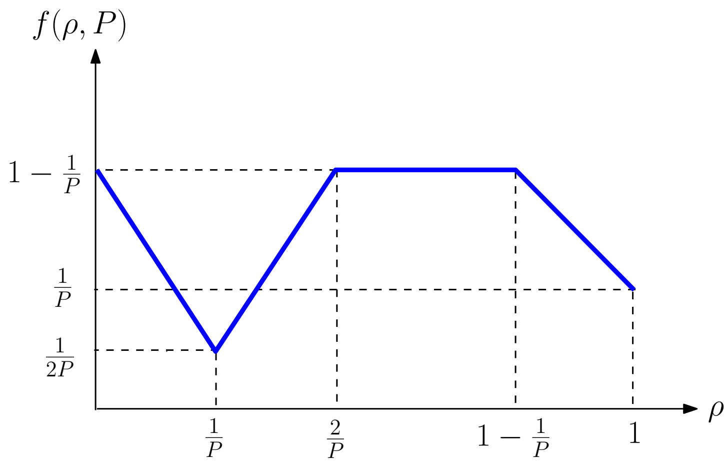


Fig. 2. An example of a function $f : [0,1] \times (1,\infty) \to [0,1]$ such that $\lim_{P\to\infty} f(\rho, P) = \mathbb{1}_{0\le\rho<1}$, however, $\arg\min_{0\le\rho\le1} f(\rho, P) = \frac{1}{P}$ which tends to 0 as P grows to infinity.

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 $0 \le r < 1$,*

$$\mathbb{P}\left(\mathcal{O}^{(\mathrm{TIN})}(1,r,P)\right) = 4P^{-(1-r)} + o\left(P^{-(1-r)}\right) \tag{13}$$

and

$$\mathbb{P}\left(\mathcal{O}^{(\mathrm{CI})}(1,r,P)\right) = \frac{4}{\sigma^2}P^{-(1-r)} + o\left(P^{-(1-r)}\right). \tag{14}$$

Moreover,

$$\mathbb{P}(\mathcal{O}^{(\mathrm{JD})}(0,r,P)) \le 8(2-r)P^{-(2-r)}\ln P + o(P^{-(2-r)}\ln P), \tag{15}$$

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

$$\mathbb{P}(\mathcal{O}^{(\mathrm{JD})}(0,r,P)) \le \frac{32}{\sigma^4}(1-r)P^{-4(1-r)}\ln P + o(P^{-4(1-r)}\ln P), \tag{16}$$

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

Proof: See Section VII. ■

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 $2P^{-(1-r)} + o(P^{-(1-r)})$ [26]. Comparing this with the performance under JD in Theorem 2 verifies the advantage of transmitting over both GICs compared to avoiding interference.

In Section VIII, we investigate the effect of difference in distribution for the direct and crossover channel coefficients on the value of ρ. Due to simplicity of the receiver structure and lower complexity of system design, we only consider a scenario where the receivers treat interference as noise. Motivated by the fact that $\rho = 0$ is optimal if the crossover channel coefficients are zero, i.e., users do not interfere with each other, we explore conditions such that $\rho = 0$ is still optimal in the presence of interference. A typical scenario for the setup considered in this article corresponds to a cellular system where distant cells cause interference to each other. We raise the following question:

Assume the interfering users are distant enough so that the crossover channel gain σ^2 and SNR level P satisfy $P^{1+\epsilon}\sigma^2 < 1$ for some fixed $\epsilon > 0$. What is the optimum value of ρ?

This question is answered in Proposition 3 in the asymptote of large P. We consider a sequence of PGICs where the SNR and the crossover channel gain in the PGIC of index n are P_n and σ_n^2, respectively, such that $P_n^{1+\epsilon}\sigma_n^2 < 1$ for any $n \geq 1$. Then we study the outage probability in terms of ρ assuming P_n grows to infinity as the index n grows. The condition $P^{1+\epsilon}\sigma^2 < 1$ in the question above can be replaced by $P\sigma^2 < o(1)$ where $o(1)$ is any function of P that vanishes as P grows to infinity. For example, one can replace $P^{1+\epsilon}\sigma^2 < 1$ by $P\sigma^2 < \frac{1}{\ln P}$. The assumption $\lim_{n\to\infty} P_n\sigma_n^2 = 0$ in Theorem 3 includes all such cases. In Proposition 3, it is shown that for $0 \leq \rho < 1$, the outage probability vanishes for any $0 \leq r < 2$ as n grows. However, if $\rho = 1$, the outage probability vanishes for $0 \leq r < 1$, but approaches 1 for any $1 < r < 2$ as n tends to infinity.

Finally, Section IX offers analogous results as in Theorem 1 and Theorem 2 for a two-user PGIC with an arbitrary number $N \geq 2$ of parallel GICs where the receivers treat interference as noise. In Proposition 2, it is shown that the probability of the achievable rate per user having a local minimum at $\rho = 0$ approaches 1 in the asymptote of large P. In Proposition 3, the outage probability is computed for $\rho = 1$ assuming the channel coefficients represent Rayleigh fading. 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 $0 \leq r < 1$. This simplifies to (13) for $N = 2$.

Remark 3- 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 [27]

$$R^{(\mathrm{TIN})}(\rho, P, H_1) = \log \frac{\det(\mathrm{cov}(\vec{\boldsymbol{y}}_1))}{\det(\mathrm{cov}(B_1\, \vec{\boldsymbol{x}}_2 + \vec{\boldsymbol{z}}_1))}. \tag{17}$$

This can be expanded as

$$R^{(\mathrm{TIN})}(\rho, P, H_1) = \log \frac{\alpha(P, H_1) - \beta(P, H_1)|\rho|^2}{\gamma(P, H_1) - \delta(P, H_1)|\rho|^2}, \tag{18}$$

where

$$\begin{aligned} \alpha(P, H_1) &= (1 + \tfrac{P}{2}(|a_1|^2 + |b_1|^2))(1 + \tfrac{P}{2}(|a_2|^2 + |b_2|^2)) \\ \beta(P, H_1) &= \tfrac{P^2}{4}|a_1 a_2^* + b_1 b_2^*|^2 \\ \gamma(P, H_1) &= (1 + \tfrac{P}{2}|b_1|^2)(1 + \tfrac{P}{2}|b_2|^2) \\ \delta(P, H_1) &= \tfrac{P^2}{4}|b_1|^2|b_2|^2 \end{aligned} . \tag{19}$$

By (18), we can assume $0 \leq \rho \leq 1$ without loss of generality. By Assumption 2, $\boldsymbol{H}_1$ and $\boldsymbol{H}_2$ are independent and hence, $\mathbb{P}(\mathcal{O}^{(\mathrm{TIN})}(\rho, r, P)) = \mathbb{P}(R^{(\mathrm{TIN})}(\rho, P, \boldsymbol{H}_1) \leq r \log P$ or $R^{(\mathrm{TIN})}(\rho, P, \boldsymbol{H}_2) \leq r \log P)$ can be expanded as in (8). Next, let us observe the following:

- If $\lim_{P\to\infty} \mathbb{P}(R^{(\mathrm{TIN})}(\rho, P, \boldsymbol{H}_i) \leq r \log P) = \mathbb{1}_{0\leq\rho<1}$ for $i = 1, 2$, then $\lim_{P\to\infty} \mathbb{P}(\mathcal{O}^{(\mathrm{TIN})}(\rho, r, P)) = \mathbb{1}_{0\leq\rho<1}$ as well.
- The function $(x, y) \mapsto x + y - xy$ for $0 \leq x, y \leq 1$ is increasing in terms of x and y separately. Therefore, if we can show that the value of ρ minimizing $\mathsf{p}_i = \mathbb{P}(R^{(\mathrm{TIN})}(\rho, P, \boldsymbol{H}_i) \leq r \log P)$ 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.

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

(i) **Proof of part (i) in Theorem 1 for** $\mathsf{S} = \mathrm{TIN}$: We consider the cases $\rho = 1$ and $\rho < 1$ separately:

- Setting $\rho = 1$ in (18),

$$R^{(\mathrm{TIN})}(1, P, H_1) = \log \left(1 + \frac{\frac{P}{2}\left(|a_1|^2 + |a_2|^2\right) + \frac{P^2}{4}|\det(H_1)|^2}{1 + \frac{P}{2}\left(|b_1|^2 + |b_2|^2\right)} \right). \tag{20}$$

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

$$\lim_{P\to\infty} \mathbb{1}_{R^{(\mathrm{TIN})}(1,P,\boldsymbol{H}_1)\le r\log P} = 0. \tag{21}$$

- For $\rho < 1$ one can write $R^{(\mathrm{TIN})}(\rho, P, H_1)$ in (18) as

$$R^{(\mathrm{TIN})}(\rho, P, H_1) = \log \frac{1 + \frac{P}{2}c + \frac{P^2}{4}\left((|a_1|^2+|b_1|^2)(|a_2|^2+|b_2|^2) - \rho^2|a_1a_2^* + b_1b_2^*|^2\right)}{1+\frac{P}{2}\left(|b_1|^2+|b_2|^2\right) + \frac{P^2}{4}|b_1|^2|b_2|^2(1-\rho^2)}, \tag{22}$$

where $c = |a_1|^2+|a_2|^2+|b_1|^2+|b_2|^2$. 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 $-r\log P$. As such,

$$\lim_{P\to\infty} \mathbb{1}_{R^{(\mathrm{TIN})}(\rho,P,\boldsymbol{H}_1)\le r\log P} = 1. \tag{23}$$

Using (21) and (23) together with dominated convergence [28], the proof of (i) is complete.

(ii) **Proof of part (ii) in Theorem 1 for** $\mathsf{S} = \mathrm{TIN}$: Let us denote the value of ρ that minimizes $\mathbb{P}(R^{(\mathrm{TIN})}(\rho, P, \boldsymbol{H}_1) \le r\log P)$ by $\rho_1^{(\mathrm{TIN})}(r, P_n)$. Based on the remark in Footnote 3, it is enough to show that

$$\lim_{n\to\infty} \rho_1^{(\mathrm{TIN})}(r, P_n) = 1, \tag{24}$$

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 [28], 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. Then

$$\begin{aligned}
&\mathbb{P}\left(R^{(\mathrm{TIN})}(\rho, P_n, \boldsymbol{H}_1) \le r\log P_n\right) \ge \mathbb{P}\left(R^{(\mathrm{TIN})}(\rho, P_n, \boldsymbol{H}_1) \le r\log P_n, \mathrm{vec}(\boldsymbol{H}_1) \in \mathcal{H}_\epsilon\right)\\
&= \mathbb{P}\left(R^{(\mathrm{TIN})}(\rho, P_n, \boldsymbol{H}_1) \le r\log P_n \,|\, \mathrm{vec}(\boldsymbol{H}_1) \in \mathcal{H}_\epsilon\right)\mathbb{P}(\mathrm{vec}(\boldsymbol{H}_1) \in \mathcal{H}_\epsilon)\\
&\ge \epsilon\, \mathbb{P}\left(R^{(\mathrm{TIN})}(\rho, P_n, \boldsymbol{H}_1) \le r\log P_n \,|\, \mathrm{vec}(\boldsymbol{H}_1) \in \mathcal{H}_\epsilon\right).
\end{aligned} \tag{25}$$

Let us define

$$p_n(\rho) = \mathbb{P}\left(R^{(\mathrm{TIN})}(\rho, P_n, \boldsymbol{H}_1) \le r\log P_n \,|\, \mathrm{vec}(\boldsymbol{H}_1) \in \mathcal{H}_\epsilon\right), \;\; 0 \le \rho \le 1. \tag{26}$$

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 (7) where we verified continuity of $\mathbb{P}(R^{(\mathrm{TIN})}(\rho, P, \boldsymbol{H}_1) \le r \log P)$ 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 $\frac{\partial}{\partial P}\left(R^{(\mathrm{TIN})}(\rho, P, \boldsymbol{H}_1) - r \log P\right) < 0$ if and only if

$$\mu_4(\rho, r, \boldsymbol{H}_1)P^4 + \mu_3(\rho, r, \boldsymbol{H}_1)P^3 + \mu_2(\rho, r, \boldsymbol{H}_1)P^2 + \mu_1(\rho, r, \boldsymbol{H}_1)P - r < 0, \tag{27}$$

where $\mu_i(\rho, r, \boldsymbol{H}_1)$ 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 (27) 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 (27) 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 terms of 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 part (i) of Theorem 1, $\lim_{n\to\infty} p_n(\rho) = 1$ 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 [29], this point-wise convergence is indeed uniform, i.e.,

$$\lim_{n\to\infty} \inf_{0 \le \rho \le \epsilon} p_n(\rho) = 1. \tag{28}$$

By (28), there exists $N'_\epsilon \ge 1$ such that if $n \ge N'_\epsilon$, then $\inf_{0 \le \rho \le \epsilon} p_n(\rho) > \epsilon$. Using this fact together with (25), we obtain

$$\inf_{0 \le \rho \le \epsilon} \mathbb{P}\left(R^{(\mathrm{TIN})}(\rho, P_n, \boldsymbol{H}_1) \le r \log P_n\right) \ge \epsilon \inf_{0 \le \rho \le \epsilon} p_n(\rho) \ge \epsilon^2, \tag{29}$$

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

$$\mathbb{P}\left(R^{(\mathrm{TIN})}(1, P_n, \boldsymbol{H}_1) \leq r \log P_n\right) < \epsilon^2. \tag{30}$$

Combining (29) and (30), we conclude that for $n \geq \max\{N'_\epsilon, N''_\epsilon\}$,

$$\rho_1^{(\mathrm{TIN})}(r, P_n) > \epsilon. \tag{31}$$

Since $0 < \epsilon < 1$ is arbitrary, (31) is equivalent to (24). This completes the proof of part (ii).

V. Proof of Theorem 1 for $\mathsf{S} = \mathrm{CI}$

Under interference cancellation, receiver 1 proceeds according to the following steps:

1) Receiver 1 decodes the message of user 2 by treating its own signal as additive Gaussian noise. To guarantee successful decoding of interference, the transmission rate of user 2 must be less than $R'(\rho, P, H_1)$ defined as

$$\begin{aligned} R'(\rho, P, H_1) &\triangleq \log \frac{\det(\mathrm{cov}(\vec{\boldsymbol{y}}_1))}{\det(\mathrm{cov}(A_1 \vec{\boldsymbol{x}}_1 + \vec{\boldsymbol{z}}_1))} \\ &= \log \frac{\alpha(P, H_1) - \beta(P, H_1)|\rho|^2}{\gamma'(P, H_1) - \delta'(P, H_1)|\rho|^2}, \end{aligned} \tag{32}$$

where $\alpha(P, H_1)$ and $\beta(P, H_1)$ are given in (19) and

$$\begin{aligned} \gamma'(P, H_1) &= (1 + \tfrac{P}{2}|a_1|^2)(1 + \tfrac{P}{2}|a_2|^2) \\ \delta'(P, H_1) &= \tfrac{P^2}{4}|a_1|^2|a_2|^2 \end{aligned}. \tag{33}$$

2) After cancelling the additive interference, receiver 1 decodes its own message. To guarantee successful decoding in this step, the transmission rate of user 1 must be less than $R''(\rho, P, H_1)$ defined as

$$\begin{aligned} R''(\rho, P, H_1) &\triangleq \log \frac{\det(\mathrm{cov}(A_1 \vec{\boldsymbol{x}}_1 + \vec{\boldsymbol{z}}_1))}{\det(\mathrm{cov}(\vec{\boldsymbol{z}}_1))} \\ &= \log\left(\gamma'(P, H_1) - \delta'(P, H_1)|\rho|^2\right). \end{aligned} \tag{34}$$

Recalling that the actual transmission rate per user is $r \log P$, receiver 1 successfully decodes the message sent by transmitter 1 if

$$r \log P < R^{(\mathrm{CI})}(\rho, P, H_1), \tag{35}$$

where

$$R^{(\mathrm{CI})}(\rho, P, H_1) = \min\left\{R'(\rho, P, H_1), R''(\rho, P, H_1)\right\}. \tag{36}$$

Note that one can restrict $0 \le \rho \le 1$. By Assumption 2, $\boldsymbol{H}_1$ and $\boldsymbol{H}_2$ are independent and hence, $\mathbb{P}(\mathcal{O}^{(\mathrm{TIN})}(\rho, r, P)) = \mathbb{P}(R^{(\mathrm{CI})}(\rho, P, \boldsymbol{H}_1) \le r \log P$ or $R^{(\mathrm{CI})}(\rho, P, \boldsymbol{H}_2) \le r \log P)$ can be expanded as in(8). Following similar lines of reasoning presented after (19) at the beginning of Section IV, it is enough to show that parts (i) and (ii) in Theorem 1 hold for $\mathbb{P}(R^{(\mathrm{CI})}(\rho, P, \boldsymbol{H}_1) \le r \log P)$ and $\mathbb{P}(R^{(\mathrm{CI})}(\rho, P, \boldsymbol{H}_2) \le r \log P)$ instead of $\mathbb{P}(\mathcal{O}^{(\mathrm{CI})}(\rho, r, P))$. Here, we only consider $i = 1$ as the case $i = 2$ is treated similarly.

(i) **Proof of part (i) in Theorem 1 for** $\mathsf{S} = \mathrm{CI}$: We consider the cases $\rho = 1$ and $\rho < 1$ separately:

- If $\rho = 1$,

$$R'(1, P, H_1) = \log\left(1 + \frac{\frac{P}{2}\left(|b_1|^2 + |b_2|^2\right) + \frac{P^2}{4}|\det(H_1)|^2}{1 + \frac{P}{2}\left(|a_1|^2 + |a_2|^2\right)}\right) \tag{37}$$

  and

$$R''(1, P, H_1) = \log\left(1 + \frac{P}{2}(|a_1|^2 + |a_2|^2)\right). \tag{38}$$

  As $\boldsymbol{a}_1, \boldsymbol{a}_2, \det(\boldsymbol{H}_1) \neq 0$, both $R'(1, P, \boldsymbol{H}_1)$ and $R''(1, P, \boldsymbol{H}_1)$ scale like $\log P$ and therefore, $R^{(\mathrm{CI})}(1, P, \boldsymbol{H}_1) - r \log P$ scales like $(1 - r) \log P$. This yields

$$\lim_{P \to \infty} \mathbb{1}_{R^{(\mathrm{CI})}(1, P, \boldsymbol{H}_1) \le r \log P} = 0. \tag{39}$$

- If $0 \le \rho < 1$, $R'(\rho, P, H_1)$ is given by the expression on the right side of (22) with a_1 and a_2, replaced by b_1 and b_2, respectively and vice versa. Following a similar reasoning to the one after (22), we conclude that $R'(\rho, P, \boldsymbol{H}_1)$ does not scale with $\log P$. Moreover,

$$R''(\rho, P, H_1) = \log\left(1 + \frac{P}{2}(|a_1|^2 + |a_2|^2) + \frac{P^2}{4}|a_1|^2|a_2|^2(1 - \rho^2)\right). \tag{40}$$

  Since $\boldsymbol{a}_1, \boldsymbol{a}_2 \neq 0$, then $R''(\rho, P, \boldsymbol{H}_1)$ scales like $2 \log P$ and hence, $R^{(\mathrm{CI})}(\rho, P, \boldsymbol{H}_1) - r \log P$ scales like $-r \log P$. This yields

$$\lim_{P \to \infty} \mathbb{1}_{R^{(\mathrm{CI})}(\rho, P, \boldsymbol{H}_1) \le r \log P} = 1. \tag{41}$$

Using (39) and (41) together with dominated convergence [28], the proof of part (i) is complete.

(ii) **Proof of part (ii) in Theorem 1 for** $\mathsf{S} = \mathrm{CI}$: We start with the following lemma:

Lemma 1 *For any* $0 < \epsilon < 1$,

$$\lim_{P\to\infty} \mathbb{P}\left(R^{(\mathrm{CI})}(\rho, P, \boldsymbol{H}_1) = R'(\rho, P, \boldsymbol{H}_1) \textit{ for any } 0 \le \rho \le \epsilon\right) = 1. \tag{42}$$

Proof: See Appendix A. ■

Let P_n be an arbitrary increasing and unbounded sequence of positive real numbers. By Lemma 1, for any $0 < \epsilon < 1$, there exists $N_\epsilon \ge 1$ such that if $n \ge N_\epsilon$, then

$$\mathbb{P}\big(R^{(\mathrm{CI})}(\rho, P_n, \boldsymbol{H}_1) = R'(\rho, P_n, \boldsymbol{H}_1) \textit{ for any } 0 \le \rho \le \epsilon\big) \ge \epsilon. \tag{43}$$

In particular, this yields $\mathbb{P}\big(R^{(\mathrm{CI})}(\rho, P_n, \boldsymbol{H}_1) = R'(\rho, P_n, \boldsymbol{H}_1)\big) \ge \epsilon$ for any $0 \le \rho \le \epsilon$ and $n \ge N_\epsilon$ and hence,

$$\inf_{0\le\rho\le\epsilon} \mathbb{P}\big(R^{(\mathrm{CI})}(\rho, P_n, \boldsymbol{H}_1) = R'(\rho, P_n, \boldsymbol{H}_1)\big) \ge \epsilon, \quad n \ge N_\epsilon. \tag{44}$$

Let us write

$$\begin{aligned}
&\mathbb{P}\big(R^{(\mathrm{CI})}(\rho, P_n, \boldsymbol{H}_1) \le r \log P_n\big) \\
&\quad\ge \mathbb{P}\big(R^{(\mathrm{CI})}(\rho, P_n, \boldsymbol{H}_1) \le r \log P_n, R^{(\mathrm{CI})}(\rho, P_n, \boldsymbol{H}_1) = R'(\rho, P_n, \boldsymbol{H}_1)\big) \\
&\quad= \mathbb{P}\big(R'(\rho, P_n, \boldsymbol{H}_1) \le r \log P_n, R^{(\mathrm{CI})}(\rho, P_n, \boldsymbol{H}_1) = R'(\rho, P_n, \boldsymbol{H}_1)\big) \\
&\quad\ge \mathbb{P}\big(R^{(\mathrm{CI})}(\rho, P_n, \boldsymbol{H}_1) \le r \log P_n\big) \\
&\qquad+\mathbb{P}\big(R^{(\mathrm{CI})}(\rho, P_n, \boldsymbol{H}_1) = R'(\rho, P_n, \boldsymbol{H}_1)\big) - 1,
\end{aligned} \tag{45}$$

where the last step follows by the bound $\mathbb{P}(\mathcal{E} \bigcap \mathcal{F}) \ge \mathbb{P}(\mathcal{E}) + \mathbb{P}(\mathcal{F}) - 1$ for any two events $\mathcal{E}$ and $\mathcal{F}$. By (45) and for any $n \ge N_\epsilon$,

$$\begin{aligned}
\inf_{0\le\rho\le\epsilon} \mathbb{P}\big(R^{(\mathrm{CI})}(\rho, P_n, \boldsymbol{H}_1) \le r \log P_n\big) &\ge \inf_{0\le\rho\le\epsilon} \mathbb{P}\big(R'(\rho, P_n, \boldsymbol{H}_1) \le r \log P_n\big) \\
&\quad + \inf_{0\le\rho\le\epsilon} \mathbb{P}\big(R^{(\mathrm{CI})}(\rho, P_n, \boldsymbol{H}_1) = R'(\rho, P_n, \boldsymbol{H}_1)\big) - 1 \\
&\ge \inf_{0\le\rho\le\epsilon} \mathbb{P}\big(R'(\rho, P_n, \boldsymbol{H}_1) \le r \log P_n\big) + \epsilon - 1,
\end{aligned} \tag{46}$$

where the last step is due to (44). Due to similarity of $R'(\rho, P_n, \boldsymbol{H}_1)$ and $R^{(\mathrm{TIN})}(\rho, P_n, \boldsymbol{H}_1)$, one can use (29) to conclude that there exists $N'_\epsilon \ge 1$ such that

$$\inf_{0\le\rho\le\epsilon} \mathbb{P}\big(R'(\rho, P_n, \boldsymbol{H}_1) \le r \log P_n\big) \ge \epsilon^2, \tag{47}$$

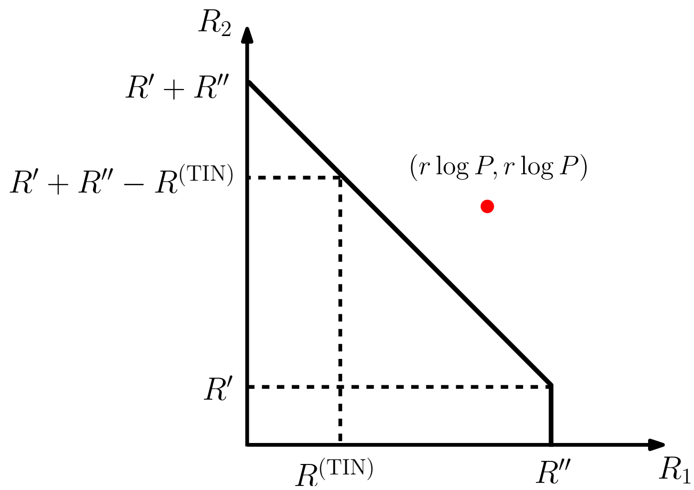


Fig. 3. Achievable region for the multiple access channel from the two transmitters to receiver 1. The constraint $R_2 < R'(\rho, P, H_1) + R''(\rho, P, H_1) - R^{(\mathrm{TIN})}(\rho, P, H_1)$ on the transmission rate of user 2 is immaterial as it concerns the event that only the message sent by transmitter 2 is decoded incorrectly by receiver 1.

for any $n \geq N'_\epsilon$. Putting (46) and (47) together,

$$\inf_{0 \leq \rho \leq \epsilon} \mathbb{P}\big(R^{(\mathrm{CI})}(\rho, P_n, \boldsymbol{H}_1) \leq r \log P_n\big) \geq \epsilon^2 + \epsilon - 1. \tag{48}$$

Let $\epsilon > \frac{\sqrt{5}-1}{2}$ so that $\epsilon^2 + \epsilon - 1 > 0$. By part (i) in Theorem 1, there exists $N''_\epsilon \geq 1$ such that

$$\mathbb{P}\big(R^{(\mathrm{CI})}(1, P_n, \boldsymbol{H}_1) \leq r \log P_n\big) < \epsilon^2 + \epsilon - 1, \tag{49}$$

for any $n \geq N''_\epsilon$. Comparing (48) and (49) and assuming $n \geq \max\{N_\epsilon, N'_\epsilon, N''_\epsilon\}$, the value of ρ that minimizes $\mathbb{P}\big(R^{(\mathrm{CI})}(\rho, P_n, \boldsymbol{H}_1) \leq r \log P_n\big)$ is larger than ϵ. Since this is true for any $\frac{\sqrt{5}-1}{2} < \epsilon < 1$, the proof is complete by letting ϵ approach 1.

VI. Proof of Theorem 1 for $\mathsf{S} = \mathrm{JD}$

Denoting the transmission rate of user i by R_i, the achievable region of the multiple access channel from the two transmitters to receiver 1 is given by

$$\begin{cases} R_1 < R''(\rho, P, H_1) \\ R_2 < R'(\rho, P, H_1) + R''(\rho, P, H_1) - R^{(\mathrm{TIN})}(\rho, P, H_1) \\ R_1 + R_2 < R'(\rho, P, H_1) + R''(\rho, P, H_1) \end{cases}. \tag{50}$$

From the viewpoint of receiver 1, the second constraint in (50) is immaterial as it concerns the event that only the message sent by transmitter 2 is decoded incorrectly. As such, the rate region under JD is

$$\begin{cases} R_1 < R''(\rho, P, H_1) \\ R_1 + R_2 < R'(\rho, P, H_1) + R''(\rho, P, H_1) \end{cases}, \quad (51)$$

as shown in Fig. 3. Since $R_1 = R_2 = r \log P$, (51) can be written as

$$\begin{aligned} r \log P < R^{(\mathrm{JD})}(\rho, P, H_1) &= \min \Big\{ R''(\rho, P, H_1), \frac{1}{2}\big(R'(\rho, P, H_1) + R''(\rho, P, H_1)\big) \Big\} \\ &= \min \big\{ \log(\gamma'(P, H_1) - \delta'(P, H_1)|\rho|^2), \frac{1}{2} \log(\alpha(P, H_1) - \beta(P, H_1)|\rho|^2) \big\}. \end{aligned} \quad (52)$$

It is seen that $R^{(\mathrm{JD})}(\rho, P, H_1)$ attains its maximum at $\rho = 0$ regardless of $P > 1$ and H_1. Setting $\rho = 0$ in (52),

$$r \log P < R^{(\mathrm{JD})}(0, P, H_1) = \min \Big\{ \log \gamma'(P, H_1), \frac{1}{2} \log \alpha(P, H_1) \Big\}. \quad (53)$$

VII. Proof of Theorem 2

Define the random matrix $\boldsymbol{W}$ by

$$\boldsymbol{W} = \begin{bmatrix} \boldsymbol{w} & \boldsymbol{w}' \\ \boldsymbol{w}'^* & \boldsymbol{w}'' \end{bmatrix} = \mathrm{diag}(1, \sigma^{-1}) \boldsymbol{H}_1^\dagger \boldsymbol{H}_1 \mathrm{diag}(1, \sigma^{-1}), \quad (54)$$

i.e.,

$$\begin{aligned} \boldsymbol{w} &= |\boldsymbol{a}_1|^2 + |\boldsymbol{a}_2|^2 \\ \boldsymbol{w}' &= \sigma^{-1} \left(\boldsymbol{a}_1^* \boldsymbol{b}_1 + \boldsymbol{a}_2^* \boldsymbol{b}_2 \right) \\ \boldsymbol{w}'' &= \sigma^{-2} \left(|\boldsymbol{b}_1|^2 + |\boldsymbol{b}_2|^2 \right) \end{aligned} . \quad (55)$$

Then $\boldsymbol{W}$ is a complex Wishart $\mathcal{W}_2(2, I_2)$ random matrix [32] with distribution

$$p_{\boldsymbol{W}}(W) = \begin{cases} \frac{1}{\pi} \exp(-(w + w'')) & w\, w'' > |w'|^2 \\ 0 & \text{otherwise} \end{cases} \quad (56)$$

A. Treating interference as noise

By (20) and noting that the elements of $\boldsymbol{H}_1$ are independent circularly symmetric complex Gaussian random variables, we conclude that $R^{(\mathrm{TIN})}(1, P, \boldsymbol{H}_1)$ and $\log\left(1 + \frac{\frac{P}{2}\boldsymbol{w} + \frac{P^2}{4}\sigma^2|\boldsymbol{w}'|^2}{1 + \frac{P}{2}\sigma^2\boldsymbol{w}''}\right)$ are identically distributed. In fact, replacing $\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{b}_1$ and $\boldsymbol{b}_2$ by $\boldsymbol{a}_1^*, \boldsymbol{a}_2^*, -\boldsymbol{b}_2$ and $\boldsymbol{b}_1$, respectively, $R^{(\mathrm{TIN})}(1, P, \boldsymbol{H}_1)$ turns into $\log\left(1+\frac{\frac{P}{2}\boldsymbol{w}+\frac{P^2}{4}\sigma^2|\boldsymbol{w}'|^2}{1+\frac{P}{2}\sigma^2\boldsymbol{w}''}\right)$. As such, as long as distribution of $R^{(\mathrm{TIN})}(1,P,\boldsymbol{H}_1)$ is concerned, we can write

$$R^{(\mathrm{TIN})}(1,P,\boldsymbol{H}_1)=\log\left(1+\frac{\frac{P}{2}\boldsymbol{w}+\frac{P^2}{4}\sigma^2|\boldsymbol{w}'|^2}{1+\frac{P}{2}\sigma^2\boldsymbol{w}''}\right). \tag{57}$$

Let us define

$$f(w,w'')=\frac{2}{\sigma^2P}\left(\frac{2}{P}(P^r-1)\Big(1+\frac{P}{2}\sigma^2w''\Big)-w\right)^+,\quad w,w''\geq 0. \tag{58}$$

Then

$$\begin{aligned}\mathbb{P}\left(R^{(\mathrm{TIN})}(1,P,\boldsymbol{H}_1)\leq r\log P\right)&=\mathbb{P}\left(|\boldsymbol{w}'|^2\leq f(\boldsymbol{w},\boldsymbol{w}'')\right)\\ &\overset{(a)}{=}\int_0^\infty\int_0^\infty\int_{-\infty}^\infty\int_{-\infty}^\infty\frac{1}{\pi}e^{-(w+w'')}\mathbb{1}_{s^2+t^2\leq\min\{ww'',f(w,w'')\}}\mathrm{d}s\,\mathrm{d}t\,\mathrm{d}w\,\mathrm{d}w''\\ &=\int_0^\infty\int_0^\infty\frac{1}{\pi}e^{-(w+w'')}\int_{-\infty}^\infty\int_{-\infty}^\infty\mathbb{1}_{s^2+t^2\leq\min\{ww'',f(w,w'')\}}\mathrm{d}s\,\mathrm{d}t\,\mathrm{d}w\,\mathrm{d}w''\\ &\overset{(b)}{=}\int_0^\infty\int_0^\infty\min\{ww'',f(w,w'')\}e^{-(w+w'')}\mathrm{d}w\,\mathrm{d}w'',\end{aligned} \tag{59}$$

where in (a), s and t represent $\mathrm{Re}(w')$ and $\mathrm{Im}(w')$, respectively, and (b) is due to

$$\iint_{(s,t):s^2+t^2\leq\min\{ww'',f(w,w'')\}}\mathrm{d}s\,\mathrm{d}t=\pi\min\{ww'',f(w,w'')\}. \tag{60}$$

We have $f(w,w'')>0$ and $\min\{ww'',f(w,w'')\}=ww''$ if and only if $w<g(w'')$ and $w\leq g(0)$, respectively, where

$$g(w'')=\frac{2}{P}(P^r-1)\Big(1+\frac{P}{2}\sigma^2w''\Big),\quad w''\geq 0. \tag{61}$$

As such, one can write (59) as

$$\begin{aligned}\mathbb{P}\left(R^{(\mathrm{TIN})}(1,P,\boldsymbol{H}_1)\leq r\log P\right)&=\int_0^\infty\int_0^{g(0)}ww''e^{-(w+w'')}\mathrm{d}w\,\mathrm{d}w''\\ &\quad+\int_0^\infty\int_{g(0)}^{g(w'')}f(w,w'')e^{-(w+w'')}\mathrm{d}w\,\mathrm{d}w''.\end{aligned} \tag{62}$$

But,

$$\int_0^\infty\int_0^{g(0)}ww''e^{-(w+w'')}\mathrm{d}w\,\mathrm{d}w''=1-(1+g(0))e^{-g(0)} \tag{63}$$

and

$$\int_0^\infty\int_{g(0)}^{g(w'')}f(w,w'')e^{-(w+w'')}\mathrm{d}w\,\mathrm{d}w''=\frac{\sigma^2(P^r-1)}{1+\sigma^2(P^r-1)}\,g(0)e^{-g(0)}. \tag{64}$$

Putting (62), (63) and (64) together,

$$\begin{aligned}\mathbb{P}\left(R^{(\mathrm{TIN})}(1,P,\boldsymbol{H}_1)\le R\right) &= 1-\left(1+\frac{\frac{2}{P}(P^r-1)}{1+\sigma^2(P^r-1)}\right)e^{-\frac{2}{P}(P^r-1)}\\ &= 2P^{-(1-r)}+o(P^{-(1-r)}).\end{aligned}\tag{65}$$

B. Interference cancellation

Following similar lines of reasoning that led us to (57), we can write

$$R'(1,P,\boldsymbol{H}_1)=\log\left(1+\frac{\frac{P}{2}\sigma^2\boldsymbol{w}''+\frac{P^2}{4}\sigma^2|\boldsymbol{w}'|^2}{1+\frac{P}{2}\boldsymbol{w}}\right)\tag{66}$$

We also have

$$R''(1,P,\boldsymbol{H}_1)=\log\Big(1+\frac{P}{2}\boldsymbol{w}\Big).\tag{67}$$

Defining

$$f(w,w'')=\frac{2}{\sigma^2P}\left(\frac{2}{P}(P^r-1)\Big(1+\frac{P}{2}w\Big)-\sigma^2w''\right)^+,\quad w,w''\ge 0.\tag{68}$$

We have $R'(1,P,\boldsymbol{H}_1)\le R$ if and only if $|\boldsymbol{w}'|^2\le f(\boldsymbol{w},\boldsymbol{w}'')$. Then

$$\begin{aligned}\mathbb{P}(R^{(\mathrm{CI})}(1,P,\boldsymbol{H}_1)\le r\log P) &= \mathbb{P}(\min\{R'(1,P,\boldsymbol{H}_1),R''(1,P,\boldsymbol{H}_1)\}\le r\log P)\\ &= \mathbb{P}(R''(1,P,\boldsymbol{H}_1)\le r\log P)\\ &\quad +\mathbb{P}(R''(1,P,\boldsymbol{H}_1)>r\log P,R'(1,P,\boldsymbol{H}_1)\le r\log P)\\ &= \mathbb{P}\Big(\boldsymbol{w}\le\frac{2}{P}(P^r-1)\Big)+\mathbb{P}\Big(\boldsymbol{w}>\frac{2}{P}(P^r-1),|\boldsymbol{w}'|^2\le f(\boldsymbol{w},\boldsymbol{w}'')\Big).\end{aligned}\tag{69}$$

The first term on the right side of (69) is given by

$$\mathbb{P}\Big(\boldsymbol{w}\le\frac{2}{P}(P^r-1)\Big)=1-\Big(1+\frac{2}{P}(P^r-1)\Big)e^{-\frac{2}{P}(P^r-1)}.\tag{70}$$

Let

$$g(w)=\frac{2}{\sigma^2P}(P^r-1)\Big(1+\frac{P}{2}w\Big),\quad w\ge 0.\tag{71}$$

Noting that $\frac{2}{P}(P^r-1)=\sigma^2 g(0)$ and by similar lines of reasoning in (59) and (62),

$$\begin{aligned}
&\mathbb{P}\Big(\boldsymbol{w} > \frac{2}{P}(P^r-1), |\boldsymbol{w}'|^2 \le f(\boldsymbol{w},\boldsymbol{w}'')\Big)\\
&= \int_0^\infty \int_0^\infty \mathbb{1}_{w>\sigma^2 g(0)} \min\{ww'', f(w,w'')\} e^{-(w+w'')} \mathrm{d}w''\,\mathrm{d}w\\
&= \int_{\sigma^2 g(0)}^\infty \int_0^{g(0)} ww'' e^{-(w+w'')} \mathrm{d}w''\,\mathrm{d}w + \int_{\sigma^2 g(0)}^\infty \int_{g(0)}^{g(w)} f(w,w'') e^{-(w+w'')} \mathrm{d}w''\,\mathrm{d}w. \qquad (72)
\end{aligned}$$

We have

$$\begin{aligned}
\int_{\sigma^2 g(0)}^\infty \int_0^{g(0)} ww'' e^{-(w+w'')} \mathrm{d}w''\,\mathrm{d}w &= \big(1-(1+\sigma^2 g(0))e^{-\sigma^2 g(0)}\big)\big(1-(1+g(0))e^{-g(0)}\big)\\
&= \frac{4}{\sigma^4} P^{-4(1-r)} + o(P^{-4(1-r)}), \qquad (73)
\end{aligned}$$

and

$$\begin{aligned}
&\int_{\sigma^2 g(0)}^\infty \int_{g(0)}^{g(w)} f(w,w'') e^{-(w+w'')} \mathrm{d}w''\,\mathrm{d}w\\
&= \left(\frac{4}{\sigma^2 P^2}\Big((P^r-1)^2 + \frac{P}{2}(P^r-\sigma^2-1)\Big) + \frac{2\sigma^2}{P(P^r+\sigma^2-1)} e^{-\frac{2}{\sigma^2 P}(P^r-1)^2}\right) e^{-\frac{2}{P}\left(1+\frac{1}{\sigma^2}\right)(P^r-1)}\\
&= \frac{2}{\sigma^2} P^{-(1-r)} + o(P^{-(1-r)}). \qquad (74)
\end{aligned}$$

By (72), (73) and (74), the scaling under CI is $\frac{2}{\sigma^2}P^{-(1-r)} + o(P^{-(1-r)})$ as desired.

C. Joint decoding

By (53) and using the union bound,

$$\mathbb{P}(R^{(\mathrm{JD})}(0,P,\boldsymbol{H}_1) \le r\log P) \le \mathbb{P}(\gamma'(P,\boldsymbol{H}_1) \le P^r) + \mathbb{P}(\alpha(P,\boldsymbol{H}_1) \le P^{2r}). \qquad (75)$$

We need the following lemmas:

Lemma 2 *Let Ψ_1 and Ψ_2 be the cumulative distribution functions of $|\boldsymbol{a}_1|^2|\boldsymbol{a}_2|^2$ and $(|\boldsymbol{a}_1|^2+|\boldsymbol{b}_1|^2)(|\boldsymbol{a}_2|^2+|\boldsymbol{b}_2|^2)$, respectively. Then*

$$\Psi_1(v) = 1 - 2\sqrt{v}\,K_1(2\sqrt{v}), \quad v>0 \qquad (76)$$

and

$$\Psi_2(v) = \begin{cases} 1 - 2vK_0(2\sqrt{v}) - 2\sqrt{v}(1+v)K_1(2\sqrt{v}) & \sigma^2 = 1\\ 1 - \frac{2}{(\sigma^2-1)^2}\Big(\sigma^2\sqrt{v}K_1\big(\tfrac{2}{\sigma^2}\sqrt{v}\big) - 2\sigma\sqrt{v}K_1\big(\tfrac{2}{\sigma}\sqrt{v}\big) + \sqrt{v}K_1(2\sqrt{v})\Big) & \sigma^2 \ne 1 \end{cases}. \qquad (77)$$

where $K_n(\cdot)$ is the modified Bessel function of the second kind of order n.

Proof: See Appendix B. ■

Lemma 3 *In the asymptote of small v,*

$$\Psi_1(v) = -v\ln v + o(v\ln v) \tag{78}$$

and

$$\Psi_2(v) = -\frac{1}{2\sigma^4}v^2\ln v + o(v^2\ln v). \tag{79}$$

Proof: See Appendix C. ■

The two terms on the right side of (75) do not admit closed form expressions. In what follows, we derive an upper bound on each of these terms:

- The term $\mathbb{P}(\gamma'(P, \boldsymbol{H}_1) \leq P^r)$: By (33),

$$\begin{aligned}\mathbb{P}(\gamma'(P, \boldsymbol{H}_1) \leq P^r) &= \mathbb{P}\left(\Big(1 + \frac{P}{2}|\boldsymbol{a}_1|^2\Big)\Big(1 + \frac{P}{2}|\boldsymbol{a}_2|^2\Big) \leq P^r\right) \\ &\leq \mathbb{P}\Big(1 + \frac{P^2}{4}|\boldsymbol{a}_1|^2|\boldsymbol{a}_2|^2 \leq P^r\Big) \\ &= \mathbb{P}\Big(|\boldsymbol{a}_1|^2|\boldsymbol{a}_2|^2 \leq \frac{4}{P^2}(P^r - 1)\Big) \\ &= \Psi_1\Big(\frac{4}{P^2}(P^r - 1)\Big),\end{aligned} \tag{80}$$

  where Ψ_1 is given by (76) in Lemma 2.

- The term $\mathbb{P}(\alpha(P, \boldsymbol{H}_1) \leq P^{2r})$: By (33),

$$\begin{aligned}\mathbb{P}(\alpha(P, \boldsymbol{H}_1) \leq P^{2r}) &= \mathbb{P}\left(\Big(1 + \frac{P}{2}(|\boldsymbol{a}_1|^2 + |\boldsymbol{b}_1|^2)\Big)\Big(1 + \frac{P}{2}(|\boldsymbol{a}_2|^2 + |\boldsymbol{b}_2|^2)\Big) \leq P^{2r}\right) \\ &\leq \mathbb{P}\Big(1 + \frac{P^2}{4}(|\boldsymbol{a}_1|^2 + |\boldsymbol{b}_1|^2)(|\boldsymbol{a}_2|^2 + |\boldsymbol{b}_2|^2) \leq P^{2r}\Big) \\ &= \mathbb{P}\Big((|\boldsymbol{a}_1|^2 + |\boldsymbol{b}_1|^2)(|\boldsymbol{a}_2|^2 + |\boldsymbol{b}_2|^2) \leq \frac{4}{P^2}(P^{2r} - 1)\Big) \\ &= \Psi_2\Big(\frac{4}{P^2}(P^{2r} - 1)\Big),\end{aligned} \tag{81}$$

  where Ψ_2 is given by (77) in Lemma 2.

Using (80) and (78), we get

$$\mathbb{P}(\gamma'(P,\boldsymbol{H}_1)\leq P^r)\leq 4(2-r)P^{-(2-r)}\ln P+o(P^{-(2-r)}\ln P), \tag{82}$$

in the asymptote of large P. Similarly, by (81) and (79),

$$\mathbb{P}(\alpha(P,\boldsymbol{H}_1)\leq P^{2r})\leq \frac{16}{\sigma^4}(1-r)P^{-4(1-r)}\ln P+o(P^{-4(1-r)}\ln P). \tag{83}$$

VIII. Interaction between P and σ^2: $\rho=0$ vs. $\rho=1$

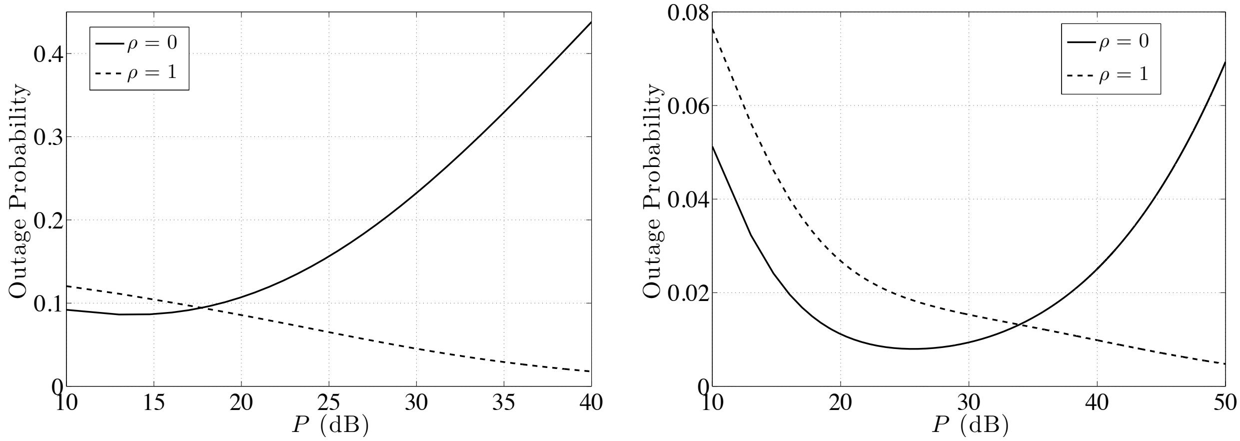
Let us consider the scenario in Theorem 2 where the direct and crossover channel coefficients are realizations of independent $\mathcal{CN}(0,1)$ and $\mathcal{CN}(0,\sigma^2)$ random variables, respectively. In this section, we study the interaction between σ^2 and P in determining the optimum choice of ρ. By Theorem 1, $\rho=1$ is the best choice as P grows to infinity as long as $\sigma>0$, whether the receivers apply TIN or CI. Moreover, it is clear that if $\sigma=0$, transmitting independent Gaussian signals over the underlying GICs maximizes the achievable rate per user, i.e., the outage probability is minimized for $\rho=0$ regardless of the value of P. Motivated by this fact, we assume σ^2 is not zero, but it is a small number. In particular, we are interested in $0<\sigma^2<1$. A typical scenario for this setup corresponds to a cellular system where distant cells cause interference to each other. We raise the question that given an arbitrary $0<\sigma^2<1$, what the largest SNR level $P_{\max}$ is such that $P<P_{\max}$ guarantees $\rho=0$ outperforms $\rho=1$, while $\rho=1$ is superior to $\rho=0$ for $P>P_{\max}$. Due to simplicity of the receiver structure and lower complexity of system design, only receivers that perform TIN are studied in this section. Moreover, cancelling interference or joint decoding require a certain level of coordination among distant cells which will be practically difficult.

To determine $P_{\max}$, the exact probability of outage for $\rho=0$ is computed in Appendix D as

$$\begin{aligned}&\mathbb{P}\left(R^{(\mathrm{TIN})}(0,P,\boldsymbol{H}_1)\leq r\log P\right)\\&=\int_1^{P^r}\left(\frac{2}{P}+\frac{\sigma^2}{1+\sigma^2(x-1)}\right)\left(1-\frac{e^{-\frac{2}{P}\left(\frac{P^r}{x}-1\right)}}{1+\sigma^2(\frac{P^r}{x}-1)}\right)\frac{e^{-\frac{2}{P}(x-1)}}{1+\sigma^2(x-1)}\mathrm{d}x.\end{aligned} \tag{84}$$

One may use (84) and the expression for $\mathbb{P}\left(R(1,P,\boldsymbol{H}_1)\leq r\log P\right)$ in (65) to determine $P_{\max}$ for given σ^2 and r. Fig. 4 presents plots of the exact value of the outage probability in terms of P (dB) for $\rho=0$ and $\rho=1$. In panel (a), $r=0.5$ and $\sigma^2=0.1$ and in panel (b), $r=0.5$ and $\sigma^2=0.01$.

Next, we consider a scenario where the optimum choice of ρ is equal to 0 in the asymptote of large P. We study a sequence of PGICs indexed by positive integers $n=1,2,\cdots$ such that the SNR and variance



(a) $r = 0.5$, $\sigma^2 = 0.1$ (b) $r = 0.5$, $\sigma^2 = 0.01$

Fig. 4. Plots of the exact value of $\mathbb{P}\left(R(\rho, P, \boldsymbol{H}_1) \leq r \log P\right)$ in terms of P (dB) for $\rho = 1$ and $\rho = 0$ given in (65) and (84), respectively. In panel (a), $r = 0.5$ and $\sigma^2 = 0.1$ and in panel (b), $r = 0.5$ and $\sigma^2 = 0.01$. It is seen that for a given σ^2, the choice of $\rho = 0$ outperforms $\rho = 1$ if SNR is sufficiently small.

of crossover channel coefficients increase and decrease by n, respectively, such that their product is $P_n \sigma_n^2$ tends to zero. The following result shows that the optimum ρ is 0 as the index n grows to infinity:

Theorem 3 *Consider a sequence of two-user PGICs with two underlying GICs where the SNR, transmission rate per user and channel matrix to receiver i in the n^{th} PGIC are given by P_n, $r \log P_n$ and $\boldsymbol{H}_{i,n} = \begin{bmatrix} \boldsymbol{a}_{i,1} & \boldsymbol{b}_{i,1} \\ \boldsymbol{a}_{i,2} & \boldsymbol{b}_{i,2} \end{bmatrix} \operatorname{diag}(1, \sigma_{i,n})$, respectively, where the matrices $\boldsymbol{H}_{i,n}$ obey Assumption 2 in Section III. The direct channel coefficients are the same in all PGICs, however, the crossover channel coefficients in the n^{th} PGIC are $\sigma_{i,n}$ times those of the PGIC of index $n = 1$. If $\lim_{n\to\infty} P_n = \infty$ and $\lim_{n\to\infty} P_n \sigma_{i,n}^2 = 0$, then $\lim_{n\to\infty} \mathbb{P}(R^{(\mathrm{TIN})}(\rho, P_n, \boldsymbol{H}_{i,n}) \leq r \log P_n) = 0$ under each of the following conditions:*

- $0 \leq \rho < 1$ *and* $0 \leq r < 2$.
- $\rho = 1$ *and* $0 \leq r < 1$.

Moreover, $\lim_{n\to\infty} \mathbb{P}(R^{(\mathrm{TIN})}(1, P_n, \boldsymbol{H}_{i,n}) \leq r \log P_n) = 1$ *if* $r > 1$.

Proof: See Appendix E. ■

As mentioned earlier, in the extreme case of $\sigma = 0$, $\rho = 0$ is the optimal choice regardless of SNR and transmission rate. We end this section with the following Proposition that offers an expression for the outage probability in this case:

Proposition 1 *If $\sigma = 0$, the outage probability is given by $1 - e^{\frac{2}{P}(P^r - 1)} - \frac{2}{P} e^{\frac{4}{P}} \int_1^{P^r} e^{-\frac{2}{P}(\frac{P^r}{x} + x)} \mathrm{d}x$ for any $0 \leq r < 2$.*

Proof: $\mathbb{P}((R^{(\mathrm{TIN})}(0,P,\boldsymbol{H}_1) \le r\log P)$ is given in (84) for arbitrary $\sigma^2 > 0$. Putting $\sigma = 0$ in the expression of $R(0,P,\boldsymbol{H}_1)$ in (22) yields the achievable rate per user in the interference-free setup where crossover channel coefficients are zero. As such, one can use dominated convergence [28] to conclude that the outage probability for $\sigma = 0$ is given by

$$\lim_{\sigma\to 0}\mathbb{P}\left(R^{(\mathrm{TIN})}(0,P,\boldsymbol{H}_1) \le r\log P\right)$$
$$= \lim_{\sigma\to 0}\int_1^{P^r}\left(\frac{2}{P}+\frac{\sigma^2}{1+\sigma^2(x-1)}\right)\left(1-\frac{e^{-\frac{2}{P}\left(\frac{P^r}{x}-1\right)}}{1+\sigma^2(\frac{P^r}{x}-1)}\right)\frac{e^{-\frac{2}{P}(x-1)}}{1+\sigma^2(x-1)}\mathrm{d}x. \tag{85}$$

If $\sigma < 1$, the integrand of the integral on the right side of (85) is bounded from above by the integrable function $(\frac{2}{P}+1)\left(1-\frac{x}{P^r}e^{-\frac{2}{P}(\frac{P^r}{x}-1)}\right)e^{-\frac{2}{P}(x-1)}$ for $1 \le x \le P^r$. Invoking dominated convergence [28] one more time,

$$\begin{aligned}\lim_{\sigma\to 0}\mathbb{P}\left(R^{(\mathrm{TIN})}(0,P,\boldsymbol{H}_1) \le r\log P\right) &= \frac{2}{P}\int_1^{P^r}\left(1-e^{-\frac{2}{P}\left(\frac{P^r}{x}-1\right)}\right)e^{-\frac{2}{P}(x-1)}\mathrm{d}x \\ &= 1-e^{\frac{2}{P}(P^r-1)}-\frac{2}{P}e^{\frac{4}{P}}\int_1^{P^r}e^{-\frac{2}{P}(\frac{P^r}{x}+x)}\mathrm{d}x. \end{aligned}\tag{86}$$

■

IX. Extension to $N > 2$ parallel GICs

In this section, we extend the result of Theorem 2 to a two-user PGIC with $N > 2$ underlying GICs. For simplicity of presentation, we only consider the case where the receivers apply TIN. The received vector in one transmission slot at receiver 1 is given by

$$\vec{\boldsymbol{y}}_1 = A\vec{\boldsymbol{x}}_1 + B\vec{\boldsymbol{x}}_2 + \vec{\boldsymbol{z}}_1, \tag{87}$$

where

$$A = \mathrm{diag}(a_1,\cdots,a_N),\ \ B = \mathrm{diag}(b_1,\cdots,b_N), \tag{88}$$

a_j and b_j are the direct and crossover channel coefficients in GIC j, $\vec{\boldsymbol{x}}_i$ is the $N\times 1$ vector of signals transmitted by user i and $\vec{\boldsymbol{z}}_1 \sim \mathcal{CN}(\vec{0}_N, I_N)$ is the ambient noise vector at the receiver of user 1. It is assumed that

$$\vec{\boldsymbol{x}}_i \sim \mathcal{CN}\left(\vec{0}_N, \frac{P}{N}C_\rho\right), \tag{89}$$

where $C_\rho = \rho\vec{1}_N\vec{1}_N^{\mathrm{t}} + (1-\rho)I_N$ is a $N \times N$ covariance matrix in which all diagonal elements are 1 and all off-diagonal elements are equal to the real number $-\frac{1}{N+1} \leq \rho \leq 1$. This specific range for ρ is to ensure all eigenvalues of C_ρ are nonnegative. Under TIN, the achievable rate by user 1 is given by

$$R^{(\mathrm{TIN})}(\rho, P, H_1) = \log \frac{\det\left(I_N + \frac{P}{N}AC_\rho A^\dagger + \frac{P}{N}BC_\rho B^\dagger\right)}{\det\left(I_N + \frac{P}{N}BC_\rho B^\dagger\right)}, \tag{90}$$

where

$$H_1 = \begin{bmatrix} a_1 & b_1 \\ \vdots & \vdots \\ a_N & b_N \end{bmatrix}. \tag{91}$$

One can make the following observations regarding $R^{(\mathrm{TIN})}(\rho, P, H_1)$ as a function of ρ:

- Had we defined C_ρ as a matrix whose diagonal elements are 1 and the elements above and below the main diagonal are ρ and ρ^* for some complex number ρ, respectively, then it is easy to see that $R^{(\mathrm{TIN})}(\rho, P, H_1)$ does not just depend on $|\rho|$ for $N > 2$. This is in contrast to the case of $N = 2$ where $R^{(\mathrm{TIN})}(\rho, P, H_1)$ is only a function of $|\rho|$ and hence, one can assume ρ is real without loss of generality. Here, we assume ρ is real only for the sake of simplicity in presentation.
- Recall the expression of $R^{(\mathrm{TIN})}(\rho, P, H_1)$ for $N = 2$ given in Section IV. It turns out that it is either a decreasing or an increasing function of $0 \leq \rho \leq 1$,[4] i.e., $R^{(\mathrm{TIN})}(\rho, P, H_1)$ is maximized either at $\rho = 0$ or $\rho = 1$ for any P and any realization H_1 of $\boldsymbol{H}_1$. This statement is no longer true for $N \geq 3$. For example, if $P = 20\mathrm{dB}$, $N = 3$ and

$$H_1 = \begin{bmatrix} -0.0583 - 1.2690\sqrt{-1} & 0.0708 - 0.4245\sqrt{-1} \\ -1.3669 + 0.5942\sqrt{-1} & -0.3850 + 0.3465\sqrt{-1} \\ -0.3104 - 0.6279\sqrt{-1} & 0.2146 + 0.5228\sqrt{-1} \end{bmatrix}, \tag{92}$$

  then $R^{(\mathrm{TIN})}(\rho, P, H_1)$ is maximized at $\rho \approx 0.89$.
- The point $\rho = 0$ is always a point of extremum for $R^{(\mathrm{TIN})}(\rho, P, H_1)$. In Appendix F, it is shown that

$$\frac{\mathrm{d}}{\mathrm{d}\rho}R^{(\mathrm{TIN})}(\rho, P, H_1) = \frac{P}{N}\left(\mathrm{tr}(\Omega^{-1}(AJA^\dagger + BJB^\dagger)) - \mathrm{tr}(\Xi^{-1}BJB^\dagger)\right), \tag{93}$$

[4] See the statement before (124).

where we have defined

$$\begin{aligned} J &= \vec{1}_N \vec{1}_N^{\mathrm{t}} - I_N \\ \Omega &= I_N + \tfrac{P}{N} A C_\rho A^\dagger + \tfrac{P}{N} B C_\rho B^\dagger \\ \Xi &= I_N + \tfrac{P}{N} B C_\rho B^\dagger \end{aligned} \quad . \tag{94}$$

At $\rho = 0$, the matrices Ω and Ξ are diagonal and hence, $\mathrm{tr}(\Omega^{-1}(AJA^\dagger + BJB^\dagger)) = \mathrm{tr}(\Xi^{-1} BJB^\dagger) = 0$ due to the fact that J is an $N \times N$ matrix with 0 on all diagonal elements and 1 on all off-diagonal elements. Using this fact in (93), $\frac{\mathrm{d}}{\mathrm{d}\rho}\big|_{\rho=0} R^{(\mathrm{TIN})}(\rho, P, H_1) = 0$, i.e., $\rho = 0$ is a point of extremum for $R^{(\mathrm{TIN})}(\rho, P, H_1)$ regardless of P and the realization H_1 of $\boldsymbol{H}_1$.

Our first observation is about the behaviour $R^{(\mathrm{TIN})}(\rho, P, H_1)$ around $\rho = 0$:

Proposition 2 *Let $N > 2$ and $\boldsymbol{a}_i \boldsymbol{b}_j - \boldsymbol{a}_j \boldsymbol{b}_i \neq 0$ for some $1 \leq i < j \leq N$. Then*

$$\lim_{P \to \infty} \mathbb{P}\left(R^{(\mathrm{TIN})}(\rho, P, \boldsymbol{H}_1) \text{ has a local minimum at } \rho = 0\right) = 1. \tag{95}$$

Proof: We already know that $\rho = 0$ is a point of extremum for $R^{(\mathrm{TIN})}(\rho, P, H_1)$ regardless of the value of P and the realization H_1 of $\boldsymbol{H}_1$. As such, it is enough to show that

$$\lim_{P \to \infty} \mathbb{P}\left(\frac{\mathrm{d}^2}{\mathrm{d}\rho^2}\Big|_{\rho=0} R^{(\mathrm{TIN})}(\rho, P, \boldsymbol{H}_1) > 0\right) = 1. \tag{96}$$

In Appendix F, it is shown that

$$\begin{aligned} \frac{\mathrm{d}^2}{\mathrm{d}\rho^2}\Big|_{\rho=0} R^{(\mathrm{TIN})}(\rho, P, H_1) \;&=\; 2\left(\frac{P}{N}\right)^2 \sum_{1 \leq i < j \leq N} \frac{|b_i b_j^*|^2}{\left(1 + \frac{P}{N}|b_i|^2\right)\left(1 + \frac{P}{N}|b_j|^2\right)} \\ &\quad -2\left(\frac{P}{N}\right)^2 \sum_{1 \leq i < j \leq N} \frac{|a_i a_j^* + b_i b_j^*|^2}{\left(1 + \frac{P}{N}(|a_i|^2 + |b_i|^2)\right)\left(1 + \frac{P}{N}(|a_j|^2 + |b_j|^2)\right)} \end{aligned} \tag{97}$$

As such,

$$\begin{aligned} \lim_{P \to \infty} \frac{\mathrm{d}^2}{\mathrm{d}\rho^2}\Big|_{\rho=0} R^{(\mathrm{TIN})}(\rho, P, H_1) \;&=\; 2 \sum_{1 \leq i < j \leq N} 1 - 2 \sum_{1 \leq i < j \leq N} \frac{|a_i a_j^* + b_i b_j^*|^2}{(|a_i|^2 + |b_i|^2)(|a_j|^2 + |b_j|^2)} \\ &=\; 2 \sum_{1 \leq i < j \leq N} \frac{|a_i b_j - a_j b_i|^2}{(|a_i|^2 + |b_i|^2)(|a_j|^2 + |b_j|^2)} \geq 0, \end{aligned} \tag{98}$$

for all realizations H_1 of $\boldsymbol{H}_1$. Let us denote the ratio on the right side of (98) by η. Using dominated convergence [28], we can write

$$\lim_{P\to\infty} \mathbb{P}\left(\frac{\mathrm{d}^2}{\mathrm{d}\rho^2}\Big|_{\rho=0} R^{(\mathrm{TIN})}(\rho, P, \boldsymbol{H}_1) > 0\right) = \mathbb{P}\left(\boldsymbol{\eta} > 0\right). \tag{99}$$

Let $1 \le i_0 < j_0 \le N$ be such that $\boldsymbol{a}_{i_0}\boldsymbol{b}_{j_0} - \boldsymbol{a}_{j_0}\boldsymbol{b}_{i_0} \neq 0$. Then $\mathbb{P}\left(\frac{|\boldsymbol{a}_{i_0}\boldsymbol{b}_{j_0}-\boldsymbol{a}_{j_0}\boldsymbol{b}_{i_0}|^2}{(|\boldsymbol{a}_{i_0}|^2+|\boldsymbol{b}_{i_0}|^2)(|\boldsymbol{a}_{j_0}|^2+|\boldsymbol{b}_{j_0}|^2)} > 0\right) = 1$ and since $\mathbb{P}\left(\boldsymbol{\eta} > 0\right) \ge \mathbb{P}\left(\frac{|\boldsymbol{a}_{i_0}\boldsymbol{b}_{j_0}-\boldsymbol{a}_{j_0}\boldsymbol{b}_{i_0}|^2}{(|\boldsymbol{a}_{i_0}|^2+|\boldsymbol{b}_{i_0}|^2)(|\boldsymbol{a}_{j_0}|^2+|\boldsymbol{b}_{j_0}|^2)} > 0\right)$, we get $\mathbb{P}\left(\boldsymbol{\eta} > 0\right) = 1$. This completes the proof of Proposition 2 in light of (99). ■

Next, we set $\rho = 1$ and compute the outage probability $\mathbb{P}(R^{(\mathrm{TIN})}(1, P, \boldsymbol{H}_1) \le r \log P)$ in the asymptote of large P under the assumption that the channel gains represent Rayleigh fading. We have the following result which is a generalization of (13) in Theorem 2:

Proposition 3 *Let all the channel coefficients in a PGIC with N underlying GICs be independent. Moreover, the direct and crossover channel coefficients are realizations of $\mathcal{CN}(0,1)$ and $\mathcal{CN}(0,\sigma^2)$ random variables, respectively. The outage probability under TIN is given by $\frac{2N^{N-1}}{(N-1)!}P^{-(N-1)(1-r)} + o(P^{-(N-1)(1-r)})$ for any $0 \le r < 1$.*

Proof: We have

$$\begin{aligned} AC_1A^\dagger + BC_1B^\dagger &\overset{(a)}{=} A\vec{1}_N(A\vec{1}_N)^\dagger + B\vec{1}_N(B\vec{1}_N)^\dagger \\ &= \begin{bmatrix} A\vec{1}_N & B\vec{1}_N \end{bmatrix} \begin{bmatrix} A\vec{1}_N & B\vec{1}_N \end{bmatrix}^\dagger \\ &\overset{(b)}{=} HH^\dagger, \end{aligned} \tag{100}$$

where H_1 is given in (91), (a) is due to $C_1 = \vec{1}_N\vec{1}_N^{\mathrm{t}}$ and (b) is due to $A\vec{1}_N = \begin{bmatrix} a_1 & \cdots & a_N \end{bmatrix}^{\mathrm{t}}$ and $B\vec{1}_N = \begin{bmatrix} b_1 & \cdots & b_N \end{bmatrix}^{\mathrm{t}}$. As such,

$$\begin{aligned} R^{(\mathrm{TIN})}(1, P, H_1) &= \log \frac{\det\left(I_N + \frac{P}{N}H_1H_1^\dagger\right)}{\det\left(I_N + \frac{P}{N}B\vec{1}_N(B\vec{1}_N)^\dagger\right)} \\ &= \log \frac{\det\left(I_2 + \frac{P}{N}H_1^\dagger H_1\right)}{1 + \frac{P}{N}(B\vec{1}_N)^\dagger B\vec{1}_N}, \end{aligned} \tag{101}$$

where the last step is due to the identity $\det(I_m + UV) = \det(I_n + VU)$ where U and V are $m \times n$ and

$n \times m$ matrices, respectively. Let

$$\boldsymbol{W} = \begin{pmatrix} \boldsymbol{w} & \boldsymbol{w}' \\ \boldsymbol{w}'^{*} & \boldsymbol{w}'' \end{pmatrix} = \mathrm{diag}(1, \sigma^{-1}) \boldsymbol{H}_1^{\dagger} \boldsymbol{H}_1 \mathrm{diag}(1, \sigma^{-1}). \tag{102}$$

Then $\boldsymbol{W}$ is a complex Wishart $\mathcal{W}_2(N, I_2)$ random matrix with distribution [32]

$$p_{\boldsymbol{W}}(W) = \frac{c_N}{\pi}(ww'' - |w'|^2)^{N-2} e^{-(w+w'')} \mathbb{1}_{ww'' > |w'|^2}, \tag{103}$$

where

$$c_N = \frac{1}{(N-1)!(N-2)!}. \tag{104}$$

By (101) and (102),

$$\begin{aligned} R^{(\mathrm{TIN})}(1, P, \boldsymbol{H}_1) &= \log \frac{\det\left(I_2 + \frac{P}{N}\mathrm{diag}(1,\sigma)\boldsymbol{W}\mathrm{diag}(1,\sigma)\right)}{1 + \frac{P}{N}\sigma^2 \boldsymbol{w}''} \\ &= \log \frac{\left(1 + \frac{P}{N}\boldsymbol{w}\right)\left(1 + \frac{P}{N}\sigma^2\boldsymbol{w}''\right) - \frac{P^2}{N^2}\sigma^2|\boldsymbol{w}'|^2}{1 + \frac{P}{N}\sigma^2\boldsymbol{w}''} \\ &= \log\left(1 + \frac{\frac{P}{N}\boldsymbol{w} + \frac{P^2}{N^2}\sigma^2(\boldsymbol{w}\boldsymbol{w}'' - |\boldsymbol{w}'|^2)}{1 + \frac{P}{N}\sigma^2\boldsymbol{w}''}\right). \end{aligned} \tag{105}$$

Let

$$f_1(w'') = \frac{N}{P}(P^r - 1)\left(1 + \frac{P}{N}\sigma^2 w''\right), \;\; w'' \geq 0, \tag{106}$$

$$f_2(w, w'') = \frac{N}{P\sigma^2}(f_1(w'') - w)^+, \;\; w, w'' \geq 0 \tag{107}$$

and

$$g(w, w'') = ww'' - f_2(w, w''), \;\; w, w'' \geq 0. \tag{108}$$

Then

$$\begin{aligned} \mathbb{P}\left(R^{(\mathrm{TIN})}(1, P, \boldsymbol{H}_1) \leq r \log P\right) &= \mathbb{P}\left(|\boldsymbol{w}'|^2 \geq g(\boldsymbol{w}, \boldsymbol{w}'')\right) \\ &= \mathbb{P}\left(g(\boldsymbol{w}, \boldsymbol{w}'') < 0\right) + \mathbb{P}\left(|\boldsymbol{w}'|^2 \geq g(\boldsymbol{w}, \boldsymbol{w}'') \geq 0\right). \end{aligned} \tag{109}$$

We have $g(w, w'') < 0$ if and only if $ww'' < f_2(w, w'')$ and $f_1(w'') > w$. These two constraints are equivalent to $w < f_1(0)$. Therefore,

$$\begin{aligned}\mathbb{P}(g(\boldsymbol{w},\boldsymbol{w}'') < 0) &= \mathbb{P}\Big(\boldsymbol{w} < f_1(0)\Big)\\ &= 1 - e^{-f_1(0)}\sum_{k=0}^{N-1}\frac{1}{k!}(f_1(0))^k\\ &= \frac{N^{N-1}}{(N-1)!}P^{-N(1-r)} + o(P^{-N(1-r)}), \qquad (110)\end{aligned}$$

where the penultimate step is due to the fact that $\boldsymbol{w}$, being the sum of N independent exponential random variables with parameter 1, is a Gamma random variable with PDF $p_{\boldsymbol{w}}(w) = \frac{1}{(N-1)!}x^{N-1}e^{-x}\mathbb{1}_{x>0}$.

To compute $\mathbb{P}\left(|\boldsymbol{w}'|^2 \geq g(\boldsymbol{w},\boldsymbol{w}'') \geq 0\right)$, note that $g(w,w'') \geq 0$ is equivalent to $w > f_1(0)$. Then

$$\begin{aligned}&\mathbb{P}\left(|\boldsymbol{w}'|^2 \geq g(\boldsymbol{w},\boldsymbol{w}'') \geq 0\right)\\ &= \int_0^\infty\int_{f_1(0)}^\infty\int_{-\infty}^\infty\int_{-\infty}^\infty \frac{c_N}{\pi}(ww'' - (s^2+t^2))^{N-2}e^{-(w+w'')}\mathbb{1}_{g(w,w'')\leq s^2+t^2\leq ww''}\mathrm{d}s\,\mathrm{d}t\,\mathrm{d}w\,\mathrm{d}w''\\ &= \int_0^\infty\int_{f_1(0)}^\infty \frac{c_N}{\pi}e^{-(w+w'')}\int_{-\infty}^\infty\int_{-\infty}^\infty(ww'' - (s^2+t^2))^{N-2}\mathbb{1}_{g(w,w'')\leq s^2+t^2\leq ww''}\mathrm{d}s\,\mathrm{d}t\,\mathrm{d}w\,\mathrm{d}w''. (111)\end{aligned}$$

Using polar coordinates,

$$\begin{aligned}\int_{-\infty}^\infty\int_{-\infty}^\infty(ww'' - (s^2+t^2))^{N-2}\mathbb{1}_{g(w,w'')\leq s^2+t^2\leq ww''}\mathrm{d}s\,\mathrm{d}t &= \frac{\pi}{N-1}(ww'' - g(w,w''))^{N-1}\\ &= \frac{\pi}{N-1}(f_2(w,w''))^{N-1}. \qquad (112)\end{aligned}$$

Hence,

$$\begin{aligned}\mathbb{P}\left(|\boldsymbol{w}'|^2 \geq g(\boldsymbol{w},\boldsymbol{w}'') \geq 0\right) &= \frac{c_N}{N-1}\int_0^\infty\int_{f_1(0)}^\infty e^{-(w+w'')}(f_2(w,w''))^{N-1}\mathrm{d}w\,\mathrm{d}w''\\ &= \frac{c_N}{N-1}\left(\frac{N}{P\sigma^2}\right)^{N-1}\int_0^\infty\int_{f_1(0)}^{f_1(w'')} e^{-(w+w'')}(f_1(w'') - w)^{N-1}\mathrm{d}w\,\mathrm{d}w'.\end{aligned}$$
(113)

But,

$$\begin{aligned}
\int_0^\infty \int_{f_1(0)}^{f_1(w'')} e^{-(w+w'')}(f_1(w'')-w)^{N-1}\mathrm{d}w\,\mathrm{d}w' &\overset{(a)}{=} \int_0^\infty e^{-(w''+f_1(w''))}\Big(\int_0^{f_1(w'')-f_1(0)} e^z z^{N-1}\mathrm{d}z\Big)\mathrm{d}w'' \\
&= \int_0^\infty e^{-(w''+f_1(w''))}\left[e^z \sum_{k=0}^{N-1}(-1)^k k!\binom{N-1}{k} z^{N-1-k}\right]_0^{f_1(w'')-f_1(0)}\mathrm{d}w'' \\
&= \int_0^\infty e^{-(w''+f_1(0))}\sum_{k=0}^{N-1}(-1)^k k!\binom{N-1}{k}(f_1(w'')-f_1(0))^{N-1-k}\mathrm{d}w'' \\
&\quad +(-1)^N(N-1)!\int_0^\infty e^{-(w''+f_1(w''))}\mathrm{d}w'' \\
&\overset{(b)}{=} e^{-f_1(0)}\sum_{k=0}^{N-1}(-1)^k k!\binom{N-1}{k}(\sigma^2(P^r-1))^{N-1-k}\int_0^\infty e^{-w''}w''^{N-1-k}\mathrm{d}w'' \\
&\quad +(-1)^N(N-1)!e^{-f_1(0)}\int_0^\infty e^{-(1+\sigma^2(P^r-1))w''}\mathrm{d}w'' \\
&\overset{(c)}{=} (N-1)!e^{-f_1(0)}\Bigg(\sum_{k=0}^{N-1}(-1)^k(\sigma^2(P^r-1))^{N-1-k}+\frac{(-1)^N}{1+\sigma^2(P^r-1)}\Bigg) \\
&= (N-1)!(\sigma^2)^{N-1}P^{r(N-1)}(1+o(1)),
\end{aligned} \tag{114}$$

where in (a) we have applied the change of variable $z = f_1(w'') - w$, (b) is due to $f_1(w'') = f_1(0) + \sigma^2(P^r-1)w''$ and (c) is due to $\int_0^\infty e^{-w''}w''^{N-1-k}\mathrm{d}w'' = (N-1-k)!$. By (113) and (114),

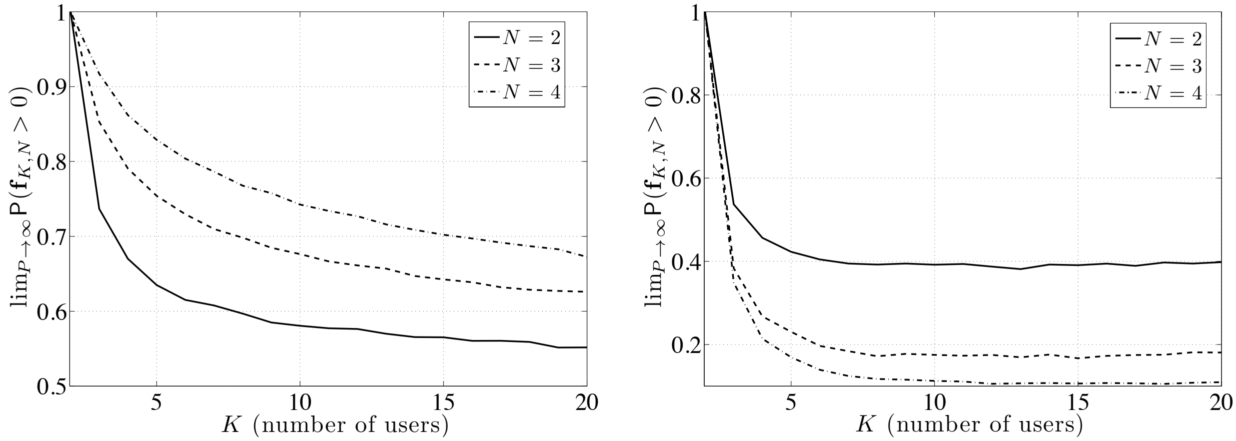
$$\begin{aligned}
\mathbb{P}\left(|\boldsymbol{w}'|^2 \geq g(\boldsymbol{w},\boldsymbol{w}'') \geq 0\right) &= \frac{c_N}{N-1}\left(\frac{N}{P\sigma^2}\right)^{N-1}(N-1)!(\sigma^2)^{N-1}P^{r(N-1)}(1+o(1)) \\
&= \frac{N^{N-1}}{(N-1)!}P^{-(N-1)(1-r)} + o(P^{-(N-1)(1-r)}).
\end{aligned} \tag{115}$$

By (109), (110) and (115), the proof is complete. ■

X. Discussion and Concluding Remarks

We studied a simple signalling scheme for a two-user PGIC consisting of $N=2$ parallel GICs where each user has a single-layer codebook and transmits Gaussian signals with identical variance $\frac{P}{2}$ and correlation ρ over the underlying GICs. Under general conditions on the fading statistics, it was shown that the value of ρ that minimizes the outage probability approaches 1 as SNR approaches infinity under both TIN and CI, while $\rho = 0$ is optimum under JD regardless of the value of SNR. Under the assumption that the direct and crossover channel coefficients are independent $\mathcal{CN}(0,1)$ and $\mathcal{CN}(0,\sigma^2)$ random variables, respectively, and the transmission rate per user is $r\log P$ for some $10 \leq r < 1$, it was demonstrated that

- The outage probability decays like $P^{-(1-r)}$ under both TIN and CI where it is assumed that $\rho = 1$.



(a) $\sigma^2 = 1$

(b) $\sigma^2 = 0.1$

Fig. 5. Plots of $\lim_{P\to\infty} \mathbb{P}(\boldsymbol{f}_{K,N} > 0)$ as given in (117) in terms of the number K of users for different values of N. The plots are produced via Monte Carlo simulation by generating 20,000 independent samples for each channel coefficient.

- The outage probability under JD is bounded from above by a term that decays like $P^{-(2-r)} \ln P$ and $P^{-4(1-r)} \ln P$ for $0 \le r < \frac{2}{3}$ and $\frac{2}{3} < r < 1$, respectively, where it is assumed that $\rho = 0$.

In particular, it was seen that one can do better in terms of achieving a smaller outage probability by letting the users share both channels in contrast to avoiding interference and having the users transmit over orthogonal GICs. It is conjectured that the upper bound on the outage probability under JD is tight in the asymptote of large P.

Let us conclude by a brief examination of a PGIC with $N \ge 2$ parallel GICs and $K > 2$ users and pointing out a significant difference in system behaviour for $K > 2$ in contrast to $K = 2$. Assume each user transmits $\mathcal{CN}(0, \frac{P}{N})$ signals with mutual correlation $\rho \in (-\frac{1}{N+1}, 1)$ over different GICs. Denoting the channel coefficient from the k^{th} transmitter to the receiver of user 1 over the n^{th} GIC by $h_{k,n}$ and treating interference as noise at the receivers, similar calculations as in Section IX show that $\rho = 0$ is a point of extremum for achievable rate per user, regardless of r, P, N, K and the realizations of channel coefficients. Moreover, $\rho = 0$ is a point of local minimum for the achievable rate of say user 1 if and only if[5]

$$f_{K,N} = \left(\frac{P}{N}\right)^2 \sum_{1\le m<n\le N} \left(\frac{|\widehat{g}_{m,n}|^2}{\left(1+\frac{P}{N}\widehat{g}_{m,m}\right)\left(1+\frac{P}{N}\widehat{g}_{n,n}\right)} - \frac{|g_{m,n}|^2}{\left(1+\frac{P}{N}g_{m,m}\right)\left(1+\frac{P}{N}g_{n,n}\right)} \right) > 0, \qquad (116)$$

where $\widehat{g}_{m,n} = \sum_{k=2}^{K} h_{k,m}h_{k,n}^*$ and $g_{m,n} = \sum_{k=1}^{K} h_{k,m}h_{k,n}^*$ for any $1 \le m \le n \le N$. In contrast to the case of $K = 2$, it turns out that the probability of $\rho = 0$ being a point of local minimum for achievable rate

[5]Note the similarity between the expressions in (116) and (97).

per user is strictly less than 1 for $K > 2$. To see this note that

$$\lim_{P\to\infty} \mathbb{P}(\boldsymbol{f}_{K,N} > 0) = \mathbb{P}\bigg(\sum_{1\le m<n\le N} \bigg(\frac{|\widehat{\boldsymbol{g}}_{m,n}|^2}{\widehat{\boldsymbol{g}}_{m,m}\widehat{\boldsymbol{g}}_{n,n}} - \frac{|\boldsymbol{g}_{m,n}|^2}{\boldsymbol{g}_{m,m}\boldsymbol{g}_{n,n}} \bigg) > 0 \bigg), \tag{117}$$

by dominated convergence [28]. Assuming the direct and crossover channel coefficients are realizations of independent $\mathcal{CN}(0,1)$ and $\mathcal{CN}(0,\sigma^2)$ random variables, respectively, Fig. 5 presents plots of $\lim_{P\to\infty}\mathbb{P}(\boldsymbol{f}_{K,N} > 0)$ in terms of $K = 2, \cdots, 20$ for different values of N. Panels (a) and (b) correspond to $\sigma^2 = 1$ and $\sigma^2 = 0.1$, respectively.

APPENDIX A; PROOF OF LEMMA 1

Fix $0 \le \epsilon < 1$. We have

$$\big\{R^{(\mathrm{CI})}(\rho, P, \boldsymbol{H}_1) = R'(\rho, P, \boldsymbol{H}_1) \text{ for any } 0 \le \rho \le \epsilon\big\} = \bigcap_{0\le\rho\le\epsilon} \big\{R'(\rho, P, \boldsymbol{H}_1) \le R''(\rho, P, \boldsymbol{H}_1)\big\}. \tag{118}$$

Let us write

$$\begin{aligned} \mathbb{P}\left(\bigcap_{0\le\rho\le\epsilon} \{R'(\rho, P, \boldsymbol{H}_1) \le R''(\rho, P, \boldsymbol{H}_1)\}\right) &= \mathbb{P}\left(\bigcap_{0\le\rho\le\epsilon}\left\{\log\frac{\alpha(P,\boldsymbol{H}_1) - \beta(P,\boldsymbol{H}_1)\rho^2}{(\gamma'(P,\boldsymbol{H}_1) - \delta'(P,\boldsymbol{H}_1)\rho^2)^2} \le 0\right\}\right) \\ &= \mathbb{P}\left(\bigcap_{0\le\rho\le\epsilon}\left\{\frac{\alpha(P,\boldsymbol{H}_1) - \beta(P,\boldsymbol{H}_1)\rho^2}{(\gamma'(P,\boldsymbol{H}_1) - \delta'(P,\boldsymbol{H}_1)\rho^2)^2} \le 1\right\}\right) \\ &= \mathbb{P}\left(\sup_{0\le\rho\le\epsilon}\frac{\alpha(P,\boldsymbol{H}_1) - \beta(P,\boldsymbol{H}_1)\rho^2}{(\gamma'(P,\boldsymbol{H}_1) - \delta'(P,\boldsymbol{H}_1)\rho^2)^2} \le 1\right) \\ &= \mathbb{P}\Big(\sup_{0\le\rho\le\epsilon} f(\rho, P, \boldsymbol{H}_1) \le 1\Big), \end{aligned} \tag{119}$$

where we have defined

$$f(\rho, P, H_1) = \frac{\alpha(P_n, H_1) - \beta(P_n, H_1)\rho^2}{(\gamma'(P, H_1) - \delta'(P, H_1)\rho^2)^2}. \tag{120}$$

To proceed, we need the following lemma:

Lemma 4 *For any* $0 \le \epsilon \le 1$*,*

$$\lim_{P\to\infty} \mathbb{P}\big(f(\rho, P, \boldsymbol{H}_1) \textit{ is an increasing function of } 0 \le \rho \le \epsilon\big) = 1. \tag{121}$$

Proof: We have

$$\frac{\mathrm{d}f}{\mathrm{d}\rho} = \frac{2\rho\,(\Delta'(P,H_1) + (\alpha(P,H_1) - \beta(P,H_1)\rho^2)\delta'(P,H_1))}{(\gamma'(P,H_1) - \delta'(P,H_1)\rho^2)^3}, \tag{122}$$

where

$$\Delta'(P,H_1) = \alpha(P,H_1)\delta'(P,H_1) - \beta(P,H_1)\gamma'(P,H_1). \tag{123}$$

Since $(\alpha(P,H_1) - \beta(P,H_1)\rho^2)\delta'(P,H_1) \geq 0$ and $\gamma'(P,H_1) - \delta'(P,H_1)\rho^2 > 0$ for all ρ, P and H_1, then f is an increasing function of ρ if $\Delta'(P,H_1) > 0$. Therefore,

$$\mathbb{P}\big(f(\rho,P,\boldsymbol{H}_1) \text{ is an increasing function of } 0 \leq \rho \leq \epsilon\big) \geq \mathbb{P}(\Delta'(P,\boldsymbol{H}_1) > 0). \tag{124}$$

$\Delta'(P,H_1)$ is a polynomial of degree at most four in terms of P. In fact,

$$\Delta'(P,H_1) = \Delta'_1(H_1)\frac{P^2}{4} + \Delta'_2(H_1)\frac{P^3}{8} + \Delta'_3(H_1)\frac{P^4}{16}, \tag{125}$$

where

$$\Delta'_1(H_1) = |a_1|^2|a_2|^2 - |b_1b_2^* + a_1a_2^*|^2, \tag{126}$$

$$\Delta'_2(H_1) = |a_1|^2|a_2|^2(|b_1|^2 + |b_2|^2 + |a_1|^2 + |a_2|^2) - (|a_1|^2 + |a_2|^2)|b_1b_2^* + a_1a_2^*|^2 \tag{127}$$

and

$$\begin{aligned}\Delta'_3(H_1) &= |a_1|^2|a_2|^2(|a_1|^2 + |b_1|^2)(|a_2|^2 + |b_2|^2) - |a_1|^2|a_2|^2|a_1a_2^* + b_1b_2^*|^2 \\ &= |a_1|^2|a_2|^2\left((|a_1|^2 + |b_1|^2)(|a_2|^2 + |b_2|^2) - |a_1a_2^* + b_1b_2^*|^2\right) \\ &= |a_1|^2|a_2|^2|a_1b_2 - a_2b_1|^2. \end{aligned} \tag{128}$$

As $\boldsymbol{a}_1\boldsymbol{b}_2 - \boldsymbol{a}_2\boldsymbol{b}_1 = \det(\boldsymbol{H}_1) \neq 0$,

$$\mathbb{P}(\Delta'_3(\boldsymbol{H}_1) > 0) = 1. \tag{129}$$

Letting $P \to \infty$,

$$\begin{aligned}\lim_{P\to\infty}\mathbb{P}(\Delta'(P,\boldsymbol{H}_1) > 0) &= \lim_{P\to\infty}\mathbb{P}\left(\Delta'_1(\boldsymbol{H}_1)\frac{P^2}{4} + \Delta'_2(\boldsymbol{H}_1)\frac{P^3}{8} + \Delta'_3(\boldsymbol{H}_1)\frac{P^4}{16} > 0\right) \\ &= \lim_{P\to\infty}\mathbb{P}\left(4\Delta'_1(\boldsymbol{H}_1)P^{-2} + 2\Delta'_2(\boldsymbol{H}_1)P^{-1} + \Delta'_3(\boldsymbol{H}_1) > 0\right) \\ &\overset{(a)}{=} \mathbb{P}(\Delta'_3(\boldsymbol{H}_1) > 0) \\ &\overset{(b)}{=} 1, \end{aligned} \tag{130}$$

where (a) is due to dominated convergence [28] and (b) follows by (129). Finally, (124) and (130) conclude the proof of Lemma 4. ■

For simplicity, let $\mathcal{F}_P$ be the event of $f(\rho, P, \boldsymbol{H}_1)$ being an increasing function of $0 \leq \rho \leq \epsilon$. Then by Lemma 4,

$$\lim_{P\to\infty} \mathbb{P}\left(\mathcal{F}_P\right) = 1. \tag{131}$$

Let us write

$$\begin{aligned} \mathbb{P}\big(\sup_{0\leq\rho\leq\epsilon} f(\rho, P, \boldsymbol{H}_1) \leq 1\big) &\geq \mathbb{P}\big(\sup_{0\leq\rho\leq\epsilon} f(\rho, P, \boldsymbol{H}_1) \leq 1, \mathcal{F}_P\big) \\ &= \mathbb{P}\big(f(\epsilon, P, \boldsymbol{H}_1) \leq 1, \mathcal{F}_P\big) \\ &\geq \mathbb{P}\big(f(\epsilon, P, \boldsymbol{H}_1) \leq 1\big) + \mathbb{P}(\mathcal{F}_P) - 1, \end{aligned} \tag{132}$$

where the penultimate step is due to $\sup_{\rho\leq\epsilon} f(\rho, P, \boldsymbol{H}_1) = f(\epsilon, P, \boldsymbol{H}_1)$ under $\mathcal{F}_P$ and the last step follows by the bound $\mathbb{P}(\mathcal{A} \bigcap \mathcal{B}) \geq \mathbb{P}(\mathcal{A}) + \mathbb{P}(\mathcal{B}) - 1$ for any two events $\mathcal{A}$ and $\mathcal{B}$. By (120) and noting that $\epsilon < 1$, $\lim_{P\to\infty} f(\epsilon, P, \boldsymbol{H}_1) = 0$. Then one can invoke dominated convergence [28] to get

$$\lim_{P\to\infty} \mathbb{P}\left(f(\epsilon, P, \boldsymbol{H}_1) \leq 1\right) = 1. \tag{133}$$

Using (131) and (133) in (132),

$$\lim_{P\to\infty} \mathbb{P}\bigg(\sup_{0\leq\rho\leq\epsilon} f(\epsilon, P, \boldsymbol{H}_1) \leq 1\bigg) = 1. \tag{134}$$

This together with (118) and (119) complete the proof of Lemma 1.

Appendix B; Proof of Lemma 2

We need several properties of $K_n(\cdot)$ given as follows [34]:

- Recursion Rule:

$$K_{n+1}(z) - K_{n-1}(z) = \frac{2n}{z} K_n(z), \tag{135}$$

- Differentiation Rule:

$$\frac{\mathrm{d}}{\mathrm{d}z} K_0(z) = -K_1(z), \quad \frac{\mathrm{d}}{\mathrm{d}z}(z^n K_n(z)) = -z^n K_{n-1}(z), \quad n \geq 1, \tag{136}$$

- Asymptotic Exapansion:

$$K_n(z) \overset{z\to\infty}{\sim} \sqrt{\frac{\pi}{2z}}e^{-z}(1+o(1)) \tag{137}$$

and

$$K_0(z) \overset{z\to 0}{\sim} -\ln z, \quad K_n(z) \overset{z\to 0}{\sim} \frac{(n-1)!}{2}\left(\frac{2}{z}\right)^n, \; n \geq 1, \tag{138}$$

where for two functions f and g and $c \in [-\infty, \infty]$, $f \overset{z\to c}{\sim} g$ means $\lim_{z\to c}\frac{f(z)}{g(z)} = 1$.

Let $\boldsymbol{u}_i$ for $1 \leq i \leq 4$ be independent exponential random variables with parameter 1. Define

$$v_t = (u_1 + tu_3)(u_2 + tu_4). \tag{139}$$

where $t > 0$. Clearly, $|\boldsymbol{a}_1|^2|\boldsymbol{a}_2|^2$ and $(|\boldsymbol{a}_1|^2 + |\boldsymbol{b}_1|^2)(|\boldsymbol{a}_2|^2 + |\boldsymbol{b}_2|^2)$ are identically distributed as $\boldsymbol{v}_0$ and $\boldsymbol{v}_{\sigma^2}$, respectively. As such, it suffices to find the cumulative distribution function of $\boldsymbol{v}_t$ for arbitrary $t > 0$. Both $p_{u_1+tu_3}(\cdot)$ and $p_{u_2+tu_4}(\cdot)$ are the convolution of the PDFs $e^{-u}\mathbb{1}_{u>0}$ and $\frac{1}{t}e^{-\frac{u}{t}}\mathbb{1}_{u>0}$. Then

$$p_{u_1+tu_3}(u) = p_{u_2+tu_4}(u) = \begin{cases} \frac{1}{t-1}\left(e^{-\frac{u}{t}} - e^{-u}\right)\mathbb{1}_{u>0} & t \neq 1 \\ ue^{-u}\mathbb{1}_{u>0} & t = 1 \end{cases}. \tag{140}$$

Using the fact that $p_{\boldsymbol{xy}}(z) = \int_{-\infty}^{\infty}\frac{1}{|x|}p_{\boldsymbol{y}}(\frac{z}{x})p_{\boldsymbol{x}}(x)\mathrm{d}x$ for any two independent random variables $\boldsymbol{x}$ and $\boldsymbol{y}$, we can write

$$\begin{aligned} p_{\boldsymbol{v}_t}(v) &= \int_0^\infty \frac{1}{v}p_{u_1+tu_3}(u)p_{u_2+tu_4}\left(\frac{v}{u}\right)\mathrm{d}v \\ &= \frac{1}{(t-1)^2}\int_0^\infty \frac{1}{u}e^{-(\frac{u}{t}+\frac{v}{tu})}\mathrm{d}u - \frac{1}{(t-1)^2}\int_0^\infty \frac{1}{u}e^{-(\frac{u}{t}+\frac{v}{u})}\mathrm{d}u \\ &\quad -\frac{1}{(t-1)^2}\int_0^\infty \frac{1}{u}e^{-(u+\frac{v}{tu})}\mathrm{d}u + \frac{1}{(t-1)^2}\int_0^\infty \frac{1}{u}e^{-(u+\frac{v}{u})}\mathrm{d}u, \end{aligned} \tag{141}$$

for $t \neq 1$ and

$$p_{\boldsymbol{v}_t}(v) = v\int_0^\infty \frac{1}{u}e^{-(u+\frac{v}{u})}\mathrm{d}u, \tag{142}$$

for $t = 1$. By identity (9.42) on page 235 in [34], $\int_0^\infty u^{-n-1}e^{-(c_1u+\frac{c_2}{u})}\mathrm{d}u = 2(\frac{c_1}{c_2})^{\frac{n}{2}}K_n(2\sqrt{c_1c_2})$. Then

$$p_{\boldsymbol{v}_t}(v) = \begin{cases} \frac{2}{(t-1)^2}\left(K_0(\frac{2\sqrt{v}}{t}) - 2K_0(2\sqrt{\frac{v}{t}}) + K_0(2\sqrt{v})\right) & t \neq 1 \\ 2vK_0(2\sqrt{v}) & t = 1 \end{cases}. \tag{143}$$

Next, let us compute $\mathbb{P}(\boldsymbol{v}_t \leq v)$. If $t = 1$,

$$\mathbb{P}(\boldsymbol{v}_t \leq v) = \int_0^v 2xK_0(2\sqrt{x})\mathrm{d}x = \frac{1}{4}\int_0^{2\sqrt{v}} y^3 K_0(y)\mathrm{d}y, \tag{144}$$

where we have applied the change of variable $y = 2\sqrt{x}$. Using the differentiation rule in (136) and integration by parts, it is easy to see that $\int y^3 K_0(y)\mathrm{d}y = -y^3K_1(y) - 2y^2K_2(y)$. Then

$$\begin{aligned}\mathbb{P}(\boldsymbol{v}_1 \leq v) &= \frac{1}{4}\Big[-\big(y^3K_1(y) + 2y^2K_2(y)\big)\Big]_0^{2\sqrt{v}} \\ &\overset{(a)}{=} 1 - 2v(\sqrt{v}K_1(2\sqrt{v}) + K_2(2\sqrt{v})) \\ &\overset{(b)}{=} 1 - 2\sqrt{v}K_0(2\sqrt{v}) - 2\sqrt{v}(1+v)K_1(2\sqrt{v}),\end{aligned} \tag{145}$$

where in (a) we have used the asymptotic behaviour for small argument in (137) to get $\lim_{z\to 0} z^3K_1(z) = 0$ and $\lim_{z\to 0} z^2K_2(z) = 2$ and (b) is due to the recursion rule in (135), i.e., $K_2(2\sqrt{v}) = K_0(2\sqrt{v}) + \frac{1}{\sqrt{v}}K_1(2\sqrt{v})$.

If $t \neq 1$, one can use a similar approach to prove

$$\mathbb{P}(\boldsymbol{v}_t \leq v) = 1 - \frac{2}{(t-1)^2}\Big(t\sqrt{v}K_1(\tfrac{2}{t}\sqrt{v}) - 2\sqrt{tv}K_1(\tfrac{2}{t}\sqrt{tv}) + \sqrt{v}K_1(2\sqrt{v})\Big). \tag{146}$$

Finally, $\lim_{t\to 0}\boldsymbol{v}_t = \boldsymbol{v}_0$ implies that $\boldsymbol{v}_t$ converges weakly to $\boldsymbol{v}_0$, i.e.,

$$\mathbb{P}(\boldsymbol{v}_0 \leq v) = \lim_{t\to 0}\mathbb{P}(\boldsymbol{v}_t \leq v) = 1 - 2\sqrt{v}K_0(2\sqrt{v}), \tag{147}$$

where in the last step we have used the asymptotic expansion for large argument in (137) to conclude $\lim_{t\to 0} tK_0(\frac{2}{t}\sqrt{v}) = \lim_{t\to 0}\sqrt{t}K_0(\frac{2}{t}\sqrt{tv}) = 0$.

APPENDIX C; PROOF OF LEMMA 3

Define

$$\phi(0) = 0, \quad \phi(k) = 1 + \frac{1}{2} + \cdots + \frac{1}{k}, \;\; k \geq 1. \tag{148}$$

Then one can expand $K_n(\cdot)$ as [33]

$$K_0(z) = \sum_{k=0}^{\infty}\frac{\left(\frac{z}{2}\right)^{2k}}{(k!)^2}\Big(\phi(k) - \gamma - \ln\frac{z}{2}\Big) \tag{149}$$

$$K_1(z) = \frac{1}{z} - \sum_{k=0}^{\infty}\frac{\left(\frac{z}{2}\right)^{2k+1}}{k!(k+1)!}\Big(\frac{\phi(k)+\phi(k+1)}{2} - \gamma - \ln\frac{z}{2}\Big) \tag{150}$$

where γ is the so-called Euler-Mascheroni constant. By (76),

$$\begin{aligned}\Psi_1(v) &= (1-2\gamma-\ln v)v+\Big(\frac{5}{4}-\gamma-\frac{1}{2}\ln v\Big)v^2+\cdots \\ &= -v\ln v+o(v\ln v). \end{aligned} \tag{151}$$

By (77) and for $\sigma^2=1$,

$$\begin{aligned}\Psi_2(v) &= \Big(\frac{1}{4}-\gamma-\frac{1}{2}\ln v\Big)v^2+\Big(\frac{7}{9}-\frac{2}{3}\gamma-\frac{1}{3}\ln v\Big)v^3+\cdots \\ &= -\frac{1}{2}v^2\ln v+o(v^2\ln v). \end{aligned} \tag{152}$$

If $\sigma^2\neq 1$,

$$\begin{aligned}\Psi_2(v) &= \frac{1}{\sigma^4}\Big(\frac{5}{4}-\gamma-\ln\sigma^4-\frac{1}{2}\ln v\Big)v^2+\frac{(\sigma^2+1)^2}{6\sigma^8}\Big(\frac{5}{3}-\gamma+\frac{1}{2}\ln v\Big)v^3+\cdots \\ &= -\frac{1}{2\sigma^4}v^2\ln v+o(v^2\ln v). \end{aligned} \tag{153}$$

APPENDIX D; DERIVATION OF (84)

We have

$$\begin{aligned}R^{(\mathrm{TIN})}(0,P,H_1) &= \log\frac{\left(1+\frac{P}{2}\left(|a_1|^2+|b_1|^2\right)\right)\left(1+\frac{P}{2}\left(|a_2|^2+|b_2|^2\right)\right)}{\left(1+\frac{P}{2}|b_1|^2\right)\left(1+\frac{P}{2}|b_2|^2\right)} \\ &= \log\left(\left(1+\frac{\frac{P}{2}|a_1|^2}{1+\frac{P}{2}|b_1|^2}\right)\left(1+\frac{\frac{P}{2}|a_2|^2}{1+\frac{P}{2}|b_2|^2}\right)\right). \end{aligned} \tag{154}$$

But,

$$\begin{aligned}\mathbb{P}\left(1+\frac{\frac{P}{2}|\boldsymbol{a}_1|^2}{1+\frac{P}{2}|\boldsymbol{b}_1|^2}<x\right) &\overset{(a)}{=} \mathbb{E}\Big[\mathbb{P}\Big(1+\frac{\frac{P}{2}|\boldsymbol{a}_1|^2}{1+\frac{P}{2}|\boldsymbol{b}_1|^2}<x\Big|\boldsymbol{b}_1=b_1\Big)\Big] \\ &= \mathbb{E}\Big[\mathbb{P}\Big(|\boldsymbol{a}_1|^2<\frac{2}{P}(x-1)+(x-1)|b_1|^2\Big)\Big] \\ &= 1-e^{-\frac{2}{P}(x-1)}\mathbb{E}\big[e^{-(x-1)|\boldsymbol{b}_1|^2}\big] \\ &\overset{(b)}{=} 1-\frac{e^{-\frac{2}{P}(x-1)}}{1+\sigma^2(x-1)},\quad x\geq 1, \end{aligned} \tag{155}$$

where (a) is by the tower property for conditional expectations [28] and (b) is due to the fact that $|\boldsymbol{b}_1|^2$ is an exponential random variable with parameter $\frac{1}{\sigma^2}$. Finally, (84) follows using the fact that $\mathbb{P}(\boldsymbol{uv}<w)=\int_{-\infty}^{\infty}\mathbb{P}(\boldsymbol{v}<\frac{w}{u})p_{\boldsymbol{u}}(u)\mathrm{d}u$ for any two independent random variables $\boldsymbol{u}$ and $\boldsymbol{v}$.

Appendix E; Proof of Theorem 3

By (22),

$$R^{(\mathrm{TIN})}(\rho, P_n, H_n) = \log \frac{1 + \frac{P_n}{2} c_n + \frac{P_n^2}{4}\left(\left(|a_1|^2 + \sigma_n^2 |b_1|^2\right)\left(|a_2|^2 + \sigma_n^2 |b_2|^2\right) - \rho^2 |a_1 a_2^* + \sigma_n^2 b_1 b_2^*|^2\right)}{1 + \frac{P_n \sigma_n^2}{2}\left(|b_1|^2 + |b_2|^2\right) + \frac{P_n^2 \sigma_n^4}{4} |b_1|^2 |b_2|^2 (1-\rho^2)}, \quad (156)$$

where $c_n = |a_1|^2 + |a_2|^2 + \sigma_n^2(|b_1|^2 + |b_2|^2)$. Since $\lim_{n\to\infty} P_n \sigma_n^2 = 0$ and P_n is an unbounded sequence, then $\lim_{n\to\infty} \sigma_n^2 = \lim_{n\to\infty} P_n^2 \sigma_n^4 = 0$. Using these facts in (156), we get

$$\lim_{n\to\infty} R^{(\mathrm{TIN})}(\rho, P_n, \boldsymbol{H}_n) - \log\left(1 + \frac{P_n}{2}(|\boldsymbol{a}_1|^2 + |\boldsymbol{a}_2|^2) + \frac{P_n^2}{4}(1-\rho^2)|\boldsymbol{a}_1|^2 |\boldsymbol{a}_2|^2\right) = 0, \quad (157)$$

almost surely for any $0 \le \rho < 1$ and

$$\lim_{n\to\infty} R^{(\mathrm{TIN})}(1, P_n, \boldsymbol{H}_n) - \log\left(1 + \frac{P_n}{2}(|\boldsymbol{a}_1|^2 + |\boldsymbol{a}_2|^2) + \frac{P_n^2 \sigma_n^2}{4}|\boldsymbol{a}_1 \boldsymbol{b}_2 - \boldsymbol{a}_2 \boldsymbol{a}_1|^2\right) = 0. \quad (158)$$

Moreover,

$$\lim_{n\to\infty} \log\left(1 + \frac{P_n}{2}(|\boldsymbol{a}_1|^2 + |\boldsymbol{a}_2|^2) + \frac{P_n^2}{4}(1-\rho^2)|\boldsymbol{a}_1|^2 |\boldsymbol{a}_2|^2\right) - r \log P_n = \infty, \quad (159)$$

for any $0 \le r < 2$ and

$$\begin{aligned}
&\lim_{n\to\infty} \log\left(1 + \frac{P_n}{2}(|\boldsymbol{a}_1|^2 + |\boldsymbol{a}_2|^2) + \frac{P_n^2 \sigma_n^2}{4}|\boldsymbol{a}_1 \boldsymbol{b}_2 - \boldsymbol{a}_2 \boldsymbol{b}_1|^2\right) - r \log P_n \\
&= \lim_{n\to\infty} \log \frac{1 + \frac{P_n}{2}\left((|\boldsymbol{a}_1|^2 + |\boldsymbol{a}_2|^2) + \frac{P_n \sigma_n^2}{2}|\boldsymbol{a}_1 \boldsymbol{b}_2 - \boldsymbol{a}_2 \boldsymbol{b}_1|^2\right)}{P_n^r} \\
&= \lim_{n\to\infty} \log \frac{\frac{P_n}{2}(|\boldsymbol{a}_1|^2 + |\boldsymbol{a}_2|^2)}{P_n^r} = \infty, \qquad (160)
\end{aligned}$$

for any $0 \le r < 1$ where the last step is due to $\lim_{n\to\infty} P_n \sigma_n^2 = 0$. Note that in writing (158) and (160), we have used the facts that $\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_1 \boldsymbol{b}_2 - \boldsymbol{a}_2 \boldsymbol{b}_1 \neq 0$. By (157) and (159) and using dominated convergence [28],

$$\lim_{n\to\infty} \mathbb{P}\left(R^{(\mathrm{TIN})}(\rho, P_n, \boldsymbol{H}_n) \le r \log P_n\right) = 0, \quad (161)$$

for any $0 \le r < 2$. Similarly, by (158) and (160),

$$\lim_{n\to\infty} \mathbb{P}\left(R^{(\mathrm{TIN})}(1, P_n, \boldsymbol{H}_n) \le r \log P_n\right) = 0, \quad (162)$$

for any $0 \leq r < 1$. Finally, noting that the value of the limit in (160) changes to $-\infty$ if $r > 1$, we conclude that $\lim_{n\to\infty} \mathbb{P}(R^{(\mathrm{NIT})}(1, P_n, \boldsymbol{H}_n) \leq r \log P_n) = 1$ if $r > 1$.

APPENDIX F; DERIVATION OF (93) AND (97)

- Computation of $\mathrm{d}(R^{(\mathrm{TIN})}(\rho, P, H_1))$:

  For any complex square matrix M with positive determinant, the complex differential [31] of $\ln\det(M)$ is given by $\mathrm{d}(\ln\det M) = \mathrm{tr}(M^{-1}\mathrm{d}M)$ where $\mathrm{d}M$ is the element-wise differential of M. Then

$$\begin{aligned}\mathrm{d}(R^{(\mathrm{TIN})}(\rho, P, H_1)) &= \mathrm{d}\left(\log\frac{\det\Omega}{\det\Xi}\right)\\ &= \mathrm{d}(\log\det\Omega) - \mathrm{d}(\log\det\Xi)\\ &= \mathrm{tr}(\Omega^{-1}\mathrm{d}\Omega) - \mathrm{tr}(\Xi^{-1}\mathrm{d}\Xi). \qquad (163)\end{aligned}$$

  But,

$$\begin{aligned}\mathrm{d}\Omega &= \mathrm{d}\Big(I_N + \frac{P}{N}AC_\rho A^\dagger + \frac{P}{N}BC_\rho B^\dagger\Big)\\ &= \frac{P}{N}A\,\mathrm{d}(C_\rho)A^\dagger + \frac{P}{N}B\,\mathrm{d}(C_\rho)B^\dagger\\ &= \frac{P}{N}\left(A\,JA^\dagger + B\,JB^\dagger\right)\mathrm{d}\rho, \qquad (164)\end{aligned}$$

  where the last step follows by $\mathrm{d}(C_\rho) = \mathrm{d}(\rho\vec{1}_N\vec{1}_N^{\mathrm{t}} + (1-\rho)I_N) = J\mathrm{d}\rho$ in which J is defined in (94). Similarly,

$$\mathrm{d}\Xi = \frac{P}{N}B\,JB^\dagger\mathrm{d}\rho. \qquad (165)$$

  By (163), (164) and (165), we obtain (93).

- Computation of $\mathrm{d}^2(R^{(\mathrm{TIN})}(\rho, P, H_1))$:

  Let M be a complex and invertible square matrix and Q be an arbitrary complex matrix such that the product MQ is defined. The complex differential of $\mathrm{tr}(M^{-1}Q)$ is given by [31] $\mathrm{d}(\mathrm{tr}(M^{-1}Q)) =$

$\mathrm{tr}(M^{-1}QM^{-1}\mathrm{d}M)$. Applying this fact to (93),

$$\begin{aligned}
\mathrm{d}\left(\frac{N}{P}\frac{\mathrm{d}(R^{(\mathrm{TIN})}(\rho,P,H_1))}{\mathrm{d}\rho}\right) &= \mathrm{d}\left(\mathrm{tr}\left(\Omega^{-1}(AJA^\dagger+BJB^\dagger)\right)\right)-\mathrm{d}\left(\mathrm{tr}\left(\Xi^{-1}BJB^\dagger\right)\right)\\
&= \mathrm{tr}\left(\Omega^{-1}(AJA^\dagger+BJB^\dagger)\Omega^{-1}\mathrm{d}\Omega\right)-\mathrm{tr}\left(\Xi^{-1}BJB^\dagger\Xi^{-1}\mathrm{d}\Xi\right)\\
&= \frac{P}{N}\mathrm{tr}\left(\Omega^{-1}(AJA^\dagger+BJB^\dagger)\Omega^{-1}(AJA^\dagger+BJB^\dagger)\right)\mathrm{d}\rho\\
&\quad -\frac{P}{N}\mathrm{tr}\left(\Xi^{-1}BJB^\dagger\Xi^{-1}BJB^\dagger\right)\mathrm{d}\rho\\
&= \frac{P}{N}\mathrm{tr}\left(\left(\Omega^{-1}(AJA^\dagger+BJB^\dagger)\right)^2\right)\mathrm{d}\rho-\frac{P}{N}\mathrm{tr}\left(\left(\Xi^{-1}BJB^\dagger\right)^2\right)\mathrm{d}\rho,
\end{aligned} \tag{166}$$

where the penultimate step is due to (164) and (165). Therefore,

$$\frac{\mathrm{d}^2}{\mathrm{d}\rho^2}R^{(\mathrm{TIN})}(\rho,P,H_1) = \left(\frac{P}{N}\right)^2\left(\mathrm{tr}\left(\left(\Omega^{-1}(AJA^\dagger+BJB^\dagger)\right)^2\right)-\mathrm{tr}\left(\left(\Xi^{-1}BJB^\dagger\right)^2\right)\right). \tag{167}$$

Noting that both Ω^{-1} and Ξ^{-1} are diagonal at $\rho=0$, we get

$$\left[\Omega_0^{-1}(AJA^\dagger+BJB^\dagger)\right]_{i,j}=\frac{a_ia_j^*+b_ib_j^*}{1+\frac{P}{N}(|a_i|^2+|b_i|^2)}\mathbb{1}_{i\neq j} \tag{168}$$

and

$$\left[\Xi_0^{-1}BJB^\dagger\right]_{i,j}=\frac{b_ib_j^*}{1+\frac{P}{N}|b_i|^2}\mathbb{1}_{i\neq j}, \tag{169}$$

where Ω_0 and Ξ_0 are Ω and Ξ evaluated at $\rho=0$, respectively. Then

$$\begin{aligned}
\mathrm{tr}\left(\left(\Omega_0^{-1}(AJA^\dagger+BJB^\dagger)\right)^2\right) &= \sum_{i=1}^{N}\left[\left(\Omega_0^{-1}(AJA^\dagger+BJB^\dagger)\right)^2\right]_{i,i}\\
&= \sum_{i=1}^{N}\sum_{j=1}^{N}\left[\Omega_0^{-1}(AJA^\dagger+BJB^\dagger)\right]_{i,j}\left[\Omega_0^{-1}(AJA^\dagger+BJB^\dagger)\right]_{j,i}\\
&= 2\sum_{1\le i<j\le N}\frac{|a_ia_j^*+b_ib_j^*|^2}{\left(1+\frac{P}{N}(|a_i|^2+|b_i|^2)\right)\left(1+\frac{P}{N}(|a_j|^2+|b_j|^2)\right)},
\end{aligned} \tag{170}$$

where the last step is due to (168). Similarly,

$$\mathrm{tr}\left(\left(\Xi^{-1}BJB^\dagger\right)^2\right)=2\sum_{1\le i<j\le N}\frac{|b_ib_j^*|^2}{\left(1+\frac{P}{N}|b_i|^2\right)\left(1+\frac{P}{N}|b_j|^2\right)}. \tag{171}$$

Finally, (167), (170) and (171) yield (97).

REFERENCES

[1] N. Seshadri and J. H. Winters, “Two signalling schemes for improving the error performance of frequency division duplex (FDD) transmission systems using transmitter antenna diversity”, *International Journal of Wireless Information Networks*, vol. 1, no. 1, pp. 49-60, Jan. 1994.

[2] J. Li, X. Wu and R. Laroia, “OFDMA Mobile Broadband Communications: A Systems Approach” *Cambridge University Press*, 2013.

[3] E. Dahlman, S. Parkvall and J. Skold, “4G: LTE/LTE-Advanced for Mobile Broadband”, *Academic Press*, 2013.

[4] A. Hiroike, F. Adachi and N. Nakajima, “Combined effects of phase sweeping transmitter diversity and channel coding”, *IEEE Trans. Veh. Tech.*, vol. 41, no. 2, pp. 170-176, Aug. 2002.

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

[6] 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.

[7] 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.

[8] 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.

[9] 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.

[10] 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.

[11] 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.

[12] 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.

[13] 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.

[14] 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.

[15] I. E. Teletar, “Capacity of multiple-antenna Gaussian channels”, *European Trans. Telecommun.* vol. 10, no. 6, pp. 585-595, Nov. 1999.

[16] M. Katz and S. Shamai (Shitz), “On the outage probability of multiple-input single-output communication link”, *IEEE Trans. Wireless Commun.* vol. 6, no. 11, pp. 4120-4128, Nov. 2007.

[17] E. Abbe, Shao-Lun Huang and E. Telatar, “Proof of the outage probability conjecture for MISO channels”, *IEEE Trans. Inf. Theory* vol. 59, no. 5, pp. 2596-2602, May 2013.

[18] L. Zheng and D. N. C. Tse, “Diversity and multiplexing: A fundamental tradeoff in multiple antenna channels”, *IEEE Trans. Inf. Theory*, vol. 49, pp. 1073–1096, 2002.

[19] R. Etkin, D. Tse, and H. Wang, “Gaussian interference channel capacity to within one bit”, *IEEE Trans. Inf. Theory*, vol. 54, no. 12, pp. 5534–5562, 2008.

[20] 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.

[21] 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.

[22] 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.

[23] Q. Li, K. Li, and K. Teh, “Diversity-multiplexing tradeoff of fading interference channels with source cooperation and partial CSIT”, *IEEE Trans. Inf. Theory*, vol. 57, no. 5, pp. 2781–2804, 2011.

[24] 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.

[25] 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.

[26] D. Tse, P. Viswanath, “Fundamentals of Wireless Communication” *Cambridge university press*, 2005.

[27] T. M. Cover and J. A. Thomas, ”Elements of information theory”, *John Wiley and Sons, Inc.*, 1991.

[28] R. M. Dudley, *Real analysis and probability*. Cambridge University Press, 2005.

[29] W. Rudin, “Principles of mathematical analysis”, *McGraw-Hill Inc.*, 1976.

[30] H. J. Malik and R. Trudel, “Probability density function of the product and quotient of two correlated exponential random variables”, *Canadian Mathematical Bulletin*, vol. 29, no. 4, pp. 413–418, 1986.

[31] A. Hjørungnes and D. Gesbert, “Complex-valued matrix differentiation: Techniques and key results”, *IEEE Trans. Signal Processing*, vol. 55, no. 6, pp. 2740-2746, June 2007.

[32] A. Tulino and S. Verdú, “Random matrix theory and wireless communications”, *Foundations and Trends in Commun. and Inf. Theory*, vol. 1, Issue 1, 2004.

[33] C. R. Wylie and L. C. Barrett, “Advanced engineering mathematics”, *McGraw-Hill Inc.*, 1982.

[34] N. M. Temme, “Special Functions: An introduction to classical functions of mathematical physics”, *John Wiley and Sons Inc.*, 1996.

[35] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*. 3rd ed., 1991.