Capacity Distribution for Interference Alignment With CSI Errors and Its Applications

Raoul F. Guiazon, Kai-Kit Wong, *Senior Member, IEEE*, and Michael Fitch

***Abstract*—Interference alignment (IA) is known to achieve the degree-of-freedom (DoF) capacity of the interference channel, if full channel state information (CSI) is available at the transmitters perfectly. Challenges, however, arise when CSI is not perfect, and the achievable capacity of IA is not well understood. In this paper, we study the achievable performance of the interference channel using perfect IA techniques based on imperfect CSI. In particular, we obtain the statistical distribution of the maximum achievable rate per stream of the channel. Utilizing our analytical results, we derive new nonasymptotic performance metrics that are then used to 1) optimize the number of streams per user for maximizing the network sum-rate and 2) assess the performance of IA in the time-varying block fading channel. Numerical results are provided to reveal the accuracy of our analytical results.**

***Index Terms*—Beamforming, capacity, channel errors, distribution, interference alignment, sum-rate.**

I. Introduction

INTERFERENCE mitigation techniques are of great importance in the design of wireless communications networks. Clearly, as radio resources are precious, the more we are able to share or reuse them the better. As of today, the conventional techniques used for sharing the frequency spectrum however have relied on orthogonalization over either time (TDMA) or frequency (FDMA). The limitation in this approach is that the resource available for each user decreases with the number of users. This puts a strong limit on the number of users one can accommodate. Reusing the same spectrum to support multiple users certainly would be much more desirable if interference can be properly controlled. This is now possible by the concept of interference alignment (IA) which was first introduced in [1] and subsequently developed in [2], [3]. Remarkably, it has been shown that IA can achieve the degree-of-freedom (DoF) capacity of the interference channel [4], [5].

Conceptually, IA is a linear precoding technique that shapes the signal at each transmitter in a way that at each receiver the interference will only occupy part of the received signal space, while leaving the remaining part free from interference for the desired signal. In this manner, getting rid of the interference becomes as easy as projecting the interference onto the null space of the signal. It is now well known that IA allows each user to access half of the bandwidth available *interference-free* regardless of the number of users in the network. As a result, rather than $\frac{1}{K}$ every user now gets $\frac{1}{2}$ of the total bandwidth [3]. This amazing result, nevertheless, comes under some strong assumptions. The first one is the assumption of having infinite diversity while the second one means the availability of perfect global channel state information (CSI) to every user.

A great deal of efforts have been spent on operating IA in a more realistic setting, e.g., [6]. One such direction to make IA more practical is to drop the assumption of infinite time or frequency diversity and use only finite spatial diversity. Under this consideration, [7] showed that the DoF would be scaling as $\frac{2}{K}$ which is still double what we get with orthogonalization techniques. Relaxing the other assumption regarding the global and perfect CSI also has led to a large body of literature, e.g., [8], [9], which considered the use of *local* CSI and exploited channel reciprocity to apply IA. Most recently in [10], IA was even considered without CSI but using only the knowledge of the network topology. Blind IA was also investigated in [11] without any knowledge of the channel coefficients. Apart from these, in order to reduce the overhead for sharing CSI globally, opportunistic IA was also studied in [12], [13]. Another issue that has been studied a lot is the feasibility of IA for different number of streams and antennas per user, e.g., [14]–[16].

Understandably, IA performs most promisingly with global CSI. In practice, nevertheless, the CSI of the crosstalk channels is likely to be far from perfect, although the CSI of the direct links may be estimated rather accurately. For this reason, there is strong desire to understand the achievable performance limit for IA under such practical scenarios. In this paper, our interest is to analyze the performance of the IA methods designed for global perfect CSI in the presence of CSI uncertainties [15], [16]. Our emphasis is on the "achievable" performance, rather than the average performance.[1] We note that there exist robust IA techniques exploiting imperfect CSI, e.g., [17]–[20] but in that case, analyzing the achievable performance is usually not possible. In [21], assuming that the CSI errors are bounded, an achievable capacity lower bound for IA was derived.

Further to [21] which provides the capacity lower bound for IA with CSI errors, this paper's aim is to provide a complete statistical characterization for the achievable rate. Specifically,

Manuscript received January 20, 2015; revised June 6, 2015; accepted August 29, 2015. This work was supported by EPSRC and BT under Grant EP/K504610/1. The associate editor coordinating the review of this paper and approving it for publication was Prof. Ali Ghrayeb.

R. F. Guiazon and K.-K. Wong are with the Department of Electronic and Electrical Engineering, University College London, London, WC1E 7JE, U.K. (e-mail: raoul.guiazon.13@ucl.ac.uk; kai-kit.wong@ucl.ac.uk).

M. Fitch is with the BT Mobility Group, Martlesham Heath, Suffolk IP5 3RE, U.K. (e-mail: michael.fitch@bt.com).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TWC.2015.2477298

[1] Achievable performance is the performance that has an operational meaning but average performance is only an average indicator for the performance. For example, an average rate is not achievable because the actual channel rate for a given error instantiation may not meet the average rate.

our main contribution is the statistical distribution for the rate per stream achievable by perfect IA based on imperfect CSI. We will also derive metrics such as outage probability and the saturating signal-to-noise ratio (SNR) that can be useful in the design of a practical system using perfect IA with imperfect CSI. Two applications are then presented to demonstrate that our result can be applied to (i) optimize the number of streams per user in the interference network for maximizing the sum-rate, and (ii) analyze the outage performance of IA in block-fading channels with degrading CSI over time.

The remaining of this paper is structured as follows. Section II presents the system model and introduces the IA method. In Section III, we then derive the probability distribution of the achievable rate per stream of IA. Section IV introduces two performance metrics, namely, the saturating SNR and the outage probability to help analyze the interference channel performance using IA, while Section V provides two applications of our results in the optimization of IA. Section VI shows our simulation results and finally, concluding remarks are given in Section VII.

Notations—Throughout, upper-case bold letters denote matrices, while lower-case bold letters denote vectors. In addition, $(\cdot)^*$ denotes the conjugate transpose operation, $\mathbb{E}\{\cdot\}$ returns the average of an input random entity, $\mathbb{P}(\cdot)$ gives the probability of an event, $(\cdot)_k$ (respectively $[\cdot]_k$) returns the kth row (respectively column) of an input matrix, $\|\cdot\|$ represents the maximum square-norm of its rows, $\|\cdot\|_2$ computes the square-norm and $\bigotimes$ represents the convolution operation. Moreover, $\Gamma(a)=\int_0^{+\infty}x^{a-1}e^{-x}dx$ is the Gamma function, $\mathbf{I}_n$ is the identity matrix of size $n\times n$, $\mathcal{CN}(\mu,\sigma^2)$ represents a circularly symmetric complex Gaussian random variable with mean μ and variance σ^2 and a realization of a random variable X is denoted by a corresponding lower case letter x.

II. IA with Imperfect Crosstalk CSI

In this paper, we study the K-user multiple-input multiple-output (MIMO) interference channel where each user k consists of a transmitter equipped with n_k antennas communicating with a receiver equipped with m_k antennas. Without loss of generality, we focus on the kth receiver which receives its intended signal sent from the kth transmitter but that signal will be corrupted by signals sent by other transmitters.

Hence, at a given time instant, the signal received at the kth user is given by

$$\mathbf{y}_k=\mathbf{H}_{k,k}\mathbf{x}_k+\sum_{\substack{\ell=1\\ \ell\neq k}}^{K}\mathbf{H}_{k,\ell}\mathbf{x}_\ell+\boldsymbol{\eta}_k, \tag{1}$$

where $\boldsymbol{\eta}_\ell$ denotes the additive white Gaussian noise (AWGN) vector with elements distributed as $\mathcal{CN}(0,\sigma_\eta^2)$, $\mathbf{H}_{k,\ell}$ denotes the deterministic MIMO channel between the ℓth transmitter and the kth receiver and $\mathbf{x}_\ell$ is the message sent by the ℓth transmitter with the power constraint $\mathbb{E}\{\|\mathbf{x}_\ell\|_2^2\}=P_0\ \forall\ell$. The sum appearing in (1) above represents the interference created by all the users except transmitter k to receiver k.

Typically, in order to remove the effects of interference, one could use multiple-access techniques such as TDMA, FDMA, etc [22] to orthogonalize users across time and/or frequency but this will lead to suboptimal use of the available bandwidth. On the other hand, IA has been devised to cope with the effect of this interference term and is well known to achieve the DoF capacity of the interference channel [4]. The working principle of IA is to align all the interference received at a given receiver in a restricted vector space which is made orthogonal to the space for the desired signal. By doing so, the desired signal can be easily extracted from the signal at each receiver.

The design of these vector spaces is achieved by designing the precoders $\mathbf{V}_k$ at each transmitter (say k) and the interference cancelling matrices $\mathbf{W}_k$ at each receiver such that

$$\begin{cases}\text{rank}(\mathbf{W}_k^*\mathbf{H}_{k,k}\mathbf{V}_k)=d_k, & \text{for } k=1,2,\dots,K,\\ \mathbf{W}_\ell^*\mathbf{H}_{\ell,k}\mathbf{V}_k=0, & \text{for all } \ell\neq k,\end{cases} \tag{2}$$

where d_k represents the number of information streams allocated to the kth user. As such, after applying the interference cancelling matrix at the kth receiver, we have

$$\mathbf{W}_k^*\mathbf{y}_k=\mathbf{W}_k^*\mathbf{H}_{k,k}\mathbf{V}_k\mathbf{x}_k+\mathbf{W}_k^*\boldsymbol{\eta}_k. \tag{3}$$

As we can see, the interference term has been neutralized, but this comes at the price of a reduced dimensional space for the desired signal which is now d_k for the kth user instead of $\min(n_k,m_k)$ in the case of a point-to-point MIMO system.

In this paper, our focus is on evaluating the performance of IA in the presence of CSI errors. Hence, we will assume that IA is always feasible with the given parameters. For feasibility conditions for IA, readers are referred to [14]–[16].

In practice, IA will operate under imperfect knowledge of the channel state, since CSI is estimated and will change over time [19]. Although the main channel CSI can be accurately estimated, the estimation of crosstalk CSI involves other user receivers, and will be less accurate and updated less frequently. Also, the bound in [21] demonstrates that the uncertainty in the main channel CSI tends to have less effects on the capacity performance than that in the crosstalk CSI. For this reason, in this paper, we consider the scenario where the main channel CSI is perfect but the crosstalk CSI is in errors.

We model the error on the channel knowledge as an additive term to the channel measurement, i.e.,

$$\mathbf{H}_{k,\ell}=\hat{\mathbf{H}}_{k,\ell}+\Delta\mathbf{H}_{k,\ell}, \tag{4}$$

where $\hat{\mathbf{H}}_{k,\ell}$ represents the measurement or the estimate of the channel matrix and $\Delta\mathbf{H}_{k,\ell}$ is the difference between the real channel and the channel estimate and will be referred to as the measurement error. We will assume that each entry of $\Delta\mathbf{H}_{k,\ell}$ is complex Gaussian distributed as $\mathcal{CN}(0,\sigma_e^2)$.

If we take into account the uncertainty on the CSI, then (1) can be rewritten as

$$\mathbf{y}_k=\mathbf{H}_{k,k}\mathbf{x}_k+\sum_{\substack{\ell=1\\ \ell\neq k}}^{K}\hat{\mathbf{H}}_{k,\ell}\mathbf{x}_\ell+\sum_{\substack{\ell=1\\ \ell\neq k}}^{K}\Delta\mathbf{H}_{k,\ell}\mathbf{x}_\ell+\boldsymbol{\eta}_k, \tag{5}$$

where we have separated the interference due to the channel estimates and those due to the measurement errors.

Using the channel estimates to design the IA precoders, the conditions on the precoders can then be reexpressed as

$$\begin{cases}\text{rank}(\mathbf{W}_k^*\mathbf{H}_{k,k}\mathbf{V}_k) = d_k, & \text{for } k = 1, 2, \ldots, K, \\ \mathbf{W}_\ell^*\hat{\mathbf{H}}_{\ell,k}\mathbf{V}_k = 0, & \text{for all } \ell \neq k.\end{cases} \tag{6}$$

Note that the second condition is now on the channel estimates instead of the real channels because of the measurement errors. The first condition remains the same as the main channel CSI is assumed to be perfect. We consider that $\mathbf{V}_k^*\mathbf{V}_k = \mathbf{I}_{d_k}\ \forall k$, and that $\mathbf{V}_k$ is given by an IA solution that does not consider the direct links $\mathbf{H}_{k,k}$ [3], [23], [24].

With these new IA conditions reflecting the practical scenarios, (3) becomes

$$\mathbf{W}_k^*\mathbf{y}_k = \mathbf{W}_k^*\mathbf{H}_{k,k}\mathbf{V}_k\mathbf{x}_k + \sum_{\substack{\ell=1 \\ \ell\neq k}}^{K} \mathbf{W}_k^*\Delta\mathbf{H}_{k,\ell}\mathbf{V}_\ell\mathbf{x}_\ell + \mathbf{W}_k^*\boldsymbol{\eta}_k. \tag{7}$$

From the above, we identify the term $\sum_{\substack{\ell=1 \\ \ell\neq k}}^{K} \mathbf{W}_k^*\Delta\mathbf{H}_{k,\ell}\mathbf{V}_\ell\mathbf{x}_\ell$ as being the interference from the other users to user k due to the imperfect knowledge of the channel.

In the next section, we focus on the effects of this term on the maximum rate achievable per stream of each user.

III. Probability Distribution of the Rates

In this section, we provide the statistical description of the achievable rate per stream in relation to the distribution of the CSI error. Thus, we focus on the interference created by the CSI error on the received signal. We will investigate this interference term prior to applying the interference cancelling matrix and we denote this term at the kth receiver by

$$\boldsymbol{i}_k = \sum_{\substack{\ell=1 \\ \ell\neq k}}^{K} \Delta\mathbf{H}_{k,\ell}\mathbf{V}_\ell\mathbf{x}_\ell. \tag{8}$$

Let us consider only the jth component of $\boldsymbol{i}_k$ given by

$$(\boldsymbol{i}_k)_j = \sum_{\substack{\ell=1 \\ \ell\neq k}}^{K} (\Delta\mathbf{H}_{k,\ell})_j\mathbf{V}_\ell\mathbf{x}_\ell, \tag{9}$$

where $(\cdot)_j$ returns the jth row of the input matrix.

We now adopt (9) to compute the instantaneous interference power contained in the jth component of $\boldsymbol{i}_k$ as

$$|(\boldsymbol{i}_k)_j|^2 = \left(\sum_{\substack{\ell=1 \\ \ell\neq k}}^{K} (\Delta\mathbf{H}_{k,\ell})_j\mathbf{V}_\ell\mathbf{x}_\ell\right)\left(\sum_{\substack{l=1 \\ l\neq k}}^{K} \mathbf{x}_l^*\mathbf{V}_l^*(\Delta\mathbf{H}_{k,l})_j^*\right) \tag{10}$$

$$= \sum_{\substack{\ell=1 \\ \ell\neq k}}^{K}\sum_{\substack{l=1 \\ l\neq k}}^{K} (\Delta\mathbf{H}_{k,\ell})_j\mathbf{V}_\ell\mathbf{x}_\ell\mathbf{x}_l^*\mathbf{V}_l^*(\Delta\mathbf{H}_{k,l})_j^*. \tag{11}$$

We now take the ensemble average of $|(\boldsymbol{i}_k)_j|^2$ over all possible transmit messages $(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_K)$ under the assumption that they can be treated as uncorrelated sources of multivariate Gaussian random variables with mean 0 and covariance matrix $P_0\mathbf{I}_{d_k}$, where d_k denotes the number of streams of the kth user. In other words, the transmit messages are independent and all the users have the same power constraint. We call that average value $\mathcal{I}_{k,j}$, which can be evaluated as

$$\begin{aligned}\mathcal{I}_{k,j} &= \mathbb{E}_{\mathbf{x}_1,\ldots,\mathbf{x}_k}\left[|(\boldsymbol{i}_k)_j|^2\right] \\ &= \mathbb{E}_{\mathbf{x}_1\ldots,\mathbf{x}_k}\left[\sum_{\substack{\ell=1 \\ \ell\neq k}}^{K}\sum_{\substack{l=1 \\ l\neq k}}^{K} (\Delta\mathbf{H}_{k,\ell})_j\mathbf{V}_\ell\mathbf{x}_\ell\mathbf{x}_l^*\mathbf{V}_l^*(\Delta\mathbf{H}_{k,l})_j^*\right],\end{aligned} \tag{12}$$

where the expectation is conditioned on $\Delta\mathbf{H}_{k,\ell}\ \forall(k,\ell)$.

The expression above can be further found as

$$\mathcal{I}_{k,j} = P_0\sum_{\substack{\ell=1 \\ \ell\neq k}}^{K} (\Delta\mathbf{H}_{k,\ell})_j\mathbf{V}_\ell\mathbf{V}_\ell^*(\Delta\mathbf{H}_{k,\ell})_j^*. \tag{13}$$

Our interest is to investigate the distribution of $\mathcal{I}_{k,j}$ but since the summation in the expression (13) contains lots of similar terms, we will focus on only one element, and for simplicity of notation we drop all the subscripts and we replace any term such as $(\Delta\mathbf{H}_{k,\ell})_j$ by $\mathbf{h}$. Furthermore, we define

$$\delta \triangleq \mathbf{h}\mathbf{V}\mathbf{V}^*\mathbf{h}^*. \tag{14}$$

For this, we denote the number of streams as d, the number of transmit antennas n and the number of receive antennas m.

The matrix $\mathbf{V}\mathbf{V}^*$ is hermitian by definition. Therefore, we can use the spectral theorem and decompose it into

$$\mathbf{V}\mathbf{V}^* = \mathbf{U}\mathbf{D}\mathbf{U}^*, \tag{15}$$

where $\mathbf{U}$ is a unitary matrix and $\mathbf{D}$ is a diagonal matrix with real entries. Using the spectral decomposition of $\mathbf{V}\mathbf{V}^*$, we can then rewrite (14) as

$$\delta = \mathbf{h}\mathbf{U}\mathbf{D}\mathbf{U}^*\mathbf{h}^*. \tag{16}$$

Now, if we define

$$\tilde{\mathbf{h}} \triangleq \mathbf{U}^*\mathbf{h}^*, \tag{17}$$

then $\tilde{\mathbf{h}}$ has a multivariate complex Gaussian distribution with covariance matrix $\sigma_e^2\mathbf{I}_n$, where σ_e^2 is the variance of each entry of $\mathbf{h}$. Recalling from (16), we now have

$$\delta = \tilde{\mathbf{h}}^*\mathbf{D}\tilde{\mathbf{h}}. \tag{18}$$

Obviously, δ is random because of the random CSI uncertainties and if we consider this as a random variable Δ, then the structure in (18) illustrates that Δ is drawn from a generalized Chi-square distribution [25]. In the following, our aim is to determine precisely the parameters of that distribution.

A. DoF of the χ^2 Distribution

Let us focus on the matrix $\mathbf{D}$ as this matrix determines the parameters of the distribution we are looking for.

Theorem 1: The matrix $\mathbf{D}$ is diagonal with exactly d ones and $n-d$ zeros on its diagonal, where d is the number of transmit streams and n is the number of transmit antennas.

Proof: $\mathbf{V}\mathbf{V}^*$ and $\mathbf{V}^*\mathbf{V}$ have the same non-zero eigenvalues and since $\mathbf{V}^*\mathbf{V} = I_d$ we deduce that $\mathbf{V}\mathbf{V}^*$ has d non-zero eigenvalues all equal to 1, which shows the desired result and completes the proof. ∎

Corollary 1: The random variable $\frac{2}{\sigma_e^2}\Delta$ is Chi-square distributed with $2d$ DoFs, i.e., $\frac{2}{\sigma_e^2}\Delta \sim \chi^2_{2d}$.

Proof: From (18), we can rewrite δ as

$$\delta = \sigma_e^2 \left(\frac{\tilde{\mathbf{h}}^*}{\sigma_e}\right) \mathbf{D} \left(\frac{\tilde{\mathbf{h}}}{\sigma_e}\right), \tag{19}$$

so that $\frac{\tilde{\mathbf{h}}^*}{\sigma_e}$ has unit variance. Now, using (19) and Theorem 1 gives the desired result. Note that the factor 2 in $\frac{2}{\sigma_e^2}\Delta$ comes from the fact that we are using complex valued numbers. ∎

Corollary 2: The probability density function (pdf) of Δ is given by

$$f_\Delta(\delta, d) = \begin{cases} \frac{1}{\sigma_e^2 \Gamma(d)} \left(\frac{\delta}{\sigma_e^2}\right)^{d-1} e^{-\frac{\delta}{\sigma_e^2}} & \text{for } \delta \geq 0, \\ 0 & \text{for } \delta < 0. \end{cases} \tag{20}$$

Proof: We recall that the pdf of a random variable X following a χ_k^2 distribution is given by

$$f_X(x, k) = \begin{cases} \frac{1}{2^{\frac{k}{2}}\Gamma\left(\frac{k}{2}\right)} x^{\frac{k}{2}-1} e^{-\frac{x}{2}} & \text{for } x \geq 0, \\ 0 & \text{for } x < 0. \end{cases} \tag{21}$$

Since $\Delta = \frac{\sigma_e^2}{2} X$, we can give the pdf of Δ as

$$f_\Delta(\delta, d) = \frac{2}{\sigma_e^2} f_X \left(\frac{2}{\sigma_e^2}\delta, 2d\right). \tag{22}$$

Therefore, we obtain (20) and complete the proof. ∎

B. Probability Distribution of the Achievable Rate per Stream

In this subsection, we will link the probability distribution of the interference to that of the maximum achievable rate for any stream of a given user in the MIMO interference channel using IA. Putting the subscripts back in the notations, δ becomes

$$\delta_{k,\ell}^j = (\Delta\mathbf{H}_{k,\ell})_j \mathbf{V}_\ell \mathbf{V}_\ell^* (\Delta\mathbf{H}_{k,\ell})_j^* \tag{23}$$

and the random variable associated is $\Delta_{k,\ell}^j$.

With this notation, we can rewrite (13) as

$$\mathcal{I}_{k,j} = P_0 \sum_{\substack{\ell=1 \\ \ell \neq k}}^{K} \delta_{k,\ell}^j. \tag{24}$$

Now, we define the following random variable

$$\Delta_k^j \triangleq \sum_{\substack{\ell=1 \\ \ell \neq k}}^{K} \Delta_{k,\ell}^j. \tag{25}$$

Therefore, if we consider $\mathcal{I}_{k,j}$ as a random variable, then

$$\mathcal{I}_{k,j} = P_0 \Delta_k^j. \tag{26}$$

At this stage, it should be reminded that we have not specified a basis at the receiver side. Thus, every result we have is true in any given basis. It is especially true in a basis in which d of the basis vectors are independent unitary vectors from the desired signal space and the remaining basis vectors are any unitary vectors that can complete the set to form a basis.

We define the basis in this manner so that the interference cancelling matrices $\mathbf{W}_k$ can be found to simply zero-force the inter-user and inter-stream interference and that the interference power received along any stream is given by $\mathcal{I}_{k,j}$. We also see that statistically Δ_k^j and therefore $\mathcal{I}_{k,j}$ are the same $\forall j$. Henceforth, we denote them, respectively, Δ_k and $\mathcal{I}_k$.

The achievable rate for the lth stream of the kth user is given by

$$R_{k,l}(P_0, \mathcal{I}_k) = \log_2 \left(1 + \frac{\left(\frac{P_0}{d_k}\right) |(\mathbf{W}_k^*)_l \mathbf{H}_{k,k} [\mathbf{V}_k]_l|^2}{\mathcal{I}_k + \sigma_n^2}\right). \tag{27}$$

From now on, we define $z_{k,l} = |(\mathbf{W}_k^*)_l \mathbf{H}_{k,k}[\mathbf{V}_k]_l|^2$ and write the achievable rate of the lth stream of the kth user as

$$R_{k,l}(P_0, \mathcal{I}_k) = \log_2 \left(1 + \frac{\left(\frac{P_0}{d_k}\right) z_{k,l}}{\mathcal{I}_k + \sigma_n^2}\right). \tag{28}$$

This rate expression will be used in the following form:

$$R_{k,l}(\rho, \Delta_k) = \log_2 \left(1 + \frac{z_{k,l}}{d_k \left(\Delta_k + \frac{1}{\rho}\right)}\right), \tag{29}$$

where $\rho = \frac{P_0}{\sigma_n^2}$ and we have used (26).

In order to obtain the pdf of $R_{k,l}$, we first express Δ_k as a function of $R_{k,l}$ with ρ fixed so that

$$\Delta_k(\rho, R_{k,l}) = \frac{z_{k,l}}{d_k(2^{R_{k,l}} - 1)} - \frac{1}{\rho}. \tag{30}$$

From this, we get

$$\frac{\partial \Delta_k(\rho, R_{k,l})}{\partial R_{k,l}} = -\frac{z_{k,l}(\log_e 2) 2^{R_{k,l}}}{d_k(2^{R_{k,l}} - 1)^2}. \tag{31}$$

We can therefore write the pdf, $f_{R_{k,l}}$, of the rate as

$$f_{R_{k,l}}(\rho, r_{k,l}) = \frac{z_{k,l}(\log_e 2) 2^{r_{k,l}}}{d_k(2^{r_{k,l}} - 1)^2} \times f_{\Delta_k} \left(\frac{z_{k,l}}{d_k(2^{r_{k,l}} - 1)} - \frac{1}{\rho}\right). \tag{32}$$

Based on this, we can characterize the pdf of $R_{k,l}$ given that of Δ_k. To do so, we know that $\Delta_k = \sum_{\substack{\ell=1 \\ \ell \neq k}}^{K} \Delta_{k,\ell}$ (we omit the superscripts since the expression is the same for all streams of the same user) and therefore, since $\frac{2}{\sigma_e^2}\Delta_{k,\ell}\ \forall(k,\ell)$ are independent and $\chi^2(2d_\ell)$ distributed, we have that $\frac{2}{\sigma_e^2}\Delta_k$ is χ^2 distributed with $2(D-d_k)$ DoFs where $D = \sum_{\ell=1}^{K} d_\ell$ denotes the total number of streams in the network.

Therefore, we have

$$f_{\Delta_k}(\delta) = \begin{cases} \frac{1}{\sigma_e^2 \Gamma(D-d_k)} \left(\frac{\delta}{\sigma_e^2}\right)^{D-d_k-1} e^{-\frac{\delta}{\sigma_e^2}} & \text{for } \delta \geq 0, \\ 0 & \text{for } \delta < 0. \end{cases} \quad (33)$$

If every user has the same number of streams d, then this expression can be written as

$$f_{\Delta_k}(\delta) = \begin{cases} \frac{1}{\sigma_e^2 \Gamma(d(K-1))} \left(\frac{\delta}{\sigma_e^2}\right)^{d(K-1)-1} e^{-\frac{\delta}{\sigma_e^2}} & \text{for } \delta \geq 0, \\ 0 & \text{for } \delta < 0. \end{cases} \quad (34)$$

IV. SATURATING SNR AND OUTAGE PROBABILITY

In this section, we study two performance metrics, namely, *saturating SNR* and outage probability. Saturating SNR was first introduced in [21] for worst-case scenarios (with bounded CSI errors). Here, we emphasize on its statistics. On the other hand, outage probability is a metric that can be used to assess the performance of each user despite the randomness.

A. *Pdf of Saturating SNR*

The saturating SNR can be seen as a non-asymptotic performance metric that accounts for estimation errors. It represents the SNR at which increasing the transmit power will not bring any gain in terms of achievable rate [21]. More precisely, it is defined as the SNR where the rate in the perfect CSI case is equal to that of the imperfect CSI case at infinite SNR. Here the saturating SNR is defined for each stream of any given user.

We will first give the pdf of the saturating SNR. Then we will show that up to the saturating SNR, the achievable rate of the IA with corrupted CSI is within 1bps/Hz of the achievable rate of the IA in the perfect CSI case.

From equation (29) we obtain that the rate at $\rho = \infty$ is given by

$$R_{k,l}(\infty, \Delta_k) = \log_2\left(1 + \frac{z_{k,l}}{d_k \Delta_k}\right). \quad (35)$$

By definition, at the saturating SNR, $\rho_s^{k,l}$, we have

$$\log_2\left(1 + \frac{\rho_s^{k,l} z_{k,l}}{d_k}\right) = R_{k,l}\left(\infty, \delta_k^{(l)}\right) \quad (36)$$

$$= \log_2\left(1 + \frac{z_{k,l}}{d_k \delta_k^{(l)}}\right), \quad (37)$$

which gives

$$\rho_s^{k,l} = \frac{1}{\delta_k^{(l)}}. \quad (38)$$

The superscript (l) in $\delta_k^{(l)}$ is there to remind that $\delta_k^{(l)}$ is one realization of Δ_k for the lth stream of the kth user.

As a result, the saturating SNR is a random variable $\wp_s^k = \frac{1}{\Delta_k}$ (note that we drop the superscript l because it is the same distribution for all the streams of the kth user) and the pdf of the saturating SNR can be derived as

$$f_{\wp_s^k}\left(\rho_s^k\right) = \frac{1}{\left(\rho_s^k\right)^2} f_{\Delta_k}\left(\frac{1}{\rho_s^k}\right). \quad (39)$$

Theorem 2: Given the saturating SNR $\rho_s^{k,l}$, we can approximate the achievable rate within 1bps/Hz by

$$\tilde{R}_{k,l}(\rho) = \begin{cases} \log_2\left(1 + \frac{\rho z_{k,l}}{d_k}\right) & \text{for } 0 \leq \rho \leq \rho_s^{k,l}, \\ \log_2\left(1 + \frac{\rho_s^{k,l} z_{k,l}}{d_k}\right) & \text{for } \rho_s^{k,l} \leq \rho, \end{cases} \quad (40)$$

where $\log_2(1 + \frac{\rho z_{k,l}}{d_k})$ is the rate in the perfect CSI case.

Proof: Define the function $G(\rho) \triangleq \tilde{R}_{k,l}(\rho) - R_{k,l}(\rho, \delta_k^{(l)})$ that represents the gap between $R_{k,l}$ and $\tilde{R}_{k,l}$. For $0 \leq \rho \leq \rho_s^{k,l}$, we have

$$G(\rho) = \log_2\left(\frac{1 + \frac{\rho}{d_k} z_{k,l}}{1 + \frac{\rho}{d_k(\delta_k^{(l)}\rho + 1)} z_{k,l}}\right). \quad (41)$$

One can notice that $G(0) = 0$ and G is an increasing function of ρ. We can now evaluate G at the saturating SNR, i.e., $\rho_s^{k,l} = \frac{1}{\delta_k^{(l)}}$. Then we have

$$G(\rho_s) = G\left(\frac{1}{\delta_k^{(l)}}\right) \quad (42)$$

$$= \log_2\left(\frac{1 + \frac{\frac{1}{\delta_k^{(l)}}}{d_k} z_{k,l}}{1 + \frac{\frac{1}{\delta_k^{(l)}}}{2d_k} z_{k,l}}\right) \quad (43)$$

$$= \log_2\left(\frac{\delta_k^{(l)} + \frac{z_{k,l}}{d_k}}{\delta_k^{(l)} + \frac{z_{k,l}}{2d_k}}\right). \quad (44)$$

The expression $\frac{\delta_k^{(l)} + \frac{z_{k,l}}{d_k}}{\delta_k^{(l)} + \frac{z_{k,l}}{2d_k}}$ is a decreasing function of $\delta_k^{(l)}$ that goes from 2 to 1 therefore $G(\rho_s^{k,l}) \in [0, 1]$ and finally $\forall \rho \in [0, \rho_s^{k,l}]$ $G(\rho) \in [0, 1]$.

On the other hand, for $\rho > \rho_s^{k,l}$, by definition of the saturating SNR, $R_{k,l}(\rho, \delta_k^{(l)}) \to \tilde{R}_{k,l}(\rho)$. This concludes that $\tilde{R}_{k,l}(\rho)$ is an approximation of $R_{k,l}(\rho, \delta_k^{(l)})$ within 1bps/Hz $\forall \rho$. ■

B. *Outage Probability*

In the case of fading channels, one often uses outage capacity as a metric to assess the performance of the communication system. Outage capacity is linked to a parameter called "outage probability", $\mathcal{P}_{\mathsf{out}}$, which represents the probability that *error-free* communications cannot be achieved at a given rate. This can be translated to a minimum SNR $\rho_{\min}$ below which the information rate is not supported, or $\mathcal{P}_{\mathsf{out}} = \mathbb{P}(\rho < \rho_{\min})$.

Define the outage probability of the ℓth stream of a user k as the probability that that stream cannot support any rate equal or above $C_{\mathsf{out}}^{k,l}$ at infinite SNR, i.e., $\mathcal{P}_{\mathsf{out}}^{k,l} = \mathbb{P}(R_{k,l}^{\infty} < C_{\mathsf{out}}^{k,l})$, with $R_{k,l}^{\infty}(\Delta_k) \triangleq R_{k,l}(\Delta_k, \infty)$, and $C_{\mathsf{out}}^{k,l}$ being the outage capacity for the ℓth stream of the kth user. Then

$$\mathbb{P}(R_{k,l}^{\infty} < C_{\mathsf{out}}^{k,l}) = \frac{z_{k,l}(\log_e 2)}{d_k} \times \int_0^{C_{\mathsf{out}}^{k,l}} \frac{2^r}{(2^r-1)^2} f_{\Delta_k}\left(\frac{z_{k,l}}{d_k(2^r-1)}, d_k\right) dr, \tag{45}$$

which can be further expressed as

$$\mathbb{P}(R_{k,l}^{\infty} < C_{\mathsf{out}}^{k,l}) = \int_0^{\frac{z_{k,l}}{d_k(2^{C_{\mathsf{out}}^{k,l}}-1)}} f_{\Delta_k}(x, d_k) dx. \tag{46}$$

If we define the outage SNR $\rho_{\mathsf{out}}^{k,l}$ so that $C_{\mathsf{out}}^{k,l} = \log_2(1 + \frac{\rho_{\mathsf{out}}^{k,l} z_{k,l}}{d_k})$, then

$$\mathcal{P}_{\mathsf{out}}^{k,l} = \int_0^{\frac{1}{\rho_{\mathsf{out}}^{k,l}}} f_{\Delta_k}(x, d_k) dx = \mathbb{P}\left(\rho_s^{k,l} < \rho_{\mathsf{out}}^{k,l}\right). \tag{47}$$

This indicates that the outage probability equals the probability that the saturating SNR is lower than the outage SNR.

We can use (33) and (47) to give the outage probability as

$$\mathcal{P}_{\mathsf{out}}^{k,l} = \frac{1}{\Gamma(D-d_k)} \gamma\left(D-d_k, \frac{1}{\rho_{\mathsf{out}}^{k,l}\sigma_e^2}\right) \tag{48}$$

$$= \tilde{\Gamma}\left(D-d_k, \frac{1}{\rho_{\mathsf{out}}^{k,l}\sigma_e^2}\right), \tag{49}$$

where

$$\tilde{\Gamma}(a, x) = \frac{\gamma(a, x)}{\Gamma(a)} \tag{50}$$

is the regularized lower incomplete Gamma function.

Note that the relation between $\mathcal{P}_{\mathsf{out}}^{k,l}$ and $\rho_{\mathsf{out}}^{k,l}$ does not involve $z_{k,l}$ but does involve d_k. Hence, that relation is the same for all the streams of the same user in accordance with the fact that all those streams transmit the same power. We can therefore define the outage probability (respectively outage SNR) of the user (say k) as $\mathcal{P}_{\mathsf{out}}^{k} = \mathcal{P}_{\mathsf{out}}^{k,l}$ (respectively $\rho_{\mathsf{out}}^{k} = \rho_{\mathsf{out}}^{k,l}$).

In (49), the effect of the other users on the performance of user k manifests itself through the total number of streams D in the network. The outage probability will decrease if the contribution of user k in the total number of streams is high, because $D - d_k$ in this case is smaller and also because there is no interference between the streams of the same user.

V. Applications

In this section, we give two application examples for utilizing our analytical results in the MIMO interference channel.

A. Degrading CSI in Block Fading Channels

For block fading channels, the channels are often considered as constant for a period of time, say T_D, but vary from one block to another. Moreover, the transmit power is normalized such that $\frac{1}{T_D}\mathbb{E}\{\|\mathbf{x}_\ell\|_2^2\} = P_0 \; \forall \ell$. A typical scenario is that the channels are estimated at the first block only and all the IA matrices are obtained from the estimated channels. The direct link channels vary from one block to the next but are tracked perfectly while the crosstalk channels are not tracked (due to high overheads), so the crosstalk CSI degrades over time.

Let us denote $\mathbf{H}(t_0)$ as the actual state of a channel matrix at time t_0. The estimated channel is denoted by $\hat{\mathbf{H}}(t_0)$, which is modelled as

$$\mathbf{H}(t_0) = \hat{\mathbf{H}}(t_0) + \Delta\mathbf{H}(t_0), \tag{51}$$

where $\Delta\mathbf{H}(t_0)$ is the deviation of the estimate from the actual channel at time t_0. For the block fading channels, we have

$$\mathbf{H}(t_0 + k \times T_D) \neq \mathbf{H}(t_0), \text{ for } k = 1, 2, \ldots, \tag{52}$$

with probability one. We define a new matrix $\mathbf{E}$ that represents the variation of the channel between two different times as

$$\mathbf{E}(k) \triangleq \mathbf{H}(t_0 + k \times T_D) - \mathbf{H}(t_0) \tag{53}$$

$$= \mathbf{H}(t_0 + k \times T_D) - \hat{\mathbf{H}}(t_0) - \Delta\mathbf{H}(t_0). \tag{54}$$

Hence,

$$\underbrace{\mathbf{E}(k) + \Delta\mathbf{H}(t_0)}_{\triangleq \Delta\mathbf{H}(k)} = \mathbf{H}(t_0 + k \times T_D) - \hat{\mathbf{H}}(t_0) \tag{55}$$

represents the deviation from the channel estimate to the actual channel at the $(k+1)$th block for the crosstalk.

We assume the model that $\Delta\mathbf{H}(k) \sim \mathcal{CN}(0, \sigma_e^2(k))$ with $\sigma_e^2(0) = \sigma_e^2$ being the power of the measurement. Based on this model, if we consider one stream of a user transmitting d_0 streams then we can express the evolution of the outage probability in each block provided that we transmit at the same outage SNR as

$$\mathcal{P}_{\mathsf{out}}(k) = \tilde{\Gamma}\left(D - d_0, \frac{1}{\rho_{\mathsf{out}}\sigma_e^2(k)}\right). \tag{56}$$

Instead of transmitting at a fixed SNR we may want to insure a preset outage probability over each block in which case the SNR at which we should transmit over that stream is given by

$$\rho_{\mathsf{out}}(k) = \frac{1}{\sigma_e^2(k)\tilde{\Gamma}^{-1}(D-d_0, \mathcal{P}_{\mathsf{out}})} \tag{57}$$

$$= \frac{\sigma_e^2}{\sigma_e^2(k)}\rho_{\mathsf{out}}, \tag{58}$$

where

$$\rho_{\mathsf{out}} \triangleq \frac{1}{\sigma_e^2\tilde{\Gamma}^{-1}(D-d_0, \mathcal{P}_{\mathsf{out}})}, \tag{59}$$

and $\tilde{\Gamma}^{-1}$ is the inverse of function $\tilde{\Gamma}$.

Employing the result of Theorem 2, the maximum achievable rate per stream for the MIMO interference channel using IA under the block fading channel at outage probability of $\mathcal{P}_{\text{out}}$ can be given within 1 bps/Hz by

$$\tilde{R}(k) = \log_2\left(1 + \frac{\rho_{\text{out}}(k)}{d_0} z(k)\right) \tag{60}$$

$$= \log_2\left(1 + \frac{\sigma_e^2}{\sigma_e^2(k)} \frac{\rho_{\text{out}}}{d_0} z(k)\right), \tag{61}$$

where $z(k)$ is the effect that the channel has on that stream over the kth block. The achievable rate for that stream over B consecutive blocks can therefore be found as

$$R_B = \sum_{k=0}^{B-1} \log_2\left(1 + \frac{\sigma_e^2}{\sigma_e^2(k)} \frac{\rho_{\text{out}}}{d_0} z(k)\right). \tag{62}$$

Under the conditions $\frac{\sigma_e^2}{\sigma_e^2(k)} \frac{\rho_{\text{out}}}{d_0} z(k) \gg 1 \; \forall k$, e.g., if the power of the channel variations over the measurement noise power is very small or if the outage SNR is sufficiently big or also if there is no overly deep fading over any block, then we have

$$R_B \approx \sum_{k=0}^{B-1} \log_2 \frac{\sigma_e^2}{\sigma_e^2(k)} \frac{\rho_{\text{out}}}{d_0} z(k) \tag{63}$$

$$= \underbrace{B \log_2 \frac{\rho_{\text{out}}}{d_0}}_{(a)} + \underbrace{\log_2 \prod_{k=0}^{B-1} z(k)}_{(b)} - \underbrace{\log_2 \prod_{k=0}^{B-1} \frac{\sigma_e^2(k)}{\sigma_e^2}}_{(c)}. \tag{64}$$

In the equation above,

(a) represents the rate achievable per stream with perfect IA and no fading at SNR $= \rho_{\text{out}}$ over B blocks.

(b) represents the effect that the channel variations of the direct link has on the rate; it could be positive or negative depending on the fading coefficients.

(c) represents the effect of the crosstalk CSI uncertainty on the rate; this effect is negative because there is at least the uncertainty from the measurement.

Let us have a look at the following example.

Recall from the definition of $\mathbf{E}(k)$ in (53), if we say $\mathbf{E}(k) \sim \mathcal{CN}(0, \sigma^2) \; \forall k$, then we can get

$$\sigma_e^2(k) = \begin{cases} \sigma_e^2 + \sigma^2 & \text{for } k > 0, \\ \sigma_e^2 & \text{for } k = 0, \end{cases} \tag{65}$$

and

$$R_B = B \log_2 \frac{\rho_{\text{out}}}{d_0} + \log_2 \prod_{k=0}^{B-1} z(k) - \log_2 \left(1 + \frac{\sigma^2}{\sigma_e^2}\right)^{B-1}, \quad \text{for } B \geq 1 \tag{66}$$

In the above, we see that if the power of the channel variation is smaller than the measurement noise power (i.e., $\frac{\sigma^2}{\sigma_e^2} \ll 1$), the channel variations have little effect on the IA performance.

B. Optimizing the Number of Streams

In an IA system with K users and imperfect crosstalk CSI, unsurprisingly, every user would want to increase the number of signal streams for enhancing its achievable rate but doing so may harm the sum-rate because of the additional interference due to imperfect IA resulting from imperfect CSI. If all users are assumed to have the same number of streams d, it would be important to determine the optimal number of streams per user of the interference network for maximizing the sum-rate, for a given measurement noise power σ_e^2.

To do so, we first set an outage probability of $\mathcal{P}_{\text{out}}$ that must remain the same for every user. Then we have the required outage SNR for all users as

$$\rho_{\text{out}} = \frac{1}{\sigma_e^2 \tilde{\Gamma}^{-1}(d(K-1), \mathcal{P}_{\text{out}})}. \tag{67}$$

We define the outage SNR per stream as $\bar{\rho}_{\text{out}} = \frac{\rho_{\text{out}}}{d}$.

For most applications, the outage probability is chosen to be a very small value, and thus the probability of the saturating SNR being lower than the outage SNR is equally small (see (47)). In that case, we can use Theorem 2 and the function $\tilde{R}$ defined in Section IV-A to approximate the rate per stream at SNR $= \rho_{\text{out}}$ with a confidence given by the choice of $\mathcal{P}_{\text{out}}$ as

$$\tilde{R}_{k,l}(\bar{\rho}_{\text{out}}) = \log_2\left(1 + \bar{\rho}_{\text{out}} z_{k,l}\right), \tag{68}$$

where $z_{k,l}$ has the same meaning as in the previous section.

Hence, the sum-rate for the network is found as

$$\bar{R}(\bar{\rho}_{\text{out}}) = \sum_{k=1}^{K} \sum_{l=1}^{d} \tilde{R}_{k,l}(\bar{\rho}_{\text{out}}). \tag{69}$$

The optimal number of streams per user can be found by

$$\max_d \bar{R}(\bar{\rho}_{\text{out}}). \tag{70}$$

We may also want to optimize the number of streams per user in average over all possible realizations of $z_{k,l}$, which means that we consider multiple realizations of IA with different channel gains and average over all possible achievable sum-rates. To do so, we assume that the direct channels $\mathbf{H}_{k,k}$ are independent across users and have their entries independent and identically distributed (i.i.d.) from $\mathcal{CN}(0, 1)$. Also, the crosstalk channels do not matter since their effects are cancelled out by IA. Under that condition we can apply Lemma 1 in [19] and derive the average sum-rate as

$$\mathcal{R}(\bar{\rho}_{\text{out}}) = \sum_{k=1}^{K} \sum_{l=1}^{d} \mathbb{E}\left[\tilde{R}_{k,l}(\bar{\rho}_{\text{out}})\right] \tag{71}$$

$$= K d \log_2(e) e^{\frac{1}{\bar{\rho}_{\text{out}}}} E_1\left(\frac{1}{\bar{\rho}_{\text{out}}}\right), \tag{72}$$

where $E_1(x) = \int_1^{\infty} t^{-1} e^{-xt} dt$ is an exponential integral.

Since we are using the approximate of the rate given by Theorem 2, the expectation does not involve the measurement error matrices. The price to pay for that is however that the result is given within 1bps/Hz precision.

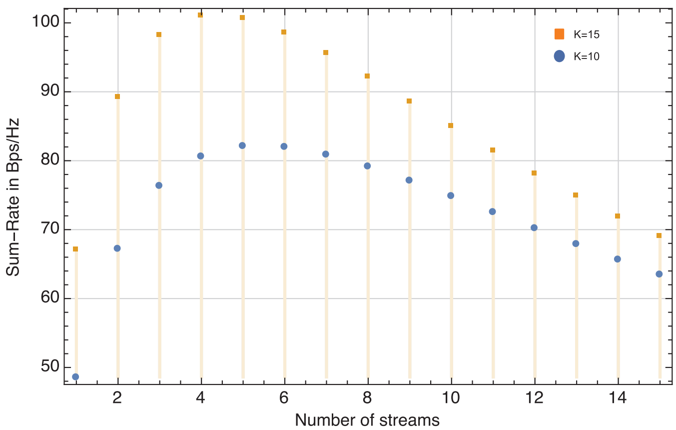


Fig. 1. The sum-rate $\mathcal{R}$ against the number of streams per user, with its maximum achieved when $d = 5$ for 10 users and $d = 4$ for 15 users.

Now the optimal number of streams is found by solving

$$\max_{d} \mathcal{R}(\bar{\rho}_{\text{out}}). \tag{73}$$

This second expression depends only on the number of users, the measurement noise power and the outage probability.

In Figure 1, we provide the numerical results for the sum-rate $\mathcal{R}$ against the number of signal streams, when $K = 10$ and $K = 15$, $\sigma_e^2 = 10^{-3}$ and $\mathcal{P}_{\text{out}} = 10^{-3}$. The results demonstrate the concavity of the sum-rate so the optimal number of streams can be easily identified to be $d = 5$ with $K = 10$ and $d = 4$ with $K = 15$. The results also imply that it is counter-productive to increase the number of streams further due to excessive interference. In addition, we note that there is a significant gain in the total rate to go from one signal stream to the optimal number of streams (up 35 bps/Hz for $K = 10$). This figure also demonstrates that the number of streams allowed per user decreases with the number of user due to CSI uncertainty.

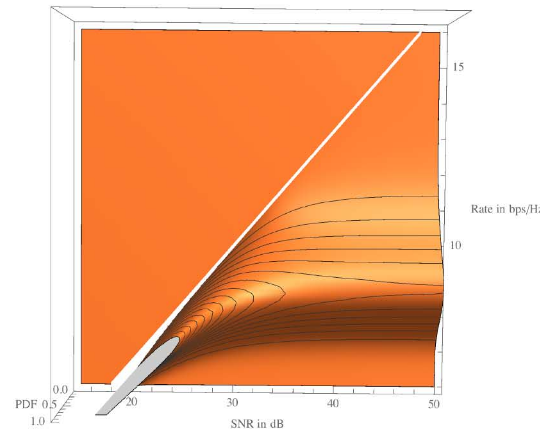
VI. Simulations versus Theory

In this section, we present the predictions of the model and compare these predictions to the results obtained from the simulations. In the simulations, we focus on the 3-user case as one can compute the perfect precoders for IA. The parameters for the simulations are set to $K = 3$ and $d = 1$.

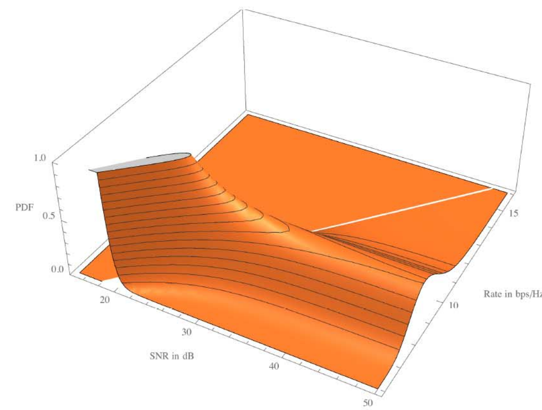
We first adapt (34) to the parameters above which yields

$$f_{\Delta_k}(\delta_k, 1) = \begin{cases} \frac{1}{\sigma_e^4} \delta_k e^{-\frac{\delta_k}{\sigma_e^2}} & \text{for } \delta_k \geq 0, \\ 0 & \text{for } \delta_k < 0. \end{cases} \tag{74}$$

Figure 2 shows the pdf of the achievable rate per stream for the case $\sigma_e^2 = 10^{-3}$ based on the theory. It is a three-dimensional plot with the x-axis showing the SNR in dB, the y-axis the rate in bps/Hz and the z-axis the pdf value. We have decided to illustrate the results from SNR = 15dB because for lower SNR, the pdf is highly localized. There also appears to have a saturation in rate due to CSI imperfection. As discussed in Section IV-A, the saturating SNR is a random variable.



(a) Top view



(b) Tilted view

Fig. 2. The pdf of the achievable rate per stream with $K = 3$, $d = 1$, and $\sigma_e^2 = 10^{-3}$. The white line represents the limiting case where there is no CSI uncertainty.

To compare the theoretical predictions to the simulations, we ran the simulations, in which channel matrices were drawn randomly from $\mathcal{CN}(0, 1)$ which represent the perfect CSI, and the erroneous channel matrices were set to be the sum of the channel matrices and the error matrices drawn randomly from $\mathcal{CN}(0, \sigma_e^2)$. Moreover, all the users were assumed to have $n = 3$ transmit antennas and $m = 2$ receive antennas, and these matrices were used to perform IA at various SNR. We also set the fading coefficients on the direct links ($z_{k,l}$) to 1 so that we only see the effects of the measurement error.

In Figure 3, we provide the results for the rates achievable by IA for a given MIMO channel for 500 independent error realisations. As can be seen, the rates appear to saturate at high SNR as predicted by the theory. In these particular results, the saturating SNR appears to be 15dB. To compare theory and simulations further, we also provide the results for the case at SNR = 80dB, as shown in Figure 4, where the theoretical pdf and the simulations fit almost perfectly.

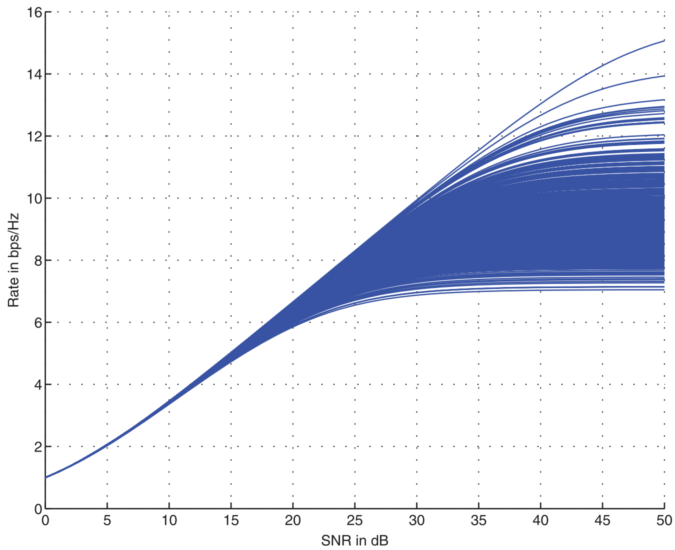


Fig. 3. The achievable rates for IA or a given MIMO interference channel with 500 independent measurement errors.

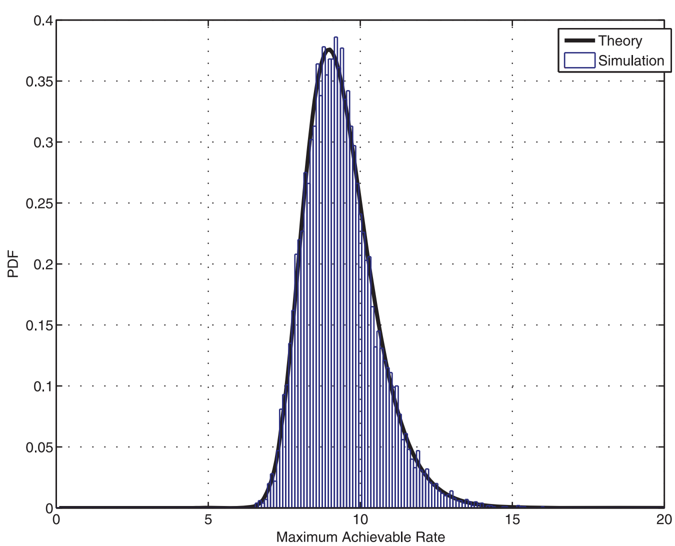


Fig. 4. The pdfs of the achievable rates when $K = 3$, $d = 1$ and $\sigma_e^2 = 10^{-3}$ from the simulations and the theory.

VII. Conclusion

This paper presented a full statistical characterization for the maximum achievable rate per user using IA in the interference channel when CSI are imperfect. We proposed two metrics to evaluate the performance of the interference network despite the randomness of CSI errors. We also applied our analytical results in two application examples. The first one investigated the performance of IA in block fading channels with degrading CSI while the second one used our results in order to optimize the number of streams per user for maximizing the sum-rate using IA in the presence of CSI errors. Simulation results have been provided to confirm the accuracy of the analysis.

References

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

[2] S. Jafar and S. Shamai, "Degrees of freedom region of the MIMO X channel," *IEEE Trans. Inf. Theory*, vol. 54, no. 1, pp. 151–170, Jan. 2008.

[3] V. Cadambe and S. Jafar, "Interference alignment and spatial degrees of freedom for the K user interference channel," in *Proc. IEEE Int. Conf. Commun.*, May 2008, pp. 971–975.

[4] S. Jafar and M. Fakhereddin, "Degrees of freedom for the MIMO interference channel," *IEEE Trans. Inf. Theory*, vol. 53, no. 7, pp. 2637–2642, Jul. 2007.

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

[6] O. El Ayach, S. Peters, and R. W. Heath Jr., "The practical challenges of interference alignment," *IEEE Wireless Commun.,* vol. 20, no. 1, pp. 35–42, Feb. 2013.

[7] G. Bresler, D. Cartwright, and D. Tse, "Feasibility of interference alignment for the mimo interference channel," *IEEE Trans. Inf. Theory*, vol. 60, no. 9, pp. 5573–5586, Sep. 2014.

[8] S. Peters and R. Heath, "Interference alignment via alternating minimization," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.,* Apr. 2009, pp. 2445–2448.

[9] K. Gomadam, V. Cadambe, and S. Jafar, "Approaching the capacity of wireless networks through distributed interference alignment," in *Proc. IEEE Global Telecommun. Conf.,* Nov. 2008, pp. 1–6.

[10] S. Jafar, "Topological interference management through index coding," *IEEE Trans. Inf. Theory*, vol. 60, no. 1, pp. 529–568, Jan. 2014.

[11] S. Jafar, "Blind interference alignment," *IEEE J. Sel. Topics Signal Process.,* vol. 6, no. 3, pp. 216–227, Jun. 2012.

[12] J. H. Lee and W. Choi, "On the achievable DOF and user scaling law of opportunistic interference alignment in 3-transmitter MIMO interference channels," *IEEE Trans. Wireless Commun.,* vol. 12, no. 6, pp. 2743–2753, Jun. 2013.

[13] J. H. Lee and W. Choi, "Interference alignment by opportunistic user selection in 3-user MIMO interference channels," in *Proc. IEEE Int. Conf. Commun.,* Jun. 2011, pp. 1–5.

[14] O. Gonzalez, C. Beltran, and I. Santamaria, "A feasibility test for linear interference alignment in MIMO channels with constant coefficients," *IEEE Trans. Inf. Theory,* vol. 60, no. 3, pp. 1840–1856, Mar. 2014.

[15] G. Bresler, D. Cartwright, and D. Tse, "Feasibility of interference alignment for the MIMO interference channel: The symmetric square case," in *Proc. IEEE Inf. Theory Workshop,* Oct. 2011, pp. 447–451.

[16] C. Yetis, T. Gou, S. Jafar, and A. Kayran, "On feasibility of interference alignment in MIMO interference networks," *IEEE Trans. Signal Process.,* vol. 58, no. 9, pp. 4771–4782, Sep. 2010.

[17] H. Shen, B. Li, Y. Luo, and F. Liu, "A robust interference alignment scheme for the MIMO X channel," in *Proc. 15th Asia-Pac. Conf. Commun.,* Oct. 2009, pp. 241–244.

[18] B. Nosrat-Makouei, J. Andrews, and R. Heath, "MIMO interference alignment over correlated channels with imperfect CSI," *IEEE Trans. Signal Process.,* vol. 59, no. 6, pp. 2783–2794, Jun. 2011.

[19] O. El Ayach, A. Lozano, and R. Heath, "On the overhead of interference alignment: Training, feedback, and cooperation," *IEEE Trans. Wireless Commun.,* vol. 11, no. 11, pp. 4192–4203, Nov. 2012.

[20] R. Tresch and M. Guillaud, "Cellular interference alignment with imperfect channel knowledge," in *Proc. IEEE Int. Conf. Commun. Workshops,* Jun. 2009, pp. 1–5.

[21] R. Guiazon, K.-K. Wong, and D. Wisely, "Capacity analysis of interference alignment with bounded CSI uncertainty," *IEEE Wireless Commun. Lett.,* vol. 3, no. 5, pp. 505–508, Oct. 2014.

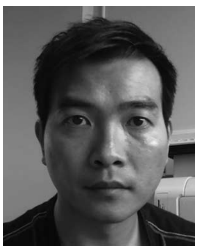
[22] M. Schwartz, Multiple Access Techniques: FDMA, TDMA, CDMA; System Capacity Comparisons. Cambridge, U.K.: Cambridge Univ. Press, 2004, pp. 137–160.

[23] S. Peters and R. Heath, "Cooperative algorithms for MIMO interference channels," *IEEE Trans. Veh. Technol.,* vol. 60, no. 1, pp. 206–218, Jan. 2011.

[24] K. Gomadam, V. Cadambe, and S. Jafar, "A distributed numerical approach to interference alignment and applications to wireless interference networks," *IEEE Trans. Inf. Theory,* vol. 57, no. 6, pp. 3309–3322, Jun. 2011.

[25] M. Koutras, "On the generalized noncentral chi-squared distribution induced by an elliptical gamma law," *Biometrika*, vol. 73, no. 2, pp. 528–532, 1986.

Raoul F. Guiazon received the M.Sc. degree in communications and signal processing from Imperial College London, London, U.K., in 2013, and the Dipôlme d'Ingénieur from the ENSEA, Cergy, France. Currently, he is pursuing the Ph.D. degree in communications and information systems with the Department of Electronic and Electrical Engineering, University College London (UCL), London, U.K. His research interests include interference alignment, cognitive radios, massive MIMO, and information theory.



Kai-Kit Wong (M'01–SM'08) received the B.Eng., M.Phil., and Ph.D. degrees in electrical and electronic engineering from the Hong Kong University of Science and Technology, Clear Water Bay, Hong Kong, in 1996, 1998, and 2001, respectively. He is a Full Professor and Chair in wireless communications with the Department of Electronic and Electrical Engineering, University College London, London, U.K. Prior to this, he took up faculty appointments as a Research Assistant Professor with the University of Hong Kong and Lecturer at the University of Hull, Hull, U.K. He also previously took up Visiting positions at the Smart Antennas Research Group, Stanford University and the Department of Wireless Communications Research, Lucent Technologies, Bell-Labs, Holmdel, NJ, USA. He has been a Senior Editor of the IEEE COMMUNICATIONS LETTERS (since 2012) and is presently also serving on the editorial boards of the IEEE WIRELESS COMMUNICATIONS LETTERS (since 2011), the IEEE COMSOC/KICS JOURNAL OF COMMUNICATIONS AND NETWORKS (since 2010), *IET Communications* (since 2009), and *Physical Communications* (Elsevier) (since 2012). He also previously served as Editor for the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS from 2005 to 2011, Review Editor for the IEEE COMMUNICATIONS LETTERS from 2009 to 2012 and an Associate Editor for the IEEE SIGNAL PROCESSING LETTERS from 2009 to 2012. He is a Fellow of IET.

Michael Fitch received the degree in mathematics and physics and the Ph.D. degree in satellite communications. He is with the Research and Technology Section, BT Innovation and Design, leading a small research team specializing in physical and systems aspects of wireless communications. Since 1989, he has been with the BT, working in various research and development roles and is currently working on a number of collaborative projects on emerging wireless technologies such as LTE and cognitive radio. In addition, he provides engineering consultancy to other parts of BT on wireless matters. His research interests include satellite systems and mobile radio systems. He is a member of the IET.