Optimal Transmission in MIMO Channels with Multiuser Interference

Vien V. Mai, Jin Sam Kwak, *Member, IEEE*, Youngmin Jeong, *Member, IEEE*, and Hyundong Shin, *Senior Member, IEEE*

***Abstract*—Cochannel interference is one of the inevitable deleterious components in designing and analyzing of a wireless network. The use of multiple antennas at both transmitting and receiving nodes is a promising technique to suppress and/or alleviate the effect of cochannel interference on capacity. In this paper, we assess the effects of both antenna correlation and cochannel interference on the ergodic capacity of multiple-input multiple-output (MIMO) channels with covariance feedback. In particular, we consider a general family of spatial fading correlation model—called *unitary-independent-unitary*—which encompasses most of zero-mean channels with arbitrary fading profiles including the popular separable correlation channel models. We derive the average minimum mean-square error and signal-to-interference-plus-noise ratio of the parallel spatial streams using Berezin's *supermathematics*. We then put forth the structure of optimal input covariance matrix maximizing the mutual information connected with the necessary and sufficient conditions as a generalization of the noise-limited case, which is tested by a simple iterative algorithm. Together with the powerful supermathematical framework, the result in the paper enables us to quantify the multiuser MIMO interference effects on the capacity in terms of spatial correlation and interference power heterogeneity.**

***Index Terms*—Achievable rate, cochannel interference, Grassmann algebra, minimum mean-square error (MMSE), multiple-input multiple-output (MIMO), power allocation, supermathematics, unitary-independent-unitary (UIU) channel model.**

I. Introduction

THE potential benefit of multiple-input multiple-output (MIMO) systems has been spurred to integrate key technologies into the future wireless communication systems [1]–[10]. The information theoretic analysis of the MIMO capacity showed that it scales linearly with the number of antennas operating on a single link with additive white Gaussian noise (AWGN) in rich-scattering wireless environments. However, the potential advantage of MIMO system may be limited by the spatial fading correlation due to closely-packed antennas and/or cochannel interference due to heterogeneous systems operating in the same spectrum [11]–[17].

When the channel state information (CSI) is only available at the receiver, the ergodic capacity of the independent and identically distributed (i.i.d.) MIMO Rayleigh fading channel was first analyzed in [2], in which it was proven that isotropic i.i.d. Gaussian inputs are capacity-achieving. In order to account for the more realistic propagation environments—more specifically, the capacity loss due to *spatial correlation*—a large wave of works has been spawned on the capacity analysis including: i) correlated MIMO Rayleigh-faded channels with one side correlation [12], [18]; ii) doubly correlated MIMO Rayleigh-faded channels [13], [19], [20]; iii) rank-deficient channels such as double scattering and keyhole channels [11], [21]; and spare channels with virtual representation [22], [23]. When the statistical CSI is available at the transmitter, it can be exploited to increase the capacity of the MIMO system. To that end, the optimal input covariance matrix structures of the covariance feedback system have been characterized for multiple-input single-output (MISO) and MIMO channels [24]–[26]. It was shown that i) the eigenvectors of the capacity-achieving input covariance matrix coincide with those of the channel covariance matrix; ii) the number of positive eigenvalues of the capacity-achieving input covariance matrix corresponds to the number of active directions to which the transmit signals are sent out; and iii) these eigenvalues (corresponding to the power allocation) can be solved by iterative algorithms. If the capacity-achieving input covariance matrix has rank one, the power allocation strategy is referred to as *beamforming*. The capacity and corresponding beamforming solutions for noise-limited MIMO systems with correlated Rayleigh-faded channels have been investigated in [27]–[29]. Most of the above works use the separable correlation model (also known as Kronecker product-form correlation model) whose accuracy is verified, especially, when scatters are locally rich at either the transmitter or receiver [30]. However, in spite of their analytic tractability and sufficient accuracy in certain environments, the separable correlation models often give a poor performance prediction of MIMO systems [30], [31]. To account for the general structure of MIMO channels with spatial correlation, the *unitary-independent-unitary* (UIU) channel model has been proposed which embraces most of zero-mean channel models including the separable correlation model [26].[1]

Another intrinsic property having an adverse impact on the

Manuscript received September 28, 2017; revised April 3, 2018; accepted August 7, 2018. Date of publication xxx xx, xxxx; date of current version xxx xx, xxxx. This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIP) (No. 2016R1A2B2014462). *(Corresponding authors: Youngmin Jeong and Hyundong Shin.)* The associate editor coordinating the review of this paper and approving it for publication was M. Cenk Gursoy.

Vien V. Mai, Y. Jeong, and H. Shin are with the Department of Electronic Engineering, Kyung Hee University, 1732 Deogyeong-daero, Giheung-gu, Yongin-si, Gyeonggi-do 17104, Korea (e-mail: yjeong@khu.ac.kr; hshin@khu.ac.kr).

Jin Sam Kwak is with the WILUS Inc., 216 Hwangsaeul-ro, Bundang-gu, Seongnam-si, Gyounggi-do 13595, Korea

[1] Note that the UIU model dose not include non-zero mean channels. Even in the absence of interference, capacity characterizations are available only few specific channels such as i.i.d. Rician fading channels (see [14] and references therein).

MIMO capacity is the presence of cochannel interference. The simulation study showed that cochannel interference can seriously degrade the overall capacity of MIMO systems [32]. Along with the impact of cochannel interference on the MIMO capacity, many studies have been devoted to the analysis of MIMO systems [15]–[17], [33]–[37]. Specifically, the optimum transmission strategies maximizing the ergodic MIMO capacity have been studied in [15], [34], and the MIMO capacity in large antenna systems over uncorrelated i.i.d. Rayleigh fading channel and separable correlation Rayleigh fading channel has been analyzed in [16] and [35], respectively. In [16], the authors also have studied the capacity-achieving input covariance in the regime of large numbers of antennas. In [15] and [34], it was found that adding more active transmit antennas at the probe transmitter can result in a lower overall system achievable rate at low aggregate signal-to-interference-plus-noise ratio (SINR) compared to the system allocating all the power budget to one transmit direction. The exact closed-form solution was studied for the capacity of uncorrelated MIMO networks in the presence of interference which is valid for an arbitrary number of interference antennas having possibly unequal power levels [17].

In the absence of cochannel interference, a set of necessary and sufficient conditions for optimal power allocation was investigated under the assumption that the statistical CSI is available at the transmitter in MISO channels [24] and UIU MIMO channels [26]. These conditions enable us to provide efficient iterative algorithms to find the input covariance matrix which maximizes the ergodic mutual information. An analytical result has been provided to reduce the computational burden in Rayleigh-faded channels with separable correlations [24]. It was shown that the power allocation algorithm requires to evaluate the average minimum mean-square error (MMSE) and SINR on the linear estimation of the spatial streams at each iteration [26]. This task resorts to the time-consuming Monte-Carlo expectation because the main difficulty in the exact analysis for general cases—unequal-power and spatially-correlated interferers—arises due to the absence of a tractable analytic framework to evaluate the random matrices relevant to the SINR.

In this paper, we develop a framework to characterize the ergodic capacity in general UIU MIMO models in the presence of multiple MIMO cochannel interferers. We consider a perfect channel knowledge at the receiver, while the transmitter has access only to the knowledge of long-term statistics (covariance). The main results of this paper can be summarized as follows.

- We first introduce the mathematical framework necessary to derive the closed-form expressions for the average MMSE and SINR on the linear estimation of the spatial streams. These streams play a central role in the optimum power allocation algorithm to maximize the capacity over correlated MIMO Rayleigh-faded interference channels with separable correlations (see Theorems 1 and 2). The key ingredient of our analysis is Berezin's supermathematics that treats the mathematical analysis and algebra for functions of both commuting and anticommuting variables on an equal footing [4], [38]–[40].
- We develop a necessary and sufficient condition for input optimization and investigate its MMSE representation. It is directly applicable to the general MIMO channels with an arbitrary power distribution of cochannel interferers, each is spatially correlated across receiving antennas (see Theorem 3). The analytical solutions for the average MMSE and SINR are valid for any number of unequal-power spatially-correlated interferers, each with an arbitrary number of transmit antennas (see Theorem 4). The simple iterative algorithm can be employed to find the optimum input power allocation without computational complexity (see Algorithm 1).
- We characterize the asymptotic capacity per receiver antenna as the antenna numbers at transmitters (probe transmitter and interferers) and receivers tend to infinity (see Theorem 5). We also introduce the fixed-point iteration method which converges very quickly to any desired accuracy level for calculating the asymptotic ergodic capacity (see Remark 5). Our result holds for a very general case when the channel gain has zero mean and arbitrary distributions. It is unveiled that the estimation of capacity in this work is powerful even for a small dimensional system.

The paper is organized as follows. In Section II, we describe the system model and then we formulate the optimum power allocation problem to maximize mutual information over input covariances for UIU channels. In Section III, we develop a mathematical methodology to analyze the average MMSE and SINR using the superanalysis framework. In Section IV, we characterize the conditions for the optimal power allocation and investigate the power optimization strategy in the presence of cochannel interference. Section V studies the asymptotic ergodic capacity of UIU MIMO channel with covariance feedback. In Section VI, we present some numerical results and finally, Section VII concludes the paper.

Throughout the paper, we shall adopt the notation: i) random variables are displayed in sans serif, upright fonts; their realizations in serif, italic fonts; and ii) vectors and matrices are denoted by bold lowercase and uppercase letters, respectively. For example, a random variable and its realization are denoted by x and x; a random vector and its realization are denoted by $\boldsymbol{\mathsf{x}}$ and $\boldsymbol{x}$; a random matrix and its realization are denoted by $\boldsymbol{\mathsf{X}}$ and $\boldsymbol{X}$, respectively. The notation and symbols used in the paper are tabulated to Table I. The basic concepts and formulae of the branch of Berezin's supermathematics can be found in [4, Table II], [16, Appendix II], [38]–[41] and references therein.

II. Models and Problem Formulation

We consider an $(n_0, [n_1, \ldots, n_L], n_{\mathrm{R}})$-MIMO interference channel where n_0 transmit and n_{R} receive antennas are equipped at the probe (desired) transmitter and receiver, respectively. There exist L cochannel interferers, each is equipped with n_ℓ transmit antennas, $\ell = 1, 2, \ldots, L$.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TWC.2018.2865939, IEEE Transactions on Wireless Communications

TABLE I
NOTATION AND SYMBOLS

$\mathbb{R}$	Real numbers
$\mathbb{R}_+$	Nonnegative real numbers
$\mathbb{R}_{++}$	Positive real numbers
$\mathbb{Z}_+$	Nonnegative integers
$\mathbb{C}$	Complex numbers
$\mathbb{G}$	Grassmann numbers
$\jmath$	Imaginary unit: $\jmath = \sqrt{-1}$
$\mathbb{1}\{\cdot\}$	Indicator function
$(\cdot)^*$	Complex conjugate
$(\cdot)^\dagger$	Transpose conjugate
$\overset{\text{a.s.}}{\to}$	Almost sure convergence
$\boldsymbol{I}_n$	$n \times n$ identity matrix
$\mathrm{tr}(\boldsymbol{A})$	Trace of a matrix $\boldsymbol{A}$
$\mathsf{rank}(\boldsymbol{A})$	Rank of a matrix $\boldsymbol{A}$
$\boldsymbol{A} \succcurlyeq \boldsymbol{B}$	Löwner partial ordering for Hermitian matrices $\boldsymbol{A}$ and $\boldsymbol{B}$
$\otimes$	Kronecker product
$\oplus$	Direct sum of matrices
$\mathsf{eig}_i(\boldsymbol{A})$	Eigenvalues of $\boldsymbol{A} \in \mathbb{C}^{n\times n} \succcurlyeq \boldsymbol{0}$ in any order, $i = 1, 2, \ldots, n$
$\varrho(\boldsymbol{A})$	Number of nonzero distinct eigenvalues of $\boldsymbol{A} \in \mathbb{C}^{n\times n} \succcurlyeq \boldsymbol{0}$
$\mathsf{eig}_{[i]}(\boldsymbol{A})$	Ordered nonzero distinct eigenvalues of $\boldsymbol{A} \in \mathbb{C}^{n\times n} \succcurlyeq \boldsymbol{0}$ in decreasing order such that $\mathsf{eig}_{[1]}(\boldsymbol{A}) > \mathsf{eig}_{[2]}(\boldsymbol{A}) > \cdots > \mathsf{eig}_{[\varrho(\boldsymbol{A})]}(\boldsymbol{A})$
$\tau_i(\boldsymbol{A})$	Multiplicity of the ith ordered nonzero distinct eigenvalue $\mathsf{eig}_{[i]}(\boldsymbol{A})$
$\mathbb{E}\{\cdot\}$	Expectation operator
$\mathbb{V}\mathrm{ar}\{\mathsf{X}\}$	Variance of X
$\phi_{\mathsf{X}}(s)$	Moment generating function of X: $\phi_{\mathsf{X}}(s) \triangleq \mathbb{E}\{e^{-s\mathsf{X}}\}$
$\mathcal{CN}(\mu, \sigma^2)$	Circularly symmetric complex Gaussian distribution with mean μ and variance σ^2
$\tilde{\mathcal{N}}_{m,n}$	Complex Gaussian matrix whose entries are independent $\mathcal{CN}(0,1)$
$\tilde{\mathcal{N}}_{m,n}(\boldsymbol{M}, \boldsymbol{\Sigma}, \boldsymbol{\Psi})$	Complex Gaussian matrix with mean matrix $\boldsymbol{M} \in \mathbb{C}^{m\times n}$, and covariance matrix $\boldsymbol{\Sigma} \otimes \boldsymbol{\Psi}$ where $\boldsymbol{\Sigma} \in \mathbb{C}^{m\times m} \succcurlyeq \boldsymbol{0}$ and $\boldsymbol{\Psi} \in \mathbb{C}^{m\times m} \succcurlyeq \boldsymbol{0}$ are Hermitian
$\Gamma(z)$	Euler gamma function [42, eq. (8.310.1)]
${}_pF_q(\cdot)$	Generalized hypergeometric function [42, eq. (9.14.1)]
$u(z)$	Heaviside step function
$\delta(z)$	Dirac's delta function
$\delta^{(n)}(z)$	nth derivative of $\delta(z)$
$\mathcal{F}^{-1}_{(\omega)}\{Y(\jmath\omega)\}(t)$	Inverse Fourier transform of $Y(\jmath\omega)$: $\mathcal{F}^{-1}_{(\omega)}\{Y(\jmath\omega)\}(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} Y(\jmath\omega)e^{\jmath\omega t}d\omega$.
$\mathcal{F}^{-1}_{(\omega_1,\omega_2)}\{Y(\jmath\omega_1, \jmath\omega_2)\}(t_1, t_2)$	Two-dimensional Inverse Fourier transform of $Y(\jmath\omega)$: $\mathcal{F}^{-1}_{(\omega_1,\omega_2)}\{Y(\jmath\omega_1, \jmath\omega_2)\}(t_1, t_2) = \frac{1}{4\pi^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} Y(\jmath\omega_1, \jmath\omega_2)e^{\jmath\omega_1 t_1 + \jmath\omega_2 t_2}d\omega_1 d\omega_2$.
$\mathcal{O}(\cdot)$	Bachmann-Landau notation: $f(x) = \mathcal{O}(g(x))$ as $x \to x_0 \Leftrightarrow \lim_{x\to x_0} \frac{f(x)}{g(x)} = c < \infty$.
$o(\cdot)$	Bachmann-Landau notation: $f(x) = o(g(x))$ as $x \to x_0 \Leftrightarrow \lim_{x\to x_0} \frac{f(x)}{g(x)} = 0$.
$\mathcal{X}_{i,j}(\boldsymbol{A})$	The (i,j)th characteristic coefficient of $\boldsymbol{A}$ [43, Definition 4]
$\boldsymbol{A}_{(k)}$	The kth constriction matrix of $\boldsymbol{A}$: $\boldsymbol{A}_{(k)} \triangleq \bigoplus_{i=1, i\neq k}^{m} \mathsf{eig}_i(\boldsymbol{A})$ where $\boldsymbol{A} \in \mathbb{C}^{m\times m} \succcurlyeq \boldsymbol{0}$.
$\boldsymbol{A}^{(k)}$	The kth dilation matrix of $\boldsymbol{A}$: $\boldsymbol{A}^{(k)} \triangleq \left(\bigoplus_{i=1}^{m} \mathsf{eig}_i(\boldsymbol{A})\right) \oplus \mathsf{eig}_k(\boldsymbol{A})$ where $\boldsymbol{A} \in \mathbb{C}^{m\times m} \succcurlyeq \boldsymbol{0}$.
$e_k(\boldsymbol{A})$	The kth eigenpolynomial of $\boldsymbol{A}$: $e_k(\boldsymbol{A}) \triangleq \sum_{i_1<i_2<\ldots<i_k} \prod_{j=1}^{k} \mathsf{eig}_{i_j}(\boldsymbol{A})$ where $\boldsymbol{A} \in \mathbb{C}^{m\times m} \succcurlyeq \boldsymbol{0}$.
$\zeta(\boldsymbol{A}_1, \boldsymbol{A}_2, \ldots, \boldsymbol{A}_n)$	The complete eigenpolynomial: $\zeta(\boldsymbol{A}_1, \boldsymbol{A}_2, \ldots, \boldsymbol{A}_n) \triangleq \sum_{k=0}^{\min_i\{\mathsf{rank}(\mathbf{Q}_i)\}} k! \prod_{i=1}^{n} e_k(\boldsymbol{A}_i)$ where $\boldsymbol{A}_i \in \mathbb{C}^{m_i\times m_i} \succcurlyeq \boldsymbol{0}$.

A. Signal Model

The n_{R}-dimensional received signal vector at the probe receiver is given by

$$\mathbf{y} = \sqrt{\frac{\mathsf{snr}}{n_0}}\grave{\mathbf{H}}_0\mathbf{x}_0 + \sum_{\ell=1}^{L}\sqrt{\frac{\mathsf{inr}_\ell}{n_\ell}}\grave{\mathbf{H}}_\ell\mathbf{x}_\ell + \mathbf{z} \tag{1}$$

where $\mathbf{x}_0 \in \mathbb{C}^{n_0}$ and $\mathbf{x}_\ell \in \mathbb{C}^{n_\ell}$ are the desired and ℓth interfering signals with $\mathbb{E}\left\{\|\mathbf{x}_0\|^2\right\} = n_0$ and $\mathbb{E}\left\{\|\mathbf{x}_\ell\|^2\right\} = n_\ell$, respectively; $\grave{\mathbf{H}}_0 \in \mathbb{C}^{n_{\mathrm{R}}\times n_0}$ and $\grave{\mathbf{H}}_\ell \in \mathbb{C}^{n_{\mathrm{R}}\times n_\ell}$ are the channel matrices for the desired and ℓth interfering links with $\mathbb{E}\{\mathrm{tr}(\grave{\mathbf{H}}_0\grave{\mathbf{H}}_0^\dagger)\} = n_0 n_{\mathrm{R}}$ and $\mathbb{E}\{\mathrm{tr}(\grave{\mathbf{H}}_\ell\grave{\mathbf{H}}_\ell^\dagger)\} = n_\ell n_{\mathrm{R}}$, respectively; snr and inr_ℓ are the signal-to-noise ratio (SNR) and the ℓth interference-to-noise ratio (INR), respectively; and $\mathbf{z}$ is the n_{R}-dimensional (circularly symmetric) complex additive white Gaussian noise vector with normalized covariance $\boldsymbol{I}_{n_{\mathrm{R}}}$. The L interfering signals are unknown to the desired receiver and all transmit signals are assumed to be zero-mean Gaussian. Note that all the random quantities $\mathbf{x}_0$, $\mathbf{x}_\ell$, $\grave{\mathbf{H}}_0$, $\grave{\mathbf{H}}_\ell$, and $\mathbf{z}$ are statistically independent. Let the interference-plus-noise covariance matrix conditioned on the channel realizations be

$$\mathsf{K} = \boldsymbol{I} + \sum_{\ell=1}^{L} \frac{\mathsf{inr}_\ell}{n_\ell} \grave{\mathsf{H}}_\ell \boldsymbol{\Sigma}_\ell \grave{\mathsf{H}}_\ell^\dagger \tag{2}$$

where $\boldsymbol{\Sigma}_\ell = \mathbb{E}\{\mathsf{x}_\ell \mathsf{x}_\ell^\dagger\}$ is the input covariance of the ℓth interfering signal with $\operatorname{tr}(\boldsymbol{\Sigma}_\ell) = n_\ell$. Then, the mutual information in bits per second per hertz (bits/s/Hz) between the input x_0 and output y for the desired link is [15], [17], [35]

$$\begin{aligned} I&\left(\mathsf{x}_0; \mathsf{y}, \{\grave{\mathsf{H}}_k\}_{k=0}^{L}\right) \\ &= \mathbb{E}\left\{\log_2 \det\left(\boldsymbol{I} + \frac{\mathsf{snr}}{n_0} \grave{\mathsf{H}}_0 \boldsymbol{\Sigma}_0 \grave{\mathsf{H}}_0^\dagger \mathsf{K}^{-1}\right)\right\} \end{aligned} \tag{3}$$

where $\boldsymbol{\Sigma}_0 = \mathbb{E}\{\mathsf{x}_0 \mathsf{x}_0^\dagger\}$ is the input covariance of the desired signal with $\operatorname{tr}(\boldsymbol{\Sigma}_0) = n_0$.

B. Channel Model

To account for the general structure of MIMO channels, we use a class of UIU MIMO models as follows (see, e.g., [14], [26], [31]):

$$\grave{\mathsf{H}}_k = \boldsymbol{R}_k \mathsf{H}_k \boldsymbol{T}_k^\dagger, \quad k = 0, 1, \ldots, L \tag{4}$$

where $\boldsymbol{R}_k$ and $\boldsymbol{T}_k$ are $n_\mathrm{R} \times n_\mathrm{R}$ and $n_k \times n_k$ deterministic unitary matrices; and $\mathsf{H}_k \in \mathbb{C}^{n_\mathrm{R} \times n_k}$ is the Karhunen–Loève transform (KLT) of $\grave{\mathsf{H}}_k$ and has independent entries whose marginal distributions are symmetric with respect to zero. Let $\boldsymbol{\Omega}_k \in \mathbb{R}_+^{n_\mathrm{R} \times n_k}$ be a gain matrix assembling the variances of the entries of H_k such that its (i,j)th entry

$$(\boldsymbol{\Omega}_k)_{i,j} = \mathbb{E}\left\{\left|(\mathsf{H}_k)_{i,j}\right|^2\right\} \tag{5}$$

determines the average power coupling between the jth transmit and ith receive antenna elements of the kth link. This UIU channel structure embraces a variety of MIMO channel models including: i) the joint transmit-receive correlation model [30] if $\boldsymbol{R}_k$ and $\boldsymbol{T}_k$ are unitary matrices whose columns are the eigenvectors of one-sided correlation matrices $\tilde{\boldsymbol{R}}_k = \mathbb{E}\{\grave{\mathsf{H}}_k \grave{\mathsf{H}}_k^\dagger\}$ and $\tilde{\boldsymbol{T}}_k = \mathbb{E}\{\grave{\mathsf{H}}_k^\dagger \grave{\mathsf{H}}_k\}$, respectively; ii) the linear virtual channel representation [44] if $\boldsymbol{R}_k$ and $\boldsymbol{T}_k$ are Fourier matrices (chosen irrespective of channel correlation) for uniform linear arrays (ULAs) at both sides; iii) the Kronecker (or separable correlation) model [19], [45] if the average power-coupling matrix $\boldsymbol{\Omega}_k$ is rank one such that $(\boldsymbol{\Omega}_k)_{i,j} = \mathsf{eig}_i(\tilde{\boldsymbol{R}}_k)\mathsf{eig}_j(\tilde{\boldsymbol{T}}_k)$; and iv) the independent and nonidentically distributed (IND) model if $\boldsymbol{R}_k = \boldsymbol{I}_{n_\mathrm{R}}$ and $\boldsymbol{T}_k = \boldsymbol{I}_{n_\ell}$ for polarization/pattern diversity or distributed MIMO. In all cases i)–iv), the entries of H_k are independent zero-mean Gaussian.

C. Problem Formulation

We consider perfect channel knowledge at the receiver, while the transmitter has access only to statistical channel knowledge. In practice, tracking channel states at the transmitter is quite challenging—especially for multi-dimensional MIMO channels. Therefore, partial channel knowledge of long-term statistics (covariance) at the transmitter is more acceptable for practical scenarios (see, e.g., [16], [24], [25], [27]–[29], [37], [46]–[48]). In this case, we can obtain the optimal structure of the input covariance $\boldsymbol{\Sigma}_0$ that maximizes the mutual information (3) as in the following proposition.

Proposition 1: Let $\boldsymbol{\Sigma}_0 = \boldsymbol{V}_0 \boldsymbol{P} \boldsymbol{V}_0^\dagger$ be the eigenvalue decomposition for the input covariance $\boldsymbol{\Sigma}_0$ of the probe transmitter, where $\boldsymbol{V}_0$ is a unitary matrix and $\boldsymbol{P} = \operatorname{diag}(p_1, p_2, \ldots, p_{n_0})$. Then, the optimal $\boldsymbol{V}_0$ that maximizes the mutual information (3) under the interference power environment $(\mathsf{inr}_\ell, \boldsymbol{\Sigma}_\ell)$, $\ell = 1, 2, \ldots, L$, is $\boldsymbol{V}_0^\star = \boldsymbol{T}_0$.

Proof: The proof is an almost verbatim copy of the proof of [26, Theorem 1] in the absence of interference. □

We suppose that each cochannel interferer (as well as the desired transmitter) employs the input covariance structure stated in Proposition 1 with knowing its own channel statistics, that is, $\boldsymbol{\Sigma}_\ell = \boldsymbol{T}_\ell \boldsymbol{Q}_\ell \boldsymbol{T}_\ell^\dagger$, $\ell = 1, 2, \ldots, L$, where $\boldsymbol{Q}_\ell = \operatorname{diag}(q_{\ell 1}, q_{\ell 2}, \ldots, q_{\ell n_\ell})$. Since the unitary matrices are unimportant in (3), we can formulate the maximization problem of mutual information over input covariances for UIU channels as equivalently the power optimization problem over IND channels as follows:

$$\boldsymbol{P}^\star = \underset{\boldsymbol{P} \succcurlyeq \boldsymbol{0}:\, \operatorname{tr}(\boldsymbol{P}) \leqslant n_0}{\arg\max}\ \mathcal{I}(\mathsf{snr}, \boldsymbol{P}, \boldsymbol{Q}) \tag{6}$$

yielding the average achievable rate in bits/s/Hz

$$R(\mathsf{snr}, \boldsymbol{Q}) = \mathcal{I}(\mathsf{snr}, \boldsymbol{P}^\star, \boldsymbol{Q}) \tag{7}$$

where $\boldsymbol{Q} = \operatorname{diag}\left(\frac{\mathsf{inr}_1}{n_1}\boldsymbol{Q}_1, \ldots, \frac{\mathsf{inr}_L}{n_L}\boldsymbol{Q}_L\right)$ denotes the interference power matrix and $\mathcal{I}(\mathsf{snr}, \boldsymbol{P}, \boldsymbol{Q})$ is given in (8). Define $N = \sum_{\ell=1}^{L} n_\ell$ and $\mathsf{inr}_\mathrm{tot} = \sum_{\ell=1}^{L} \mathsf{inr}_\ell$ as the total number of interferers and the aggregate INR, respectively. When $\mathsf{inr}_\mathrm{tot} \gg 1$ and $N \geqslant n_\mathrm{R}$, the MIMO system operates in an interference-rich environment in which the effect of thermal noise is negligible [35], [37].

III. Mathematical Methodology

In this section, we introduce the mathematical framework necessary to derive the analytic expressions for MMSE and SINR in the general structure of MIMO channels. We begin by providing closed-form formulas for the detquotients of one and two complex Gaussian matrices which will be used throughout the paper.

$$\mathcal{I}(\mathsf{snr}, \boldsymbol{P}, \boldsymbol{Q}) = \mathbb{E}\left\{\log_2 \det\left[\boldsymbol{I} + \frac{\mathsf{snr}}{n_0} \mathsf{H}_0 \boldsymbol{P} \mathsf{H}_0^\dagger \left(\boldsymbol{I} + \sum_{\ell=1}^{L} \frac{\mathsf{inr}_\ell}{n_\ell} \mathsf{H}_\ell \boldsymbol{Q}_\ell \mathsf{H}_\ell^\dagger\right)^{-1}\right]\right\} \tag{8}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TWC.2018.2865939, IEEE Transactions on Wireless Communications

Theorem 1 (Detquotient of Two Complex Gaussian Matrices): For $\boldsymbol{A} \in \mathbb{C}^{m\times m} \succ \mathbf{0}$, $\boldsymbol{B} \in \mathbb{C}^{n\times n} \succcurlyeq \mathbf{0}$, and $\boldsymbol{C} \in \mathbb{C}^{\ell\times\ell} \succcurlyeq \mathbf{0}$, the detquotient of the complex matrices $\mathbf{X} \sim \tilde{\mathcal{N}}_{m,n}$ and $\mathbf{Y} \sim \tilde{\mathcal{N}}_{m,\ell}$, denoted by $\mathcal{K}_1(\boldsymbol{A},\boldsymbol{B},\boldsymbol{C})$, is defined as

$$\mathcal{K}_1(\boldsymbol{A},\boldsymbol{B},\boldsymbol{C}) \triangleq \mathbb{E}_{\mathbf{X},\mathbf{Y}}\left\{\frac{\det\left(\boldsymbol{I}_m + \boldsymbol{A}\mathbf{X}\boldsymbol{B}\mathbf{X}^\dagger\right)}{\det\left(\boldsymbol{I}_m + \boldsymbol{A}\mathbf{X}\boldsymbol{B}\mathbf{X}^\dagger + \boldsymbol{A}\mathbf{Y}\boldsymbol{C}\mathbf{Y}^\dagger\right)}\right\}. \tag{9}$$

Then, we have

$$\mathcal{K}_1(\boldsymbol{A},\boldsymbol{B},\boldsymbol{C}) = \zeta(\boldsymbol{A},\boldsymbol{B})\,\Delta(\boldsymbol{B},\boldsymbol{C},\boldsymbol{A}) - \sum_{k=1}^{m}\sum_{\ell=1}^{n} \mathsf{eig}_k^2(\boldsymbol{A})\,\mathsf{eig}_\ell^2(\boldsymbol{B})\,\zeta\left(\boldsymbol{A}_{(k)},\boldsymbol{B}_{(\ell)}\right)\Delta\left(\boldsymbol{B}^{(\ell)},\boldsymbol{C},\boldsymbol{A}^{(k)}\right) \tag{10}$$

where $\zeta(\boldsymbol{A}_1,\boldsymbol{A}_2,\ldots,\boldsymbol{A}_n)$ is the complete eigenpolynomial [4, Definition 3]; $\boldsymbol{A}_{(k)}$ and $\boldsymbol{A}^{(k)}$ denote the kth constriction and dilation matrices of $\boldsymbol{A}$, respectively [4, Definition 1]; and

$$\Delta(\boldsymbol{B},\boldsymbol{C},\boldsymbol{A}) = \sum_{u=1}^{\varrho(\boldsymbol{C})}\sum_{v=1}^{\tau_u(\boldsymbol{C})}\sum_{p=1}^{\varrho(\boldsymbol{B})}\sum_{q=1}^{\tau_p(\boldsymbol{B})}\sum_{i=1}^{\varrho(\boldsymbol{A})}\sum_{j=1}^{\tau_i(\boldsymbol{A})}\left\{\frac{\mathcal{X}_{i,j}(\boldsymbol{A})\,\mathcal{X}_{p,q}(\boldsymbol{B})\,\mathcal{X}_{u,v}(\boldsymbol{C})}{\mathsf{eig}_{[i]}^j(\boldsymbol{A})} \times \mathcal{J}_{j,q}^v\left(\mathsf{eig}_{[p]}(\boldsymbol{B}),\,\mathsf{eig}_{[u]}(\boldsymbol{C}),\,\mathsf{eig}_{[i]}(\boldsymbol{A})\right)\right\} \tag{11}$$

for $a>0$, $b>0$, $c>0$, and $j\in\mathbb{Z}_+$; where $\mathcal{X}_{i,j}(\boldsymbol{A})$ denotes the (i,j)th characteristic coefficient of $\boldsymbol{A}$ [43, Definition 4]; and $\mathcal{J}_{j,q}^v(a,b,c)$ for $a\neq b$ is given in (12) and

$$\mathcal{J}_{j,q}^v(a,b,c) = c^j\,{}_2F_0(j,q+v;-ac) \tag{13}$$

for $a=b$.

Proof: See Appendix A. □

Theorem 2 (Detquotient of One Complex Gaussian Matrix): For $\boldsymbol{A}\in\mathbb{C}^{m\times m}\succcurlyeq\mathbf{0}$ and $\boldsymbol{B}\in\mathbb{C}^{n\times n}\succcurlyeq\mathbf{0}$, the detquotient of the complex Gaussian matrix $\mathbf{X}\sim\tilde{\mathcal{N}}_{m,n}$, is defined as

$$\mathcal{K}_2(\boldsymbol{A},\boldsymbol{B};s) \triangleq \mathbb{E}_{\mathbf{X}}\left\{\frac{\det\left(\boldsymbol{I}_m + \boldsymbol{A}\mathbf{X}\boldsymbol{B}\mathbf{X}^\dagger\right)}{\det\left(\boldsymbol{I}_m + s\boldsymbol{A} + \boldsymbol{A}\mathbf{X}\boldsymbol{B}\mathbf{X}^\dagger\right)}\right\}. \tag{14}$$

Then, we have

$$\mathcal{K}_2(\boldsymbol{A},\boldsymbol{B};s) = \zeta(\boldsymbol{A},\boldsymbol{B})\,\Theta(\boldsymbol{A},\boldsymbol{B};s) - \sum_{k=1}^{m}\sum_{\ell=1}^{n}\mathsf{eig}_k^2(\boldsymbol{A})\,\mathsf{eig}_\ell^2(\boldsymbol{B})\,\zeta\left(\boldsymbol{A}_{(k)},\boldsymbol{B}_{(\ell)}\right)\Theta\left(\boldsymbol{A}^{(\ell)},\boldsymbol{B}^{(\ell)};s\right) \tag{15}$$

where

$$\Theta(\boldsymbol{A},\boldsymbol{B};s) = \sum_{i=1}^{\varrho(\boldsymbol{A})}\sum_{j=1}^{\tau_i(\boldsymbol{A})}\sum_{p=1}^{\varrho(\boldsymbol{B})}\sum_{q=1}^{\tau_p(\boldsymbol{B})}\left\{\frac{\mathcal{X}_{i,j}(\boldsymbol{A})\,\mathcal{X}_{p,q}(\boldsymbol{B})}{\left(s\,\mathsf{eig}_{[i]}(\boldsymbol{A})+1\right)^j} \times {}_2F_0\left(j,q;-\frac{\mathsf{eig}_{[p]}(\boldsymbol{B})\,\mathsf{eig}_{[i]}(\boldsymbol{A})}{s\,\mathsf{eig}_{[i]}(\boldsymbol{A})+1}\right)\right\}. \tag{16}$$

Proof: Using similar steps in the proof of Theorem 1, we get the desired result. □

Remark 1: For $\boldsymbol{A}=\boldsymbol{I}_m$, the detquotient $\mathcal{K}_2(\boldsymbol{I}_m,\boldsymbol{B};s)$ can be evaluated in determinantal form using the generalized Binet-Cauchy formulas [13], [17], [43]. However, this approach cannot be applicable to $\boldsymbol{A}\neq\boldsymbol{I}_m$ since there is no analytically tractable form of the joint eigenvalue distribution for this case. The problem of finding the closed-form expression for $\mathcal{K}_1(\boldsymbol{A},\boldsymbol{B},\boldsymbol{C})$ is even more challenging because it requires taking double expectations over two random matrices. Berezin's supermathematics enables to derive the analytical solutions for $\mathcal{K}_1(\boldsymbol{A},\boldsymbol{B},\boldsymbol{C})$ and $\mathcal{K}_2(\boldsymbol{A},\boldsymbol{B};s)$ by resorting to superintegrals, Grassmann integrals, and Fourier representations.

Corollary 1: For $\boldsymbol{A}=\sigma\boldsymbol{I}_m$, $\boldsymbol{B}=\rho\boldsymbol{I}_n$, and $\boldsymbol{C}=\varrho\boldsymbol{I}_\ell$, the detquotient in (10) boils down to (17) where

$$T_{m,n}(s) = \sum_{k=0}^{\min\{m,n\}} k!s^k\binom{m}{k}\binom{n}{k}. \tag{18}$$

Proof: It follows readily from Theorem 1 along with [4, eq. (10)] and the characteristic coefficients of the identity matrix [43, eq. (130)]. □

$$\mathcal{J}_{j,q}^v(a,\,b,\,c) = \sum_{k=1}^{q}\left(-\frac{b}{a}\right)^{q-k}\frac{(v+q-k-1)!}{(q-k)!\,(v-1)!}\left(1-\frac{b}{a}\right)^{k-v-q}c^j\,{}_2F_0(j,k;-ac) + \left(-\frac{b}{a}\right)^{q}\sum_{i=1}^{v}\frac{(v+q-i-1)!}{(q-1)!\,(v-i)!}\left(1-\frac{b}{a}\right)^{i-v-q}c^j\,{}_2F_0(j,i;-bc) \tag{12}$$

$$\mathcal{K}_1(\sigma\boldsymbol{I}_m,\rho\boldsymbol{I}_n,\varrho\boldsymbol{I}_\ell) = \begin{cases}\sigma^{-m}\,T_{m,n}(\sigma\rho)\,\mathcal{J}_{m,n}^l(\rho,\,\varrho,\,\sigma) - mn\sigma^{1-m}\rho^2\,T_{m-1,n-1}(\sigma\rho)\,\mathcal{J}_{m+1,n+1}^l(\rho,\,\varrho,\,\sigma), & \varrho\neq\rho\\ T_{m,n}(\sigma\varrho)\,{}_2F_0(m,n+\ell;-\varrho) - mn\sigma^2\varrho^2\,T_{m-1,n-1}(\sigma\varrho)\,{}_2F_0(m+1,n+\ell+1;-\varrho), & \varrho=\rho\end{cases} \tag{17}$$

Corollary 2: For $\boldsymbol{A} = \sigma \boldsymbol{I}_m$ and $\boldsymbol{B} = \rho \boldsymbol{I}_n$, the detquotient in (15) boils down to

$$\mathcal{K}_2\left(\sigma \boldsymbol{I}_m, \rho \boldsymbol{I}_n; s\right) = \frac{T_{m,n}\left(\sigma\rho\right)}{\left(s\sigma+1\right)^m}\, {}_2F_0\left(m, n; -\frac{\sigma\rho}{s\sigma+1}\right) - mn\sigma^2\rho^2\, T_{m-1,n-1}\left(\sigma\rho\right) {}_2F_0\left(m+1, n+1; -\frac{\sigma\rho}{s\sigma+1}\right). \tag{19}$$

Proof: It follows immediately from Theorem 2 along with [4, eq. (10)] and (18). □

IV. Power Allocation Policies

A. A Necessary and Sufficient Condition

In a noise-limited system, the capacity achieving power allocation $\boldsymbol{P}_0^\star$ has been characterized for MIMO channels with antenna correlation [14], [26]. We consider a MMSE linear estimator at the receiver in the presence of cochannel interference. Then, the SINR on the jth spatial stream with power p_j along with a direction of corresponding eigenvector after MMSE detector can be expressed as

$$\mathsf{sinr}_j\left(p_j\right) = p_j \frac{\mathsf{snr}}{n_0} \mathbf{h}_{0j}^\dagger \boldsymbol{\Psi}_j^{-1} \mathbf{h}_{0j} \tag{20}$$

where

$$\boldsymbol{\Psi}_j = \boldsymbol{I} + \frac{\mathsf{snr}}{n_0}\sum_{i\neq j} p_i \mathbf{h}_{0i}\mathbf{h}_{0i}^\dagger + \sum_{\ell=1}^{L} \frac{\mathsf{inr}_\ell}{n_\ell} \mathbf{H}_\ell \boldsymbol{Q}_\ell \mathbf{H}_\ell^\dagger. \tag{21}$$

With a one-to-one relationship between the MMSE and SINR [3], [49], we get the MMSE of the jth transmit signal at the receiver as

$$\mathsf{mmse}_j\left(\boldsymbol{P}, \boldsymbol{Q}\right) = \frac{1}{1+\mathsf{sinr}_j\left(p_j\right)} = \frac{\det\left(\boldsymbol{\Psi}_j\right)}{\det\left(p_j \frac{\mathsf{snr}}{n_0}\mathbf{h}_{0j}\mathbf{h}_{0j}^\dagger + \boldsymbol{\Psi}_j\right)}. \tag{22}$$

Using (20) and (22), we obtain an optimum condition for input power allocation which maximizes the average mutual information in (8) for the MIMO interference channels as follows.

Theorem 3 (Optimum Condition for Input Power Allocation): Let

$$\gamma_j\left(\boldsymbol{P}, \boldsymbol{Q}\right) = \frac{1}{p_j}\mathbb{E}\left\{1 - \mathsf{mmse}_j\left(\boldsymbol{P}, \boldsymbol{Q}\right)\right\} \tag{23}$$

$$\gamma_{\text{th}}\left(\boldsymbol{P}, \boldsymbol{Q}\right) = \max_{j:p_j=0}\mathbb{E}\left\{\mathsf{sinr}_j\left(1\right)\right\}. \tag{24}$$

Then, a necessary and sufficient condition for the optimal power allocation $\boldsymbol{P}^\star$ in the MIMO interference channels is given by

$$\gamma_i\left(\boldsymbol{P}^\star, \boldsymbol{Q}\right) = \gamma_j\left(\boldsymbol{P}^\star, \boldsymbol{Q}\right), \quad \forall\ p_i > 0 \text{ and } p_j > 0 \tag{25}$$

$$\gamma_i\left(\boldsymbol{P}^\star, \boldsymbol{Q}\right) \geqslant \gamma_{\text{th}}\left(\boldsymbol{P}^\star, \boldsymbol{Q}\right), \quad \forall\ p_i > 0 \text{ and } p_j = 0. \tag{26}$$

Proof: See Appendix B. □

Note that $\gamma_i\left(\boldsymbol{P}^\star, \boldsymbol{Q}\right)$ in (23) can be interpreted as the average recovered signal level relative to the transmit power p_j along with the direction of the jth eigenvector, while $\gamma_{\text{th}}\left(\boldsymbol{P}^\star, \boldsymbol{Q}\right)$ measures the maximum average signal level achieved by allocating a unit power to the one of unused eigenvectors. Theorem 3 reveals that the non-zero power should be allocated to the directions providing higher recovered signal level than the maximum average SINR with unit power allocation along the unused eigenvectors. In addition, the transmit power should be also allocated to satisfy normalized signal level achievable on each eigenvector to be fairly balanced. Hence, $\gamma_i\left(\boldsymbol{P}^\star, \boldsymbol{Q}\right)$ for all $p_j > 0$ satisfies

$$\gamma_i\left(\boldsymbol{P}^\star, \boldsymbol{Q}\right) = \frac{1}{n_0}\sum_{k=1}^{n_0} p_k \gamma_k\left(\boldsymbol{P}^\star, \boldsymbol{Q}\right) = \frac{1}{n_0}\sum_{k=1}^{n_0}\left(1 - \mathbb{E}\left\{\mathsf{mmse}_k\left(\boldsymbol{P}^\star, \boldsymbol{Q}\right)\right\}\right). \tag{27}$$

Substituting (27) into (25) and (26), yields that the optimum condition in Theorem 3 has the same MMSE representation for the noise-limited MIMO channel in [26]. It unveils that an iterative algorithm for input power allocation in [26] can be applied to the MIMO interference channels with an arbitrary power distribution of cochannel interferers, each is spatially correlated across receiving antennas.

B. Power Optimization Strategy

In order to find an optimum power allocation $\boldsymbol{P}^\star$, it is crucial to evaluate (23) and (24). For the noise-limited MIMO channels, the knowledge of fading channel distribution at the transmitter has been exploited to evaluate the MMSE in [26]. The analytic expressions have been also derived for the MMSE to reduce its computation burden in Rayleigh-faded MISO channels [24]. In the following theorem, we investigate the closed-form expressions for the average MMSE and SINR in Theorem 3 over the MIMO interference channel with separable correlation structure.

Theorem 4 (Average MMSE and SINR): We consider MIMO interference channels where $\grave{\mathbf{H}}_k \sim \tilde{\mathcal{N}}_{n_{\mathrm{R}}, n_k}\left(\mathbf{0}, \boldsymbol{\Sigma}, \boldsymbol{\Phi}_k\right)$ for $k = 0, \ldots, L$. Let $\boldsymbol{\Lambda}_{\mathrm{T},k}$ and $\boldsymbol{\Lambda}_{\mathrm{R}}$ be the diagonal matrices containing eigenvalues of $\boldsymbol{\Phi}_k$ and $\boldsymbol{\Sigma}$, respectively, and

$$\boldsymbol{W}_j = \left(\bigoplus_{i=1, i\neq j}^{n_0} \frac{\mathsf{snr}}{n_0} p_i\right) \oplus \left(\bigoplus_{\ell=1}^{L} \frac{\mathsf{inr}_\ell}{n_\ell}\boldsymbol{Q}_\ell\right) \tag{28}$$

$$\boldsymbol{T}_j = \left(\bigoplus_{i=1, i\neq j}^{n_0} \left(\boldsymbol{\Lambda}_{\mathrm{T},0}\right)_{i,i}\right) \oplus \left(\bigoplus_{\ell=1}^{L} \boldsymbol{\Lambda}_{\mathrm{T},\ell}\right) \tag{29}$$

$$\boldsymbol{C}_j = \boldsymbol{T}_j^{1/2}\boldsymbol{W}_j\boldsymbol{T}_j^{1/2}. \tag{30}$$

Then, for a given input power matrix $\boldsymbol{P} = \operatorname{diag}\left(p_1, \ldots, p_{n_0}\right)$, the analytic expressions for the expected values in (23) and (24) are

$$\mathbb{E}\left\{\mathsf{mmse}_j\left(\boldsymbol{P}, \boldsymbol{Q}\right)\right\} = \mathcal{K}_1\left(\boldsymbol{\Sigma}, \boldsymbol{C}_j, p_j \mathsf{eig}_j(\boldsymbol{\Lambda}_{\mathrm{T},0})\frac{\mathsf{snr}}{n_0}\right) \tag{31}$$

$$\mathbb{E}\left\{\mathsf{sinr}_j\left(1\right)\right\} = \zeta\left(\boldsymbol{\Lambda}_{\mathrm{R}}, \boldsymbol{C}_j\right)\dot{\Theta}\left(\boldsymbol{C}_j, \boldsymbol{\Lambda}_{\mathrm{R}}; \mathsf{eig}_j(\boldsymbol{\Lambda}_{\mathrm{T},0})\right) - \sum_{k=1}^{m}\sum_{\ell=1}^{n}\left\{\mathsf{eig}_k^2(\boldsymbol{C}_j)\,\mathsf{eig}_\ell^2(\boldsymbol{\Lambda}_{\mathrm{R}})\,\zeta\left(\boldsymbol{C}_{j(k)}, \boldsymbol{\Lambda}_{\mathrm{R}(\ell)}\right) \times \dot{\Theta}\left(\boldsymbol{C}_j^{(k)}, \boldsymbol{\Lambda}_{\mathrm{R}}^{(\ell)}; \mathsf{eig}_j(\boldsymbol{\Lambda}_{\mathrm{T},0})\right)\right\} \tag{32}$$

where

$$\begin{aligned}\dot{\Theta}\left(\boldsymbol{A},\boldsymbol{B};s\right) &= \sum_{i=1}^{\varrho(\boldsymbol{B})}\sum_{j=1}^{\tau_i(\boldsymbol{B})}\sum_{p=1}^{\varrho(\boldsymbol{A})}\sum_{q=1}^{\tau_p(\boldsymbol{A})}\left\{\frac{\mathcal{X}_{i,j}(\boldsymbol{B})\,\mathcal{X}_{p,q}(\boldsymbol{A})\,s\,\mathsf{eig}_{[i]}(\boldsymbol{B})j\mathsf{snr}}{n_0}\right.\\ &\qquad\left.\times\,{}_2F_0\left(j+1,q;-\mathsf{eig}_{[p]}(\boldsymbol{A})\mathsf{eig}_{[i]}(\boldsymbol{B})\right)\right\}. \quad (33)\end{aligned}$$

Proof: The analytical solution for $\mathbb{E}\left\{\mathsf{mmse}_j\left(\boldsymbol{P},\boldsymbol{Q}\right)\right\}$ in (31) follows immediately from Theorem 1. In order to find the solution for $\mathbb{E}\left\{\mathsf{sinr}_j\left(1\right)\right\}$, we derive the moment generating function of $\mathsf{sinr}_j\left(1\right)$ using (20) as

$$\begin{aligned}&\phi_{\mathsf{sinr}_j(1)}\left(s\right)\\ &= \mathbb{E}_{\mathbf{h}_{0j},\mathbf{H}}\left\{\exp\left(-s\frac{\mathsf{snr}}{n_0}\mathbf{h}_{0j}^{\dagger}\left(\boldsymbol{I}+\mathbf{H}\boldsymbol{W}_j\mathbf{H}^{\dagger}\right)^{-1}\mathbf{h}_{0j}\right)\right\}\\ &= \mathbb{E}_{\mathbf{H}}\left\{\frac{\det\left(\boldsymbol{I}+\mathbf{H}\boldsymbol{W}_j\mathbf{H}^{\dagger}\right)}{\det\left(\boldsymbol{I}+s\frac{\mathsf{snr}}{n_0}\mathsf{eig}_j(\boldsymbol{\Lambda}_{\mathrm{T},0})\boldsymbol{\Sigma}+\mathbf{H}\boldsymbol{W}_j\mathbf{H}^{\dagger}\right)}\right\}\\ &= \mathcal{K}_2\left(\boldsymbol{\Sigma},\boldsymbol{C}_j;s\frac{\mathsf{snr}}{n_0}\mathsf{eig}_j\left(\boldsymbol{\Lambda}_{\mathrm{T},0}\right)\right) \quad (34)\end{aligned}$$

where $\mathbf{H} = [\mathbf{H}_{0j}\ \mathbf{H}_1\ \ldots\ \mathbf{H}_L]$. Examining the first order derivative of $\phi_{\mathsf{sinr}_j(1)}\left(s\right)$ with respect to s, we arrive at the desired result (32) at $s=0$. □

Remark 2: It is worth noting that the joint transmit-receive correlation model is not subsumed by our framework. The main difficulty arises due to the absence of a method to decouple the matrix $\mathbf{H}_k$ in (4) into the product of an i.i.d. zero mean Gaussian matrix and the nonnegative semidefinite matrices. To circumvent the difficulty, the large random matrix theory is applied as the numbers of antennas tend to infinity in the next section.

Corollary 3: For $\frac{\mathsf{snr}}{n_0} = \frac{\mathsf{inr}_1}{n_1} = \ldots = \frac{\mathsf{inr}_L}{n_L} = \gamma$ and spatially uncorrelated across receiving antennas, we have $\boldsymbol{Q}_\ell = \boldsymbol{I}_{n_\ell}$ and $\boldsymbol{P} = \boldsymbol{I}_{n_0}$. Then, the average MMSE and SINR reduce to (35) and (36), respectively.

Proof: It follows immediately from Theorem 4, (17) and (19). □

Invoking [26, Algorithm 1] and Theorem 4, we introduce a procedure to find the optimum power allocation $\boldsymbol{P}^{\star}$ maximizing the ergodic mutual information in Algorithm 1.

Remark 3: If p_j is initialized to zero, then it remains at zero for all future iterations. The convergence of the algorithm is determined by the convergence of the step of computing $p_j^{(n)}$, which is in turn a fixed-point iteration procedure. Since the functions defining fixed-point equations for p_j are continuous, a sufficient condition for such fixed-point iterations to converge is that the functions are Lipschitz with Lipschitz constant less than 1. The empirical verification of the sufficient condition is available for several special cases of the MIMO interference channel in [26].

Algorithm 1: Power Allocation Algorithm

Initialize: Power matrix $\boldsymbol{P}^{(0)}$ such that $\mathrm{tr}\left(\boldsymbol{P}^{(0)}\right) > 0$
Convergence thresholds $\epsilon \in \mathbb{R}_{++}$ and $\epsilon' \in \mathbb{R}_{++}$
$p_j^{(0)}$ to be the arbitrary large number for $j \in \{1,\ldots,n_0\}$

for $n = 1,2,\ldots$ **do**
  for $j = 1$ **to** n_0 **do**
    $p_j^{(n)} = \frac{1-\mathbb{E}\left\{\mathsf{mmse}_j\left(\boldsymbol{P}^{(n-1)},\boldsymbol{Q}\right)\right\}}{\frac{1}{n_0}\sum_{k=1}^{n_0}\left(1-\mathbb{E}\left\{\mathsf{mmse}_k\left(\boldsymbol{P}^{(n-1)},\boldsymbol{Q}\right)\right\}\right)}$
  end
  if $\max_j\left|p_j^{(n)}-p_j^{(n-1)}\right| < \epsilon$ **then** $\boldsymbol{P}^{\infty} = \boldsymbol{P}^{(n)}$
    if $\mathbb{E}\left\{\mathsf{sinr}_j^{(\infty)}\left(1\right)\right\} \le \frac{1}{n_0}\sum_{k=1}^{n_0}\left(1-\mathbb{E}\left\{\mathsf{mmse}_k\left(\boldsymbol{P}^{(\infty)},\boldsymbol{Q}\right)\right\}\right)$ *for* j *such that* $p_j^{(\infty)} < \epsilon'$ **then** $\boldsymbol{P}^{\star} = \boldsymbol{P}^{(\infty)}$;
      break ;
    else $p_j^{(0)} = 0$ for $j = \arg\min_k \mathbb{E}\left\{\mathsf{sinr}_k^{(\infty)}\left(1\right)\right\}$;
  end
end

Remark 4: Recalling the noise-limited MIMO channel [26], we remark that some observations for the uniform power allocation in the presence of cochannel interference as: i) the isotropic input achieves the maximum rate in (7) if the columns of $\mathbf{H}_0$ are marginally identically distributed or the channel gain matrix $\boldsymbol{\Omega}_0$ is *column-regular* [14, Definition 4] in a Rayleigh fading channel; ii) In a regime of large numbers of antennas, if the gain matrix $\boldsymbol{\Omega}_0$ is *asymptotic column-regular* [14], then the asymptotic capacity-achieving power input is isotropic regardless of the marginal distribution of $\mathbf{H}_0$; and iii) For the Rayleigh-faded channels with separable correlations, the isotropic input is optimal if the transmit antennas are uncorrelated.

V. Asymptotic Capacity

Using the power allocation algorithm developed in Sec. IV, it is now possible to characterize the impact of correlation on the ergodic capacity of MIMO channels in the presence of cochannel interference when the channel knowledge of long-term statistics is available at the transmitter. For a channel matrix $\mathbf{H}_0$ whose entries are arbitrarily distributed with uniformly bounded variances associated with the gain matrix $\boldsymbol{\Omega}_0$ with ratio $\beta = n_0/n_{\mathrm{R}}$, the *gain profile*

$$\mathcal{G}_0^{(n_{\mathrm{R}})} : [0,1) \times [0,\beta) \to \mathbb{R} \quad (37)$$

$$\begin{aligned}\mathbb{E}\left\{\mathsf{mmse}_j\left(\boldsymbol{P},\boldsymbol{Q}\right)\right\} &= T_{n_{\mathrm{R}},N+n_0-1}\left(\gamma\right){}_2F_0\left(n_{\mathrm{R}},N+n_0;-\gamma\right)\\ &\quad - n_{\mathrm{R}}\left(N+n_0-1\right)\gamma^2 T_{n_{\mathrm{R}}-1,N+n_0-2}\left(\gamma\right){}_2F_0\left(n_{\mathrm{R}}+1,N+n_0+1;-\gamma\right) \quad (35)\\ \mathbb{E}\left\{\mathsf{sinr}_j\left(1\right)\right\} &= n_{\mathrm{R}}\gamma\, T_{n_{\mathrm{R}},N+n_0-1}\left(\gamma\right){}_2F_0\left(n_{\mathrm{R}}+1,N+n_0-1;-\gamma\right)\\ &\quad - n_{\mathrm{R}}\left(n_{\mathrm{R}}+1\right)\left(N+n_0-1\right)\gamma^3\, T_{n_{\mathrm{R}}-1,N+n_0-2}\left(\gamma\right){}_2F_0\left(n_{\mathrm{R}}+2,N+n_0;-\gamma\right) \quad (36)\end{aligned}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TWC.2018.2865939, IEEE Transactions on Wireless Communications

maps the entries of $\boldsymbol{\Omega}_0$ onto a two-dimensional piece-wise constant function such that

$$\mathcal{G}_0^{(n_{\mathrm{R}})}(r,t) \triangleq (\boldsymbol{\Omega}_0)_{i,j}\,, \quad \frac{i-1}{n_{\mathrm{R}}} \leqslant r < \frac{i}{n_{\mathrm{R}}}, \quad \frac{j-1}{n_{\mathrm{R}}} \leqslant t < \frac{j}{n_{\mathrm{R}}}. \tag{38}$$

The variables r and t can be interpreted as the normalized receive and transmit antenna indices taking values in $[0,1)$ and $[0,\beta)$, respectively. Similarly, associated with the power matrix $\boldsymbol{P}$ with uniformly bounded diagonal entries, we define an input *power profile* given by

$$\mathcal{P}^{(n_{\mathrm{R}})}(t) \triangleq p_j, \quad \frac{j-1}{n_{\mathrm{R}}} \leqslant t < \frac{j}{n_{\mathrm{R}}}. \tag{39}$$

We assume that, in the regime of large numbers of antennas, the gain profile $\mathcal{G}(r,t)$ and power profile $\mathcal{P}(t)$ converge uniformly to bounded functions

$$\mathcal{G}(r,t) \triangleq \lim_{n_{\mathrm{R}}\to\infty} \mathcal{G}^{(n_{\mathrm{R}})}(r,t) \tag{40}$$

$$\mathcal{P}(t) \triangleq \lim_{n_{\mathrm{R}}\to\infty} \mathcal{P}^{(n_{\mathrm{R}})}(t) \tag{41}$$

referred to as an *asymptotic gain profile* and an *asymptotic power profile*, respectively.

Theorem 5 (Asymptotic Mutual Information): Let $\mathbf{H}_k$, $k = 0,\ldots,L$ be INDs which satisfy $n_{\mathrm{R}} \to \infty$, $\frac{n_0}{n_{\mathrm{R}}} = \beta \in (0,\infty)$ and $\frac{N}{n_{\mathrm{R}}} = \beta_{\mathrm{I}} \in (0,\infty)$. Let $\mathcal{G}_0(r,t)$, $\mathcal{G}_{\mathrm{I}}(r,t)$, and $\mathcal{G}(r,t)$ be the asymptotic gain profiles associated with $\boldsymbol{\Omega}_0$, $\boldsymbol{\Omega}_{\mathrm{I}} = [\boldsymbol{\Omega}_1 \ldots \boldsymbol{\Omega}_L]$, and $\boldsymbol{\Omega} = [\boldsymbol{\Omega}_0\ \boldsymbol{\Omega}_{\mathrm{I}}]$, respectively, and $\mathcal{P}_0(t)$, $\mathcal{P}_{\mathrm{I}}(t)$, and $\mathcal{P}(t)$ be the asymptotic power profiles associated with $\boldsymbol{P}$, $\boldsymbol{P}_2 = \bigoplus_{\ell=1}^{L} \frac{\mathsf{inr}_\ell}{\mathsf{snr}} \frac{n_0}{n_\ell} \boldsymbol{Q}_\ell$, and $\boldsymbol{P}_1 = \mathbf{P}^\star \oplus \boldsymbol{P}_2$, respectively. Then, the asymptotic mutual information per receive antenna, denoted by $\mathcal{I}^\star(\mathsf{snr}, \boldsymbol{P}^\star, \boldsymbol{Q})$, is given in (42) where $\mathcal{B}_{\mathrm{I}}(r,\mathsf{snr})$ and $\mathcal{B}(r,\mathsf{snr})$, satisfying the equations

$$\mathcal{B}_{\mathrm{I}}(r,\mathsf{snr}) = \frac{1}{1 + \mathsf{snr}\beta_{\mathrm{I}}\,\mathbb{E}_{\mathsf{t}_{\mathrm{I}}}\left\{\frac{\mathcal{G}_{\mathrm{I}}(r,\mathsf{t}_{\mathrm{I}})\mathcal{P}_{\mathrm{I}}(\mathsf{t}_{\mathrm{I}})}{\beta + \mathsf{snr}\,\mathbb{E}_{\mathsf{r}_{\mathrm{I}}}\{\mathcal{G}_{\mathrm{I}}(\mathsf{r}_{\mathrm{I}},\mathsf{t}_{\mathrm{I}})\mathcal{P}_{\mathrm{I}}(\mathsf{t}_{\mathrm{I}})\mathcal{B}_{\mathrm{I}}(\mathsf{r}_{\mathrm{I}},\mathsf{snr})|\mathsf{t}_{\mathrm{I}}\}}\right\}} \tag{43}$$

$$\mathcal{B}(r,\mathsf{snr}) = \frac{1}{1 + \mathsf{snr}(\beta+\beta_{\mathrm{I}})\,\mathbb{E}_{\mathsf{t}}\left\{\frac{\mathcal{G}(r,\mathsf{t})\mathcal{P}(\mathsf{t})}{\beta + \mathsf{snr}\,\mathbb{E}_{\mathsf{r}}\{\mathcal{G}(\mathsf{r},\mathsf{t})\mathcal{P}(\mathsf{t})\mathcal{B}(\mathsf{r},\mathsf{snr})|\mathsf{t}\}}\right\}} \tag{44}$$

and the expectations are over r, t_0, t_{I}, and t, independent and uniform in $[0,1]$, $[0,\beta]$, $[0,\beta_{\mathrm{I}}]$ and $[0,\beta+\beta_{\mathrm{I}}]$, respectively.

Proof: The average mutual information in (8) can be manipulated into

$$\begin{aligned}\mathcal{I}(\mathsf{snr},\boldsymbol{P}^\star,\boldsymbol{Q}) &= \mathbb{E}\left\{\log_2\det\left[\boldsymbol{I} + \frac{\mathsf{snr}}{n_0}\mathbf{H}_0\boldsymbol{P}^\star\mathbf{H}_0^\dagger + \sum_{\ell=1}^{L}\frac{\mathsf{inr}_\ell}{n_\ell}\mathbf{H}_\ell\boldsymbol{Q}_\ell\mathbf{H}_\ell^\dagger\right]\right\} \\ &\quad - \mathbb{E}\left\{\log_2\det\left[\boldsymbol{I} + \sum_{\ell=1}^{L}\frac{\mathsf{inr}_\ell}{n_\ell}\mathbf{H}_\ell\boldsymbol{Q}_\ell\mathbf{H}_\ell^\dagger\right]\right\}.\end{aligned} \tag{45}$$

Let us define block matrices $\boldsymbol{P}_2 = \bigoplus_{\ell=1}^{L} \frac{\mathsf{inr}_\ell}{\mathsf{snr}} \frac{n_0}{n_\ell} \boldsymbol{Q}_\ell$ and $\boldsymbol{P}_1 = \mathbf{P}^\star \oplus \boldsymbol{P}_2$. Then, the average mutual information in (45) can be further simplified into

$$\begin{aligned}\mathcal{I}(\mathsf{snr},\boldsymbol{P}^\star,\boldsymbol{Q}) &= \underbrace{\mathbb{E}\left\{\log_2\det\left[\boldsymbol{I} + \frac{\mathsf{snr}}{n_0}\mathbf{H}\boldsymbol{P}_1\mathbf{H}^\dagger\right]\right\}}_{\triangleq \mathcal{I}_1(\mathsf{snr},\boldsymbol{P}^\star,\boldsymbol{Q})} \\ &\quad \underbrace{- \mathbb{E}\left\{\log_2\det\left[\boldsymbol{I} + \frac{\mathsf{snr}}{n_0}\mathbf{H}_{\mathrm{I}}\boldsymbol{P}_2\mathbf{H}_{\mathrm{I}}^\dagger\right]\right\}}_{\triangleq \mathcal{I}_2(\mathsf{snr},\boldsymbol{Q})}.\end{aligned} \tag{46}$$

The term $\mathcal{I}_1(\mathsf{snr},\boldsymbol{P}^\star,\boldsymbol{Q})$ can be interpreted as the ergodic capacity of the cooperative network where the receiver can decode all the desired and interfering signals, while the term $\mathcal{I}_2(\mathsf{snr},\boldsymbol{Q})$ is the ergodic capacity of the channel between the interferers and receiver. The difference between both terms gives the ergodic capacity of the desired MIMO channel in the presence of cochannel interference. Since

$$\mathbb{V}\mathrm{ar}\left\{\left(\mathbf{H}\boldsymbol{P}_1^{1/2}\right)_{i,j} n_0^{-1/2}\right\} = \frac{(\boldsymbol{\Omega})_{i,j}\,p_{1j}}{n_{\mathrm{R}}\beta}, \tag{47}$$

the *asymptotic variance profile* associated with $\frac{1}{\sqrt{n_0}}\mathbf{H}\boldsymbol{P}_1^{1/2}$ and $\frac{1}{\sqrt{n_0}}\mathbf{H}_{\mathrm{I}}\boldsymbol{P}_2^{1/2}$ is given by [14, Theorem 2.53]

$$\begin{aligned}v(r,t) &\triangleq \lim_{n_{\mathrm{R}}\to\infty} n_{\mathrm{R}}\,\mathbb{V}\mathrm{ar}\left\{\left(\mathbf{H}\boldsymbol{P}_1^{1/2}\right)_{i,j} n_0^{-1/2}\right\} \\ &= \frac{\mathcal{G}(r,t)\,\mathcal{P}(t)}{\beta}\end{aligned} \tag{48}$$

$$\begin{aligned}v_{\mathrm{I}}(r,t) &\triangleq \lim_{n_{\mathrm{R}}\to\infty} n_{\mathrm{R}}\,\mathbb{V}\mathrm{ar}\left\{\left(\mathbf{H}_{\mathrm{I}}\boldsymbol{P}_2^{1/2}\right)_{i,j} n_0^{-1/2}\right\} \\ &= \frac{\mathcal{G}_{\mathrm{I}}(r,t)\,\mathcal{P}_{\mathrm{I}}(t)}{\beta}.\end{aligned} \tag{49}$$

After some manipulations, we get (50) and (51), where $\mathcal{B}_{\mathrm{I}}(r,\mathsf{snr})$ and $\mathcal{B}(r,\mathsf{snr})$ are the unique solutions to the fixed-point equations in (43) and (44), respectively. Finally, substituting (50) and (51) into (46), we complete the proof. □

$$\begin{aligned}\mathcal{I}^\star(\mathsf{snr},\boldsymbol{P}^\star,\boldsymbol{Q}) &\overset{\text{a.s.}}{\to} \beta\,\mathbb{E}_{\mathsf{t}_0}\left\{\log_2\left(1 + \frac{\mathsf{snr}}{\beta}\mathbb{E}_{\mathsf{r}}\left\{\mathcal{G}_0(\mathsf{r},\mathsf{t}_0)\,\mathcal{P}_0(\mathsf{t}_0)\,\mathcal{B}(\mathsf{r},\mathsf{snr})\,|\mathsf{t}_0\right\}\right)\right\} \\ &\quad + \beta_{\mathrm{I}}\,\mathbb{E}_{\mathsf{t}_{\mathrm{I}}}\left\{\log_2\left(\frac{\beta + \mathsf{snr}\,\mathbb{E}_{\mathsf{r}}\left\{\mathcal{G}_{\mathrm{I}}(\mathsf{r},\mathsf{t}_{\mathrm{I}})\,\mathcal{P}_{\mathrm{I}}(\mathsf{t}_{\mathrm{I}})\,\mathcal{B}(\mathsf{r},\mathsf{snr})\,|\mathsf{t}_{\mathrm{I}}\right\}}{\beta + \mathsf{snr}\,\mathbb{E}_{\mathsf{r}}\left\{\mathcal{G}_{\mathrm{I}}(\mathsf{r},\mathsf{t}_{\mathrm{I}})\,\mathcal{P}_{\mathrm{I}}(\mathsf{t}_{\mathrm{I}})\,\mathcal{B}_{\mathrm{I}}(\mathsf{r},\mathsf{snr})\,|\mathsf{t}_{\mathrm{I}}\right\}}\right)\right\} \\ &\quad + \mathbb{E}\left\{\log_2\frac{\mathcal{B}_{\mathrm{I}}(\mathsf{r},\mathsf{snr})}{\mathcal{B}(\mathsf{r},\mathsf{snr})}\right\} + \left(\mathbb{E}\left\{\mathcal{B}(\mathsf{r},\mathsf{snr}) - \mathcal{B}_{\mathrm{I}}(\mathsf{r},\mathsf{snr})\right\}\right)\log_2 e\end{aligned} \tag{42}$$

The key mathematical tool used in Theorem 5 is Girko's theorem [50, Corollary 10.1.2] along with the Shannon transform [51]. Theorem 5 provides the asymptotic capacity for a very general case when the fading has zero mean and arbitrary distributions.[2]

Remark 5: Since the asymptotic normalized ergodic capacity in (42) is difficult to compute, we can approximate all the involved functions in (42), (43), and (44) with linear-piecewise curves [19]. By replacing $\mathcal{B}(r,\mathsf{snr})=\mathcal{B}_i(\mathsf{snr})$ and $\mathcal{B}_{\mathrm{I}}(r,\mathsf{snr})=\mathcal{B}_{\mathrm{I}i}(\mathsf{snr})$ for all $\frac{i-1}{n_{\mathrm{R}}}\le r<\frac{i}{n_{\mathrm{R}}}$, and solving the expectation operations in (42), (43) and (44), the capacity of an IND channel for arbitrary numbers of antennas as a functional of the discrete set of values of these functions is approximated by

$$\begin{aligned}
&\mathcal{I}^{\star}(\mathsf{snr},\boldsymbol{P}^{\star},\boldsymbol{Q})\\
&\approx\frac{1}{n_{\mathrm{R}}}\sum_{k=1}^{n_0}\log_2\left(1+\frac{\mathsf{snr}}{\beta n_{\mathrm{R}}}\sum_{i=1}^{n_{\mathrm{R}}}p_k\left(\boldsymbol{\Omega}_0\right)_{i,j}\mathcal{B}_i(\mathsf{snr})\right)\\
&+\frac{1}{n_{\mathrm{R}}}\sum_{j=1}^{N}\log_2\left(\frac{\beta n_{\mathrm{R}}+\mathsf{snr}\sum_{i=1}^{n_{\mathrm{R}}}p_{2j}\left(\boldsymbol{\Omega}_{\mathrm{I}}\right)_{i,j}\mathcal{B}_i(\mathsf{snr})}{\beta n_{\mathrm{R}}+\mathsf{snr}\sum_{i=1}^{n_{\mathrm{R}}}p_{2j}\left(\boldsymbol{\Omega}_{\mathrm{I}}\right)_{i,j}\mathcal{B}_{\mathrm{I}i}(\mathsf{snr})}\right)\\
&-\frac{1}{n_{\mathrm{R}}}\sum_{i=1}^{n_{\mathrm{R}}}\log_2\left(\frac{\mathcal{B}_i(\mathsf{snr})}{\mathcal{B}_{\mathrm{I}i}(\mathsf{snr})}\right)\\
&+\frac{1}{n_{\mathrm{R}}}\sum_{i=1}^{n_{\mathrm{R}}}\left(\mathcal{B}_i(\mathsf{snr})-\mathcal{B}_{\mathrm{I}i}(\mathsf{snr})\right)\log_2 e
\end{aligned}\tag{52}$$

where $\mathcal{B}_{\mathrm{I}j}(\mathsf{snr})$, for $j\in\{1,\ldots,n_{\mathrm{R}}\}$, and $\mathcal{B}_i(\mathsf{snr})$, for $i\in\{1,\ldots,n_{\mathrm{R}}\}$, are satisfying the equations

$$\begin{aligned}
&\mathcal{B}_{\mathrm{I}j}(\mathsf{snr})\\
&=\left(1+\frac{\mathsf{snr}}{n_{\mathrm{R}}}\sum_{k=1}^{N}\frac{\left(\boldsymbol{\Omega}_{\mathrm{I}}\right)_{j,k}p_{2k}}{\beta+\frac{\mathsf{snr}}{n_{\mathrm{R}}}\sum_{\ell=1}^{n_{\mathrm{R}}}\left(\boldsymbol{\Omega}_{\mathrm{I}}\right)_{\ell,k}p_{2k}\mathcal{B}_{\mathrm{I}\ell}(\mathsf{snr})}\right)^{-1},
\end{aligned}\tag{53}$$

$$\begin{aligned}
&\mathcal{B}_i(\mathsf{snr})\\
&=\left(1+\frac{\mathsf{snr}}{n_{\mathrm{R}}}\sum_{k=1}^{n_0+N}\frac{\left(\boldsymbol{\Omega}\right)_{i,k}p_{1k}}{\beta+\frac{\mathsf{snr}}{n_{\mathrm{R}}}\sum_{\ell=1}^{n_{\mathrm{R}}}\left(\boldsymbol{\Omega}\right)_{\ell,k}p_{1k}\mathcal{B}_{\ell}(\mathsf{snr})}\right)^{-1}.
\end{aligned}\tag{54}$$

Although the functions $\mathcal{B}_j(\mathsf{snr})$ and $\mathcal{B}_{\mathrm{I}j}(\mathsf{snr})$ do not possess closed-form representations, they easily can be solved by fixed-point iteration in $2n_{\mathrm{R}}$ variables which converge very quickly to any desired accuracy level.

[2]Note that the channel matrices consist of i.i.d. elements with zero mean and unit variance, and the number of probe transmit antennas is equal to the number of transmit antennas at each interferer, Theorem 5 can be reduced to the result derived in [35].

VI. Numerical Results

In this section, we provide some numerical examples to illustrate our analysis. Specifically, we use the nth-order positive-definite *exponential* correlation matrix $\boldsymbol{\Phi}_n^{(\exp)}(\rho)=\left(\rho^{|i-j|}\right)_{i,j=1,2,\ldots,n}$ with a correlation coefficient $\rho\in[0,1)$. In all examples associated with the exponential correlation, we set $\boldsymbol{\Phi}_\ell=\boldsymbol{\Phi}_{n_\ell}^{(\exp)}(\rho_\ell)$, $\ell=0,\ldots,L$, and $\boldsymbol{\Sigma}=\boldsymbol{\Phi}_{n_{\mathrm{R}}}^{(\exp)}(\rho_{\mathrm{R}})$. To characterize the effect of interference heterogeneity, we use the *intraclass* power profile matrix

$$\boldsymbol{Q}_{\mathrm{intra}}(\eta,\mathsf{inr}_{\mathrm{tot}})=\frac{\mathsf{inr}_{\mathrm{tot}}}{N}\begin{bmatrix}1&\eta&\eta&\cdots&\eta\\ \eta&1&\eta&\cdots&\eta\\ \vdots&\vdots&\vdots&\ddots&\vdots\\ \eta&\eta&\eta&\cdots&1\end{bmatrix}_{N\times N},\tag{55}$$

for $\eta\in[0,1]$. The intraclass power profile implies a interference network consists of one dominant strong interferer and $N-1$ relatively weak equal-power interferers. For example, if the receiver in a femtocell is affected by one strong interferer from the macrocell and two weak interferers from the other femtocell, we can set $\boldsymbol{Q}_{\mathrm{intra}}(0.5,3\mathsf{inr})$ which implies that $\mathsf{inr}_1=2\mathsf{inr}$, $\mathsf{inr}_2=0.5\mathsf{inr}$, and $\mathsf{inr}_3=0.5\mathsf{inr}$. Note that the interference network power is homogeneous when $\eta=0$ and its heterogeneity increases as η tends to 1.

A. Verification

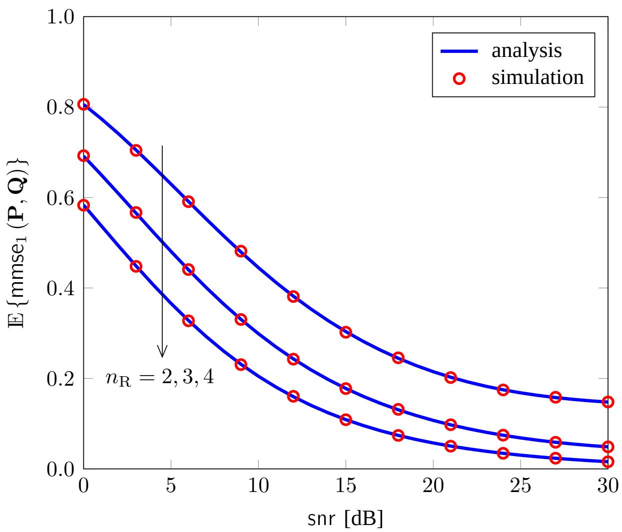
We first verify our analysis on the average MMSE and SINR derived in Theorem 4. Figs. 1(a) and Figs. 1(b) show the average MMSE and SINR for the first parallel stream of the $(4,[2,3],n_{\mathrm{R}})$-MIMO interference channel as a function of snr at $\mathsf{inr}_{\mathrm{tot}}=10$ dB when $n_{\mathrm{R}}=2,3,4$. The corresponding input power matrices are set to $\boldsymbol{P}=\mathrm{diag}(1.8,1.2,0.7,0.3)$ and $\boldsymbol{Q}=\boldsymbol{Q}_{\mathrm{intra}}(0.5,\mathsf{inr}_{\mathrm{tot}})$, respectively. We set the relevant correlation matrices to $\boldsymbol{\Sigma}=\boldsymbol{\Phi}_{n_{\mathrm{R}}}^{(\exp)}(0.2)$, $\boldsymbol{\Phi}_0=\boldsymbol{\Phi}_{n_0}^{(\exp)}(0.5)$, $\boldsymbol{\Phi}_1=\boldsymbol{\Phi}_{n_1}^{(\exp)}(0.1)$, and $\boldsymbol{\Phi}_2=\boldsymbol{\Phi}_{n_2}^{(\exp)}(0.2)$, respectively. It can be seen that theoretical results are perfectly matched with the simulation results, which verify the accuracy of our analysis.

B. Asymptotic Ergodic Capacity

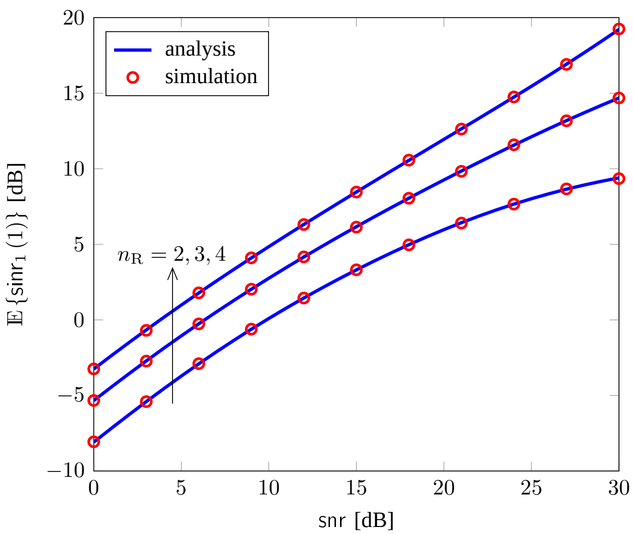
The effectiveness of power allocation policies can be ascertained by referring to Fig. 2 where the average mutual information $\mathcal{I}(\mathsf{snr},\boldsymbol{P},\boldsymbol{Q})$ for the $(4,[2,3],4)$-MIMO interference channel is depicted as a function of snr. In this figure,

$$\begin{aligned}
\frac{1}{n_{\mathrm{R}}}\mathcal{I}_1(\mathsf{snr},\boldsymbol{P}^{\star},\boldsymbol{Q})&\xrightarrow{\text{a.s.}}(\beta+\beta_{\mathrm{I}})\,\mathbb{E}_{\mathsf{t}}\left\{\log_2\left(1+\frac{\mathsf{snr}}{\beta}\mathbb{E}_{\mathsf{r}}\left\{\mathcal{G}(\mathsf{r},\mathsf{t})\,\mathcal{P}(\mathsf{t})\,\mathcal{B}(\mathsf{r},\mathsf{snr})\,|\mathsf{t}\right\}\right)\right\}\\
&\quad-\mathbb{E}\left\{\log_2\mathcal{B}(\mathsf{r},\mathsf{snr})\right\}+\left(\mathbb{E}\left\{\mathcal{B}(\mathsf{r},\mathsf{snr})-1\right\}\right)\log_2 e
\end{aligned}\tag{50}$$

$$\begin{aligned}
\frac{1}{n_{\mathrm{R}}}\mathcal{I}_2(\mathsf{snr},\boldsymbol{Q})&\xrightarrow{\text{a.s.}}\beta_{\mathrm{I}}\,\mathbb{E}_{\mathsf{t}_{\mathrm{I}}}\left\{\log_2\left(1+\frac{\mathsf{snr}}{\beta}\mathbb{E}_{\mathsf{r}_{\mathrm{I}}}\left\{\mathcal{G}_{\mathrm{I}}(\mathsf{r}_{\mathrm{I}},\mathsf{t}_{\mathrm{I}})\,\mathcal{P}_{\mathrm{I}}(\mathsf{t}_{\mathrm{I}})\,\mathcal{B}_{\mathrm{I}}(\mathsf{r}_{\mathrm{I}},\mathsf{snr})\,|\mathsf{t}_{\mathrm{I}}\right\}\right)\right\}\\
&\quad-\mathbb{E}\left\{\log_2\mathcal{B}_{\mathrm{I}}(\mathsf{r}_{\mathrm{I}},\mathsf{snr})\right\}+\left(\mathbb{E}\left\{\mathcal{B}_{\mathrm{I}}(\mathsf{r}_{\mathrm{I}},\mathsf{snr})-1\right\}\right)\log_2 e
\end{aligned}\tag{51}$$



(a) $\mathbb{E}\{\mathsf{mmse}_1(\boldsymbol{P}, \boldsymbol{Q})\}$



(b) $\mathbb{E}\{\mathsf{sinr}_1(1)\}$

Fig. 1. (a) Average MMSE $\mathbb{E}\{\mathsf{mmse}_1(\boldsymbol{P}, \boldsymbol{Q})\}$ and (b) average SINR $\mathbb{E}\{\mathsf{sinr}_1(1)\}$ in dB for the first parallel streams of the $(4, [2, 3], n_{\mathrm{R}})$-MIMO interference channel as a function of snr at $\mathsf{inr}_{\mathrm{tot}} = 10$ dB with $\boldsymbol{Q} = \boldsymbol{Q}_{\mathrm{intra}}(0.5, \mathsf{inr}_{\mathrm{tot}})$ when $n_{\mathrm{R}} = 2, 3, 4$.

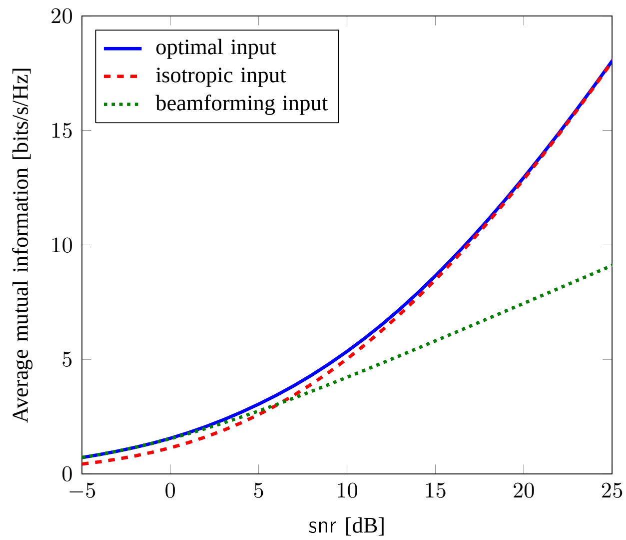


Fig. 2. Average mutual information $\mathcal{I}(\mathsf{snr}, \boldsymbol{P}, \boldsymbol{Q})$ of the $(4, [2, 3], 4)$-MIMO interference channel archived by different input schemes: optimal input, isotropic input, and beamforming input at $\mathsf{inr}_{\mathrm{tot}} = 10$ dB with $\boldsymbol{Q} = \boldsymbol{Q}_{\mathrm{intra}}(0.5, \mathsf{inr}_{\mathrm{tot}})$.

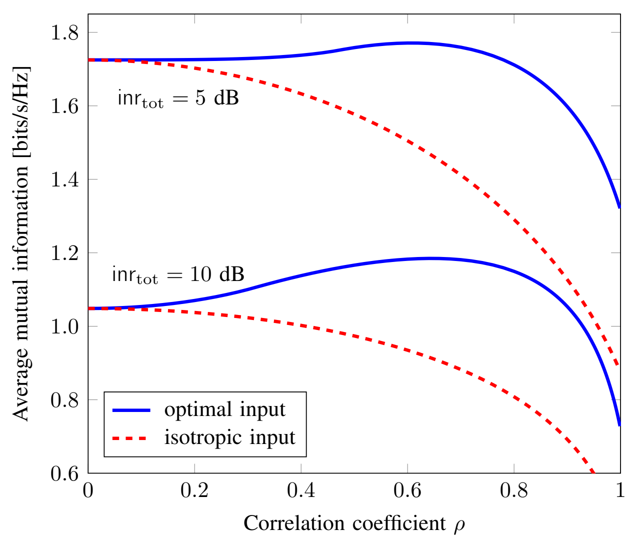


Fig. 3. Average mutual information $\mathcal{I}(\mathsf{snr}, \boldsymbol{P}, \boldsymbol{Q})$ of the $(2, [1, 1, 1], 2)$-MIMO interference channel as a function of the correlation coefficient ρ at $\mathsf{snr} = 5$ dB with $\boldsymbol{\Sigma} = \boldsymbol{\Phi}_{n_{\mathrm{R}}}^{(\exp)}(\rho)$, $\boldsymbol{\Phi}_0 = \boldsymbol{\Phi}_{n_0}^{(\exp)}(\rho)$, and $\boldsymbol{Q} = \boldsymbol{Q}_{\mathrm{intra}}(0, \mathsf{inr}_{\mathrm{tot}})$ when $\mathsf{inr}_{\mathrm{tot}} = 5$ dB and 10 dB.

the correlation matrices and their coefficients are the same as in Fig. 1. It is obvious that the input optimization gives higher achievable rate than that of isotropic input, especially, at low-SNR regime. However, the isotropic input approaches the maximum achievable rate as SNR increases. We also plot the achievable rate obtained by using beamforming input along with the best eigenvector direction. It can be seen that, the beamforming input gives a good performance in low-SNR regimes. To further ascertain the efficacy of signaling strategy, we plot the average mutual information $\mathcal{I}(\mathsf{snr}, \boldsymbol{P}, \boldsymbol{Q})$ of the $(2, [1, 1, 1], 2)$-MIMO interference channel in Fig. 3 as a function of the correlation coefficient ρ at $\mathsf{snr} = 5$ dB with $\boldsymbol{\Sigma} = \boldsymbol{\Phi}_{n_{\mathrm{R}}}^{(\exp)}(\rho)$, $\boldsymbol{\Phi}_0 = \boldsymbol{\Phi}_{n_0}^{(\exp)}(\rho)$, and $\boldsymbol{Q} = \boldsymbol{Q}_{\mathrm{intra}}(0, \mathsf{inr}_{\mathrm{tot}})$ when $\mathsf{inr}_{\mathrm{tot}} = 5$ dB and 10 dB. We can see that the statistical channel knowledge can be exploited to boost the achievable rate when the transmit antennas are highly correlated. We can also observe that the antenna correlation can improve the capacity of MIMO interference channels, especially at the low SNR regime.

C. Effect of Interference Power Profile

Fig. 4 demonstrates the effect of interference power profile on the ergodic capacity, where the approximate ergodic capacity as a function of $\mathsf{inr}_{\mathrm{tot}}$ is plotted for the intraclass power profile $\boldsymbol{Q}_{\mathrm{intra}}(\eta, \mathsf{inr}_{\mathrm{tot}})$. In this figure, we set $\mathsf{snr} = 15$ dB, $n_{\mathrm{R}} = 2$, $n_0 = 3$, $N = 3$, $\boldsymbol{P}^\star = \boldsymbol{I}_{n_0}$, and $\eta = 0$ (equal-power interferers), 0.5, 0.7, 0.8, 0.9, 0.95, 1 (dominant strong interferer). All the fading channels are uniformly distributed

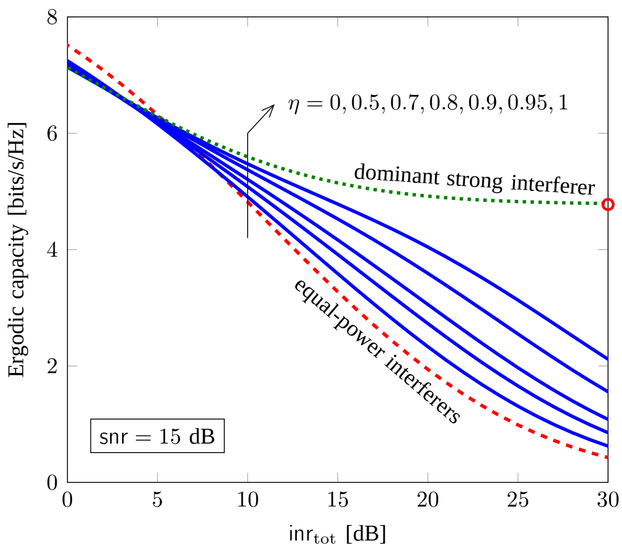


Fig. 4. Ergodic capacity as a function of $\mathsf{inr}_{\mathrm{tot}}$ at $\mathsf{snr} = 15$ dB with $n_{\mathrm{R}} = 2$, $n_0 = 3$, $N = 3$, $\boldsymbol{P}^{\star} = \boldsymbol{I}_{n_0}$, $\boldsymbol{Q} = \boldsymbol{Q}_{\mathrm{intra}}(\eta, \mathsf{inr}_{\mathrm{tot}})$ when $\eta = 0$ (equal-power interferers), 0.5, 0.7, 0.8, 0.9, 0.95, 1 (dominant strong interferer).

around zero with the all-one matrix $\boldsymbol{\Omega}_0$, and

$$\boldsymbol{\Omega}_{\mathrm{I}} = \begin{bmatrix} 3.6 & 0.4 & 0.5 \\ 1 & 0.3 & 0.2 \end{bmatrix}.$$

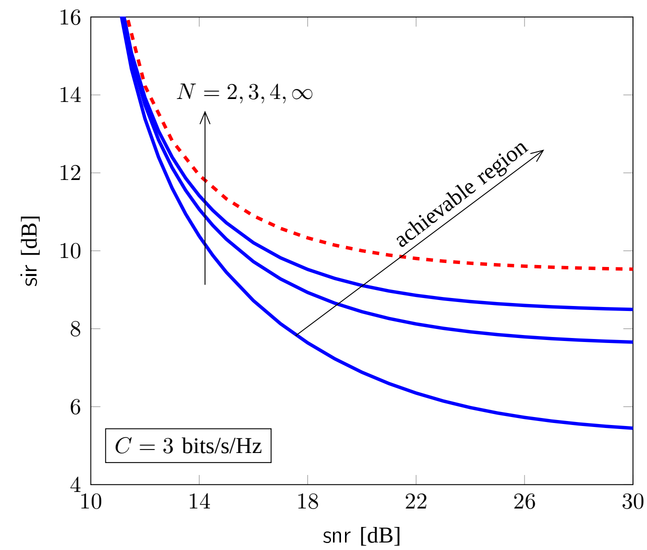
We can observe that the homogeneity of the interference power profile gives less severe interfering effect at low-INR regimes, while the more heterogeneous interference power profile results in the less severe interfering effect at high-INR regimes in terms of ergodic capacity because it helps to increase the ability of the receiver to mitigate the interfering signals. We can also see that the approximate ergodic capacity converges to a floor given by the ergodic capacity of a single user $n_0 \times n_{\mathrm{R}} - 1$ MIMO system (the red circle on the same figure) for $\eta = 1$. This can be attributed to the fact that for small value of η, the receiver uses one degree of freedom to mitigate the interfering signal and uses the remain ones for boosting the data rate, and hence it can be seen as a single user $n_0 \times n_{\mathrm{R}} - 1$ MIMO system.

D. Achievable Region

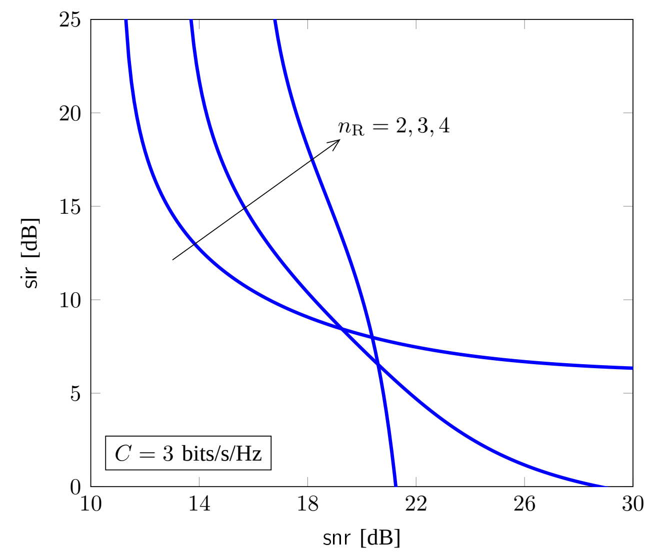
Fig. 5 shows the achievable region of (SIR, SNR) for the desired ergodic capacity per receive antenna $C = 3$ bits/s/Hz of the polarization system when (a) $n_{\mathrm{R}} = 2, n_0 = 3$, and $N = 2, 3, 4, \infty$, (b) $n_0 = 2$, $N = 2$, and $n_{\mathrm{R}} = 2, 3, 4$. In this example, we use the polarization gain matrix model as [14]

$$\boldsymbol{\Omega}^{\mathrm{polar}}_{n \times m}(\mathcal{X}) \triangleq \frac{2}{1+\mathcal{X}} \begin{bmatrix} 1 & \mathcal{X} & 1 & \cdots & \mathcal{X} \\ \mathcal{X} & 1 & \mathcal{X} & \cdots & 1 \\ 1 & \mathcal{X} & 1 & \cdots & \mathcal{X} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathcal{X} & 1 & \mathcal{X} & \cdots & 1 \end{bmatrix}.$$

We assume that the fading has independent real and imaginary parts uniformly distributed around zero with the relevant channel gain matrices $\boldsymbol{\Omega}_0 = \boldsymbol{\Omega}^{\mathrm{polar}}_{n_{\mathrm{R}} \times n_0}(0.5)$ and $\boldsymbol{\Omega}_{\mathrm{I}} = \boldsymbol{\Omega}^{\mathrm{polar}}_{n_{\mathrm{R}} \times N}(0.3)$, respectively. In Fig. 5(a), each curve corresponds to the required aggregate SIR as a function of snr, at which the desired normalized ergodic capacity $C = 3$ bits/s/Hz is attained for the total number of interferers. For example, at the average received SNR of 20 dB, the MMSE receiver requires at least the values of the aggregate SIR of about 7 dB, 8.5 dB, 9 dB, and 10 dB for $N = 2, 3, 4$ and ∞, respectively, while there is no SIR value achieving the desired normalized ergodic capacity for SNRs below 11 dB. As the SNR tends to infinity, the required SIR approaches 5.5 dB, 7.7 dB, 8.5 dB and 9.5 dB for $N = 2, 3, 4$ and ∞, respectively, which are same as the corresponding SINRs required in the interference-limited regime. The effect of the number of receiving antenna on the achievable region of (SIR, SNR) is further ascertained in Fig. 5(b). As can be seen from the figure that for $n_{\mathrm{R}} = 2$, when snr grows large, the SIR is saturated, and it can not be improved even when snr tends to infinity. This is due to the



(a) $n_{\mathrm{R}} = 2$, $n_0 = 3$



(b) $n_0 = 2$, $N = 2$

Fig. 5. Achievable region of (SNR,SIR) for the desired ergodic capacity per receive antenna $C = 3$ bits/s/Hz of the polarization system with isotropic inputs for both the probe transmitter and the interferers when (a) $n_{\mathrm{R}} = 2$, $n_0 = 3$, and $N = 2, 3, 4, \infty$, (b) $n_0 = 2$, $N = 2$, and $n_{\mathrm{R}} = 2, 3, 4$.

lack of number of degrees of freedom such that the MMSE receiver can completely suppress the interference, and thus the ergodic capacity is mostly determined by the SIR. When $n_{\mathrm{R}} = 3$, the number of receive antennas exceeds the number of the probe transmit antennas, but smaller than the number of probe transmit antennas plus the interferer's antennas. Therefore, the receiver must compromise between mitigating the interfering effect and boosting the data rate. For the case $n_{\mathrm{R}} = 4$, the number of receive antennas exceeds the number of transmit plus interferers's antennas, thus the receiver can simultaneously totally suppress the outside interference and detect the desired signal. Therefore, the normalized ergodic capacity is independent of the SIR as the SNR grows large and is determined only by the SNR. For example, at the average received SNR of 24 dB, the MMSE receiver required at least the values of the aggregate SIR of about 7 dB and 2.6 dB for $n_{\mathrm{R}} = 2$ and $n_{\mathrm{R}} = 3$, respectively, while we can attain higher desired capacity for any values of the aggregate SIR for $n_{\mathrm{R}} = 4$.

VII. Conclusion

Using key results of Berezin's supermathematics and large random matrix theory, we developed a framework to characterize the effects of cochannel interference and antenna correlation on the capacity in of MIMO channels with multiuser interference. In particular, we studied the optimal transmission strategy and the ergodic capacity of a MIMO system, taking into account the heterogeneity of interferer powers and spatial fading correlation falling inside the UIU model. We characterized the optimal power allocation in terms of a necessary and sufficient condition that encompasses the noise limited case. As a result, we provided a simple but efficient algorithm to find the optimum power allocation for a covariance feedback system. To reduce the computational complexity, we used Berezin's supermathematics to derive the analytical solutions for the average MMSE and SINR on the linear estimations of the parallel spatial streams. We derived the asymptotic capacity per receive antenna of UIU MIMO channels as the numbers of antennas grow large. We showed that the asymptotic analysis gives a very good accuracy and it only required to solve the fixed point equations of $2n_{\mathrm{R}}$ variables which converge very quickly to any target of accuracy level. Finally, in terms of ergodic capacity, it is more favorable to have a dominant strong interferer rather than to have equal-power interferers as INR increases, while equal-power interferers are more preferable at low INR regimes.

Appendix A
Proof of Theorem 1

Let $\boldsymbol{g}_1 \in \mathbb{G}^m$, $\boldsymbol{y}_1 \in \mathbb{C}^m$, $\boldsymbol{g}_2 \in \mathbb{G}^n$, $\boldsymbol{y}_2 \in \mathbb{C}^n$, $\boldsymbol{g}_3 \in \mathbb{G}^\ell$, $\boldsymbol{y}_3 \in \mathbb{C}^\ell$, and

$$\boldsymbol{\mathfrak{s}}_1 = \begin{bmatrix} \boldsymbol{g}_1 \\ \boldsymbol{y}_1 \end{bmatrix}, \quad \boldsymbol{\mathfrak{s}}_2 = \begin{bmatrix} \boldsymbol{g}_2 \\ \boldsymbol{y}_2 \end{bmatrix}, \quad \boldsymbol{\mathfrak{s}}_3 = \begin{bmatrix} \boldsymbol{g}_3 \\ \boldsymbol{y}_3 \end{bmatrix}. \tag{56}$$

Then, using basic formulas of supermathematics and [43, Lemma 2], the detquotient in (9) can be decoupled in terms of superintegrals as (57). Exploiting the Fourier representation of the delta function and introducing an integral over the Grassmann variables $\hbar_1$ and $\hbar_2$, we decouple the exponents in (57) in a similar way as in [4]. Note that alternative expression for $\mathcal{K}_1(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C})$ can be obtained from the results derived in [16] using the replica method. Next, we combine the results and evaluate the integral first over the anticommuting vectors, and then over the commuting ones, we obtain (58) where

$$\boldsymbol{V}_1 = \boldsymbol{A}\left(\boldsymbol{I}_m - \jmath\omega_1\boldsymbol{A}\right)^{-1}\boldsymbol{A}\left(\boldsymbol{I}_m + \jmath\left(\omega_2 + \omega_3\right)\boldsymbol{A}\right)^{-1}, \tag{59}$$

$$\boldsymbol{V}_2 = \boldsymbol{B}\left(\boldsymbol{I}_n + t_1\boldsymbol{B}\right)^{-1}\boldsymbol{B}\left(\boldsymbol{I}_n + t_2\boldsymbol{B}\right)^{-1}. \tag{60}$$

For a Grassmann variable $g \in \mathbb{G}$ and a complex matrix $\boldsymbol{\Omega} \in \mathbb{C}^{m\times m}$, we have [4, eq. (43)]

$$\det\left(\boldsymbol{I}_m + g\boldsymbol{\Omega}\right)^{-1} = \prod_{i=1}^{m}\left(1 + g\,\mathsf{eig}_i(\boldsymbol{\Omega})\right)^{-1} \tag{61}$$

from which the integral $\mathcal{I}_1$ can be evaluated as

$$\mathcal{I}_1 = 1 - \operatorname{tr}\left(\boldsymbol{V}_1\right)\operatorname{tr}\left(\boldsymbol{V}_2\right). \tag{62}$$

Hence, we get

$$\mathcal{K}_1\left(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}\right) = \mathcal{L}_1\left(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}\right) - \mathcal{L}_2\left(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}\right) \tag{63}$$

where $\mathcal{L}_1(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C})$ and $\mathcal{L}_2(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C})$ are given in (64) and (65), respectively.

$$\begin{aligned}\mathcal{K}_1\left(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}\right) = \iiint & \exp\left(\boldsymbol{g}_1^\dagger\boldsymbol{A}\boldsymbol{g}_1\boldsymbol{g}_2^\dagger\boldsymbol{B}\boldsymbol{g}_2 + \boldsymbol{y}_1^\dagger\boldsymbol{A}\boldsymbol{g}_1\boldsymbol{g}_2^\dagger\boldsymbol{B}\boldsymbol{y}_2 - \boldsymbol{g}_1^\dagger\boldsymbol{A}\boldsymbol{y}_1\boldsymbol{y}_2^\dagger\boldsymbol{B}\boldsymbol{g}_2 - \boldsymbol{y}_1^\dagger\boldsymbol{A}\boldsymbol{y}_1\boldsymbol{y}_2^\dagger\boldsymbol{B}\boldsymbol{y}_2\right) \\ & \times \exp\left(-\boldsymbol{y}_1^\dagger\boldsymbol{A}\boldsymbol{y}_1\boldsymbol{y}_3^\dagger\boldsymbol{C}\boldsymbol{y}_3\right)\exp\left(-\boldsymbol{\mathfrak{s}}_1^\dagger\boldsymbol{\mathfrak{s}}_1 - \boldsymbol{\mathfrak{s}}_2^\dagger\boldsymbol{\mathfrak{s}}_2 - \boldsymbol{\mathfrak{s}}_3^\dagger\boldsymbol{\mathfrak{s}}_3\right)d\boldsymbol{\mathfrak{s}}_3^\dagger d\boldsymbol{\mathfrak{s}}_3 d\boldsymbol{\mathfrak{s}}_2^\dagger d\boldsymbol{\mathfrak{s}}_2 d\boldsymbol{\mathfrak{s}}_1^\dagger d\boldsymbol{\mathfrak{s}}_1\end{aligned} \tag{57}$$

$$\begin{aligned}\mathcal{K}_1\left(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}\right) = \left(\frac{1}{2\pi}\right)^3 \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} \Bigg\{ & \overbrace{\iint \frac{e^{-\left(\hbar_1^*\hbar_1 + \hbar_2^*\hbar_2\right)}}{\det(\boldsymbol{I}_m + \hbar_2\hbar_1\boldsymbol{V}_1)\det\left(\boldsymbol{I}_n + \hbar_2^*\hbar_1^*\boldsymbol{V}_2\right)} d\hbar_1^* d\hbar_1 d\hbar_2^* d\hbar_2}^{\triangleq \mathcal{I}_1} \\ & \times \underbrace{\frac{e^{\jmath(\omega_1 t_1 + \omega_2 t_2 + \omega_3 t_3)}\det(\boldsymbol{I}_m - \jmath\omega_1\boldsymbol{A})\det(\boldsymbol{I}_n + t_1\boldsymbol{B})}{\det(\boldsymbol{I}_m + \jmath(\omega_2 + \omega_3)\boldsymbol{A})\det(\boldsymbol{I}_n + t_2\boldsymbol{B})\det(\boldsymbol{I}_l + t_3\boldsymbol{C})}}_{\triangleq f(\omega_1, \omega_2, \omega_3, t_1, t_2, t_3)}\Bigg\} d\omega_1 dt_1 d\omega_2 dt_2 d\omega_3 dt_3\end{aligned} \tag{58}$$

$$\mathcal{L}_1(\boldsymbol{A},\boldsymbol{B},\boldsymbol{C}) = \left(\frac{1}{2\pi}\right)^3 \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} f(\omega_1,\omega_2,\omega_3,t_1,t_2,t_3)\, d\omega_1 dt_1 d\omega_2 dt_2 d\omega_3 dt_3 \tag{64}$$

$$\mathcal{L}_2(\boldsymbol{A},\boldsymbol{B},\boldsymbol{C}) = \left(\frac{1}{2\pi}\right)^3 \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} f(\omega_1,\omega_2,\omega_3,t_1,t_2,t_3)\,\mathrm{tr}(\boldsymbol{V}_1)\,\mathrm{tr}(\boldsymbol{V}_2)\, d\omega_1 dt_1 d\omega_2 dt_2 d\omega_3 dt_3 \tag{65}$$

1) Evaluation of $\mathcal{L}_1(\boldsymbol{A},\boldsymbol{B},\boldsymbol{C})$: Let

$$\mathcal{L}_{1,1} = \int_{-\infty}^{\infty} \mathcal{F}^{-1}_{(\omega_1)}\{\det(\boldsymbol{I}_m - \jmath\omega_1\boldsymbol{A})\}(t_1)\det(\boldsymbol{I}_n + t_1\boldsymbol{B})\,dt_1 \tag{66}$$

$$\begin{aligned}\mathcal{L}_{1,2} &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathcal{F}^{-1}_{(\omega_2,\,\omega_3)}\left\{\det(\boldsymbol{I}_m + \jmath(\omega_2+\omega_3)\boldsymbol{A})^{-1}\right\}(t_2,t_3)\\ &\quad\times \det(\boldsymbol{I}_n + t_2\boldsymbol{B})^{-1}\det(\boldsymbol{I}_l + t_3\boldsymbol{C})^{-1}\,dt_2 dt_3.\end{aligned} \tag{67}$$

Then, $\mathcal{L}_1(\boldsymbol{A},\boldsymbol{B},\boldsymbol{C})$ in (64) can be written as $\mathcal{L}_1(\boldsymbol{A},\boldsymbol{B},\boldsymbol{C}) = \mathcal{L}_{1,1}\mathcal{L}_{1,2}$. Expanding $\det(\boldsymbol{I}_n + t_1\boldsymbol{B})$ with the eigenpolynomials and using [4, Lemma 1-ii)], we obtain $\mathcal{L}_{1,1} = \zeta(\boldsymbol{A},\boldsymbol{B})$. Again using the characteristic coefficients expansion, we obtain

$$\begin{aligned}&\mathcal{F}^{-1}_{(\omega_2,\,\omega_3)}\left\{\det(\boldsymbol{I}_m + \jmath(\omega_2+\omega_3)\boldsymbol{A})^{-1}\right\}(t_2,t_3)\\ &= \frac{1}{2\pi}\sum_{i=1}^{\varrho(\boldsymbol{A})}\sum_{j=1}^{\tau_i(\boldsymbol{A})}\Bigg\{\mathcal{X}_{i,j}(\boldsymbol{A})\\ &\quad\times\int_{-\infty}^{\infty}\frac{t_2^{j-1}}{\Gamma(j)\mathsf{eig}^j_{[i]}(\boldsymbol{A})}e^{\jmath\omega_3 t_3 - \left(\frac{1}{\mathsf{eig}_{[i]}(\boldsymbol{A})}+\jmath\omega_3\right)t_2}u(t_2)\,d\omega_3\Bigg\}\\ &= \sum_{i=1}^{\varrho(\boldsymbol{A})}\sum_{j=1}^{\tau_i(\boldsymbol{A})}\frac{\mathcal{X}_{i,j}(\boldsymbol{A})}{\Gamma(j)\mathsf{eig}^j_{[i]}(\boldsymbol{A})}e^{\frac{-t_2}{\mathsf{eig}_{[i]}(\boldsymbol{A})}}t_2^{j-1}u(t_2)\,\delta(t_3-t_2)\end{aligned} \tag{68}$$

where the last equality follows the integral identity [42, eq. (3.382.7)]:

$$\int_{-\infty}^{\infty}(a+\jmath\omega)^{-\nu}e^{\jmath\omega t}d\omega = \frac{2\pi t^{\nu-1}}{\Gamma(\nu)}e^{-at}u(t) \tag{69}$$

for $\Re(a) > 0, \Re(\nu) > 0$. Expanding $\det(\boldsymbol{I}_n + t_2\boldsymbol{B})^{-1}$ and $\det(\boldsymbol{I}_l + t_3\boldsymbol{C})^{-1}$ with the characteristic coefficients, along with (68), we obtain (70), where the last equality follows from the integral

$$\mathcal{J}^v_{j,q}(a,\,b,\,c) \triangleq \frac{1}{(j-1)!}\int_0^{\infty}(1+ax)^{-q}(1+bx)^{-v}x^{j-1}e^{-\frac{x}{c}}dx \tag{71}$$

for $a > 0,\ b > 0,\ c > 0$, and $j \in \mathbb{Z}_+$. For $a \neq b$, in order to find the closed-form expression for $\mathcal{J}^v_{j,q}(a,\,b,\,c)$, we first decompose the term $(1+ax)^{-q}(1+bx)^{-v}$ as

$$\begin{aligned}&(1+ax)^{-q}(1+bx)^{-v}\\ &= \sum_{k=1}^{q}\left(-\frac{b}{a}\right)^{q-k}\frac{(v+q-k-1)!}{(q-k)!\,(v-1)!}\left(1-\frac{b}{a}\right)^{k-v-q}(1+ax)^{-k}\\ &\quad+\left(-\frac{b}{a}\right)^{q}\sum_{i=1}^{v}\frac{(v+q-i-1)!}{(q-1)!\,(v-i)!}\left(1-\frac{b}{a}\right)^{i-v-q}(1+bx)^{-i}.\end{aligned} \tag{72}$$

Then, substituting (72) into (71), we arrive at the desired result in (12), in which we use the integral identity:

$$\frac{1}{(n-1)!}\int_0^{\infty}(1+ax)^{\mu-1}x^{n-1}e^{-x/b}dx = b^n{}_2F_0(n,-\mu+1\,;-ab) \tag{73}$$

for $a > 0,\ b > 0$, and $n \in \mathbb{Z}_+$ [13, Appendix A]. For $a = b$, we get $\mathcal{J}^v_{j,q}(a,\,b,\,c) = c^j{}_2F_0(j,q+v\,;-ac)$. Therefore, we arrive at the desired result

$$\mathcal{L}_1(\boldsymbol{A},\boldsymbol{B},\boldsymbol{C}) = \zeta(\boldsymbol{A},\boldsymbol{B})\,\Delta(\boldsymbol{B},\boldsymbol{C},\boldsymbol{A}). \tag{74}$$

$$\begin{aligned}\mathcal{L}_{1,2} &= \sum_{p=1}^{\varrho(\boldsymbol{B})}\sum_{q=1}^{\tau_p(\boldsymbol{B})}\sum_{u=1}^{\varrho(\boldsymbol{C})}\sum_{v=1}^{\tau_u(\boldsymbol{C})}\sum_{i=1}^{\varrho(\boldsymbol{A})}\sum_{j=1}^{\tau_i(\boldsymbol{A})}\Bigg\{\frac{\mathcal{X}_{i,j}(\boldsymbol{A})\,\mathcal{X}_{p,q}(\boldsymbol{B})\,\mathcal{X}_{u,v}(\boldsymbol{C})}{\Gamma(j)\,\mathsf{eig}^j_{[i]}(\boldsymbol{A})}\\ &\qquad\times\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left(1+t_2\mathsf{eig}_{[p]}(\boldsymbol{B})\right)^{-q}\left(1+t_3\mathsf{eig}_{[u]}(\boldsymbol{C})\right)^{-v}e^{\frac{-t_2}{\mathsf{eig}_{[i]}(\boldsymbol{A})}}t_2^{j-1}u(t_2)\,\delta(t_3-t_2)\,dt_2dt_3\Bigg\}\\ &= \sum_{p=1}^{\varrho(\boldsymbol{B})}\sum_{q=1}^{\tau_p(\boldsymbol{B})}\sum_{u=1}^{\varrho(\boldsymbol{C})}\sum_{v=1}^{\tau_u(\boldsymbol{C})}\sum_{i=1}^{\varrho(\boldsymbol{A})}\sum_{j=1}^{\tau_i(\boldsymbol{A})}\frac{\mathcal{X}_{i,j}(\boldsymbol{A})\,\mathcal{X}_{p,q}(\boldsymbol{B})\,\mathcal{X}_{u,v}(\boldsymbol{C})}{\Gamma(j)\,\mathsf{eig}^j_{[i]}(\boldsymbol{A})}\int_0^{\infty}\left(1+t\mathsf{eig}_{[p]}(\boldsymbol{B})\right)^{-q}\left(1+t\mathsf{eig}_{[u]}(\boldsymbol{C})\right)^{-v}e^{\frac{-t}{\mathsf{eig}_{[i]}(\boldsymbol{A})}}t^{j-1}dt\\ &= \Delta(\boldsymbol{B},\boldsymbol{C},\boldsymbol{A})\end{aligned} \tag{70}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TWC.2018.2865939, IEEE Transactions on Wireless Communications

2) Evaluation of $\mathcal{L}_2(\boldsymbol{A},\boldsymbol{B},\boldsymbol{C})$: Let

$$\begin{aligned} & f_1(\omega_1,t_2,t_3) \\ &= \mathcal{F}^{-1}_{(\omega_2,\omega_3)}\left\{\frac{\operatorname{tr}(\boldsymbol{V}_1)}{\det(\boldsymbol{I}_m+\jmath(\omega_2+\omega_3)\boldsymbol{A})}\right\}(t_2,t_3) \qquad (75) \\ & f_2(t_1,t_2,t_3) \\ &= \mathcal{F}^{-1}_{(\omega_1)}\left\{\det(\boldsymbol{I}_m-\jmath\omega_1\boldsymbol{A}) f_1(\omega_1,t_2,t_3)\right\}(t_1) \qquad (76) \\ & f_3(t_1) \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{\operatorname{tr}(\boldsymbol{V}_2) f_2(t_1,t_2,t_3)}{\det(\boldsymbol{I}_n+t_2\boldsymbol{B})^{-1}\det(\boldsymbol{I}_l+t_3\boldsymbol{C})^{-1}}dt_2 dt_3. \qquad (77) \end{aligned}$$

Then, we can write $\mathcal{L}_2(\boldsymbol{A},\boldsymbol{B},\boldsymbol{C})$ as

$$\mathcal{L}_2(\boldsymbol{A},\boldsymbol{B},\boldsymbol{C}) = \int_{-\infty}^{\infty}\det(\boldsymbol{I}_n+t_1\boldsymbol{B}) f_3(t_1)\, dt_1. \qquad (78)$$

Since

$$\operatorname{tr}(\boldsymbol{V}_1) = \sum_{k=1}^{m}\frac{\mathsf{eig}_k^2(\boldsymbol{A})}{(1-\jmath\omega_1\mathsf{eig}_k(\boldsymbol{A}))(1+\jmath(\omega_2+\omega_3)\mathsf{eig}_k(\boldsymbol{A}))},$$

we have

$$\begin{aligned} & f_1(\omega_1,t_2,t_3) \\ &= \sum_{k=1}^{m}\sum_{i=1}^{\varrho(\boldsymbol{A})}\sum_{j=1}^{\tau_i(\boldsymbol{A}^{(k)})}\left\{\frac{\mathcal{X}_{i,j}\left(\boldsymbol{A}^{(k)}\right)\mathsf{eig}_k^2(\boldsymbol{A})t_2^{j-1}}{\Gamma(j)\mathsf{eig}_{[i]}^{j}(\boldsymbol{A})(1-\jmath\omega_1\mathsf{eig}_k(\boldsymbol{A}))}\right. \\ &\qquad \left.\times e^{\frac{-t_2}{\mathsf{eig}_{[i]}(\boldsymbol{A})}}u(t_2)\delta(t_3-t_2)\right\}. \qquad (79) \end{aligned}$$

From (79), $f_2(t_1,t_2,t_3)$ in (76) is given by

$$\begin{aligned} & f_2(t_1,t_2,t_3) \\ &= \sum_{k=1}^{m}\mathcal{F}^{-1}_{(\omega_1)}\left\{\det\left(\boldsymbol{I}_m-\jmath\omega_1\boldsymbol{A}_{(k)}\right)\right\}(t_1) \\ &\times\sum_{i=1}^{\varrho(\boldsymbol{A})}\sum_{j=1}^{\tau_i(\boldsymbol{A}^{(k)})}\left\{\frac{\mathcal{X}_{i,j}\left(\boldsymbol{A}^{(k)}\right)\mathsf{eig}_k^2(\boldsymbol{A})t_2^{j-1}}{\Gamma(j)\mathsf{eig}_{[i]}^{j}(\boldsymbol{A})}\right. \\ &\qquad \left.\times e^{\frac{-t_2}{\mathsf{eig}_{[i]}(\boldsymbol{A})}}u(t_2)\delta(t_3-t_2)\right\}. \qquad (80) \end{aligned}$$

Again, from (70), (71), (80), and [4, Lemma 1-ii)] along with (81), we obtain (82). Finally, we get (83), from which we complete the proof.

Appendix B
Proof of Theorem 3

We sketch the proof as follows. Using a perturbing transmit power $\epsilon\in\mathbb{R}_+$, we first parameterize the average achievable rate as a function of ϵ and exploit its derivative with respect to ϵ to show the necessary condition for the optimal power allocation. We then prove the sufficiency of the necessary condition by evaluating the difference of the ergodic mutual information between the interference MIMO using the input power matrix satisfying necessary condition and an arbitrary input power matrix.

A. Necessity for Optimal Power Allocation

Let us reallocate the perturbing power ϵ from the ith transmit power p_i to the jth transmit power p_j for the power

$$\begin{aligned} & \det(\boldsymbol{I}_n+t_2\boldsymbol{B})^{-1}\det(\boldsymbol{I}_l+t_2\boldsymbol{C})^{-1}\operatorname{tr}(\boldsymbol{V}_2) \\ &= \sum_{l=1}^{n}\sum_{p=1}^{\varrho(\boldsymbol{B})}\sum_{q=1}^{\tau_p(\boldsymbol{B}^{(l)})}\sum_{u=1}^{\varrho(\boldsymbol{C})}\sum_{v=1}^{\tau_u(\boldsymbol{C})}\frac{\mathcal{X}_{p,q}\left(\boldsymbol{B}^{(l)}\right)\mathcal{X}_{u,v}(\boldsymbol{C})\,\mathsf{eig}_l^2(\boldsymbol{B})}{(1+t_1\mathsf{eig}_l(\boldsymbol{B}))\left(1+t_2\mathsf{eig}_{[p]}(\boldsymbol{B})\right)^q\left(1+t_3\mathsf{eig}_{[p]}(\boldsymbol{C})\right)^v} \end{aligned} \qquad (81)$$

$$\begin{aligned} f_3(t_1) &= \sum_{k=1}^{m}\sum_{l=1}^{n}\sum_{p=1}^{\varrho(\boldsymbol{B})}\sum_{q=1}^{\tau_p(\boldsymbol{B}^{(l)})}\sum_{u=1}^{\varrho(\boldsymbol{C})}\sum_{v=1}^{\tau_u(\boldsymbol{C})}\sum_{i=1}^{\varrho(\boldsymbol{A})}\sum_{j=1}^{\tau_i(\boldsymbol{A}^{(k)})}\left\{\frac{\mathcal{X}_{i,j}\left(\boldsymbol{A}^{(k)}\right)\mathcal{X}_{p,q}\left(\boldsymbol{B}^{(l)}\right)\mathcal{X}_{u,v}(\boldsymbol{C})\mathsf{eig}_l^2(\boldsymbol{B})\mathsf{eig}_k^2(\boldsymbol{A})}{\mathsf{eig}_{[i]}^{j}(\boldsymbol{A})\left(1+t_1\mathsf{eig}_l(\boldsymbol{B})\right)\Gamma(j)}\right. \\ &\quad\left.\times\sum_{k_1=0}^{m-1}(-1)^{k_1}e_{k_1}\left(\boldsymbol{A}_{(k_1)}\right)\delta^{(k_1)}(t_1)\int_0^{\infty}\left(1+t\mathsf{eig}_{[p]}(\boldsymbol{B})\right)^{-q}\left(1+t\mathsf{eig}_{[u]}(\boldsymbol{C})\right)^{-v}e^{\frac{-t}{\mathsf{eig}_{[i]}(\boldsymbol{A})}}t^{j-1}dt\right\} \end{aligned} \qquad (82)$$

$$\mathcal{L}_2(\boldsymbol{A},\boldsymbol{B},\boldsymbol{C}) = \sum_{k=1}^{m}\sum_{l=1}^{n}\mathsf{eig}_k^2(\boldsymbol{A})\,\mathsf{eig}_l^2(\boldsymbol{B})\,\Delta\left(\boldsymbol{B}^{(l)},\boldsymbol{C},\boldsymbol{A}^{(k)}\right)\underbrace{\sum_{k_1=0}^{m-1}(-1)^{k_1}e_{k_1}\left(\boldsymbol{A}_{(k)}\right)\int_{-\infty}^{\infty}\det\left(\boldsymbol{I}_n+t_1\boldsymbol{B}_{(l)}\right)\delta^{(k_1)}(t_1)\,dt_1}_{=\zeta\left(\boldsymbol{A}_{(k)},\boldsymbol{B}_{(l)}\right)} \qquad (83)$$

allocation matrix $\boldsymbol{P} = \mathrm{diag}\left(p_1, p_2, \ldots, p_{n_0}\right)$, leading to

$$\boldsymbol{P}_\epsilon = \mathrm{diag}\left(p_1, \ldots, p_i - \epsilon, \ldots, p_j + \epsilon, \ldots, p_{n_0}\right) \tag{84}$$

for $0 \leqslant \epsilon \leqslant p_i$. The parameterized achievable rate can be then written as

$$\begin{aligned}\mathcal{I}\left(\mathsf{snr}, \boldsymbol{P}_\epsilon, \boldsymbol{Q}\right) = \mathbb{E}\Bigg\{ \log_2 \det \Bigg[& p_j \mathsf{snr} \mathbf{h}_{0j} \mathbf{h}_{0j}^\dagger \\ & + \epsilon\, \mathsf{snr} \left(\mathbf{h}_{0j}\mathbf{h}_{0j}^\dagger - \mathbf{h}_{0i}\mathbf{h}_{0i}^\dagger\right) + \boldsymbol{\Psi}_j \Bigg] \Bigg\} + \text{const.}\end{aligned} \tag{85}$$

Invoking [13, Lemma 1] and the dominated convergence theorem for interchanging the order of differentiation and integration along with the fact that the argument matrix of the determinant in (85) is positive definite, it is clear that $\partial^2 \mathcal{I}\left(\mathsf{snr}, \boldsymbol{P}_\epsilon, \boldsymbol{Q}\right) / \partial \epsilon^2 \leqslant 0$ and hence, $\mathcal{I}\left(\mathsf{snr}, \boldsymbol{P}_\epsilon, \boldsymbol{Q}\right)$ as a function of the perturbation ϵ is concave in $\epsilon \in [0, p_i]$. Therefore, it is sufficient to examine $\left[\partial \mathcal{I}\left(\mathsf{snr}, \boldsymbol{P}_\epsilon, \boldsymbol{Q}\right) / \partial \epsilon\right]\big|_{\epsilon=0} \leqslant 0$ in order to obtain a necessary condition for optimal power allocation. Specifically, we have (86). Since

$$\begin{aligned}\mathrm{tr}&\left(\left(p_j \frac{\mathsf{snr}}{n_0} \mathbf{h}_{0j}\mathbf{h}_{0j}^\dagger + \boldsymbol{\Psi}_j\right)^{-1} \mathbf{h}_{0j}\mathbf{h}_{0j}^\dagger\right) \\ &= \mathrm{tr}\left(\left(\boldsymbol{I} + p_j \frac{\mathsf{snr}}{n_0} \mathbf{h}_{0j}\mathbf{h}_{0j}^\dagger \boldsymbol{\Psi}_j^{-1}\right)^{-1} \mathbf{h}_{0j}\mathbf{h}_{0j}^\dagger \boldsymbol{\Psi}_j^{-1}\right) \\ &= \frac{1 - \mathsf{mmse}_j\left(\boldsymbol{P}, \boldsymbol{Q}\right)}{p_j \frac{\mathsf{snr}}{n_0}},\end{aligned} \tag{87}$$

the inequality $\left[\partial \mathcal{I}\left(\mathsf{snr}, \boldsymbol{P}_\epsilon, \boldsymbol{Q}\right) / \partial \epsilon\right]\big|_{\epsilon=0} \leqslant 0$ is equivalent to $\gamma_i\left(\boldsymbol{P}, \boldsymbol{Q}\right) \geqslant \gamma_j\left(\boldsymbol{P}, \boldsymbol{Q}\right)$. Changing the indices between i and j for $p_i > 0$ and $p_j > 0$ establishes the necessary condition in (25) while the necessary condition in (26) is established from the fact that $\epsilon \geq 0$ and $\gamma_i\left(\boldsymbol{P}, \boldsymbol{Q}\right) \geqslant \gamma_j\left(\boldsymbol{P}, \boldsymbol{Q}\right)$.

B. Sufficiency for Optimal Power Allocation

Let $\boldsymbol{P} = \mathrm{diag}\left(p_1, \ldots, p_{n_0}\right)$ be the power allocation matrix satisfying the necessary condition. Then, for an arbitrary power allocation matrix $\hat{\boldsymbol{P}} = \mathrm{diag}\left(\hat{p}_1, \ldots, \hat{p}_{n_0}\right)$, the difference of ergodic mutual information between $\mathcal{I}\left(\mathsf{snr}, \boldsymbol{P}, \boldsymbol{Q}\right)$ and $\mathcal{I}\left(\mathsf{snr}, \hat{\boldsymbol{P}}, \boldsymbol{Q}\right)$ is given in (88), where (a) follows from the matrix inversion identity $\left(\boldsymbol{I} + \boldsymbol{A}^{-1}\right)^{-1} = \boldsymbol{A}\left(\boldsymbol{I} + \boldsymbol{A}\right)^{-1}$ [52, eq. (161)]; (b) follows from the arithmetic-geometric inequality [42]; and (c) follows from the facts that

$$\frac{1}{n_0} \sum_{i=1}^{n_0} \hat{p}_i \gamma_i\left(\boldsymbol{P}, \boldsymbol{Q}\right) \leqslant \max_i \gamma_i\left(\boldsymbol{P}, \boldsymbol{Q}\right), \tag{89}$$

$$\frac{1}{n_0} \sum_{i=1}^{n_0} p_i \gamma_i\left(\boldsymbol{P}, \boldsymbol{Q}\right) = \max_i \gamma_i\left(\boldsymbol{P}, \boldsymbol{Q}\right). \tag{90}$$

This shows that a transmit power allocation satisfying the necessary condition in (25) and (26) maximizes ergodic mutual information, which completes the proof.

References

[1] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Commun.*, vol. 6, no. 3, pp. 311–335, Mar. 1998.

[2] İ. E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Trans. Telecommun.*, vol. 10, no. 6, pp. 585–595, Nov./Dec. 1999.

[3] Y. Jeong, T. Q. S. Quek, and H. Shin, "Beamforming optimization for multiuser two-tier networks," *J. Commun. and Networks*, vol. 13, no. 4, pp. 327–338, Aug. 2011.

[4] Y. Jeong, H. Shin, and M. Z. Win, "Superanalysis of optimum combining with application to femtocell networks," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 3, pp. 509–524, Apr. 2012.

[5] T. M. Nguyen, Y. Jeong, T. Q. S. Quek, W. P. Tay, and H. Shin, "Interference alignment in a Poisson field of MIMO femtocells," *IEEE Trans. Wireless Commun.*, vol. 12, no. 6, pp. 2633–2645, Jun. 2013.

[6] C. Wang, T. Gou, and S. A. Jafar, "Subspace alignment chains and the degrees of freedom of the three-user MIMO interference channel," *IEEE Trans. Inf. Theory*, vol. 60, no. 5, pp. 2432–2479, May 2014.

$$\frac{\partial \mathcal{I}\left(\mathsf{snr}, \boldsymbol{P}_\epsilon, \boldsymbol{Q}\right)}{\partial \epsilon}\bigg|_{\epsilon=0} = \log_2(e)\, \mathbb{E}\left\{ \mathrm{tr}\left(\left(p_j \frac{\mathsf{snr}}{n_0}\mathbf{h}_{0j}\mathbf{h}_{0j}^\dagger + \boldsymbol{\Psi}_j\right)^{-1} \frac{\mathsf{snr}}{n_0}\left(\mathbf{h}_{0j}\mathbf{h}_{0j}^\dagger - \mathbf{h}_{0i}\mathbf{h}_{0i}^\dagger\right)\right)\right\} \tag{86}$$

$$\begin{aligned}\frac{1}{n_\mathrm{R}}\left(\mathcal{I}\left(\mathsf{snr}, \boldsymbol{P}, \boldsymbol{Q}\right) - \mathcal{I}\left(\mathsf{snr}, \hat{\boldsymbol{P}}, \boldsymbol{Q}\right)\right) &\overset{\text{(a)}}{=} \frac{1}{n_\mathrm{R}} \mathbb{E}\left\{ \log_2 \det\left(\boldsymbol{I} + \left(\frac{\mathsf{snr}}{n_0}\mathbf{H}_0\hat{\boldsymbol{P}}\mathbf{H}_0^\dagger - \frac{\mathsf{snr}}{n_0}\mathbf{H}_0\boldsymbol{P}\mathbf{H}_0^\dagger\right)\left(p_j\frac{\mathsf{snr}}{n_0}\mathbf{h}_{0j}\mathbf{h}_{0j}^\dagger + \boldsymbol{\Psi}_j\right)^{-1}\right)\right\} \\ &\overset{\text{(b)}}{\leqslant} \mathbb{E}\Bigg\{ \log_2\Bigg(\frac{1}{n_\mathrm{R}}\Bigg(n_\mathrm{R} + \sum_{i=1}^{n_0}\frac{\mathsf{snr}}{n_0}\hat{p}_i\, \mathrm{tr}\left(\left(p_j\frac{\mathsf{snr}}{n_0}\mathbf{h}_{0j}\mathbf{h}_{0j}^\dagger + \boldsymbol{\Psi}_j\right)^{-1}\mathbf{h}_{0j}\mathbf{h}_{0j}^\dagger\right) \\ &\qquad - \sum_{i=1}^{n_0}\frac{\mathsf{snr}}{n_0}p_i\, \mathrm{tr}\left(\left(p_j\frac{\mathsf{snr}}{n_0}\mathbf{h}_{0j}\mathbf{h}_{0j}^\dagger + \boldsymbol{\Psi}_j\right)^{-1}\mathbf{h}_{0j}\mathbf{h}_{0j}^\dagger\right)\Bigg)\Bigg)\Bigg\} \\ &\leqslant \log_2\left(\frac{1}{n_\mathrm{R}}\left(n_\mathrm{R} + \sum_{i=1}^{n_0}\hat{p}_i\gamma_i\left(\boldsymbol{P}, \boldsymbol{Q}\right) - \sum_{i=1}^{n_0}p_i\gamma_i\left(\boldsymbol{P}, \boldsymbol{Q}\right)\right)\right) \\ &\overset{\text{(c)}}{\leqslant} 0\end{aligned} \tag{88}$$

[7] J. Hoydis, S. ten Brink, and M. Debbah, "Massive MIMO in the UL/DL of cellular networks: How many antennas do we need?" *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 160–171, Feb. 2013.

[8] Y. Jeong, T. Q. S. Quek, J. S. Kwak, and H. Shin, "Multicasting in stochastic MIMO networks," *IEEE Trans. Wireless Commun.*, vol. 13, no. 4, pp. 1729–1741, Apr. 2014.

[9] D. P. Trinh, Y. Jeong, and H. Shin, "MIMO capacity in Binomial fields networks," *IEEE Access*, vol. 5, pp. 12 545–12 551, 2017.

[10] Y. Jeong, D. P. Trinh, and H. Shin, "Cutset bounds on the capacity of MIMO relay channels," *IEEE Access*, vol. 5, Sep. 2017.

[11] H. Shin and J. H. Lee, "Capacity of multiple-antenna fading channels: Spatial fading correlation, double scattering, and keyhole," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2636–2467, Oct. 2003.

[12] M. Chiani, M. Z. Win, and A. Zanella, "On the capacity of spatially correlated MIMO Rayleigh-fading channels," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2363–2371, Oct. 2003.

[13] H. Shin, M. Z. Win, J. H. Lee, and M. Chiani, "On the capacity of doubly correlated MIMO channels," *IEEE Trans. Wireless Commun.*, vol. 5, no. 8, pp. 2253–2265, Aug. 2006.

[14] A. M. Tulino, A. Lozano, and S. Verdú, "Impact of antenna correlation on the capacity of multiantenna channels," *IEEE Trans. Inf. Theory*, vol. 51, no. 7, pp. 2491–2509, Jul. 2005.

[15] R. S. Blum, "MIMO capacity with interference," *IEEE J. Sel. Areas Commun.*, vol. 21, no. 5, pp. 793–801, Jun. 2003.

[16] A. L. Moustakas, S. H. Simon, and A. M. Sengupta, "MIMO capacity through correlated channels in the presence of correlated interferers and noise: A (not so) large N analysis," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2545–2561, Oct. 2003.

[17] M. Chiani, M. Z. Win, and H. Shin, "MIMO networks: The effects of interference," *IEEE Trans. Inf. Theory*, vol. 56, no. 1, pp. 336–349, Jan. 2010.

[18] P. J. Smith, S. Roy, and M. Shafi, "Capacity of MIMO systems with semicorrelated flat fading," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2781–2788, Oct. 2003.

[19] C.-N. Chuah, D. N. C. Tse, J. M. Kahn, and R. A. Valenzuela, "Capacity scaling in MIMO wireless systems under correlated fading," *IEEE Trans. Inf. Theory*, vol. 48, no. 3, pp. 637–650, Mar. 2002.

[20] A. Ghaderipoor, C. Tellambura, and A. Paulraj, "On the application of character expansions for MIMO capacity analysis," *IEEE Trans. Inf. Theory*, vol. 58, no. 5, pp. 2950–2962, May 2012.

[21] C. Zhong, S. Jin, K.-K. Wong, and M. R. McKay, "Ergodic mutual information analysis for multi-keyhole MIMO channels," *IEEE Trans. Wireless Commun.*, vol. 10, no. 6, pp. 1754–1763, Jun. 2011.

[22] V. Raghavanand and A. M. Sayeed, "Sublinear capacity scaling laws for sparse MIMO channels," *IEEE Trans. Inf. Theory*, vol. 57, no. 1, pp. 345–364, Jan. 2011.

[23] A. M. Sayeed and V. Raghavan, "Maximizing MIMO capacity in sparse multipath with reconfigurable antenna arrays," *IEEE J. Sel. Topics Signal Process.*, vol. 1, no. 1, pp. 156–166, Jun. 2007.

[24] E. A. Jorswieck and H. Boche, "Optimal transmission strategies and impact of correlation in multiantenna systems with different types of channel state information," *IEEE Trans. Signal Process.*, vol. 52, no. 12, pp. 3440–3453, Dec. 2004.

[25] V. V. Veeravalli, Y. Liang, and A. M. Sayeed, "Correlated MIMO wireless channels: Capacity, optimal signaling, and asymptotics," *IEEE Trans. Inf. Theory*, vol. 51, no. 6, pp. 2058–2072, Jun. 2005.

[26] A. M. Tulino, A. Lozano, and S. Verdú, "Capacity-achieving input covariance for single-user multi-antenna channels," *IEEE Trans. Wireless Commun.*, vol. 5, no. 3, pp. 662–671, Mar. 2006.

[27] S. A. Jafar and A. Goldsmith, "Transmitter optimization and optimality of beamforming for multiple antenna systems," *IEEE Trans. Wireless Commun.*, vol. 3, no. 4, pp. 1165–1174, Jul. 2004.

[28] E. A. Jorswieck and H. Boche, "Channel capacity and capacity-range of beamforming in mimo wireless systems under correlated fading with covariance feedback," *IEEE Trans. Wireless Commun.*, vol. 3, no. 5, pp. 1543–1553, Sep. 2004.

[29] S. H. Simon and A. L. Moustakas, "Optimizing MIMO antenna systems with channel covariance feedback," *IEEE J. Sel. Areas Commun.*, vol. 21, no. 3, pp. 406–417, Apr. 2003.

[30] W. Weichselberger, M. Herdin, H. Özelik, and E. Bonek, "A stochastic MIMO channel model with joint correlation of both link ends," *IEEE Trans. Wireless Commun.*, vol. 5, no. 1, pp. 90–100, Jan. 2006.

[31] Y. Jeong and H. Shin, "Effect of joint spatial correlation on the diversity performance of space–time block codes," *IEEE Commun. Lett.*, vol. 13, no. 7, pp. 477–479, Jul. 2009.

[32] S. Catreux, P. F. Driessen, and L. J. Greenstein, "Simulation results for an interference-limited multiple-input multiple-output cellular system," *IEEE Commun. Lett.*, vol. 4, no. 11, pp. 334–336, Nov. 2000.

[33] J. G. Andrews, W. Choi, and R. W. Heath Jr., "Overcoming interference in spatial multiplexing MIMO cellular networks," *IEEE Wireless Commun. Magazine*, vol. 14, no. 6, pp. 95 – 104, Dec. 2006.

[34] R. S. Blum, J. H. Winters, and N. R. Sollenberger, "On the capacity of cellular systems with MIMO," *IEEE Commun. Lett.*, vol. 6, no. 6, pp. 242–244, Jun. 2002.

[35] A. Lozano and A. M. Tulino, "Capacity of multiple-transmit multiple-receive antenna architectures," *IEEE Trans. Inf. Theory*, vol. 48, no. 12, pp. 3117–3128, Dec. 2002.

[36] A. Lozano, A. M. Tulino, and S. Verdú, "Multiple-antenna capacity in the low-power regime," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2527–2544, Oct. 2003.

[37] A. Lozano, A. M. Tulino, and S. Verdú, "High-SNR power offset in multiantenna communication," *IEEE Trans. Inf. Theory*, vol. 51, no. 12, pp. 4134–4151, Dec. 2005.

[38] F. A. Berezin, *The Method of Second Quantization*. New York: Academic Press, 1966.

[39] K. Efetov, *Supersymmetry in Disorder and Chaos*. Cambridge, United Kingdom: Cambridge University Press, 1997.

[40] F. A. Berezin, *Introduction to Superanalysis*, ser. Mathematical Physics and Applied Mathematics, A. A. Kirillov, Ed. New York: Springer, 1987, vol. 9.

[41] F. Haake, *Quantum Signatures of Chaos*, ser. Springer Series in Synergetics. New York: Springer, 2010, vol. 54.

[42] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*, 7th ed. San Diego, CA: Academic, 2007.

[43] H. Shin and M. Z. Win, "MIMO diversity in the presence of double scattering," *IEEE Trans. Inf. Theory*, vol. 54, no. 7, pp. 2976–2996, Jul. 2008.

[44] A. M. Sayeed, "Deconstructing multiantenna fading channels," *IEEE Trans. Signal Process.*, vol. 50, no. 10, pp. 2563–2579, Oct. 2002.

[45] D.-S. Shiu, G. J. Foschini, M. J. Gans, and J. M. Kahn, "Fading correlation and its effect on the capacity of multielement antenna systems," *IEEE Trans. Commun.*, vol. 48, no. 3, pp. 502–513, Mar. 2000.

[46] E. Visotsky and U. Madhow, "Space-time transmit precoding with imperfect feedback," *IEEE Trans. Inf. Theory*, vol. 47, no. 6, pp. 2632–2639, Sep. 2001.

[47] E. A. Jorswieck and H. Boche, "Optimal transmission with imperfect channel state information at the transmit antenna array," *Wireless Personal Commun.*, vol. 27, no. 1, pp. 33–56, Oct. 2003.

[48] Y. Li and Z. Ye, "On optimum transmission power strategy in MIMO systems with covariance feedback," *IEEE Commun. Lett.*, vol. 9, no. 5, pp. 444–456, May 2005.

[49] S. Shi, M. Schubert, and H. Boche, "Dowinlink MMSE transceiver optimization for multiuser MIMO systems: MMSE balancing," *IEEE Trans. Signal Process.*, vol. 56, no. 8, pp. 3702–3712, Aug. 2008.

[50] V. L. Girko, *Theory of Random Determinants*. Norwell, MA: Kluwer (Academic), 1990.

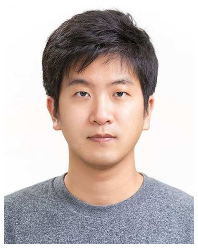
[51] A. M. Tulino and S. Verdú, *Random matrix theory and wireless communications*. Delft, The Netherlands: Now Publishers Inc, 2004, vol. 1.

[52] K. B. Petersen and M. S. Pedersen, "The matrix cookbook," Technical University of Denmark, Nov. 2012.

Vien V. Mai received the B.E. degree in Electronic and Telecommunication from Posts and Telecommunications Institute of Technology, Ho Chi Minh, Vietnam, in 2013, the M.E. degree in Electronics and Radio Engineering from Kyung Hee University, Yongin, South Korea, in 2015. He is currently working toward the Ph.D. degree at the KTH Royal Institute of Technology, Stockholm, Sweden. His research interests include communication theory, distributed detection and estimation, and design and complexity analysis of efficient randomized algorithms suitable for scalable convex optimization.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TWC.2018.2865939, IEEE Transactions on Wireless Communications

Jin Sam Kwak (S'98-M'04) received his B.S. M.S. and Ph.D. degrees in Electrical Engineering and Computer Science from Seoul National University, Seoul, Korea, in 1998, 2000, and 2004, respectively. From 2004 to 2005, he was a postdoctoral research associate in the School of Electrical and Computer Engineering, Georgia Institute of Technology. During 2006, he was also with the Wireless Networks and Communications Group (WNCG), the University of Texas at Austin as a post-doctoral research fellow. From 2007 to 2012, he was with LG Electronics as a chief research engineer. During this time, he carried out research tasks focused on the IMT-Advanced and also led the standards activities for wireless communications such as IEEE 802 (especially, IEEE 802.11/15/16/19), Wi-Fi Alliance, and WiMAX Forum, as well as served as an alternative board member in Wi-Fi Alliance. Since 2012, he has been with WILUS Inc., where he is currently CEO and co-founder. His main research interests focus on wireless communications and immersive A/V technologies including 5G, LTE, Wi-Fi, and MPEG standards.



Youngmin Jeong (S'09-M'15) received the B.E. M.E. and Ph.D. degrees in Electronics and Radio Engineering from Kyung Hee University, Yongin, Korea, in 2009, 2011 and 2015, respectively. He is currently a Research Professor at the Kyung Hee University, Korea. He was a Postdoctoral Fellow at the Kyung Hee University, Korea, from 2015 to 2018, and a Postdoctoral Associate at the Massachusetts Institute of Technology, USA, from 2016 to 2018. His research interests include wireless communications and information theory with current emphasis on MIMO systems, network interference, vehicular communication networks, molecular communications, and quantum communication networks.

Hyundong Shin (S'01-M'04-SM'11) received the B.S. degree in Electronics Engineering from Kyung Hee University, Korea, in 1999, and the M.S. and Ph.D. degrees in Electrical Engineering from Seoul National University, Seoul, Korea, in 2001 and 2004, respectively. During his postdoctoral research at the Massachusetts Institute of Technology (MIT) from 2004 to 2006, he was with the Wireless Communication and Network Sciences Laboratory within the Laboratory for Information Decision Systems (LIDS).

In 2006, Dr. Shin joined Kyung Hee University, Korea, where he is now a Professor at the Department of Electronic Engineering. His research interests include wireless communications and information theory with current emphasis on MIMO communication systems, cooperative communication networks, cognitive radio and networks, network interference, vehicular communication networks, physical-layer security, location-aware radios and networks, molecular communications, and quantum information science.

He served as a Technical Program Co-Chair of the IEEE WCNC (PHY Track 2009) and the IEEE Globecom (Communication Theory Symposium 2012 and Cognitive Radio and Networks Symposium 2016). He was an Editor of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS (2007–2012) and the IEEE COMMUNICATIONS LETTERS (2013–2015). He received the Knowledge Creation Award in the field of computer science from the Korean Ministry of Education, Science and Technology (2010). He also received the IEEE Communications Society's Guglielmo Marconi Prize Paper Award (2008) and the William R. Bennett Prize Paper Award (2012).