<http://www.diva-portal.org>

Postprint

This is the accepted version of a paper published in *IEEE Transactions on Communications*. This paper has been peer-reviewed but does not include the final publisher proof-corrections or journal pagination.

Citation for the original published paper (version of record):

Stathakis, E., Rasmussen, L. K., Skoglund, M. (2015)
Closed-Form Capacity Result for Interference-Limited Environments With Mixed Fading.
IEEE Transactions on Communications, 63(6): 2374-2387
<http://dx.doi.org/10.1109/TCOMM.2015.2427351>

Access to the published version may require subscription.

N.B. When citing this work, cite the original published paper.

Permanent link to this version:

<http://urn.kb.se/resolve?urn=urn:nbn:se:kth:diva-170689>

Closed-form Capacity Result For Interference Limited Environments With Mixed Fading

Efthymios Stathakis, *Student Member, IEEE*, Lars K. Rasmussen, *Senior Member, IEEE*, and Mikael Skoglund, *Senior Member, IEEE*

Abstract—We study a multi-node network, where a multi-antenna transmitter T_x communicates with its desired receiver R_x , while a cluster $\mathcal{P} \triangleq \{P_{x,n}, n = 1, \dots, N\}$ of unintended nodes is disturbed by the T_x - R_x (TR) communication. To prevent severe performance degradation, we impose a constraint on the total interference that is inflicted at the nodes of $\mathcal{P}$. The TR link contains a line-of-sight component whereas the propagation environment for each T_x - $P_{x,n}$ link is shadowed. The T_x node is preprocessing the information sequence by means of a precoding matrix that is optimized to achieve the ergodic capacity under a constraint on the maximum admissible ergodic interference power, arriving on $\mathcal{P}$. In this paper, we show that the optimum precoding strategy involves transmission of a single stream over the precoding direction, i.e., the eigenvector of the precoding matrix, which corresponds to beamforming along the instantaneous direction of the TR link channel. Solution of the remaining power allocation problem yields the optimal precoding matrix. For this setup, we provide an efficient stochastic characterization of the network, that allows us to obtain an *analytical* expression for the TR link ergodic-capacity; this problem has been previously open even for the case of a single-antenna node T_x and a single-element set $\mathcal{P}$. We complement the analysis by deriving the TR link signal-to-noise ratio and the average bit-error-rate, which are associated with our transmission scheme. Numerical results corroborate the theoretical analysis and reveal an interplay between the network parameters and their impact on the TR link performance.

I. INTRODUCTION

The *interference channel (IC)* describes the scenario where multiple independent transmitters, sharing the same resources, wish to simultaneously convey information to their associated receivers [1]. These parallel transmissions interfere with each other such that every receiver sees a superposition of the desired signal and the unwanted signals. This situation arises frequently in wireless cellular networks due to frequency reuse and the intrinsically broadcast nature of the wireless medium [2] and cannot be remedied by simply increasing the transmit powers; while this naive approach is a solution against noise perturbations, it will only lead to more interference in IC systems. Thus, the design of interference-aware strategies is an imperative step towards the delivery of information with high

fidelity because the need for ubiquitous coverage of the increasing number of end-users was addressed by denser deployment of base stations, resulting in interference-limited systems.

According to experimental studies [3], based on real data, the signal-to-interference-plus-noise ratio (SINR) in a cell is typically 2dB, and *less than 0dB* for a fraction of the users ($\sim 10\%$). These findings advocate interference-power constrained rather than transmit-power constrained designs [4] because systems will not operate with full power, in interference-limited environments (ILEs) to avoid inflicting excessive interference. An approach that is suitable for ILEs, see Fig. 1, involves the optimization of certain quality-of-service (QoS) metrics, e.g., rate, for a particular link subject to *sum interference constraints* that are imposed to protect a cluster of cross-link receivers [5]. This framework is simple and captures the main shortcoming that is responsible for performance degradation in modern interference networks (INs). It is also fortuitous in that it has wide application in the area of underlay cognitive radio (CR) systems [6], where licensed (primary) users are sharing the spectrum with unlicensed (secondary) users, under constraints that are imposed to guarantee the provision of a minimum QoS level at the primary receiver. Under this paradigm, the ergodic capacity of the Gaussian channel, under an ergodic interference power constraint (EIPC) was found in [7], while more sophisticated schemes were studied in [8], [9].

A common denominator in existing studies is that channel modeling is limited to the case of shadowed environments. However, cellular architectures, especially femtocells [10], admit more generic propagation characterization, i.e., line-of-sight (LoS) channel modeling. The area of INs and underlay CR networks with LoS links, i.e., Rician fading [11], is largely uncharted even for elementary architectures. Therefore, a review of fundamental network structures under a general channel model will be instructive for escalation to more sophisticated designs. In the context of CR systems, few interesting studies include [12]–[14], however analysis was limited to the provision of numerical [12] or asymptotical results [13]. Characterization of statistics, pertaining to LoS fading, was provided in [14] though the setup, considered therein, addressed the single-antenna case and it was not motivated by a practical scenario. This is settled in [15] where characterization of the link ergodic capacity, with transmit beamforming and optimized power allocation, was presented. The authors

The authors are with the Communication Theory department and the ACCESS Linnaeus Center, at The Royal Institute of Technology (KTH), Stockholm, Sweden (e-mail: eftsta,lkra,skoglund@kth.se).

Lars K. Rasmussen is also affiliated with the Institute for Telecommunications Research, University of South Australia, Australia.

exploited the choice of channel model and identified the F -distribution [16], [17] in order to obtain closed-form results for the ergodic capacity under interference power constraints.

In this paper, we generalize the original study [15] and refine some of the results presented therein. We expand the scope of the network model, studied in [15], to address general cellular networks with clusters of neighboring-cell edge-users, see Fig. 1, that are protected, by means of an ergodic sum-interference power constraint (ESIPC), from an active link between a multi-antenna base station and its associated receiver (TR link). The goal is to find the optimal precoder that achieves the ergodic capacity of the TR link while conforming to an ESIPC. The main contributions of our work can be summarized in the following points:

- (Po1): We show that the capacity-achieving linear precoder is rank-one, i.e., beamforming is the optimal strategy for the considered network utility function. We further argue that the receiver needs to know only the instantaneous signal-to-noise ratio (SNR), obviating the need for channel acquisition at the receiver since it does not increase capacity;
- (Po2): We derive a closed-form expression for the long-term capacity. The solution presented here is more compact and computationally efficient than [15, Eqs. (2),(3)];
- (Po3): From our main analytical result, we obtain algebraic approximations for certain cases such as operation at high SNR or nearly shadowed environments with weak LoS components. We further show how our result subsumes several existing models in literature [12], [14], [18].
- (Po4): We complement our analysis with the derivation of metrics, which are relevant for the assessment of short-term and long-term performance, such as the distribution of the receive SNR and the average bit-error-rate (BER). These metrics are pertinent for practical schemes with fixed-rate transmission from finite constellation schemes.

In particular, we capitalize on the rank-one optimality result (Po1) to reduce the problem to power allocation. We emphasize an insightful mathematical result, i.e., how the F -distribution arises in the stochastic characterization of the network, because of its gravity in analytical derivations that were previously inhibited without this observation [12]. Moreover, we introduce a mathematical technique that allows us to *reshape a certain class of infinite series into a finite one* and we utilize it to further enhance [15, Eqs. (2),(3)], which are in the form of infinite series and whose numerical evaluation can only be approximate, since one has to truncate the series and characterize the behavior of the truncation error. Herein, as stated in (Po2), we are able to cast the power allocation and the associated capacity formulas in a mathematical form whose numerical evaluation is *exact and can be carried out in predictable computation time*, in contrast to the

Fig. 1. An interference network with a base station communicating with its desired receiver R_X and interfering with a cluster $\mathcal{P}$ of unintended nodes.

truncated counterparts [15, Eqs. (2),(3)]. Several existing results follow as special cases from our model, see (Po3).

Due to the various properties of the wireless channel, different QoS metrics can serve as performance indicators, see (Po4). Therefore, apart from the ergodic capacity, which is our main design guideline, we derive the distribution of the TR link SNR and the average BER, since these metrics are good proxies for practical performance under known constellations, e.g., M-PSK. It is noted that the average BER expression is given in terms of infinite series but we are able to show, using previous results [15, Lemma 1], how one can properly choose the series truncation index in order to approximate the theoretical BER value with arbitrarily low truncation error. Numerical results are congruent with our theoretical analysis and further shed light on the impact of system parameters into performance.

The paper is organized as follows. The network model and the system architecture are presented in Section II, whereas theoretical analysis of the scheme is carried out in Section III. Further system aspects and complementary performance metrics are investigated in Section IV. Numerical results for various setups are available in Section V and, finally, in Section VI we discuss our findings along with directions for future work.

Notation and functions: Bold lower (upper) case letters stand for vectors (matrices). We denote the N -length all-zero (all-one) vector as $\mathbf{0}_N$ ($\mathbf{1}_N$) and the $N \times N$ identity matrix as $\mathbf{I}_N$. The operators $\|\cdot\|$, $(\cdot)^H$ stand for the Euclidean norm and Hermitian form of their (vector/matrix) arguments. The matrix trace is denoted as $\text{tr}(\cdot)$. We define the operators $\{x\}^+ \triangleq \max\{x, 0\}$ and $[f(x)]_a^b \triangleq f(b) - f(a)$. Similarly we use the notation $f(x) \approx g(x)$ and $f(x) \leq g(x)$ to indicate that in the limit, as x tends to zero, $f(x)$ can be approximated by $g(x)$ ($\approx$) and $f(x)$ is lower than $g(x)$ ($\leq$), respectively. The logarithm base is e .

II. SYSTEM SETUP AND PROBLEM STATEMENT

A. Network Model

We consider the multi-node communication network, depicted in Fig. 1. Our setup comprises a link between a source node T_x , with M antennas, and its single-antenna destination node R_x , as well as undesirable interference inflicted to an external cluster of single-antenna nodes $P_{x,n}$ where $n = 1, \dots, N$. The T_x - R_x (TR) transmission has to adhere to certain operational requirements, e.g., it should not severely degrade the aggregate QoS at the cluster of nodes. To guarantee a minimum service level, the TR transceiver architecture is designed with account for these inherent limitations.

Let $\mathbf{h} = \sqrt{\frac{K}{K+1}}\mathbf{h}_{\text{LoS}} + \mathbf{h}_{\text{NLoS}}$ denote the TR link channel, with Rician factor K , LoS component $\mathbf{h}_{\text{LoS}}$ that satisfies $\|\mathbf{h}_{\text{LoS}}\|^2 = M$ [11] and non-LoS (NLoS) component $\mathbf{h}_{\text{NLoS}} \sim \mathcal{CN}(\mathbf{0}_M, \sigma_h^2 \mathbf{I}_M)$, where $\sigma_h^2 = \frac{1}{K+1}$. The Rician factor quantifies the contribution of the LoS component, to the channel, relative to its NLoS counterpart. In particular, the higher K is the more concentrated the channel will be around the mean and in the limit, as K tends to infinity, the channel becomes deterministic. Moreover, we define the matrix $\mathbf{G} = [\mathbf{g}_1 \dots \mathbf{g}_N]^H \in \mathbb{C}^{N \times M}$ is the collection of all T_x - $P_{x,n}$ (TP $_n$) link channels, with $\mathbf{g}_n \sim \mathcal{CN}(\mathbf{0}_M, \sigma_g^2 \mathbf{I}_M) \forall n = 1, \dots, N$. All the stochastic quantities in the network are assumed to be independent in both time and space domain.

B. Transceiver architecture and power control

In our system, the secondary transmitter employs linear precoding; to communicate its message during the k -th block, the node T_x i) rotates the information-bearing vector $\mathbf{t}[k]$ using a matrix $\mathbf{U}[k]$ (not full-rank in general) with orthonormal columns; and ii) it allocates certain power to each stream using a diagonal matrix $\mathbf{\Lambda}^{\frac{1}{2}}[k] = \text{diag}(\sqrt{\lambda_1[k]}, \dots, \sqrt{\lambda_M[k]})$.¹ The composite vector that is sent over the channel can be expressed as $\mathbf{x} \triangleq \mathbf{\Lambda}^{\frac{1}{2}} \mathbf{U} \mathbf{t}$ with covariance $\mathbf{Q} \triangleq E\{\mathbf{x}\mathbf{x}^H\}$. Note that the elements of $\mathbf{t}[k]$ are assumed to be independent *along all dimensions*, i.e., time and space. The received signal y_s at the destination node R_x is written as

$$y_s[k] = \mathbf{h}^H[k] \mathbf{x}[k] + n[k], \quad (1)$$

where k is the time index, in symbol intervals, and $n[k] \sim \mathcal{CN}(0, \sigma^2)$ denotes the additive white noise process. We stress here that indexing the sequence $\mathbf{x}[k]$ implies that both the information sequence $\mathbf{t}[k]$ and the precoding strategy, i.e., the matrix $\mathbf{Q}[k]$, are time-variant. Therefore, the transmission strategy is dynamic and adapted on a slot-by-slot basis to serve a certain design target. To simplify the ensuing exposition, we shall occasionally drop the indices unless explicit account is needed to clarify parts of the analysis.

The choice of the transmit precoding strategy should reflect the $P_{x,n}$ QoS guarantees, mentioned earlier, which

are imposed in order to protect the unintended receivers from excessive interference. These guarantees essentially translate into constraints on the instantaneously inflicted sum-interference temperature $I_i \triangleq \text{tr}(\mathbf{G}[k] \mathbf{Q}[k] \mathbf{G}[k]^H)$ or its ergodic counterpart $I_a \triangleq E\{I_i\}$, where the average is over time. In other words, a design has to always satisfy $I_i \leq Q_{\text{th}}$ (correspondingly $I_a \leq Q_{\text{th}}$), where Q_{th} represents the highest admissible interference level. In this work we adopt the *ESIPC* for our design, i.e., $I_a = E\{\text{tr}(\mathbf{G}[k] \mathbf{Q}[k] \mathbf{G}[k]^H)\} \leq Q_{\text{th}}$, yet our analysis is applicable to the case with $I_i \leq Q_{\text{th}}$.

The design is based on channel side-information (CSI), which is available at the nodes T_x and R_x . We assume that the transmitter knows the instantaneous TR-link channel $\mathbf{h}[k]$. To satisfy the ESIPC, the system further needs information for the cross-link channels. Since learning $\mathbf{G}$ is a challenging task that requires many resources, we assume that the system does not know $\mathbf{G}$ but instead learns a lower-dimensional (less than $N \times M$) function of it $z = f(\mathbf{G})$. In particular, if $\mathbf{G}$ is not known then T_x designs the rotating matrix $\mathbf{U}$ independently of it. Once the spatial strategy $\mathbf{U}[k]$ is determined, we can determine the *directional* sum-interference $z[k] \triangleq \sum_{n=1}^N \|\mathbf{U}^H[k] \mathbf{g}_n[k]\|^2$, which is used for power control.² The receiver knows only the instantaneous SNR.

We assume that the stochastic nature of the environment is known, i.e., the distributions of the various quantities such that the expectation can be evaluated. To characterize the ergodic capacity under an ESIPC, we rely on the following lemma.

Lemma 1. *Under the given CSI model, the capacity-achieving covariance matrix $\mathbf{Q}_o[k]$ is rank-one, i.e., $\mathbf{Q}_o[k] = P(y[k], z[k]) \mathbf{w}_o[k] \mathbf{w}_o^H[k]$, with principal eigenvector given as $\mathbf{w}_o[k] = \mathbf{h}[k] / \|\mathbf{h}[k]\|$ and $y[k] \triangleq \|\mathbf{h}[k]\|^2, z[k] \triangleq \|\mathbf{G}[k] \mathbf{w}_o[k]\|^2$. Then the optimal power allocation strategy $P^*(y[k], z[k])$ is found by solving the following optimization problem*

$$\begin{aligned} & \underset{P(y[k], z[k])}{\text{maximize}} && E_{y,z} \left\{ \log \left[1 + \frac{y[k] P(y[k], z[k])}{\sigma^2} \right] \right\} \\ & \text{subject to} && E_{y,z} \{ z[k] P(y[k], z[k]) \} \leq Q_{\text{th}}. \end{aligned} \quad (\text{P0})$$

where $E_{y,z}\{\cdot\}$ denotes expectation with respect to the distribution of y and z . The ergodic capacity does not increase even if the receiver has the same CSI as the transmitter.

Proof. Given in Appendix A. $\square$

It follows from this result that the optimal vector $\mathbf{w}_o$ is essentially maximum-ratio transmission (MRT) in the direction of the TR link channel [19] and that the optimal precoding vectors *are not coupled in time*. Therefore the aforementioned CSI scheme, where $z[k]$ is measured at the nodes of $\mathcal{P}$ and then fed back, can be easily realized. Once $\mathbf{w}_o[k]$ is available, each of the unintended nodes $P_{x,n}$ can measure, during one-slot training, $z_n[k] = \|\mathbf{g}_n^H[k] \mathbf{w}_o[k]\|^2$ and then simultaneously feed the value back to T_x . We

¹The cardinality of $\mathbf{t}$, i.e., the number of streams, equals the number of non-zero diagonal elements of $\mathbf{\Lambda}$.

² z can be obtained by measuring $z_n \triangleq \|\mathbf{U} \mathbf{g}_n\|^2$ at every node $P_{x,n}$ and then feeding it back simultaneously.

stress that the CSI assumptions define the domain of the transmit strategy; under the current CSI model, the stochastic quantities $\mathbf{h}[k], z[k]$ are the causally acquired input at the precoder. However, if side-information about the cross-link channel matrix $\mathbf{G}[k]$ was available, then the domain of $\mathbf{w}_o[k]$ could be extended to include the extra information.³

Lemma 1 reduces the precoder search space to that of one dimensional functionals $P(y, z)$ and the optimization problem (P0) is convex in the functional $P(y, z)$. However, this observation does not imply that it can be solved efficiently using numerical techniques, such as interior-point methods that are implemented in existing numerical software packages, e.g., CVX [20]. The reason is that the optimal policy $P^*(y, z)$ is a functional, implying that (P0), when viewed in the vector space, is an infinite-dimensional problem.⁴

III. ANALYSIS OF THE SCHEME

In this section we elaborate on the solution of (P0) by introducing a mathematical result, which is crucial for our analysis, and then present the main result.

A. Preliminaries

The power allocation functional $P^*(y, z)$, which solves (P0) for any continuous distribution, is found by the Karush-Kuhn-Tucker (KKT) conditions and it is $P^*(y, z) = \left\{ \frac{1}{z\mu} - \frac{\sigma^2}{y} \right\}^+$ [7], where μ is the Lagrange dual variable associated with the ESIPC. Given that the inequality is tight at the optimum, the Lagrange multiplier is found by solving the equation

$$E_{y,z}\{zP^*(y, z)\} = Q_{\text{th}}. \quad (2)$$

If we plug the expression for $P^*(y, z)$ into the objective function of (P0) and Eq. (2) then, after trivial manipulations, we find that the resulting expressions are functions of $\bar{\gamma} \triangleq \frac{y}{z}$. We conclude that stochastic characterization of $\bar{\gamma}$ is an indispensable task prior to proceeding with the theoretical analysis of our problem.

B. An insightful mathematical result

Let $\chi_{2N}^2(\lambda)$ denote a (non-)central chi-square distribution with $2N$ degrees-of-freedom and non-centrality parameter λ and further consider the following quantities

$$\tilde{y} \triangleq \frac{y}{\sigma_h^2/2}, \tilde{z} \triangleq \frac{z}{\sigma_g^2/2} = \frac{\mathbf{w}_o^H \mathbf{G}^H \mathbf{G} \mathbf{w}_o}{\sigma_g^2/2} = \sum_{k=1}^N \frac{|\mathbf{g}_k^H \mathbf{w}_o|^2}{\sigma_g^2/2}.$$

To obtain the distribution of $\bar{\gamma}$, defined above, we exploit the following two observations,

³For instance, if the rows of $\mathbf{G}[k]$ had known (and fixed) covariance then this side-information could be incorporated in the precoder domain.

⁴Albeit this deficiency one can still solve (P0) *approximately*, by exploiting the ergodicity of the channel. For example, one can generate a “large” block of i.i.d. channels and solve (P0) by averaging over many such realizations. In the next section we settle this disparity by providing a closed-form expression for the integrals involved in (P0).

(O1): $\tilde{z} \sim \chi_{2N}^2(0)$ and $\tilde{y} \sim \chi_{2M}^2(2MK)$, where the non-centrality is determined after the definition of $\mathbf{h}$.

(O2): The random variables (RVs) $\tilde{y}$ and $\tilde{z}$ are *independent*.

Since $\tilde{y}$ and $\tilde{z}$ are both functions of $\mathbf{h}$, (O2) seems counter-intuitive. To see why it holds we define the vector $\bar{\mathbf{z}} \triangleq [\tilde{z}_1 \dots \tilde{z}_N]$ with elements $\tilde{z}_n \triangleq \frac{\mathbf{g}_n^H \mathbf{w}_o}{\sigma_g^2/2}$, where $\mathbf{g}_n^H$ are the rows of $\mathbf{G}$. Let us denote the density functions of a general RV X and a $\mathcal{CN}(\mu, \sigma^2)$ -distributed RV as $f_X(t)$ and $\phi(t, \mu, \sigma^2)$, respectively. We can then determine the density function of $\tilde{z}_n$ as

$$\begin{aligned} f_{\tilde{z}_n}(t) &= \int_{-\infty}^{\infty} f_{\tilde{z}_n|\mathbf{h}}(t, \mathbf{h}|\mathbf{h}) f_{\mathbf{h}}(\mathbf{h}) d\mathbf{h} \\ &\stackrel{(a)}{=} \int_{-\infty}^{\infty} \phi(t, 0, \sigma_g^2) f_{\mathbf{h}}(\mathbf{h}) d\mathbf{h} = \phi(t, 0, \sigma_g^2) \end{aligned}$$

where (a) holds because the distribution of $\tilde{z}_n$ conditioned on $\mathbf{h}$ is zero-mean Gaussian with variance σ_g^2 , hence it is independent of $\mathbf{h}$. Since the elements of $\bar{\mathbf{z}}$ are independent of $\mathbf{h}$ it follows that $\tilde{z} = (2/\sigma_g^2)\|\bar{\mathbf{z}}\|^2$ is independent of $\mathbf{h}$ and consequently of $\tilde{y}$.

Next we define the quantity $c \triangleq \frac{M\sigma_h^2}{N\sigma_g^2}$ that allows us to express the stochastic quantity $\bar{\gamma}$ as $\bar{\gamma} = \frac{2M}{2N} \frac{\sigma_h^2/2}{\sigma_g^2/2} \frac{\tilde{y}}{\tilde{z}} = c\gamma$, where $\gamma \triangleq \frac{\tilde{y}/(2M)}{\tilde{z}/(2N)}$. This characterization is insightful because in this form we recognize that $\gamma \sim \mathcal{F}_{2M, 2N}(\lambda, 0)$ [16], i.e., γ has a non-central F -distribution whose probability density function (pdf) $f_{\gamma}(x)$ is given by

$$f_{\gamma}(x) = e^{-\frac{\lambda}{2}} \sum_{n=0}^{\infty} \frac{(\lambda/2)^n (\frac{M}{N})^{M_n}}{B(M_n, N) n!} \frac{x^{M_n-1}}{[1 + \frac{M}{N}x]^{M_n+N}}, \quad (3)$$

where, for brevity, we have defined $M_n \triangleq M + n$. This characterization simplifies the technical part since: i) it settles any ambiguities around the independence of $\tilde{y}, \tilde{z}$, which is a critical point in the analysis and ii) we have to deal with one stochastic quantity. Moreover, our result generalizes the single-antenna ($M = N = 1$) counterpart, derived in [14, Eq. (19)]. One can further tailor our result to the reverse scenario, with a NLoS TR link and LoS TP_n links. In that case the corresponding random variable $\gamma_{\text{inv}} \triangleq \frac{1}{\gamma}$, can be recovered directly from Eq. (3), through a transformation of variables.

It is worthwhile pointing out that a key element, that led to Eq. (3), was the selection of the channel model itself. In other words, the analysis is less intuitive if one adopts, similarly to [12], the Rician distribution to characterize the channel gains $\{|h_k|\}_{k=1}^M$. In that case the density of each individual component would be given as [22, Ch. 2]

$$f_{|h_k|}(x) = 2(K+1)xe^{-K-(K+1)x^2}I_0(2x\sqrt{K(K+1)}). \quad (4)$$

Strictly speaking, working with the above model should eventually lead to Eq. (3). However, tailoring an appropriately parametrized non-zero mean Gaussian random vector to the LoS link channel is insightful in introducing the non-central F -distribution. This conclusion is otherwise obscure, under the umbrella of Eq. (4), since one has to view γ as the ratio of independent (non-)central chi-square

random variables in order to proceed with a tractable stochastic characterization. This was also observed in [14] but our particular architecture, with MRT beamforming, has an inherent connection with these statistics. Therefore, we can judiciously argue that a non-zero mean Gaussian model is more appropriate for LoS communication modeling while it is possible to translate it into the equivalent Rician model of Eq. (4), to draw conclusions on practical aspects.

C. Capacity results

As pointed out in Section III-A, recovery of μ relies on solving Eq. (2). To further assist the mathematical exposition we define $\gamma_0 \triangleq \frac{\mu\sigma^2}{c}$, $\tilde{Q} \triangleq \frac{cQ_{\text{th}}}{\sigma^2}$ and the following function $h : \mathcal{R} \rightarrow \mathcal{R}$

$$h(\gamma_0) \triangleq E_\gamma \left\{ \frac{1}{\gamma_0} - \frac{1}{\gamma} \right\}^+. \quad (5)$$

Using the definitions above we can recast Eq. (2) as $h(\gamma_0) = \tilde{Q}$, a form which is tailored to the stochastic profile described in the previous section. Solution of the integral, involved in $h(\gamma_0) = \tilde{Q}$, is provided in the following Proposition.

Proposition 1. *Let $\delta(m)$ denote the Kronecker delta, i.e., $\delta(m) = 1$ for $m = 0$ and is zero otherwise. Then, the optimal γ_0 (equivalently μ) that solves (P0) is found as the solution to Eq. (6) (see top of next page), where $L \triangleq M + N - 1 - k + m$, $\tilde{x}_0 \triangleq \frac{\lambda}{2} \frac{\frac{M}{N}\gamma_0}{1 + \frac{M}{N}\gamma_0}$, $\delta_L \triangleq \delta(m)$, $\tilde{\delta}_L \triangleq 1 - \delta_L$ and $\text{Ei}(x) \triangleq -\int_{-x}^{\infty} \frac{e^{-t}}{t} dt$ is the exponential integral [23, 8.211.1].*

Proof. Given in Appendix B. $\square$

Several methods can be implemented for solving Eq. (6) but, prior to enumerating few, it is useful to look into its monotonicity properties. If we rewrite $h(\gamma_0)$ in its integral form as $h(\gamma_0) = \int_{\gamma_0}^{\infty} (\frac{1}{\gamma_0} - \frac{1}{\gamma}) f_\gamma(x) dx$ then its first order derivate $h^{(1)}(\gamma_0) \triangleq \frac{\partial h(\gamma_0)}{\partial \gamma_0}$ is found as [24]

$$h^{(1)}(\gamma_0) = -\frac{1}{\gamma_0^2} \Pr(\gamma > \gamma_0) < 0 \quad \forall \gamma_0. \quad (7)$$

The above equation asserts that $h(\gamma_0)$ is a strictly monotone-decreasing function of its argument. Capitalizing on the monotonicity property, one can implement Newton's method to obtain the root of Eq. (6), by starting from a point close to zero and using Eq. (7) to update the running estimate. Alternative methods would be bisection or high-resolution grid search within a given interval $\mathcal{S}_{\gamma_0}^{(k)} \triangleq [(k-1)\delta, k\delta]$, where δ is some positive quantity. One can initiate the search from the set $\mathcal{S}_{\gamma_0}^{(1)}$ and further update k in an incremental fashion, as $k \rightarrow k+1$, until the *unique* root is retrieved. To settle the numerical solution of Eq. (6) we note that, albeit its complicated form, it involves elementary functions⁵ and thus can be easily evaluated.

⁵The only non-trivial component is the exponential integral term but, as long as M is greater than one, this term will not be considered in the evaluation.

Since Q_{th} is indirectly related to the objective of (P0) it is useful to analyze the sensitivity of $h(\gamma_0)$ on perturbations of γ_0 . A qualitative sensitivity analysis can assist in assessing how shifts in the threshold-to-noise (TNR) ratio $q \triangleq \frac{Q_{\text{th}}}{\sigma^2}$, translate into shifts in γ_0 , which is important because γ_0 controls the waterfilling level and subsequently the power allocation. For sufficiently large values of q , i.e., in the “high-TNR” region, γ_0 will be very small and $|h^{(1)}(\gamma_0)|$ will be very large. Therefore $h(\gamma_0)$ decays rapidly with increasing γ_0 , in this regime, implying that γ_0 changes slightly in order to compensate for a linear increase in q (equivalently Q_{th}). On the contrary, when the system operates in the low TNR region γ_0 attains large values and the derivative converges to zero very fast,⁶ which implies that the decremental steps of $h(\gamma_0)$ are very small. Thus, at the “low TNR” regime, γ_0 should decrease significantly such that linear shifts in q induce similar changes in $h(\gamma_0)$.

Given the value of γ_0 , that solves Eq. (6), we can proceed to characterize the long-term capacity of the TR link by plugging this value into the objective of (P0). The closed-form expression for the ergodic capacity is established in the following Proposition.

Proposition 2. *For a given γ_0 , the objective function of (P0) is given in closed-form by Eq. (8) (see top of next page), where $L_1 \triangleq M + N - 1 - k + m$, $L_2 \triangleq N - 1 + m$ and $\tilde{\gamma}_0 \triangleq \frac{M}{N}\gamma_0$. The terms $B(a, b)$, $\mathcal{B}(x; a, b)$ correspond to the Beta function [23, 6.2.1] and the incomplete Beta function [26, 6.6.1], respectively, while $\mathcal{I}(x; a, b) \triangleq \frac{\mathcal{B}(x; a, b)}{B(a, b)}$ denotes the regularized incomplete Beta function [26, 6.6.2]. Finally, $C_{\text{Eu}} \approx 0.57721$ is the Euler constant [23, 8.367.1].*

Proof. Given in Appendix C. $\square$

It is possible to further simplify Eq. (8) when certain network parameters, e.g., the TNR q or the Rician factor K , attain extreme values. To see how large values of q influence ergodic capacity we factor in the monotonicity properties of $h(\gamma_0)$, described earlier. Recalling that γ_0 is very small when the system operates in the high-TNR regime, we conclude that the dominating term in Eq. (8) will be $\log \gamma_0$ and that it can be approximated by Eq. (9) (where we include more terms for better algebraic approximation). Another asymptotic case arises when the LoS component is weak, i.e., the Rician factor K is sufficiently low, such that the original expression can be simplified into Eq. (10). When the network propagation environment is completely shadowed, then by setting $\lambda = 0$ in Eq. (10) we recover Eq. [18, Eq. (8)] (the equivalence follows after some algebra) which is a special case of our general capacity formula (8). Moreover, if we set $M = N = 1$, which corresponds to the case of a single-antenna source $\mathcal{T}_x$ and single element set $\mathcal{P}$, then the generic capacity formula reduces to

$$C_{\text{sa}} = \log K - \log \gamma_0 - \text{Ei}\left(\frac{K}{1 + \gamma_0}\right) + C_{\text{Eu}}. \quad (11)$$

⁶In fact the derivative $h^{(1)}(\gamma_0)$ converges to zero exponentially fast, owing to the properties of the tail probability of the F -distribution [25].

$$\begin{aligned}
\tilde{Q} = & \frac{e^{-\lambda/2} \gamma_0^{-1}}{(\frac{\lambda}{2})^M \Gamma(N)} \sum_{k=0}^N \sum_{m=0}^{N-1} \binom{N}{k} \frac{(M+N-1)!}{(M+N-1-k)!} \binom{N-1}{m} \frac{(-1)^m}{(\frac{\lambda}{2})^m} \left\{ \sum_{l=0}^L \frac{L!}{l!} (-1)^{L-l} \left[\left(\frac{\lambda}{2} \right)^l e^{\frac{\lambda}{2}} - \tilde{x}_0^l e^{\tilde{x}_0} \right] \right\} \\
& - \frac{e^{-\lambda/2} (M/N)}{(\frac{\lambda}{2})^{M-1} \Gamma(N)} \sum_{k=0}^N \sum_{m=0}^N \binom{N}{k} \frac{(M+N-1)!}{(M+N-1-k)!} \binom{N}{m} \frac{(-1)^m}{(\frac{\lambda}{2})^m} \left(\delta_L \left[\text{Ei} \left(\frac{\lambda}{2} \right) - \text{Ei}(\tilde{x}_0) \right] \right. \\
& \left. + \delta_L \left\{ \sum_{l=0}^{L-1} \frac{(L-1)!}{l!} (-1)^{L-1-l} \left[\left(\frac{\lambda}{2} \right)^l e^{\frac{\lambda}{2}} - \tilde{x}_0^l e^{\tilde{x}_0} \right] \right\} \right)
\end{aligned} \tag{6}$$

$$\begin{aligned}
C = & \frac{e^{-\frac{\lambda}{2}}}{\Gamma(N)} \sum_{k=0}^N \binom{N}{k} \frac{(M+N-1)!}{(M+N-1-k)!} \left\{ \sum_{m=0}^{N-1} \binom{N-1}{m} \frac{(-1)^m}{(\frac{\lambda}{2})^{M+m}} \left(\sum_{l=1}^{L_1} \sum_{j=0}^{L_1-l} \frac{L_1! (-1)^{L_1-j}}{j! (L_1-l+1)!} [x^j e^x] \right)^{\frac{\lambda}{2}} \right. \\
& \left. + \sum_{l=0}^{L_1} \frac{L_1! (-1)^l}{(L_1-l)!} \left[x^{L_1-l} e^x \log \left(\frac{x}{\gamma_0} \right) \right]^{\frac{\lambda}{2}} \frac{\gamma_0}{1+\gamma_0} - (L_1)! (-1)^{L_1} \left[\text{Ei} \left(\frac{\lambda}{2} \right) - \text{Ei} \left(\frac{\lambda}{2} \frac{\gamma_0}{1+\gamma_0} \right) \right] \right\} \\
& + e^{\frac{\lambda}{2}} \sum_{m=0}^{M+N-1-k} \binom{M+N-1-k}{m} \frac{(-1)^m}{(\frac{\lambda}{2})^{m+k}} \left(\sum_{l=1}^{L_2} \frac{(L_2)!}{(L_2-l+1)!} \left[1 - \sum_{j=0}^{L_2-l} \frac{(L_2-l)!}{j!} \left(\frac{\lambda/2}{1+\gamma_0} \right)^j e^{-\frac{\lambda/2}{1+\gamma_0}} \right] \right. \\
& \left. - e^{-\frac{\lambda/2}{1+\gamma_0}} \log \left(\frac{\lambda/2}{1+\gamma_0} \right) \sum_{l=0}^{L_2} \frac{L_2!}{(L_2-l)!} \left(\frac{\lambda/2}{1+\gamma_0} \right)^{L_2-l} - L_2! \left[\text{Ei} \left(\frac{\lambda/2}{1+\gamma_0} \right) + C_{\text{Eu}} \right] \right\}
\end{aligned} \tag{8}$$

$$\begin{aligned}
C_{\text{high}} = & \frac{e^{-\frac{\lambda}{2}} \gamma_0^M \log \gamma_0}{(1+\gamma_0)^{M+N-1}} \left[\frac{1 - (1+\gamma_0)^{-1}}{MB(M, N)} \right] + \log \left(\frac{1+\gamma_0}{\gamma_0} \right) \left[1 + \frac{1}{(1+\gamma_0)^{N-1}} \right] + \text{Ei} \left[-\frac{\lambda}{2} \frac{\gamma_0}{1+\gamma_0} \right] - \text{Ei} \left[-\frac{\lambda}{2} \right] \\
& - \sum_{j=2}^N \frac{(1+\gamma_0)^{-(j-1)}}{(j-1)!} - e^{-\frac{\lambda}{2}} \sum_{k=0}^{M-2} \left[\frac{(-1)^{k+1} k!}{(\frac{\lambda}{2})^{k+1}} + e^{\frac{\lambda}{2} \frac{\gamma_0}{1+\gamma_0}} \sum_{j=0}^k \frac{k!}{j!} \frac{(-1)^{k-j} \gamma_0^j}{(\frac{\lambda}{2})^{k-j+1}} \right]
\end{aligned} \tag{9}$$

$$\begin{aligned}
C_{\text{weak}} = & \left(1 - \frac{\lambda}{2} \right) \sum_{m=1}^{M-1} \left\{ \frac{\gamma_0^k \log \gamma_0}{(1+\gamma_0)^{k+N}} + \mathcal{B} \left(\frac{1}{1+\gamma_0}; N, k \right) \right\} + \frac{\log \gamma_0}{(1+\gamma_0)^N} + \log \left(\frac{1+\gamma_0}{\gamma_0} \right) - \sum_{j=2}^N \frac{(1+\gamma_0)^{-(j-1)}}{(j-1)!} \\
& - \log(\gamma_0) \left[1 - \left(1 - \frac{\lambda}{2} \right) \mathcal{I} \left(\frac{\gamma_0}{1+\gamma_0}; M, N \right) \right]
\end{aligned} \tag{10}$$

Eq. (11) settles a previously open question, i.e., the solution of [12, Eq. (23)]. Note here that [12], [18] are studies for cognitive radio networks, implying that our network model is particularly relevant in this area and subsumes existing results as special cases.

We stress that Eqs. (6) and (8) yield the exact solution of the problem of achieving the ergodic capacity (P0), which can otherwise be solved approximately with numerical integration since the expectations, involved in (P0), are performed over $f_\gamma(x)$ that contains an infinite series (see Eq. (3)). Thus, our main result, apart from its mathematical merit, is a tractable alternative for implementation compared to other methods, e.g., numerical integration or simulation via channel realization (see footnote 4), which result in approximate solutions.

It is worth pointing out here that the expressions (6) and (8) have *predictable computational complexity*. To see the benefit of this we compare our results with [15, Eqs. (2),(3)], which are the infinite series counterparts of Eqs. (6),(8) for solving (P0). In the case of [15, Eqs. (2),(3)], one needs to characterize the truncation error that arises in the implementation of the infinite series since only a finite number T of terms will be considered eventually.

It was established in [15, Proposition 3] that selecting the truncation index T such that it satisfies $T > eKM$ guarantees an exponentially decreasing trend in the truncation error. However the linear dependence of T on K implies that evaluation of the truncated versions of [15, Eqs. (2),(3)], albeit accurate, *has a variable computational cost* that depends on K , i.e., the physical properties of the propagation environment.

As a figure of merit, we report in Table I the absolute execution time, averaged over 10^3 runs, that is spent in MATHEMATICA [27] for the evaluation of Eq. (8) and [15, Eq. (3)], respectively. Overall, we observe that Eq. (8) is evaluated at least 50 times faster than [15, Eq. (3)] when $T = 10^2$, up to more than 5×10^3 times faster when $T = 10^3$. The execution time of [15, Eq. (3)] increases polynomially with T , which indicates that for setups where large values of T are indispensable, e.g., due to large K , the time savings from Eq. (8) will be substantial.

IV. ON PRACTICAL APPLICATIONS

In this section, we investigate system aspects that are interesting from a practical point-of-view. To realize our transceiver strategy one needs to implement Gaussian

coding schemes that span infinitely long transmission blocks, with independent channel realizations. However, in practical cellular systems the propagation environments can possibly be quasi-static and the information symbols are chosen from finite-size alphabets. Therefore, metrics such as the outage probability or the average BER for known constellations become relevant indicators of the scheme's performance.

A. Outage Probability

The short-term (within the channel coherence-time window) QoS is assessed through the probability $P_{\text{out}}(x) \triangleq \Pr(\rho \leq x)$ that the instantaneous received SNR $\rho \triangleq \frac{P^*(y,z)\|\mathbf{h}\|^2}{\sigma^2}$ cannot support the transmission rate R , which satisfies $x = e^R - 1$. By definition there is an equivalence between the outage probability and the distribution function $F_\rho(x)$ of ρ . The latter is characterized in the following proposition.

Proposition 3. Let $F_\gamma(x) \triangleq \int_0^x f_\gamma(x)dx$, where $f_\gamma(x)$ is defined in Eq. (3), be the distribution function of the random variable γ . Then the distribution function of the received SNR ρ can be expressed as follows

$$F_\rho(x) = F_\gamma[\gamma_0(x+1)], \quad x \geq 0. \quad (12)$$

Proof. Given in Appendix C. $\square$

A mathematical subtlety, hidden in Eq. (12), is the presence of a mass point at $x = 0$, since $F_\rho(0) > 0$. Therefore, the density function $f_\rho(x)$ cannot be obtained in the conventional fashion, through differentiation of $F_\rho(x)$. To see why, one has to reiterate the proof of Proposition 3 and zoom in on Eq. (D1), where precisely the mass point arises and this is shown by the indicator function at $x = 0$. This observation, along with Eq. (D1) itself, implies that we should express the pdf $f_\rho(x)$, in a similar fashion to [28, Eq. (66)], using the Dirac delta operator as follows,

$$f_\rho(x) = \gamma_0 f_\gamma[\gamma_0(x+1)] + \delta(x) \int_0^{\gamma_0} f_\gamma(t)dt.$$

Since $f_\rho(x)$ is not differentiable at $x = 0$, we cannot determine the diversity order using well-known methods such as Taylor expansion of $f_\rho(x)$ around zero [29]. This framework is not applicable in our case because it requires $f_\rho(x)$ to be well-behaved around zero. Therefore, we need to use conventional techniques, i.e., characterize the behavior of $F_\rho(x)$ for extreme values of its argument x . Another dissimilarity with traditional outage analysis is that Q_{th} , which can be seen as the equivalent to transmit power, is only indirectly related to ρ , by means of γ_0 . However, this disparity is reconciliated through our

discussion around the monotonicity of $h(\gamma_0)$ and Eq. (8) (see Section III-C). The high-SNR approximation of the SNR distribution is given in the following result,

Proposition 4. Let $g(x) \triangleq \frac{\frac{M}{N}\gamma_0(x+1)}{1+\frac{M}{N}\gamma_0(x+1)}$. Then, for sufficiently large values of the TNR q , it holds that $g(x) \rightarrow 0$ and the SNR distribution is approximated as

$$F_\rho(x) \approx e^{-KM} \frac{[g(x)]^M [1-g(x)]^{N-1}}{MB(M, N)}. \quad (13)$$

Proof. Given in Appendix D. $\square$

In the context of outage probability, we interpret the above result as follows; for a fixed Rician factor K , the probability that the system cannot support the transmission rate decreases polynomially, with order M , which can be viewed as the counterpart to diversity order. We also see that for small x and fixed M , $F_\rho(x)$ decays exponentially fast in K .

B. Average error probability

The average BER of a transmission scheme is another performance metric, which captures practical aspects of the implementation since it is derived on the basis of practical modulation schemes. There are various methods for obtaining the BER of a specific constellation scheme, see [19, Ch. 3], [30, Ch. 6], and here we use the following (approximate) formula,

$$P_e = E_\rho \left[\alpha Q \left(\sqrt{\beta \rho} \right) \right], \quad (14)$$

where $E_\rho(\cdot)$ denotes expectation with respect to ρ , $Q(x)$ denotes the Q -function (tail of the Gaussian distribution) and α, β are constellation-specific values. The average, over the transmission block, BER is established in the following result.

Proposition 5. Let $\bar{x}_0 \triangleq \frac{\frac{M}{N}\gamma_0}{1+\frac{M}{N}\gamma_0}$ and $L_0 \triangleq M+n+N-1-k+l$. Then the average BER of the TR link, which is designed according to (P0), is given by

$$P_e = \alpha \frac{e^{-\frac{\lambda}{2}}}{2} \sqrt{\frac{\beta}{2\bar{x}_0}} \sum_{n=0}^{\infty} \sum_{k=M_n}^{M+N-1} \sum_{l=0}^k \frac{(\lambda/2)^n}{n!} \binom{M+N-1}{k} \binom{k}{l} (-1)^l \frac{\Psi[\frac{1}{2}, \frac{3}{2} - L_0; \frac{\beta}{2\bar{x}_0}]}{(1 + \frac{M}{N}\gamma_0)^{L_0}}, \quad (15)$$

where $\Psi[\frac{1}{2}, \frac{3}{2} - L_0; \frac{\beta}{2\bar{x}_0}] \triangleq \frac{\Gamma(L_0 - \frac{1}{2})}{\Gamma(L_0)} \Phi(\frac{1}{2}, \frac{3}{2} - L_0; \frac{\beta}{2\bar{x}_0}) + \frac{\Gamma(\frac{1}{2} - L_0)}{\Gamma(\frac{1}{2})} \left(\frac{\beta}{2\bar{x}_0} \right)^{L_0 - \frac{1}{2}} \Phi(L_0, L_0 + \frac{1}{2}; \frac{\beta}{2\bar{x}_0})$ [23, 9.21-2] and $\Phi(a, b; x)$ is Kummer's confluent hypergeometric function [23, 9.210-1].

Proof. Given in Appendix E. $\square$

We should point out here the infinite series in Eq. (15). Due to the presence of the Gaussian tail function, the resulting integrals are not amenable to the manipulations that facilitated reshaping of the infinite series into finite, as in Eqs. (6) and (8). However, it is possible to characterize

T	Capacity [15, Eq. (3)]				Eq. (8)
	100	250	500	1000	
Time (msec)	275.63	1713.5	7323.78	35850.12	5.63

TABLE I
ABSOLUTE CALCULATION TIME (MSEC) FOR
 $M = 3, N = 2, K = 10\text{dB}$ AND VARIOUS VALUES OF T . THE
FORMULAS ARE EVALUATED AT $x = 0.1$.

the behavior of the truncation error using the following lemma [15], stated below without proof.

Lemma 2. *The series $S_T \triangleq e^{-\theta} \sum_{n=T}^{\infty} \frac{\theta^n}{n!}$, where θ is a parameter, satisfies $S_T < C_T(e\theta/T)^T$ for some constant C_T and index T that satisfies $T > e\theta$.*

Using Eqs. (12) and (D1) (see Appendix D) we obtain the distribution function of ρ as

$$F_{\rho}(x) = e^{-\frac{\lambda}{2}} \sum_{n=0}^{\infty} \frac{(\lambda/2)^n}{n!} \mathcal{I}\left[M_n, N; \frac{\frac{M}{N}\gamma_0(x+1)}{1 + \frac{M}{N}\gamma_0(x+1)}\right],$$

Noting that $\mathcal{I}(a, b; x) \leq 1$ for any non-negative x , one can exploit Lemma 2 ($\theta \equiv \frac{\lambda}{2}$) to show that truncating Eq. (15) to the first T terms will induce an approximation error that decays exponentially fast with T , when $T > MKe$. Thus, one can rely on this partial result to make a proper selection of the truncation parameter T , such that Eq. (15) can be approximated with very high accuracy. Numerical results support this statement (see Section V).

V. NUMERICAL RESULTS

In this section we investigate numerically different setups in order to quantify the impact of the system parameters into its performance, assessed by the metrics derived in the previous sections. To maintain consistency across the different scenarios we set $\mathbf{h}_{\text{LoS}} = \frac{1}{\sqrt{2}}(1 + i)\mathbf{1}_N$. The simulated curves, corresponding to the solution of (P0), are generated with CVX [20], assuming a block length of 2.5×10^3 independent channel states and averaging over ten transmission blocks. The theoretical curves are obtained with MATHEMATICA [27] using a gradient search method (function **FindRoot**). To obtain the BER results we truncated the infinite series to the first $T = 100$ terms, unless otherwise stated. This choice guarantees, for the setups considered herein, that the truncation error will be infinitesimal.

We begin the numerical investigation with focus on capacity and the associated waterlevel γ_0 , through Figs. 2–5. It follows from Fig. 2 that ergodic capacity scales slowly with increasing TNR q , since the multiplexing gain is bounded by the single antenna at R_x . One can see that for small values of the Rician factor, i.e., $K \in \{-20\text{dB}, -10\text{dB}\}$ the curves practically coincide. A small gain is observed between the aforementioned cases and $K = 0\text{dB}$, yet the gap becomes more salient when the contribution of the LoS component increases further to 10dB and 20dB, because randomness and dispersion of the power around the desired steering direction becomes smaller. The dual to ergodic capacity is the waterlevel γ_0 , depicted in Fig. 3 where one can observe precisely the trend described in Section III-C.⁷ In the low TNR regime, the waterlevel exhibits the predicted superlinear trend, which becomes more pronounced as K grows large,

⁷We skip the curve corresponding to $K = 20\text{dB}$ because it clutters the figure, yet for the sake of completeness we mention for $q \in \{0, 3, 6, 12\}\text{dB}$, the waterlevel γ_0 takes on the values $\{40.73, 25.27, 14.47, 7.816\}$.

Fig. 2. Capacity in nats per channel use (npcu) for a network with $M = 3, N = 2$ and different values of the Rician factor K . Solid curves correspond to the equivalent theoretical results (Eq. (8)).

Fig. 3. Simulated (CVX) versus theoretical (Eq. (6)) results for the water level for a network with $(M, N) = (3, 2)$ and different values of the Rician factor K .

whereas in the high TNR regime one can see that γ_0 converges to zero rather slowly. Finally, we observe that the asymptotic expressions (9) and (10), for small K and high SNR, approximate the exact result quite well in the regime of interest.

The impact of M and N on capacity is illustrated through Figs. 4,5. Two main conclusions are drawn from Fig. 4 with respect to the number N of nodes in $\mathcal{P}$. In particular, we see that as N increases there is a rate penalty which, however, is diminishing: i) among setups with the same Rician factor (marked by the big ellipse);

Fig. 4. Capacity vs number of transmit antennas M for a network with TNR $q = 10\text{dB}$ and various pairs (K, N) . Dotted curves, plotted on top of the associated solid lines, correspond to theoretical results (Eq. (8)).

Fig. 5. Capacity versus number of transmit antennas M for a network with TNR $q = 10\text{dB}$, various pairs (K, d) and N given as $N = M - d$. Dotted curves, plotted on top of the associated solid lines, correspond to theoretical results (Eq. (8)).

and ii) between the $K = 0\text{dB}$ and $K = 10\text{dB}$ cases, when N is fixed (small ellipses). The explanation lies in the tail of the F -distribution, which becomes lighter as N grows and hence the probability of successful transmission, i.e., a non-zero SNR ρ , within the k -th time-slot decreases. However this tail-thinning effect slows down with each increment of N and this is reflected into the diminishing rate back-off (gap), observed for fixed K and increasing N . In Fig. 5 we see the effect of simultaneously increasing both the T_x array length and the number of nodes in $\mathcal{P}$,

Fig. 6. Outage probability for a network with $M = 3$, rate $R = 2$ npcu and various (N, K) pairs. Dotted curves, plotted on top of the associated solid lines, correspond to theoretical results (Eq. (12)) and dashed curves correspond to asymptotic results (Eq. (13)).

Fig. 7. Outage probability for a network with $N = 2$, rate $R = 2$ npcu and various (M, K) pairs. Dotted curves, plotted on top of the associated solid lines, correspond to theoretical results (Eq. (12)).

using the rule $N = M - d$. Adding one element at both ends entails a penalty into system capacity because the power back-off induced by the introduced node in $\mathcal{P}$ is not compensated by adding an extra transmit antenna at the source node. The penalty is more severe for small values of M and it increases with d , however, for fixed d the suffered capacity loss becomes negligible in the regime where d is small compared to M and N .

A different point-of-entry into the system is provided via numerical evaluation of short-term performance, shown in Figs. 6 and 7. In Fig. 6 one can see that with increasing

Fig. 8. Average BER for a network with $N = 2$, QPSK modulation $(\alpha, \beta) = (1, 1)$ and various (M, K) pairs. Dotted curves, plotted on top of the associated solid lines, correspond to theoretical results (Eq. (15)).

Fig. 9. Simulated (CVX) and theoretical (Eq. (15)) average BER versus the Rician factor K for a network with $(M, N) = (2, 1)$, QPSK modulation $(\alpha, \beta) = (1, 1)$, truncation index $T = 500$ and different values of the TNR q .

Rician factor the probability of outage exhibits a consistent improvement, i.e., for each value of N we observe a gap of approximately 1dB between the $K = 0$ dB and $K = 10$ dB cases. However, for fixed K and increasing N the performance gap is diminishing. The major conclusion to draw here is that high-TNR performance conforms with the insights of Eq. (13), i.e., diversity is not affected by the size of $\mathcal{P}$ and all curves exhibit the same slope at high TNR. We note here that, in the high TNR region, there is good agreement between the exact formula-driven

curves (Eq. (12)) and their asymptotic counterparts (Eq. (13)), yet the latter do not exhibit a monotonic trend. The reason is that the dominant summand depends not only on the argument $g(x)$ (see Eq. (13)) but also on the non-centrality parameter λ (see condition (E2) in Appendix E). For moderate values of the argument $g(x)$, the condition (E2) might not be satisfied therefore approximation of the summation with the first term might not be valid, resulting in the observed behavior. The impact of K on diversity is highlighted in Fig. 7, where we see that for $K = 10$ dB the outage probability curve has a very steep slope. For fixed M we can argue, on the grounds of Figs. 6 and 7, that the system greatly benefits, in terms of outage, from large K .

Figs. 8 and 9 contain results pertaining to the BER of the system. From Fig. 8 it can be inferred that the rate at which the probability of error drops depends both on M and K . When the TR link is nearly shadowed, i.e., $K = -10$ dB, the system can still achieve sufficient detection performance, though a substantial improvement takes place when either of the parameters, M or K , increases. Keeping K fixed and varying M indicates that the array size has a strong impact on detection performance since the improvement is dual, both in the slope and shift of the curves. The alternative system snapshot, obtained by fixing M and letting K vary, shows that the Rician factor influences the BER slope only when the LoS components outweighs its shadowed counterpart ($K = 10$ dB). Fig. 9 provides another visual aid in fleshing out the role of K . The main message, embedded in this figure, is that K is mostly influential in reducing the average error probability, at high TNR. We further observe that the BER saturates after a certain threshold for K . The threshold indicates that the channel LoS component is sufficiently dominant such that the BER performance is dictated by the beam-former and not the power allocation policy. This interpretation also explains why the threshold increases with TNR. Finally, we note that for all considered scenarios we have practically perfect agreement between theoretical and numerical results even in the case of BER, where we truncated the series in Eq. (15) under the guidelines of Lemma 2 to properly select the truncation index T .

VI. SUMMARY AND CONCLUSIONS

In this work we studied a multi-node interference network employing a transmit precoding strategy that is shown to achieve the ergodic capacity under an ESIPC. For the network under consideration, we showed MRT beamforming optimality and provided an efficient characterization of the optimized utility function. The rest of the results can be briefly summarized in the following points: i) we generalized and unified in a single scenario existing results [12], [18]; ii) we improved previous mathematical results [15], which involved infinite series representations, by exploiting a technique for reshaping series summations into a finite number of terms; and iii) we derived “practical” QoS metrics such as the distribution of the TR link SNR and its associated average BER. When the series representation could not be obviated we

supplemented the analysis with approximation guarantees for the truncation error. Numerical simulations were in line with the theoretical predictions and further assisted in quantifying the impact of the network parameters into the TR link's performance.

Our numerical experiments captured several properties of the network under consideration, yet there are more aspects to be examined. Our network structure, being elementary, could serve as the baseline model in order to proceed with more advanced setups and design metrics. Some straightforward follow-up research ideas belong to the following span: i) analyze our network under its mirrored channel model, i.e., a shadowed TR link and TP_n links with LoS fading; ii) consider more involved topologies such as the relay channel; and iii) replace the ESIPC with its instantaneous counterpart or a peak transmit power constraint and further experiment with other QoS metrics, such as the mean-square-error, the outage probability or combinations thereof. Mathematical treatment of the abovementioned research problems is, in many cases, feasible because of the infinite series-resaping technique. This tool is actually applicable whenever the modified Bessel function of the first kind arises in the stochastic characterization because Bessel functions of this family are the core of the χ^2 - and the F -distribution. These mathematical tools could be useful for exploring and conducting performance analysis of network architectures with LoS fading, an area that has received far less attention in literature.

VII. ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers whose comments and suggestions helped improve the quality of the paper. The research leading to these results has received funding from the European Research Council under the European Community's Seventh Framework Programme (FP7/2007-2013) / ERC grant agreement number 228044.

APPENDIX

A. Proof of Lemma 1

We carry out the proof in reverse order, i.e., we first derive capacity for the case where the receiver has the same CSI level as the transmitter and then show that the capacity does not decrease when the receiver CSI is reduced to SNR only. Our proof relies on previous work [31] and we shall appeal to the corresponding mathematical results, wherever necessary. For consistency and ease of presentation, we adopt the notation of [31], i.e., we define:

- a) (u_k, v_k) as the i.i.d. sequence of CSI tuples at the transmitter and the receiver;
- b) $\Pr(\hat{\omega} \neq \omega)$ as the probability of error, where $\omega, \hat{\omega}$ denote the information and the estimated messages, respectively;
- c) $\mathbf{x}[k]$ as the precoded codeword that is a function of ω and the transmitter CSI $\{u_n\}_{n=1}^k$ up

to instant k and which satisfies the constraint $E\{\text{tr}(\mathbf{G}[k]\mathbf{x}[k]\mathbf{x}^H[k]\mathbf{G}[k])\} \leq Q_{\text{th}}$.

Our CSI model can then be expressed as $u_k = \{\mathbf{h}[k], z[k]\}$. According to [31, Lemma 1], one can upper-bound, by means of the Fano-inequality, the achievable rate R as follows

$$R \leq \frac{1}{K} \sum_{k=1}^K E\{\log \det(\mathbf{I}_M + \frac{1}{\sigma^2} E[\mathbf{H}[k]x[k]x^H[k]\mathbf{H}^H[k] | \{v_n\}_{n=1}^k])\} + \epsilon_K,$$

where K is the transmission block length and ϵ_K is a term that depends on $\Pr(\hat{\omega} \neq \omega)$ and satisfies $\lim_{K \rightarrow \infty} \epsilon_K = 0$. This bound holds under general CSI assumptions and codeword constraints and can be further simplified depending on the CSI availability. We consider below two separate CSI models.

For the first part of the proof we assume that the receiver and the transmitter CSI sequences are equivalent, i.e., $v_k \equiv u_k$. Then, our setup fulfils the conditions of [31, Theorem 1] according to which the ergodic capacity is achieved with Gaussian codebooks and is obtained via the solution of the following optimization problem,

$$\begin{aligned} & \underset{\mathbf{W}(u_k)}{\text{maximize}} E_{u_k} \{ \log [1 + (1/\sigma^2) \mathbf{h}[k]^H \mathbf{W}(u_k) \mathbf{W}(u_k)^H \mathbf{h}[k]] \} \\ & \text{subject to } E_{u_k} \{ \text{tr}(\mathbf{G}[k] \mathbf{W}(u_k) \mathbf{W}(u_k)^H \mathbf{G}[k]) \} \leq Q_{\text{th}}, \end{aligned} \quad (\text{P1})$$

where the index k is used to emphasize the time-varying nature of the channel, as stated in Section II-B and $\mathbf{W}(u_k) = [\mathbf{w}_1(u_k) \dots \mathbf{w}_M(u_k)]$ is a weight matrix. Note that the only difference with the setup in [31, Theorem 1] is on the codeword constraint, but it is straightforward to see that this does not affect the achievability part and converse [31, Appendix B], hence no changes are required in the respective proof and it can be invoked in its entirety. To proceed with the solution of (P1) we define the random vectors $\tilde{\mathbf{g}}_m[k] \triangleq \mathbf{G}[k] \mathbf{w}_m(u_k)$. Then, Problem (P1) can be recast as

$$\begin{aligned} & \underset{L_s, \{\mathbf{w}_m\}_{m=1}^{L_s}}{\text{maximize}} E\{ \log [1 + (1/\sigma^2) \sum_{m=1}^{L_s} |\mathbf{h}^H[k] \mathbf{w}_m(u_k)|^2] \} \\ & \text{subject to } E\{ \sum_{m=1}^{L_s} \|\tilde{\mathbf{g}}_m[k]\|^2 \} \leq Q_{\text{th}}, \end{aligned} \quad (\text{P1a})$$

where $L_s (\leq M)$ is the number of streams and the expectation is taken jointly over the distribution $\mathbf{h}[k]$ and the set of vectors $\tilde{\mathbf{g}}_m[k] \sim \mathcal{CN}(\mathbf{0}_N, \sigma_g^2 \|\mathbf{w}_m[k]\|^2 \mathbf{I}_N) \forall m$. Assume that the optimal vectors $\mathbf{w}_m^0(u_k)$ are not aligned, in general, with $\mathbf{h}[k]$. Then we can construct a new set of vectors $\mathbf{w}_m^1(u_k) = \frac{\mathbf{h}}{\|\mathbf{h}\|} \|\mathbf{w}_m^0(u_k)\|$ that satisfy the ESIPC, which depends only on the norms of $\mathbf{w}_m$'s and $\|\mathbf{w}_m^1\| = \|\mathbf{w}_m^0\|$, and yield a higher objective value, since they maximize the inner product $|\mathbf{h}^H[k] \mathbf{w}_m(u_k)|$ according to the Cauchy-Schwarz inequality. It follows that the optimal vectors should be aligned with $\mathbf{h}[k]$ implying that the optimal covariance matrix is rank-one, i.e., $L_s = 1$, and it can be decomposed as $Q(u_k) = P(u_k) \mathbf{w}_o(u_k) \mathbf{w}_o^H(u_k)$, where

$\mathbf{w}_o(u_k) = \mathbf{h}[k]/\|\mathbf{h}[k]\|^2$ is the eigenvector that corresponds to beamforming in the direction of the channel and the eigenvalue $P(u_k)$ is the power to be optimized. Invoking this result we obtain the lower-dimensional counterpart of (P1a), which is exactly Problem (P0).

For the second part of the proof, we adopt a reduced receiver CSI model, where the receiver knows *only the instantaneous SNR*, i.e., the receiver CSI sequence is $v'_k = (1/\sigma^2)\|\mathbf{h}^H[k]\mathbf{W}(u_k)\|^2$. Moreover we constrain the transmit covariance $\mathbf{Q}(u_k) = \mathbf{W}(u_k)\mathbf{W}^H(u_k)$ to lie in the set of *rank-one matrices*, i.e., we consider $\mathbf{Q}(u_k) = \mathbf{v}(u_k)\mathbf{v}^H(u_k)$.⁸ Upon these premises, the conditions of [31, Theorem 2] are satisfied and the capacity is given by

$$\begin{aligned} & \underset{\mathbf{v}(u_k)}{\text{maximize}} \quad E_{u_k} \left\{ \log \left[1 + (1/\sigma^2) |\mathbf{h}^H[k]\mathbf{v}(u_k)|^2 \right] \right\} \\ & \text{subject to} \quad E_{u_k} \left\{ \|\mathbf{G}[k]\mathbf{v}(u_k)\|^2 \right\} \leq Q_{\text{th}}. \end{aligned} \quad (\text{P2})$$

The achievability and the converse proofs are available in [31, Appendix D]. It can be easily shown, following the line-of-proof for (P1a), that the optimal beamformer is $\mathbf{v}_o(u_k) = \sqrt{P(u_k)}\mathbf{w}_o(u_k)$, $\mathbf{w}_o(u_k) = \mathbf{h}[k]/\|\mathbf{h}[k]\|$, such that (P2) becomes essentially equivalent to (P0) and in turn to Problem (P1). We observe that: a) (P2) is equivalent to (P1), i.e., the optimization problem that corresponds to the reduced receiver CSI model v'_k is equivalent to the capacity result when $v_k \equiv u_k$; and b) when $v_k \equiv u_k$ the instantaneous SNR is implicitly available at the receiver since it can be calculated from u_k , therefore the capacity for the CSI model (u_k, v_k) cannot be lower than that of the reduced CSI model (u_k, v'_k) . From these two observations we infer that (P2) is indeed the capacity.

B. Proof of Proposition 1

Evaluation of the left-hand-side (LHS) of the equation $E_\gamma \left\{ \frac{1}{\gamma_0} - \frac{1}{\gamma} \right\}^+ = \tilde{Q}$ involves the solution of the following integral

$$\begin{aligned} I_0 &= \int_{\gamma_0}^{\infty} \left(\frac{1}{\gamma_0} - \frac{1}{z} \right) f_\gamma(z) dz = \frac{\Pr(\gamma \geq \gamma_0)}{\gamma_0} - \int_{\gamma_0}^{\infty} \frac{f_\gamma(z)}{z} dz \\ &= \frac{1}{\gamma_0} \int_{\gamma_0}^{\infty} f_\gamma(z) dz - e^{-\frac{\lambda}{2}} \sum_{n=0}^{\infty} \frac{(\lambda/2)^n (M/N)}{B(M_n, N)n!} \\ &\quad \left(\frac{M}{N} \right)^{M_n-1} \int_{\gamma_0}^{\infty} \frac{1}{z} \left(\frac{1}{1 + \frac{M}{N}z} \right)^{M_n+N} z^{M_n-1} dz, \end{aligned}$$

where we used the definition $M_n \triangleq M + n$. The equation of interest involves two integrals; the first actually corresponds to the complementary distribution function $\Pr(\gamma > \gamma_0) \triangleq 1 - F_\gamma(\gamma_0)$. We begin with the second summand which is analyzed as

$$\begin{aligned} I_1 &= e^{-\frac{\lambda}{2}} \sum_{n=0}^{\infty} \frac{(\lambda/2)^n (M/N)^{M_n}}{B(M_n, N)n!} \int_{\gamma_0}^{\infty} \left(\frac{1}{1 + \frac{M}{N}z} \right)^{M_n+N} z^{M_n-1} \frac{dz}{z} \\ &\stackrel{(a)}{=} e^{-\frac{\lambda}{2}} \sum_{n=0}^{\infty} \frac{(\lambda/2)^n (M/N)}{B(M_n, N)n!} \int_{x_0}^1 (1-x)^N x^{M_n-2} dx \end{aligned}$$

⁸This constraint is imposed such that the number of transmit dimensions does not exceed that of receive dimensions as required for the application of the respective result.

$$= \frac{e^{-\frac{\lambda}{2}}}{\left(\frac{\lambda}{2}\right)^{M-1}} \int_{x_0}^1 \frac{(1-x)^N}{x\Gamma(N)} \sum_{n=0}^{\infty} \frac{\left(\frac{x\lambda}{2}\right)^{M_n-1}}{n!} \prod_{m=0}^{N-1} (M_n + m) dx, \quad (\text{B1})$$

where (a) follows by the change of variables $x \triangleq \frac{\frac{M}{N}z}{1 + \frac{M}{N}z}$, $x_0 \triangleq \frac{(M/N)\gamma_0}{1 + (M/N)\gamma_0}$, respectively. In order to reach the desired result we need to reshape the infinite series into a finite summation. This is achieved by generalizing [32, Eqs. (74),(75)] as follows

$$\begin{aligned} \frac{\partial x^n e^x}{\partial x^m} &= \sum_{k=0}^{\infty} \frac{1}{k!} \frac{\partial x^{n+k}}{\partial x^m} \\ &= \sum_{k=0}^{\infty} \frac{x^{n+k-m}}{k!} (n+k) \dots (n+k-m+1), \end{aligned} \quad (\text{B2})$$

where we have implicitly assumed that $n \geq m$, similar to the case considered in this work. As pointed out in [32], exchanging the order of the summation and the differential is allowed because the resulting series is uniformly convergent. Moreover it can be shown by mathematical induction that

$$\frac{\partial x^n e^x}{\partial x^m} = e^x \sum_{k=0}^m \binom{m}{k} \frac{n!}{(n-k)!} x^{n-k}. \quad (\text{B3})$$

If we now set $n \triangleq M + N - 1$ and $m \triangleq N$ in Eqs. (B2),(B3) and plug the result back in (B1) we can obtain the desired expression of $F_x(x)$ by following the steps below

$$\begin{aligned} I_1 &= \frac{e^{-\lambda/2} (M/N)}{\left(\frac{\lambda}{2}\right)^{M-1} \Gamma(N)} \sum_{k=0}^N \binom{N}{k} \frac{(M+N-1)!}{(M+N-1-k)!} \\ &\quad \int_{x_0}^1 \frac{(1-x)^N}{x} \left(\frac{x\lambda}{2} \right)^{M+N-1-k} \exp\left(\frac{x\lambda}{2}\right) dx \\ &\stackrel{(a)}{=} \frac{e^{-\lambda/2} (M/N)}{\left(\frac{\lambda}{2}\right)^{M-1} \Gamma(N)} \sum_{k=0}^N \sum_{m=0}^N \binom{N}{k} \frac{(M+N-1)!}{(M+N-1-k)!} \binom{N}{m} \\ &\quad \frac{(-1)^m}{(\lambda/2)^m} \int_{\frac{\lambda}{2}x_0}^{\frac{\lambda}{2}} x^{M+N-2-k+m} e^x dx \\ &= \frac{e^{-\lambda/2} (M/N)}{\left(\frac{\lambda}{2}\right)^{M-1} \Gamma(N)} \sum_{k=0}^N \sum_{m=0}^N \binom{N}{k} \frac{(M+N-1)!}{(M+N-1-k)!} \binom{N}{m} \\ &\quad \frac{(-1)^m}{(\lambda/2)^m} \left\{ \sum_{l=0}^{L-1} \frac{(L-1)!}{l!} (-1)^{L-1-l} \left[\left(\frac{\lambda}{2} \right)^l e^{\frac{\lambda}{2}} - \tilde{x}_0^l e^{\tilde{x}_0} \right] \right\}, \end{aligned}$$

where (a) follows by the binomial expansion of the term $(1-x)^N$ and setting $v \triangleq \frac{\lambda}{2}x$. The last equation follows by setting $L \triangleq M + N - 1 - k + m$, $\tilde{x}_0 \triangleq \frac{\lambda}{2}x_0 \left(= \frac{\lambda}{2} \frac{(M/N)\gamma_0}{1 + (M/N)\gamma_0} \right)$ and further using integration by parts. A special case arises when $L = 0$ and it requires separate treatment.⁹ In that case, the final integral becomes

$$\int_{\tilde{x}_0}^{\frac{\lambda}{2}} \frac{e^x}{x} dx = \text{Ei}\left(\frac{\lambda}{2}\right) - \text{Ei}(\tilde{x}_0),$$

where $\text{Ei}(x) \triangleq -\int_{-x}^{\infty} \frac{e^{-t}}{t} dt$ is the exponential integral. Since $\Pr(\gamma > \gamma_0)$ has the same form as Eq. (B1), but

⁹This event occurs only if $M = 1$ and affects the solution of just one of the terms of the outer summations, namely the one that corresponds to the triplet $(M, k, m) = (1, N, 0)$.

without the x at the denominator, its solution can be found following the same steps. For brevity, we do not repeat the calculations and simply mention that no special case for L arises here. Therefore, as a byproduct we obtain a finite summation formula for the distribution function of an F -distributed random variable, which applies to the case where both nominator and denominator have an even number of degrees-of-freedom.

Assembling the above results leads to Eq. (6). Thus, we have been able to express previously established results [15], which involved infinite series, in terms of finite summations, using (B3) and this is a novel technical contribution. The finite representation is especially useful for solving the associated Eq. (6) since one does not have to compromise accuracy by resorting to some approximation.¹⁰

C. Proof of Proposition 2

The ergodic capacity is found by averaging over all possible channel states as follows

$$C = E\gamma \left\{ \log \left[1 - \left\{ \frac{c\gamma}{\lambda_0 \sigma^2} - 1 \right\}^+ \right] \right\} = \int_{\gamma_0}^{\infty} \log \left(\frac{z}{\gamma_0} \right) f_{\gamma}(z) dz \\ = \int_{\gamma_0}^{\infty} \log \left(\frac{M}{N} z \right) f_{\gamma}(z) dz - \log \left(\frac{M}{N} \gamma_0 \right) \Pr(\gamma > \gamma_0).$$

The term $\Pr(\gamma > \gamma_0)$ has already been considered in Appendix B. Therefore, in order to evaluate the capacity we need to resolve only the first summand S (the integral) which has the following form

$$S = e^{-\frac{\lambda}{2}} \sum_{n=0}^{\infty} \frac{(\lambda/2)^n}{B(M+n, N)n!} \left(\frac{M}{N} \right)^{M+n} \\ \int_{\gamma_0}^{\infty} \log \left(\frac{M}{N} z \right) \left(\frac{1}{1 + \frac{M}{N} z} \right)^{M+N+n} z^{M-1+n} dz \\ = \frac{e^{-\lambda/2}}{(\frac{\lambda}{2})^{M-1} \Gamma(N)} \sum_{k=0}^N \binom{N}{k} \frac{(M+N-1)!}{(M+N-1-k)!} \frac{1}{(\frac{\lambda}{2})^N} \\ \int_{\hat{x}_0}^{\frac{\lambda}{2}} \left(\frac{\lambda}{2} - x \right)^{N-1} x^{M+N-1-k} e^x \log \left(\frac{x}{\frac{\lambda}{2} - x} \right) dx, \quad (C1)$$

where the second equality follows from the exact same manipulations used for the derivation of (B3). To facilitate the solution of the above integral we split the logarithm into two terms in order to deal with two simpler integrals. For ease of exposition we set $L_1 \triangleq M+N-1-k+m$ and $L_2 \triangleq N-1+m$. Then the first resulting integral is expressed as

$$S_1 = \sum_{m=0}^{N-1} \binom{N-1}{m} \left(\frac{\lambda}{2} \right)^{L_2} (-1)^m \underbrace{\int_{\hat{x}_0}^{\frac{\lambda}{2}} x^{L_1} e^x \log x dx}_{S'_1}. \quad (C2)$$

¹⁰The term approximation here refers to the truncation of the sequence in a practical implementation since only a finite number of terms can be considered. We stress this to differentiate from numerical approximation which stems from insufficient arithmetic precision.

To reach the solution of S'_1 we perform integration by parts (viewing e^t and $t^{L_1} \log t$ as two separate functions) and using induction we obtain the following result

$$S'_1 = \sum_{l=0}^{L_1-1} \frac{L_1!}{(L_1-l)!} (-1)^l x^{L_1-l} e^x \log x \Big|_{\hat{x}_0}^{\frac{\lambda}{2}} + \\ \sum_{l=1}^{L_1} \sum_{j=0}^{L_1-l} \frac{L_1! (-1)^{L_1-j}}{j! (L_1-l+1)!} x^j e^x \Big|_{\hat{x}_0}^{\frac{\lambda}{2}} + (L_1)! (-1)^{L_1} \\ \left[e^{\frac{\lambda}{2}} \log \left(\frac{\lambda}{2} \right) - e^{\hat{x}_0} \log \hat{x}_0 + \text{Ei}(\hat{x}_0) - \text{Ei} \left(\frac{\lambda}{2} \right) \right]. \quad (C3)$$

In the same spirit we derive the second term of the integral. Performing the change of variables $t \triangleq \frac{\lambda}{2} - x$ and defining $\hat{x}_0 \triangleq \frac{\lambda}{2} - \hat{x}_0$ transforms the second term as follows

$$S_2 = \int_0^{\hat{x}_0} x^{N-1} \left(\frac{\lambda}{2} - x \right)^{M+N-1-k} e^{\frac{\lambda}{2}-x} \log x dx \\ = e^{\frac{\lambda}{2}} \sum_{m=0}^{M+N-1-k} \binom{M+N-1-k}{m} \\ \left(\frac{\lambda}{2} \right)^{M+N-1-k-m} (-1)^m \underbrace{\int_0^{\hat{x}_0} t^{L_2} e^{-t} \log t dt}_{S'_2}. \quad (C4)$$

Again we employ integration by parts (now viewing e^{-t} and $t^{L_2} \log t$ as the two functions) and further make use of [23, Eqs. 4.331.1-2]. These two combined steps yield the following result

$$S'_2 = -e^{-\hat{x}_0} \log \hat{x}_0 \sum_{l=0}^{L_2-1} \frac{L_2!}{(L_2-l)!} \hat{x}_0^{L_2-l} \\ + \sum_{l=1}^{L_2} \frac{(L_2)!}{(L_2-l+1)!} \left[1 - \sum_{j=0}^{L_2-l} \frac{(L_2-l)!}{j!} \hat{x}_0^j e^{-\hat{x}_0} \right] \\ + L_2! [\text{Ei}(-\hat{x}_0) - C_{\text{Eu}} - e^{-\hat{x}_0} \log \hat{x}_0], \quad (C5)$$

where $C_{\text{Eu}} \approx 0.57721$ is the Euler-Mascheroni constant. Plugging Eqs. (C3) and (C5) into (C2) and (C4), respectively, and the resulting expressions back in (C1) yields Eq. (8) and concludes the proof.

D. Proof of Proposition 3

The proof of this result relies on proper partitioning of the event $\Omega_{\rho} \triangleq \{\rho \leq x\}$. Recalling that $\rho = \{\frac{\gamma}{\gamma_0} - 1\}^+$ we can rewrite Ω_{ρ} as

$$\Omega_{\rho} = \{\gamma \leq \gamma_0(x+1) \cap \{\gamma \leq \gamma_0 \cup \gamma > \gamma_0\}\} \\ = \{\gamma \leq \gamma_0(x+1) \cap \gamma \leq \gamma_0\} \cup \{\gamma \leq \gamma_0(x+1) \cap \gamma > \gamma_0\} \\ = \mathbb{I}_{\{x=0\}} \{\gamma \leq \gamma_0 \cup \emptyset\} \\ + \mathbb{I}_{\{x>0\}} \underbrace{\{\gamma \leq \gamma_0 \cup \gamma_0 < \gamma \leq \gamma_0(x+1)\}}_{\gamma \leq \gamma_0(x+1)} \\ = \mathbb{I}_{\{x \geq 0\}} \{\gamma \leq \gamma_0(x+1)\}. \quad (D1)$$

Using this decomposition for Ω_{ρ} one can now express the distribution function $F_{\rho}(x)$ as

$$F_{\rho}(x) = \Pr(\Omega_{\rho}) = \Pr[\gamma \leq \gamma_0(x+1)] = F_{\gamma}[\gamma_0(x+1)],$$

which holds for $x \geq 0$. This completes the proof.

E. Proof of Proposition 4

We start from the following general representation of the distribution function $F_x(x) = \int_0^x f_x(y)dy$ of an F -distributed random variable x ,

$$F_x(x) = e^{-\frac{\lambda}{2}} \sum_{n=0}^{\infty} \underbrace{\frac{(\lambda/2)^n}{n!} \mathcal{I}(h(x); M+n, N)}_{c_n}. \quad (\text{E1})$$

The argument $h(x) \triangleq \frac{Mx}{N+Mx}$ is smaller than one for any non-negative x and $\mathcal{I}(x; a, b) \triangleq \frac{\mathcal{B}(x; a, b)}{\mathcal{B}(a, b)}$ is the *regularized incomplete beta function* [23, 8.392] that, for integer arguments a and b , can be expanded into a summation formula as $\mathcal{I}(x; a, b) = \sum_{m=a}^{a+b-1} \binom{a+b}{m} x^m (1-x)^{a+b-1-m}$. Prior to analyzing the behavior of $F_x(x)$, for small values of its argument, one has to establish conditions on the ordering of the summands c_n . In particular, the inequality $c_n \geq c_{n+1}$ translates to

$$\begin{aligned} \lambda &\stackrel{(a)}{\leq} 2(n+1) \frac{\mathcal{I}(h(x); M+n, N)}{\mathcal{I}(h(x); M+n, N) - \frac{[h(x)]^{M_n} [1-h(x)]^{N-n}}{M_n B(M_n, N)}} \\ &\leq 2(n+1) \frac{1}{h(x)}, \end{aligned} \quad (\text{E2})$$

where (a) follows by the property $\mathcal{I}(x; M+1, N) = \mathcal{I}(x; M, N) - \frac{x^M (1-x)^N}{MB(M, N)}$ and the asymptotic inequality follows after the fact that $\mathcal{I}(h(x); M+n, N) \approx \frac{[h(x)]^{M_n} [1-h(x)]^{N-1}}{M_n B(M_n, N)}$ and further algebraic steps. For sufficiently small values of x , this inequality will always be satisfied, implying that c_0 will be the dominant term and $F_x(x) \approx e^{-\frac{\lambda}{2}} \frac{[h(x)]^M [1-h(x)]^{N-1}}{MB(M, N)}$. Substituting $\lambda = 2MK$ and replacing $h(x)$ with $g(x)$ yields the desired result and completes the proof.

F. Proof of Proposition 5

To obtain the error probability one needs to evaluate the expectation involved in Eq. (14). Some steps, towards to the solution, are shown below

$$\begin{aligned} P_e &= \alpha \int_0^\infty Q(\sqrt{\beta x}) f_p(x) dx \\ &\stackrel{(a)}{=} \alpha \left[\int_0^\infty Q(\sqrt{\beta x}) f_\gamma[\gamma_0(x+1)] \gamma_0 dx + Q(0) F_\gamma(\gamma_0) \right] \\ &= \alpha \left[Q(\sqrt{\beta x}) F_\gamma[\gamma_0(x+1)] \Big|_0^\infty \right. \\ &\quad \left. + \frac{1}{2} \sqrt{\frac{\beta}{2\pi}} \int_0^\infty \frac{e^{-\frac{\beta}{2}x}}{\sqrt{x}} F_\gamma[\gamma_0(x+1)] dx + Q(0) F_\gamma(\gamma_0) \right] \\ &= \frac{\alpha}{2} \sqrt{\frac{\beta}{2\pi}} e^{-\frac{\lambda}{2}} \sum_{n=0}^{\infty} \sum_{k=M_n}^{M_n+N-1} \binom{M_n+N-1}{k} \frac{(\lambda/2)^n}{n!} \\ &\quad \int_0^\infty \left[\frac{\frac{M}{N} \gamma_0(x+1)}{1 + \frac{M}{N} \gamma_0(x+1)} \right]^k \frac{(e^{-\frac{\beta}{2}x}/\sqrt{x}) dx}{[1 + \frac{M}{N} \gamma_0(x+1)]^{M_n+N-1-k}} \\ &= \frac{\alpha}{2} \sqrt{\frac{\beta}{2\pi}} e^{-\frac{\lambda}{2}} \sum_{n=0}^{\infty} \sum_{k=M_n}^{M_n+N-1} \sum_{l=0}^k \binom{M_n+N-1}{k} \\ &\quad \binom{k}{l} \frac{(\lambda/2)^n}{n!} (-1)^l \int_0^\infty \frac{e^{-\frac{\beta}{2}x}}{\sqrt{x}} \frac{dx}{[1 + \frac{M}{N} \gamma_0(x+1)]^{L_0}} \end{aligned}$$

$$\stackrel{(b)}{=} \frac{\alpha}{2} \sqrt{\frac{\beta}{2\pi}} e^{-\frac{\lambda}{2}} \sum_{n=0}^{\infty} \sum_{k=M_n}^{M_n+N-1} \sum_{l=0}^k \binom{M_n+N-1}{k} \binom{k}{l} \frac{(\lambda/2)^n}{n!} (-1)^l \sqrt{\frac{\pi}{x_0}} \frac{\Psi(\frac{1}{2}, \frac{3}{2} - L_0; \frac{\beta}{2x_0})}{(1 + \frac{M}{N} \gamma_0)^{L_0}},$$

where (a) follows by plugging (12) into (14) and the last equation follows by some algebraic steps which give rise to the integral of the Ψ function.

REFERENCES

- [1] A. B. Carleial, "Interference channels," *IEEE Trans. Inf. Theory*, vol. 24, no. 1, pp. 60–70, Jan. 1978.
- [2] J. Andrews, "Interference cancellation for cellular systems: a contemporary overview," *IEEE Wirel. Commun. Magazine*, vol. 12, no. 2, pp. 19–29, Apr. 2005.
- [3] R. Valenzuela, "Fundamental limits and evolution of broadband wireless access networks," *Texas Wirel. Symp.*, Austin, TX; <http://www.ece.utexas.edu/wncg/events/symposium2005/>, Nov. 2005.
- [4] H. Huh, H. C. Papadopoulos, and G. Caire, "Multiuser MISO transmitter optimization for intercell interference mitigation," *IEEE Trans. Sig. Proc.*, vol. 58, no. 8, Aug. 2010.
- [5] M. Gastpar, "On capacity under receive and spatial spectrum-sharing constraints," *IEEE Trans. Inf. Theory*, vol. 53, no. 2, pp. 471–487, Feb. 2007.
- [6] A. Goldsmith, S. A. Jafar, I. Maric, and S. Srinivasa, "Breaking spectrum gridlock with cognitive radios: An information theoretic perspective," *Proceedings of the IEEE*, vol. 97, no. 5, pp. 894–914, May 2009.
- [7] A. Ghasemi and E. S. Sousa, "Fundamental limits of spectrum-sharing in fading environments," *IEEE Trans. Wirel. Commun.*, vol. 6, no. 2, pp. 649–658, Feb. 2007.
- [8] L. Musavian and S. Aissa, "Ergodic and outage capacities of spectrum sharing systems in fading channels," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Washington, DC, USA, 2007, pp. 3327–3331.
- [9] X. Kang, Y. C. Liang, A. Nallanathan, H. K. Garg, and R. Zhang, "Optimal power allocation for fading channels in cognitive radio networks: Ergodic capacity and outage capacity," *IEEE Trans. Wirel. Commun.*, vol. 8, no. 2, pp. 940–950, Feb. 2009.
- [10] V. Chandrasekhar, J. Andrews, and A. Gatherer, "Femtocell networks: A survey," *IEEE Commun. Mag.*, vol. 46, no. 9, pp. 59–67, Sep. 2008.
- [11] F. R. Farrokhi, G. J. Foschini, A. Lozano, and R. Valenzuela, "Link optimal space-time processing with multiple transmit and receive antennas," *IEEE Commun. Lett.*, vol. 5, no. 3, pp. 85–87, Mar. 2001.
- [12] H. Suraweera, J. Gao, P. Smith, M. Shafi, and M. Faulkner, "Channel capacity limits of cognitive radio in asymmetric fading environments," in *Proc. IEEE Int. Conf. on Commun. (ICC)*, Beijing, China, May 2008, pp. 4048–4053.
- [13] E. Nekouei, H. Inaltekin and S. Dey, "Throughput scaling in cognitive multiple access with average power and interference constraints," in *Proc. IEEE Int. Conf. on Commun. (ICC)*, Kyoto, Japan, Jun. 2011, pp. 1–6.
- [14] M. F. Hanif, and P. J. Smith, "On the statistics of cognitive radio capacity in shadowing and fast fading environments," *IEEE Trans. Wirel. Commun.*, vol. 9, no. 2, pp. 844–852, Feb. 2010.
- [15] E. Stathakis, L. K. Rasmussen, and M. Skoglund, "Closed-form capacity formula for multi-antenna cognitive radio networks with asymmetric fading," in *Proc. IEEE Int. Conf. on Commun. (ICC)*, Sydney, Australia, Jun. 2014.
- [16] S. M. Kay, *Fundamentals of Statistical Signal Processing: Detection Theory*, Prentice-Hall, Upper Saddle River, NJ, 1998.
- [17] P. B. Patnaik, "The non-central χ^2 - and F -distributions and their applications," *Biometrika*, vol. 36, no. 1/2, pp. 202–232, Jun. 1949.

- [18] R. Duan, M. Elmusrati, R. Jantti, and R. Virrankoski, "Capacity for spectrum sharing cognitive radios with MRC diversity at the secondary receiver under asymmetric fading," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Miami, USA, December 2010, pp. 1–5.
- [19] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge University Press, 2005.
- [20] M. Grant and S. Boyd, CVX: Matlab Software for Disciplined Convex Programming, version 1.21, <http://cvxr.com/cvx>, Apr. 2011.
- [21] G. Grimmett and D. Stirzaker, *Probability and Random Processes*. third ed., Oxford University Press, Oxford, 2001.
- [22] J. Proakis, *Digital Communications*, 4th ed., New York: McGraw-Hill, 2000.
- [23] I. S. Gradshteyn, and I. M. Ryzhik, *Table of integrals, series, and products*, seventh ed., Academic Press, London 2007.
- [24] W. Kaplan, *Advanced Calculus*. 5th ed., Reading: Addison-Wesley, 2002.
- [25] S. Wang, and H. L. Gray, "Approximating tail probabilities of noncentral distributions," *Computational Statistics and Data Analysis*, vol. 15, no. 3, pp. 343–352, Mar. 1993.
- [26] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, tenth pr., New York: Dover, 1972.
- [27] Wolfram Research, Inc., *Mathematica*, Version 9.0, Champaign, Illinois, 2012.
- [28] S. K. Ryhove, G. E. Øien, and F. Bøhagen, "Subchannel SNR distributions in dual-branch MIMO systems," in *Proc. IEEE/ITG Intern. Workshop on Smart Antennas*, Duisburg, Germany, Apr. 2005.
- [29] Z. Wang and G. B. Giannakis, "A simple and general parameterization quantifying performance in fading channels," *IEEE Trans. Commun.*, vol. 51, no. 8, pp. 1389–1398, Aug. 2003.
- [30] A. Goldsmith, *Wireless Communications*. New York: Cambridge Univ. Press, 2005.
- [31] M. Skoglund and G. Jørgen, "On the capacity of a multiple-antenna communication link with channel side information," *IEEE J. Select. Areas Commun.*, vol. 21, no. 3, pp. 395–405, Apr. 2003.
- [32] M. J. Gans, "The effect of Gaussian error in maximal ratio combiners," *IEEE Trans. Commun.*, vol. 19, no. 4, pp. 492–500, Aug. 1971.