

Finite-SNR Bounds on the Sum-Rate Capacity of Rayleigh Block-Fading Multiple-Access Channels with no a Priori CSI

Rahul Devassy, *Student Member, IEEE*, Giuseppe Durisi, *Senior Member, IEEE*, Johan Östman, Wei Yang, *Student Member, IEEE*, Tome Eftimov, Zoran Utkovski, *Member, IEEE*

Abstract—We provide nonasymptotic upper and lower bounds on the sum-rate capacity of Rayleigh block-fading multiple-access channels for the set up where *a priori* channel state information is not available. The upper bound relies on a dual formula for channel capacity and on the assumption that the users can cooperate perfectly. The lower bound is derived assuming a noncooperative scenario where each user employs unitary space-time modulation (independently from the other users). Numerical results show that the gap between the upper and the lower bound is small already at moderate SNR values. This suggests that the sum-rate capacity gains obtainable through user cooperation are minimal for the scenarios considered in the paper.

I. INTRODUCTION

The multiple-access channel (MAC) models a scenario where two or more noncooperating users communicate with a single receiver. This scenario is relevant for the uplink of wireless cellular networks, where the users may be mobile terminals and the receiver may be a cellular base station. In this paper, we consider the setup where neither the users nor the receiver have *a priori* information on the realization of the fading process. Such a situation arises in high mobility scenarios where it is not desirable for the receiver to feed back channel state information (CSI) to the users because it may be outdated [1]. It may also arise in the initialization phase of a communication link, e.g., when a mobile terminal joins a cellular network. Capacity analyses under assumption of no *a priori* CSI have the advantage of capturing the cost of estimating the fading channel, and, hence, yield more realistic throughput estimates than analyses based on the assumptions of perfect CSI [2]–[4]. Indeed, under the assumption of no feedback from the receiver to the user terminals,

This work was supported in part by the Swedish Research Council under grant 2012-4571 and in part by the German Research Society (Deutsche Forschungsgemeinschaft), via the project Li 659/13-1. The simulations were performed on resources at Chalmers Centre of Computational Science and Engineering (C3SE) provided by the Swedish National Infrastructure for Computing (SNIC).

R. Devassy, G. Durisi and W. Yang are with the Department of Signals and Systems, Chalmers University of Technology, Gothenburg, Sweden (e-mail: {devassy,durisi,ywei}@chalmers.se).

J. Östman is with the Department of Electronic and Computer Engineering, Hong Kong University of Science and Technology, Hong Kong, China (email: eejohan@ust.hk).

T. Eftimov is with the Laboratory for Complex Systems and Networks, Macedonian Academy of Sciences and Arts, R. Macedonia. (email: tomeeftimov@gmail.com).

Z. Utkovski is with the Faculty of Computer Science, University Goce Delčev Štip, and with the Laboratory for Complex Systems and Networks, Macedonian Academy of Sciences and Arts, R. Macedonia. (e-mail: zoran.utkovski@ugd.edu.mk).

these capacity analyses provide a fundamental limit on the performance of every communication scheme, irrespectively of whether it relies on explicit channel estimation or not.

In this paper, we consider the case where the users as well as the receiver are equipped with one or more antennas. We shall focus on the so-called Rayleigh block-fading model [5], [6]. The two key features of this model are that i) the fading coefficients associated to the channels between each transmit and receive antenna pair are independent and identically distributed (i.i.d.) circularly symmetric complex Gaussian random variables; ii) each fading coefficient remains constant over t_c channel uses before changing to a new independent realization. The parameter t_c , which is the ratio between the channel coherence time and the symbol duration, will be referred to in this paper as *coherence interval*.

The capacity of Rayleigh block-fading channels under the assumption of no *a priori* CSI has been studied extensively in the point-to-point case. Specifically, considering a system with n_t transmit and n_r receive antennas, Marzetta and Hochwald proved that the capacity-achieving input matrix $\mathbf{X} \in \mathbb{C}^{n_t \times t_c}$ can be expressed as [6, Thm. 2]

$$\mathbf{X} = \mathbf{D}\mathbf{Q} \quad (1)$$

where \mathbf{D} is a real $n_t \times n_t$ diagonal matrix whose diagonal entries have a joint probability density function (p.d.f.) that is invariant to permutation of its arguments, and $\mathbf{Q} \in \mathbb{C}^{n_t \times t_c}$, independent of \mathbf{D} , is an isotropically distributed matrix with orthonormal rows (truncated unitary matrix).

In spite of the partial characterization of the capacity-achieving input distribution provided in [6, Thm. 2], no closed form expression for capacity is available to date. However, the high-SNR capacity behavior is well understood. Indeed, extending a result obtained for the single-input single-output case in [5], Zheng and Tse [7] proved that in the high-SNR regime, the capacity \mathcal{C} of a $n_t \times n_r$ multiple-input multiple-output (MIMO) Rayleigh block-fading channel behaves as

$$\mathcal{C}(\rho) = n^* \left(1 - \frac{n^*}{t_c}\right) \log(\rho) + \mathcal{O}(1). \quad (2)$$

Here, ρ stands for the SNR, $n^* = \min(n_t, n_r, \lfloor t_c/2 \rfloor)$, and $\mathcal{O}(1)$ indicates a function whose magnitude is upper-bounded by a finite constant for sufficiently large SNR values. The asymptotic expression (2) can be tightened to [7], [8]

$$\mathcal{C}(\rho) = n^* \left(1 - \frac{n^*}{t_c}\right) \log(\rho) + c + o(1) \quad (3)$$

where c is the constant given in [8, Eq. (9)] and $o(1)$ denotes a function that vanishes as $\rho \rightarrow \infty$. For the case when $t_c \geq n_t + n_r$, one can achieve (3) by choosing \mathbf{D} in (1) to be a scaled identity matrix [7]. The resulting probability distribution on \mathbf{X} is commonly referred to as unitary space-time modulation (USTM) [5]. When $t_c < n_t + n_r$, the matrix \mathbf{D} must be chosen so that its diagonal entries are distributed as the square root of the eigenvalues of a Beta-distributed random matrix of appropriate size [8]. The resulting probability distribution on \mathbf{X} is referred to in [8] as Beta-variate space-time modulation (BSTM).

Nonasymptotic (i.e., finite-SNR) lower bounds on the capacity of point-to-point MIMO Rayleigh block-fading channels have been obtained for specific probability distributions on \mathbf{X} . Specifically, an i.i.d. Gaussian input distribution is considered in [9], USTM in [10], and BSTM in [11], where the analysis is also extended to Rician block-fading and to land mobile satellite channels. A key tool in the derivation of these nonasymptotic lower bounds is the Itzykson-Zuber integral [12, Eq. (3.4)], which allows one to obtain a closed-form expression for the conditional probability distribution of the channel output given the diagonal input matrix \mathbf{D} in (1). The method employed so far to assess the tightness of the bounds obtained in [9]–[11] is to compare these lower bounds with the asymptotic expansion (3) (with the $o(1)$ term omitted). Unfortunately, this method is not conclusive because the error incurred by omitting the $o(1)$ term in (3) is not quantified. A simple capacity upper bound can be obtained by assuming that a genie provides the receiver with perfect CSI. However, this bound is tight only when t_c is large and the channel estimation overhead negligible.

Leaving the point-to-point case and moving to the MAC, we note that the independence constraint on the signals transmitted by the various users implies that the partial characterization of the capacity-achieving input distribution obtained for the point-to-point case in [6, Thm. 2], as well as the asymptotic capacity expansions in [7], [8] and the nonasymptotic capacity lower bounds in [10], [11] do not carry over to the MAC sum-rate capacity. The only exception is the i.i.d. Gaussian lower bound obtained in [9], which also applies to the MAC because the transmission of i.i.d. Gaussian signals does not require coordination among the users.

Coarse upper and lower bounds on the MAC sum-rate capacity for the case when n_u single-antenna users communicate with a receiver equipped with n_r antennas are provided in [13]. By examining these bounds in two different asymptotic regimes (high SNR and large t_c , for a fixed n_u/t_c ratio) the authors conjecture that the sum-rate capacity is maximized when $n_u = t_c$. In the same paper, the authors pose the question of whether the constraint that the users transmit independent signals yields a sum-rate capacity *prelog*¹ that is strictly lower than the one achievable when the users can cooperate perfectly, and the MAC reduces to a point-to-point MIMO channel, for which (2) holds. It follows from [7, Sec. V] that this is not the case, provided that the users are able to transmit orthogonal pilot signals, used at the receiver to estimate the channel.

¹The *prelog* (a.k.a. multiplexing gain) is the asymptotic ratio between the (sum-rate) capacity and $\log(\rho)$ in the limit $\rho \rightarrow \infty$.

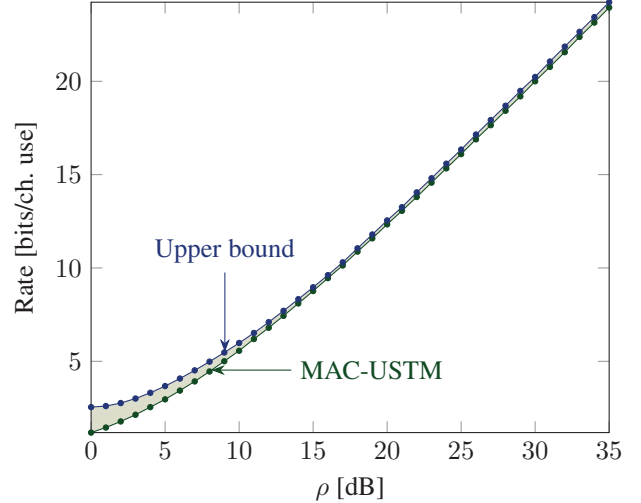


Fig. 1. Upper and lower bounds on the MAC sum-rate capacity: 4 single-antenna users; receiver with 4 antennas; coherence interval of 10 channel uses.

Lin and Moser [14] characterized the high-SNR behavior of the sum-rate capacity of an i.i.d. Rician-fading MAC (block-fading channel with coherence interval $t_c = 1$). They showed that the sum-rate capacity grows double-logarithmically in SNR, and that the sum-rate capacity maximizing strategy at high SNR is to switch off all users but one.

Contributions: We present nonasymptotic (i.e., finite-SNR) upper and lower bounds on the sum-rate capacity of Rayleigh block-fading MACs. Similarly to the nonasymptotic capacity lower bounds previously reported in [9]–[11], our bounds are not in closed form, but they can be evaluated numerically. Our upper bound is obtained by assuming that the users can perfectly cooperate, which turns the MAC into an equivalent MIMO point-to-point channel. In addition, we use the duality upper bound on mutual information reported in [15, Eq. (186)]. As *auxiliary* output distribution in the duality step, we choose the one induced by USTM inputs in the absence of additive noise. This method was used in [8] to establish the asymptotic expansion (3) for the case $n_t \leq \min(n_r, \lfloor t_c/2 \rfloor)$. Here, we provide a finite-SNR analysis, which generalizes to MIMO the one reported in [4] for the single-input single-output (SISO) case.

The nonasymptotic lower bound is obtained by allowing each user to transmit a USTM signal over the available antennas. We call the resulting input distribution MAC-USTM. Note that MAC-USTM does not yield a global USTM input distribution, because independence among users imply that orthogonality among them cannot be enforced. Numerical evidence (see Fig. 1 and Fig. 2) suggests that for the scenarios considered in the paper:

- Our upper and lower bounds characterize accurately the sum-rate capacity (see Fig. 1).
- The gain in sum-rate capacity obtainable by allowing user co-operation is minimal. This follows because the upper bound is obtained under the assumption of perfect cooperation between users. A similar observation has been recently reported in [16] for the nonfading asynchronous Gaussian MAC.

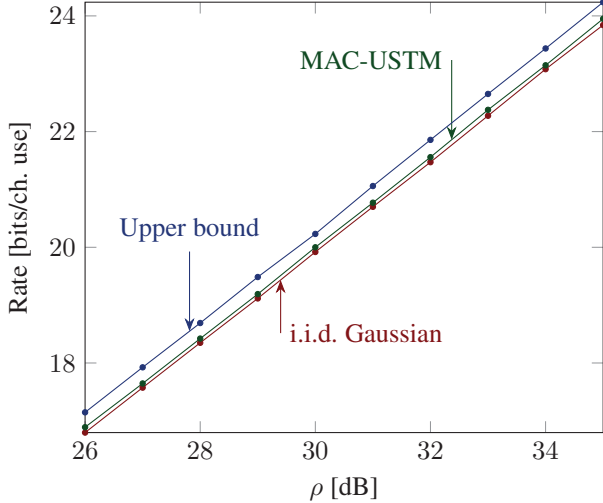


Fig. 2. Upper bound, MAC-USTM lower bound, and i.i.d. Gaussian lower bound on the MAC sum-rate capacity: 4 single-antenna users; receiver with 4 antennas; coherence interval of 10 channel uses.

- Orthogonal pilot transmission is not required to obtain rates close to the full-cooperation case (cf. [13], [7, Sec. V]).
- The MAC-USTM lower bound is tighter than the i.i.d. Gaussian lower bound, although the gain is marginal when the users are equipped with a single antenna (see Fig. 2). This suggests that i.i.d. Gaussian inputs are almost sum-rate capacity optimal already at moderate SNR values, which confirms an observation reported in [9] for the point-to-point MIMO case.

Notation: Uppercase letters denote matrices,² lowercase letters designate scalars, and boldface letters denote random quantities. The superscript \dagger stands for Hermitian transposition. For a random variable \mathbf{x} with p.d.f. $f_{\mathbf{x}}(x)$, we write $\mathbf{x} \sim f_{\mathbf{x}}(x)$. When two random variables \mathbf{x} and \mathbf{y} have the same p.d.f., we write $\mathbf{x} \sim \mathbf{y}$. For a matrix X , we let X_{ij} denote its entry on the i th row and the j th column. With $\det(f(i, j))$ we indicate the determinant of a matrix whose entry on the i th row and the j th column is given by $f(i, j)$, for some arbitrary function $f(\cdot, \cdot)$. It will turn out convenient to define the following function

$$\mathcal{P}(n) = \begin{cases} \prod_{i=1}^n \Gamma(i), & n \in \mathbb{N} \\ 1, & n = 0 \end{cases} \quad (4)$$

where $\Gamma(\cdot)$ is the Gamma function [17, Sec. 6.1]. With $\beta(\cdot, \cdot)$ we denote the Beta function [17, Sec. 6.2] and $\psi(\cdot)$ denotes the Digamma function [17, Sec. 6.3]. The identity matrix of dimension $n \times n$ is denoted by I_n . We let $\text{diag}\{d_1, d_2, \dots, d_n\}$ be the diagonal matrix with entries d_1, d_2, \dots, d_n on its diagonal. For a Hermitian matrix $A \in \mathbb{C}^{n \times n}$ with ordered eigenvalues $a_1 > a_2 > \dots > a_n$ we denote the determinant of the Vandermonde matrix constructed from a_1, a_2, \dots, a_n as

$$\mathcal{V}(A) = \prod_{1 \leq i < j \leq n} (a_i - a_j). \quad (5)$$

²We do not distinguish between vectors and matrices. We treat n -dimensional vectors as $1 \times n$ or $n \times 1$ matrices.

We shall often use the following two functions

$$\mathcal{K}(A, k) = \mathcal{V}(A) \det(A)^k \quad (6)$$

where $A \in \mathbb{C}^{n \times n}$ is Hermitian and $k \in \mathbb{N}$, and

$$\gamma(x, n) = e^x - \sum_{k=0}^{n-1} \frac{x^k}{k!} \quad (7)$$

where $x \in \mathbb{R}$ and $n \in \mathbb{N}$. The set of unitary matrices in $\mathbb{C}^{n \times n}$ is denoted by $\mathcal{U}(n)$ (*unitary group*) and the set of matrices $U \in \mathbb{C}^{n \times m}$, $m \geq n$ with $UU^\dagger = I_n$ is denoted by $\mathcal{S}(n, m)$ (*Stiefel manifold*). With $\mathbb{E}_{\mathbf{x}}[f(\mathbf{x})]$, we denote the expectation of the function $f(\mathbf{x})$ over the random variable \mathbf{x} . We let $h(\mathbf{x})$ denote the differential entropy of a continuous random variable \mathbf{x} ; furthermore, $\mathcal{I}(\mathbf{y}; \mathbf{z})$ stands for the mutual information between the random variables \mathbf{y} and \mathbf{z} . The set of all diagonal matrices in $\mathbb{R}^{n \times m}$ with ordered and distinct positive entries on their main diagonal is denoted by $\mathbb{D}^{n \times m}$. With $0_{n \times m}$ we indicate the $n \times m$ zero matrix. We use $\mathcal{CN}(0, \sigma^2)$ to denote a circularly symmetric complex Gaussian random variable with zero mean and variance σ^2 , and $\text{Beta}(a, b)$ to denote a Beta-distributed random variable with parameters a and b .

II. SYSTEM MODEL

We consider a Rayleigh block-fading MAC where n_u users communicate with a receiver having n_r antennas, and the channel coherence interval is t_c (same for all users, which corresponds to a scenario where users with similar mobility requirements are scheduled together). We assume that each user is equipped with one or more antennas and denote by n_i the number of antennas at user i , $i = 1, \dots, n_u$. The received signal $\mathbf{Y} \in \mathbb{C}^{n_r \times t_c}$ over an arbitrary coherence interval can be compactly written in matrix notation as follows:

$$\mathbf{Y} = \sum_{i=1}^{n_u} \mathbf{S}_i \mathbf{X}_i + \mathbf{W}. \quad (8)$$

Here, $\mathbf{X}_i \in \mathbb{C}^{n_i \times t_c}$ denotes the signal transmitted by user i over the coherence interval, and the matrix $\mathbf{S}_i \in \mathbb{C}^{n_r \times n_i}$ contains the fading coefficients associated to the channels between each transmit antenna of user i and the receive antennas, within the coherence interval. We assume that \mathbf{S}_i has i.i.d. $\mathcal{CN}(0, 1)$ entries and that the channel matrices $\{\mathbf{S}_i\}_{i=1}^{n_u}$ are independent. Finally, the matrix $\mathbf{W} \in \mathbb{C}^{n_r \times t_c}$, whose entries are i.i.d. $\mathcal{CN}(0, 1)$ -distributed, denotes the additive noise. Let

$$n_t = \sum_{i=1}^{n_u} n_i \quad (9)$$

be the total number of transmit antennas. We can rewrite (8) as

$$\mathbf{Y} = \mathbf{S} \mathbf{X} + \mathbf{W} \quad (10)$$

where

$$\mathbf{S} = [\mathbf{S}_1 \quad \mathbf{S}_2 \quad \dots \quad \mathbf{S}_{n_u}] \in \mathbb{C}^{n_r \times n_t} \quad (11)$$

and

$$\mathbf{X} = \begin{bmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \\ \vdots \\ \mathbf{X}_{n_u} \end{bmatrix} \in \mathbb{C}^{n_t \times t_c}. \quad (12)$$

We assume that \mathbf{W} and \mathbf{S} are independent, and that their probability law does not depend on \mathbf{X} . Throughout the paper, we also assume that $t_c \geq \max(n_t, n_r)$ and focus on the scenario where neither the transmitter nor the receiver have prior knowledge of matrix \mathbf{S} (no *a priori* CSI). It will turn out convenient to define the following two constants

$$l = \min(n_t, n_r) \quad (13)$$

$$p = \max(n_t, n_r). \quad (14)$$

The sum-rate capacity of the MAC in (10) is given by

$$\mathcal{C}(\rho) = \frac{1}{t_c} \sup \mathcal{I}(\mathbf{X}; \mathbf{Y}) \quad (15)$$

where the supremum is over all probability distributions on \mathbf{X} for which

- 1) $\{\mathbf{X}_i\}_{i=1}^{n_u}$ are independent;
- 2) the per-user power constraint

$$\mathbb{E} \left[\text{Tr} \{ \mathbf{X}_i \mathbf{X}_i^\dagger \} \right] \leq \frac{t_c n_i \rho}{n_t}, \quad i = 1, 2, \dots, n_u \quad (16)$$

is satisfied.

Here, ρ can be thought of as the total energy per channel use available over all users. The particular form of average power constraint in (16) allows all users to transmit at the same average power per antenna. As reviewed in Section I, $\mathcal{C}(\rho)$ is not known in closed form.

In the next section, we provide nonasymptotic upper and lower bounds on the sum-rate capacity $\mathcal{C}(\rho)$ in (15). These bounds will be numerically evaluated in Section IV, and conclusions be drawn in Section V.

III. BOUNDS ON CAPACITY

Theorem 1 below provides a nonasymptotic upper bound on $\mathcal{C}(\rho)$. The upper bound is obtained by dropping the requirement that the $\{\mathbf{X}_i\}_{i=1}^{n_u}$ are independent (which enlarges the set of distributions over which the maximization in (15) is performed), and by bounding the mutual information in (15) using the duality bound [15, Eq. (186)]. The duality approach requires the specification of an auxiliary p.d.f. on the channel output \mathbf{Y} , which, following [8], we choose so that i) \mathbf{Y} is isotropic; ii) the largest l singular values of \mathbf{Y} are distributed as the singular values of the noiseless channel output $\mathbf{S}\mathbf{X}$, with \mathbf{X} following a USTM distribution; iii) the remaining $n_r - l$ singular values are distributed as the singular values of an additive noise matrix of appropriate dimension.

Theorem 1: The sum-rate capacity $\mathcal{C}(\rho)$ in (15) of the Rayleigh block-fading MAC (8) is upper-bounded by

$$\mathcal{C}(\rho) \leq u(\rho) + \frac{1}{t_c} \inf_{\lambda \geq 0} \sup_{D \in \mathbb{D}^{n_t \times n_t}} g(D, \lambda). \quad (17)$$

Here,

$$\begin{aligned} u(\rho) = & -n_r + \frac{n_r n_t}{t_c} \log \left(\frac{t_c \rho}{n_t} \right) \\ & + \frac{1}{t_c} \log \left(\frac{\mathcal{P}(n_r - l) \mathcal{P}(t_c - l) \mathcal{P}(n_t)}{\mathcal{P}(t_c) \mathcal{P}(p - l)} \right) + \frac{1}{t_c} \log(\kappa(\rho)) \\ & + \frac{n_t n_r}{t_c \rho} + \frac{(n_r - l)(t_c - l)}{t_c} \\ & - \frac{(t_c - n_t)(n_r - l)}{t_c} \log(\bar{\mu}) \end{aligned} \quad (18)$$

and

$$\begin{aligned} g(D, \lambda) = & \frac{n_t n_r \text{Tr} \{ D^2 \}}{t_c \rho} \\ & + (t_c - n_t) \mathbb{E}_{\mathbf{G}} \left[\log \det (\mathbf{G}(I_{n_t} + D^2) \mathbf{G}^\dagger + \bar{\mu} I_{n_r}) \right] \\ & - n_r \log \det (I_{n_t} + D^2) + \lambda (t_c \rho - \text{Tr} \{ D^2 \}) \end{aligned} \quad (19)$$

where $\mathcal{P}(\cdot)$ is given in (4), \mathbf{G} is an $n_r \times n_t$ complex random matrix with i.i.d. $\mathcal{CN}(0, 1)$ entries and $\bar{\mu}$ is the expected value of the square of the largest singular value of an $n_r \times (t_c - n_t)$ complex random matrix with i.i.d. $\mathcal{CN}(0, 1)$ entries. Finally, with $\kappa(\rho)$ we denote the probability that the lowest nonzero singular value of an $n_r \times n_t$ complex random matrix with i.i.d. $\mathcal{CN}(0, t_c \rho / n_t)$ entries is greater than the largest singular value of an independent $(n_r - l) \times (t_c - l)$ complex random matrix with i.i.d. $\mathcal{CN}(0, 1)$ entries.

Proof: See Appendix B. ■

When $n_t = n_r$, the sum-rate upper bound (17) can be tightened. This result is given in the following corollary.

Corollary 2: The sum-rate capacity $\mathcal{C}(\rho)$ in (15) of the Rayleigh block-fading MAC (8) for the case $n_t = n_r = n$ is upper-bounded by

$$\mathcal{C}(\rho) \leq u^*(\rho) + \frac{1}{t_c} \inf_{\lambda \geq 0} \sup_{D \in \mathbb{D}^{n \times n}} g^*(D, \lambda). \quad (20)$$

Here,

$$\begin{aligned} u^*(\rho) = & -n + \frac{n^2}{t_c} \log \left(\frac{t_c \rho}{n} \right) + \frac{n^2}{t_c \rho} \\ & + \frac{1}{t_c} \log \left(\frac{\mathcal{P}(n) \mathcal{P}(t_c - n)}{\mathcal{P}(t_c)} \right) \end{aligned} \quad (21)$$

and

$$\begin{aligned} g^*(D, \lambda) = & \frac{n^2 \text{Tr} \{ D^2 \}}{t_c \rho} - n \log \det (I_n + D^2) \\ & + \lambda (t_c \rho - \text{Tr} \{ D^2 \}) \\ & + \frac{(t_c - n) \sum_{k=1}^n \det (R_k (I_n + D^2))}{\mathcal{K}(D^2, t_c - n) \det (I_n + D^2)^{-t_c + n + 1}} \end{aligned} \quad (22)$$

where $\mathcal{P}(\cdot)$ is given in (4) and $\mathcal{K}(\cdot, \cdot)$ is given in (6). The matrix $R_k(A)$ in (22), which is a function of the $n \times n$ diagonal matrix A , is defined as follows. Let $a_1 > a_2 > \dots > a_n$ denote the ordered diagonal entries of A . Let the $n \times n$ real matrix $P_k(A)$ and the $(t_c - n) \times n$ real matrix T_k , $k = 1, 2, \dots, n$ be defined as follows:

$$[P_k(A)]_{ij} = \begin{cases} a_i^{n-k+1} (\log(a_i) + \psi(n-k+1)), & 1 \leq i \leq n, j = k \\ a_i^{n-j+1}, & 1 \leq i \leq n, 1 \leq j \leq n, j \neq k \end{cases} \quad (23)$$

and

$$[T_k]_{ij} = \begin{cases} \frac{1}{\beta(n-j+1, t_c-n-i)}, & 1 \leq i \leq n, 1 \leq j \leq n, j \neq k \\ \frac{\psi(t_c-i-k+1)}{\beta(n-k+1, t_c-i)}, & 1 \leq i \leq t_c-n-1, j = k \\ 1, & i = t_c-n, 1 \leq j \leq n, j \neq k \\ \psi(n-k+1), & i = t_c-n, j = k \end{cases} \quad (24)$$

where $\beta(\cdot, \cdot)$ is defined in the notation section. Finally, let $Q(A) \in \mathbb{R}^{n \times (t_c-n)}$ and $S \in \mathbb{R}^{(t_c-n) \times (t_c-n)}$ be given by

$$[Q(A)]_{ij} = (-a_i)^{j+n-t_c}, \quad 1 \leq i \leq n, 1 \leq j \leq t_c-n \quad (25)$$

and

$$S_{ij} = \begin{cases} \frac{(-1)^{i-j}}{\beta(i-j+1, t_c-n-i)}, & 1 \leq j \leq i \leq t_c-n-1 \\ (-1)^{i-j}, & 1 \leq j \leq i = t_c-n \\ 0, & 1 \leq i < j \leq t_c-n. \end{cases} \quad (26)$$

Then,

$$R_k(A) = P_k(A) - Q(A)S^{-1}T_k. \quad (27)$$

Proof: See Appendix C. ■

In Theorem 3 below we provide a lower bound on $\mathcal{C}(\rho)$. This lower bound is obtained by evaluating the mutual information in (15) for the MAC-USTM input distribution introduced in Section I. Specifically, we set

$$\mathbf{X}_i = \sqrt{\frac{t_c \rho}{n_t}} \mathbf{V}_i \quad (28)$$

with \mathbf{V}_i uniformly distributed on $\mathcal{S}(n_i, t_c)$. With this choice, the power constraint in (16) is satisfied with equality. Similar to the bounds developed in [9]–[11], our lower bound relies on the Itzykson-Zuber integral [12, Eq. (3.4)]. In order to give a compact expression for our lower bound, we shall focus on the setup where each user is equipped with a single antenna i.e., $n_i = 1, i = 1, 2, \dots, n_u$. We will address the case when the users have multiple antennas at the end of this section.

Theorem 3: The sum-rate capacity $\mathcal{C}(\rho)$ in (15) of the Rayleigh block-fading MAC (8) for the case when all transmitters have a single antenna, i.e., $n_i = 1, i = 1, 2, \dots, n_u$ is lower-bounded as follows:

$$\begin{aligned} \mathcal{C}(\rho) \geq & \log\left(\frac{\mathcal{P}(t_c-l)}{\mathcal{P}(t_c)}\right) + n_r t_c \rho \\ & - n_r \mathbb{E}[\log \det(I_{n_t} + \mathbf{D}^2)] \\ & + \mathbb{E}[\log(\mathcal{K}(\mathbf{Y}\mathbf{Y}^\dagger, t_c - n_r))] \\ & - \mathbb{E}\left[\log\left(\mathbb{E}_{\mathbf{D}}\left[\frac{\det(M(\mathbf{Y}\mathbf{Y}^\dagger, \mathbf{E}))}{\det(I_{n_t} + \mathbf{D}^2)^{n_r} \mathcal{K}(\mathbf{E}, t_c - n_t)}\right]\right)\right]. \end{aligned} \quad (29)$$

Here, \mathbf{D} is an $n_t \times n_t$ real diagonal matrix with entries containing the singular values of the input matrix \mathbf{X} , which is assumed to follow the MAC-USTM distribution (28). Furthermore, the $n_t \times n_t$ matrix \mathbf{E} is defined as $\mathbf{E} = (\mathbf{D}^{-2} + I_{n_t})^{-1}$, $\mathcal{P}(\cdot)$ is given in (4), and $\mathcal{K}(\cdot, \cdot)$ is given in (6). Finally, the $p \times p$ real matrix $M(A, B)$ is defined as follows:

$$[M(A, B)]_{ij} = \begin{cases} \gamma(a_i b_j, t_c - p), & 1 \leq i \leq n_r, 1 \leq j \leq n_t \\ b_j^{t_c-i}, & n_r + 1 \leq i \leq p, 1 \leq j \leq n_t \\ a_i^{t_c-j}, & 1 \leq i \leq n_r, n_t + 1 \leq j \leq p. \end{cases} \quad (30)$$

Here, $a_1 > a_2 > \dots > a_{n_r}$ are the ordered eigenvalues of the positive-definite matrix $A \in \mathbb{C}^{n_r \times n_r}$; similarly $b_1 > b_2 > \dots > b_{n_t}$ are the ordered eigenvalues of the positive-definite matrix $B \in \mathbb{C}^{n_t \times n_t}$; the function $\gamma(\cdot, \cdot)$ in (30) is given in (7).

Proof: See Appendix D. ■

The lower bound (29) involves expectations that are not known in closed form. Hence, we resort to Monte-Carlo methods for the evaluation of (29). One exception is the two-user case $n_u = 2$, for which the expectation over \mathbf{D} in (29) admits a closed-form integral expression. This result is presented in Corollary 4 below. The proof of this corollary exploits that for the two-user case, the probability distribution of the eigenvalues of $\mathbf{X}\mathbf{X}^\dagger$ can be obtained from the p.d.f. of $\det(\mathbf{X}\mathbf{X}^\dagger)$, which can be calculated through Bartlett's decomposition [18, Prop. 2.1]. This approach does not extend to the case $n_u > 2$ or $n_t > 2$.

Corollary 4: The sum-rate capacity $\mathcal{C}(\rho)$ in (15) of the Rayleigh block-fading MAC (8) with $n_u = 2$ and $n_i = 1, i = 1, 2$ is lower-bounded as follows:

$$\begin{aligned} \mathcal{C}(\rho) \geq & \log\left(\frac{\mathcal{P}(t_c-l)}{\mathcal{P}(t_c)}\right) + n_r t_c \rho + \log\left(\frac{t_c \rho}{2}\right) \\ & - n_r (t_c - 1) \left[\int_0^1 \log(\mu^2 - \alpha) (1 - \alpha)^{t_c-2} d\alpha \right] \\ & + \mathbb{E}[\log(\mathcal{K}(\mathbf{Y}\mathbf{Y}^\dagger, t_c - n_r))] - \log(t_c - 1) \\ & - \mathbb{E}\left[\log\left(\int_0^1 \det(M(\mathbf{Y}\mathbf{Y}^\dagger, E(\alpha))) (\mu^2 - \alpha^2)^{t_c-n_r-1} d\alpha\right)\right]. \end{aligned} \quad (31)$$

Here, $\mu = 1 + 2/t_c \rho$ and

$$E(\alpha) = \text{diag}\left\{\frac{1+\alpha}{1+\mu(1+\alpha)}, \frac{1-\alpha}{1+\mu(1-\alpha)}\right\}. \quad (32)$$

Furthermore, $M(\cdot, \cdot)$ is given in (30), $\mathcal{P}(\cdot)$ is given in (4), and $\mathcal{K}(\cdot, \cdot)$ is given in (6).

Proof: See Appendix E. ■

For the case when the users are equipped with multiple antennas, the singular values of the input matrix \mathbf{X} are no longer distinct, and (29) needs to be further simplified using L'Hôpital's rule. The final expression of the resulting lower bound is omitted because it is involved. Instead, numerical results are provided in Section IV.

IV. NUMERICAL RESULTS

Focusing on the setup where the total number of transmit antennas n_t is equal to the number of receive antennas n_r , we

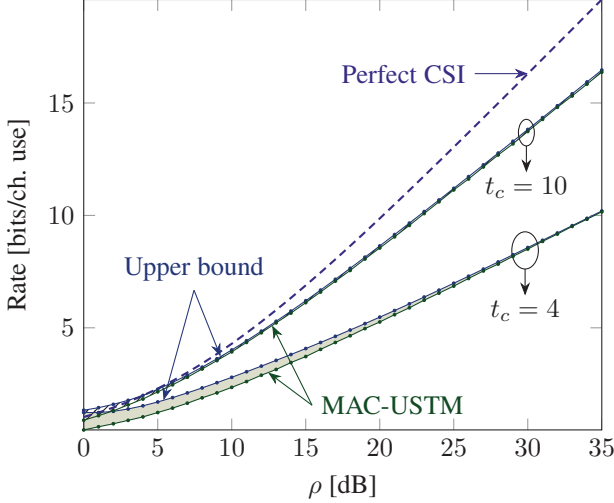


Fig. 3. Upper bounds (20) and (33), and lower bound (29) on the sum-rate capacity of the MAC (8); $n_u = 2$ single-antenna users, $n_r = 2$, $t_c \in \{4, 10\}$.

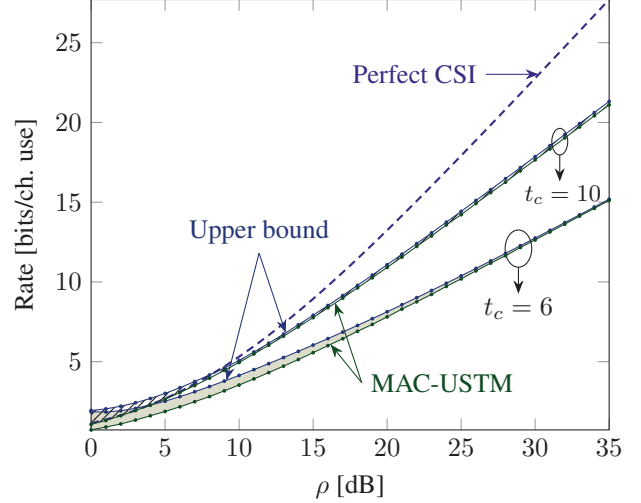


Fig. 4. Upper bounds (20) and (33), and lower bound (29) on the sum-rate capacity of the MAC (8); $n_u = 3$ single-antenna users, $n_r = 3$, $t_c \in \{6, 10\}$.

numerically evaluate in this section the upper bound (20) and the lower bound (29) on the sum-rate capacity (15).³

We also evaluate the i.i.d. Gaussian lower bound developed by Rusek *et al.* [9], which is obtained by setting $\mathbf{X} = \sqrt{\rho/n_t}\mathbf{G}$ in (15), with \mathbf{G} having i.i.d. $\mathcal{CN}(0, 1)$ entries. For this choice of \mathbf{X} , the expectation over \mathbf{D} in (29) can be evaluated in closed form using the properties of Wishart matrices and the integral formula [19, Cor. 2]. See [9], [11] for details.

We also consider the simple capacity upper bound obtained by assuming that the receiver has perfect knowledge of the channel matrix \mathbf{S} in (11). This bound is given by [20]

$$\mathcal{C}(\rho) \leq \mathbb{E} \left[\log \det \left(I_{n_r} + \frac{\rho}{n_t} \mathbf{S} \mathbf{S}^\dagger \right) \right]. \quad (33)$$

We start by considering the case of single-antenna users. In Figs. 3–5, we depict the upper bounds (20) and (33), and the lower bound (29) on the sum-rate capacity of the MAC (8) for different values of number of users n_u , number of receive antennas n_r , and coherence interval t_c . For the choice of parameters in Figs. 3–5, the upper bound (20) and the lower bound (29) characterize the sum-rate capacity accurately already at SNR values as low as 10 dB. The tightness of the bounds increases as the coherence interval t_c or the SNR increases. This last observation comes as no surprise, since the choice of the auxiliary distribution in the derivation of the upper bound (20) is dictated by high-SNR considerations.

The tightness of our bounds implies that the sum-rate capacity of the MAC (8) is well-approximated by the capacity of an equivalent point-to-point MIMO channel with the same SNR and the same total number of transmit and receive antennas. This follows because the upper bound is derived under the assumption of perfect cooperation among the users. Furthermore, we observe that the perfect-CSI sum-rate capacity upper bound (33) is loose for the relatively small t_c values considered in this section.

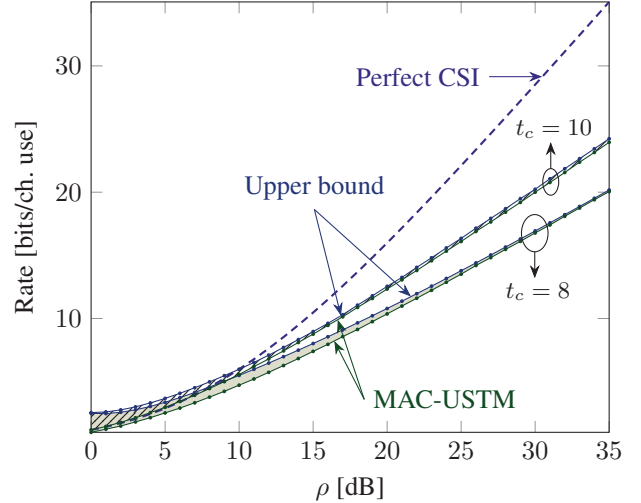


Fig. 5. Upper bounds (20) and (33), and lower bound (29) on the sum-rate capacity of the MAC (8); $n_u = 4$ single-antenna users, $n_r = 4$, $t_c \in \{8, 10\}$.

As shown in Fig. 2 (see Section I) for the case $n_u = n_r = 4$ and $t_c = 10$, the MAC-USTM lower bound is tighter than the i.i.d. Gaussian lower bound in the high-SNR regime, although marginally so. The same consideration holds for the case $n_u = n_r = 2$ and $n_u = n_r = 3$ considered in Fig. 3 and Fig. 4. Furthermore, the gap between the two bounds gets smaller for t_c values smaller than 10.

In Fig. 6, we consider the case when $n_t = n_r = 4$ and $t_c = 10$, but the available 4 antennas are divided unevenly among the users. Specifically, we assume $n_u = 2$ (two users) and that the first user has one antenna ($n_1 = 1$) whereas the second user has three antennas ($n_2 = 3$). Note that the upper bound and the i.i.d. Gaussian lower bound depend only on the total number of antennas n_t and not on the way the antennas are divided among users. Hence, these two curves coincide with the curves drawn in Fig. 1 and Fig. 2. On the contrary, the MAC-USTM bound

³Numerical routines implementing the upper bound (20) and lower bound (29) can be downloaded at https://github.com/infotheorychalmers/mac_capacity_bounds

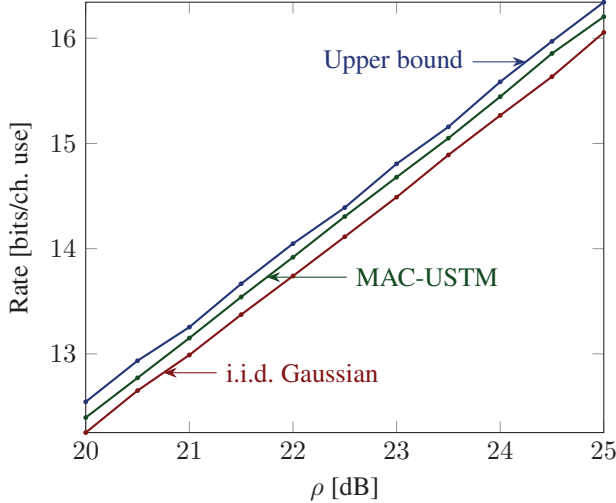


Fig. 6. Upper bound (20), MAC-USTM lower bound (29), and i.i.d. Gaussian lower bound [9]; $n_u = 2$ with $n_1 = 1$ and $n_2 = 3$; furthermore, $n_r = 4$ and $t_c = 10$.

depends on n_1 and n_2 , and, for the asymmetric antenna scenario considered in Fig. 6, is tighter than the i.i.d. Gaussian bound, although marginally so.

V. CONCLUSION

We presented finite-SNR upper and lower bounds on the sum-rate capacity of Rayleigh block-fading channels for the scenario where neither the transmitters nor the receiver have access to *a priori* CSI. The upper bound, which is based on duality, is derived under the assumption that the transmitters can cooperate perfectly. This transforms the MAC into an equivalent point-to-point MIMO channel. We obtain the lower bound by assuming that a USTM signal is independently transmitted by each user over its available antennas (MAC-USTM). The gap between the upper and the lower bounds is less than 8% at 10 dB for the case of 4 single-antenna transmitters communicating with a 4-antenna receiver over a MAC with coherence interval 10. The gap reduces further when the SNR or the coherence interval get larger. This implies that the capacity gains obtainable by allowing cooperation among transmitters are minimal in this case. Our numerical results show also that MAC-USTM yields rates that are only marginally larger than the ones obtainable using i.i.d. Gaussian inputs. This suggests that i.i.d. Gaussian inputs are almost sum-rate capacity optimal for the Rayleigh block-fading MACs considered in Section IV.

APPENDIX A PRELIMINARY RESULTS

A. An Integral Formula

The following lemma is useful in the evaluation of integrals that involve matrix determinants. We shall need this lemma in the proof of Corollary 2.

Lemma 5: Let $a, b \in \mathbb{R}$, with $a < b$. Let $\{f_i(\cdot)\}_{i=1}^n$ and $\{g_i(\cdot)\}_{i=1}^n$ be arbitrary integrable functions over $[a, b]$. Let $A = A(x_1, x_2, \dots, x_n)$ and $B = B(x_1, x_2, \dots, x_n)$ be $n \times n$ and

$m \times m$ matrices ($m \geq n$) whose entries depend on the scalars x_1, x_2, \dots, x_n as follows:

$$A_{ij} = \begin{cases} f_i(x_j), & 1 \leq i \leq m, 1 \leq j \leq n \\ c_{ij}, & 1 \leq i \leq m, n+1 \leq j \leq m \end{cases} \quad (34)$$

$$B_{ij} = g_i(x_j), \quad 1 \leq i \leq n, 1 \leq j \leq n \quad (35)$$

where c_{ij} are arbitrary scalar real constants. Finally, let $E \in \mathbb{R}^{m \times m}$ be defined as

$$E_{ij} = \begin{cases} \int_a^b f_i(x)g_j(x)dx, & 1 \leq i \leq m, 1 \leq j \leq n \\ c_{ij}, & 1 \leq i \leq m, n+1 \leq j \leq m. \end{cases} \quad (36)$$

Then

$$\int_{a \leq x_n \leq x_{n-1} \leq \dots \leq x_1 < b} \det(A) \det(B) dx_1 dx_2 \dots dx_n = \det(E). \quad (37)$$

Proof: See [21, Lem. 2]. ■

B. Limits of Determinants

The following lemma, which characterizes the limiting behavior of the ratio between the determinant of a certain matrix and a Vandermonde determinant, will be needed in the proof of Corollary 2.

Lemma 6: Let $A \in \mathbb{C}^{m \times m}$ be a positive-definite matrix with ordered eigenvalues $a_1 > a_2 > \dots > a_m$. Let $C \in \mathbb{C}^{m \times m}$ be a matrix with entries $C_{i,j} = f_i(a_j)$, for some differentiable functions $\{f_i(\cdot)\}_{i=1}^m$. Then for every integer $n < m$, and every scalar real constant a_0 ,

$$\lim_{\{a_{n+1}, a_{n+2}, \dots, a_m\} \rightarrow a_0} \frac{\det(C)}{\mathcal{V}(A)} = \frac{\det(E)}{\mathcal{P}(m-n)\mathcal{K}(A_0 - a_0 I_n, m-n)} \quad (38)$$

where $A_0 = \text{diag}\{a_1, a_2, \dots, a_n\}$, $\mathcal{P}(\cdot)$ is given in (4), $\mathcal{V}(\cdot)$ is given in (5), $\mathcal{K}(\cdot, \cdot)$ is given in (6), and the entries of the $m \times m$ matrix E are

$$E_{i,j} = \begin{cases} f_i(a_j), & 1 \leq i \leq m, 1 \leq j \leq n \\ f_i^{(m-j)}(a_0), & 1 \leq i \leq m, n+1 \leq j \leq m. \end{cases} \quad (39)$$

Here, $f_i^{(k)}(\cdot)$ denotes the k th-order derivative of $f_i(\cdot)$.

Proof: See [22, Lem. 5]. ■

C. Expectation of the Log Determinant of a Gaussian Quadratic Form

The following lemma gives a closed-form expression for $\mathbb{E}[\log \det(\mathbf{X}L\mathbf{X}^\dagger)]$ where \mathbf{X} has i.i.d. $\mathcal{CN}(0, 1)$ entries and L is a certain positive-definite matrix. A closed-form expression for the case when the eigenvalues of L are distinct is provided in [23, Lem. 2]. Here, we derive a different closed-form expression, which does not require the eigenvalues to be distinct, and appears better suited for numerical evaluations. This lemma is used in the proof of Corollary 2.

Lemma 7: Let $\mathbf{X} \in \mathbb{C}^{n \times m}$, where $n < m$, be a random matrix with i.i.d. $\mathcal{CN}(0, 1)$ entries. Let $L \in \mathbb{C}^{m \times m}$ be a positive-definite matrix whose largest n eigenvalues satisfy $l_1 > l_2 > \dots > l_n > 1$, and whose lowest $m - n$ eigenvalues are equal to 1. Then

$$\begin{aligned} \mathbb{E}[\log \det(\mathbf{X}L\mathbf{X}^\dagger)] &= \frac{\det(L_0)^{m-n-1}}{\mathcal{K}(L_0 - I_n, m-n)} \sum_{k=1}^n \det(R_k(L_0)) \quad (40) \end{aligned}$$

where $\mathcal{K}(\cdot, \cdot)$ is given in (6), $L_0 = \text{diag}\{l_1, l_2, \dots, l_n\}$, and $R_k(\cdot)$ is defined as in Corollary 2.

Proof: We first obtain the joint p.d.f. of the eigenvalues of $\mathbf{S}^* = \mathbf{X}L^*\mathbf{X}^\dagger$, where $L^* \in \mathbb{C}^{m \times m}$ is a positive-definite matrix whose largest n eigenvalues coincide with the eigenvalues of L , and whose remaining $m - n$ eigenvalues are distinct, i.e., $l_{n+1}^* > l_{n+2}^* > \dots > l_m^*$. We then derive the p.d.f. of the eigenvalues of $\mathbf{S} = \mathbf{X}L\mathbf{X}^\dagger$ by letting $\{l_{n+1}^*, l_{n+2}^*, \dots, l_m^*\} \rightarrow 1$. Knowledge of this p.d.f. allows us to obtain the moment generating function of $\log \det(\mathbf{S})$ in closed form, from which (40) follows by evaluating the derivative of this moment generating function at zero. The p.d.f. of \mathbf{S}^* is [24, Sec. 2]⁴

$$\begin{aligned} f_{\mathbf{S}^*}(S) &= \frac{1}{\pi^{n(n-1)/2}} \frac{\det(F(-(L^*)^{-1}, S))}{\mathcal{V}(S)\mathcal{V}(-(L^*)^{-1})\det(L^*)^n}, \quad S \in \mathbb{C}^{n \times n} \quad (41) \end{aligned}$$

where $F(A, B)$ is an $m \times m$ matrix that depends on the ordered eigenvalues $\{a_i\}_{i=1}^m$ and $\{b_i\}_{i=1}^n$ of the Hermitian matrices $A \in \mathbb{C}^{m \times m}$ and $B \in \mathbb{C}^{n \times n}$ as follows

$$[F(A, B)]_{ij} = \begin{cases} e^{a_i b_j}, & 1 \leq i \leq m, 1 \leq j \leq n \\ a_i^{m-j}, & 1 \leq i \leq m, n+1 \leq j \leq m. \end{cases} \quad (42)$$

Let $\mathbf{E}_{\mathbf{S}^*}$ denote the $n \times n$ real diagonal matrix having the ordered eigenvalues of \mathbf{S}^* on its main diagonal. The p.d.f. of $\mathbf{E}_{\mathbf{S}^*}$ is [26, Thm. 3.2]

$$\begin{aligned} f_{\mathbf{E}_{\mathbf{S}^*}}(E) &= \frac{1}{\mathcal{P}(n)} \frac{\det(F(-(L^*)^{-1}, E))\mathcal{V}(E)}{\mathcal{V}(-(L^*)^{-1})\det(L^*)^n}, \quad E \in \mathbb{D}^{n \times n}. \quad (43) \end{aligned}$$

Finally, let $\mathbf{E}_{\mathbf{S}}$ denote the $n \times n$ real diagonal matrix having the ordered eigenvalues of $\mathbf{S} = \mathbf{X}L\mathbf{X}^\dagger$ on its main diagonal. We obtain the p.d.f. of $\mathbf{E}_{\mathbf{S}}$ by computing the limit $\{l_{n+1}^*, l_{n+2}^*, \dots, l_m^*\} \rightarrow 1$ in (43):

$$f_{\mathbf{E}_{\mathbf{S}}}(E) = \lim_{\{l_{n+1}^*, l_{n+2}^*, \dots, l_m^*\} \rightarrow 1} f_{\mathbf{E}_{\mathbf{S}^*}}(E) \quad (44)$$

$$= \frac{\mathcal{V}(E) \det(G(L_0, E))}{\mathcal{P}(n) \det(L_0)^n \mathcal{P}(m-n) \mathcal{K}(-L_0^{-1} + I_n, m-n)} \quad (45)$$

$$= \frac{\det(L_0)^{m-n-1} \det(G(L_0, E))\mathcal{V}(E)}{\mathcal{P}(n)\mathcal{P}(m-n)\mathcal{K}(L_0 - I_n, m-n)}, \quad E \in \mathbb{D}^{n \times n}. \quad (46)$$

⁴There seems to be a typo in the expression given in [24, Sec. 2]. The term $\pi^{n(n-1)/2}$ should be in denominator (see [25, Eq. (50) and Eq. (57)]).

Here, $L_0 \in \mathbb{D}^{n \times n}$ is a diagonal matrix, which contains the n largest eigenvalues $l_1 > l_2 > \dots > l_n$ of L on its diagonal. Furthermore, the $m \times m$ matrix $G(L_0, E)$ is defined as follows

$$[G(L_0, E)]_{ij} = \begin{cases} \exp(-l_i^{-1}e_j), & 1 \leq i \leq n, 1 \leq j \leq n \\ e_j^{m-i} \exp(-e_j), & n+1 \leq i \leq m, 1 \leq j \leq n \\ (-l_i)^{j-m}, & 1 \leq i \leq n, n+1 \leq j \leq m \\ \frac{\Gamma(m-j+1)}{\Gamma(i-j+1)} (-1)^{i-j}, & n+1 \leq j \leq i \leq m \\ 0, & n+1 \leq i < j \leq m \end{cases} \quad (47)$$

where $e_1 > e_2 > \dots > e_n$ denote the diagonal entries of E . To obtain (45), we used Lemma 6, and (46) follows because for all $A \in \mathbb{C}^{n \times n}$, we have that $\mathcal{V}(-A) = \mathcal{V}(A)$ and that $\mathcal{V}(A^{-1}) = \mathcal{V}(A) / \det(A)^{n-1}$.

Next, we evaluate the moment generating function of $\log \det(\mathbf{S})$:

$$g(t) = \mathbb{E}[\exp(t \log \det(\mathbf{S}))] = \mathbb{E}[\det(\mathbf{E}_{\mathbf{S}})^t] \quad (48)$$

$$= \int \dots \int_{\mathbb{D}^{n \times n}} \det(E)^t f_{\mathbf{E}_{\mathbf{S}}}(E) dE \quad (49)$$

$$= \frac{\det(L_0)^{m-n-1} \det(H(t, L_0))}{\mathcal{P}(n)\mathcal{P}(m-n)\mathcal{K}(L_0 - I_n, m-n)}. \quad (50)$$

Here, the $m \times m$ matrix $H(t, L_0)$ is defined as follows:

$$[H(t, L_0)]_{ij} = \begin{cases} \Gamma(t+n-j+1)l_i^{t+n-j+1}, & 1 \leq i \leq n, 1 \leq j \leq n \\ \Gamma(t+m-i+n-j+1), & n+1 \leq i \leq m, 1 \leq j \leq n \\ (-l_i)^{j-m}, & 1 \leq i \leq n, n+1 \leq j \leq m \\ \frac{\Gamma(m-j+1)}{\Gamma(i-j+1)} (-1)^{i-j}, & n+1 \leq j \leq i \leq m \\ 0, & n+1 \leq i < j \leq m. \end{cases} \quad (51)$$

To obtain (50), we used the integral formula (37) in Lemma 5. To establish (40), we now proceed as follows. Let $g'(t)$ denote the first derivative of $g(t)$. Then

$$\mathbb{E}[\log \det(\mathbf{Y})] = \lim_{t \rightarrow 0} g'(t) \quad (52)$$

$$= \frac{\det(L_0)^{m-n-1} \sum_{k=1}^n \det(H_k(L_0))}{\mathcal{P}(n)\mathcal{P}(m-n)\mathcal{K}(L_0 - I_n, m-n)} \quad (53)$$

where the $m \times m$ matrices $H_k(L_0)$, $k = 1, 2, \dots, n$, are given by

$$[H_k(L_0)]_{ij} = \begin{cases} [H(0, L_0)]_{ij}, & 1 \leq i \leq m, 1 \leq j \leq m, j \neq k \\ \Gamma(n-k+1)l_i^{n-k+1}(\log(l_i) + \psi(n-k+1)), & 1 \leq i \leq n, j = k \\ \Gamma(m-i+n-k+1)\psi(m-i+n-k+1), & n+1 \leq i \leq m, j = k. \end{cases} \quad (54)$$

In (53), we have written the derivative of the determinant of the matrix $H(t, L_0)$ as a sum of n determinants by using that the

matrix $H(t, L_0)$ depends on t only through its first n columns. Through algebraic manipulation one can show that

$$\det(H_k(L_0)) = \mathcal{P}(n)\mathcal{P}(m-n)\det(R_k(L_0)) \quad (55)$$

where $R_k(\cdot)$ is defined in (27). The equality in (55) relates the determinant of an $m \times m$ matrix to the determinant of a smaller $n \times n$ matrix. Substituting (55) into (53) we obtain (40). ■

APPENDIX B PROOF OF THEOREM 1

We upper-bound the sum-rate capacity $\mathcal{C}(\rho)$ by enlarging the set over which the supremum in (15) is computed. Specifically, we drop the assumption that the $\{\mathbf{X}_i\}_{i=1}^{n_u}$ are independent, and we substitute (16) with the ‘‘global’’ power constraint

$$\mathbb{E}[\text{Tr}\{\mathbf{X}\mathbf{X}^\dagger\}] \leq t_c \rho. \quad (56)$$

Let $\mathcal{J}(\rho)$ denote this enlarged set. Then

$$\mathcal{C}(\rho) \leq \frac{1}{t_c} \sup_{\mathcal{J}(\rho)} \mathcal{I}(\mathbf{X}; \mathbf{Y}). \quad (57)$$

Note that the right-hand side (RHS) of (57) is the capacity of an $n_t \times n_r$ MIMO Rayleigh block-fading channel. Hence, it follows from [6, Thm. 2] that we can restrict the supremum in (57) to input distributions for which $\mathbf{X} = \mathbf{D}\mathbf{Q}$, where \mathbf{D} is an $n_t \times n_t$ real diagonal matrix and \mathbf{Q} is independent of \mathbf{D} , and uniformly distributed over the Stiefel manifold $\mathcal{S}(n_t, t_c)$.

Next, we upper-bound the RHS of (57) using the duality bound [15, Eq. (186)]. This requires the specification of an auxiliary p.d.f. $r_{\mathbf{Y}}(\cdot)$ on \mathbf{Y} , which we choose as follows. Let us denote the singular value decomposition (SVD) of \mathbf{Y} as

$$\mathbf{Y} = \mathbf{U}\Sigma\mathbf{V}. \quad (58)$$

Here, \mathbf{U} belongs to the set $\mathcal{U}(n_r)$ and has real and positive entries on the main diagonal,⁵ Σ is an $n_r \times n_r$ diagonal matrix that contains the ordered singular values of \mathbf{Y} on its main diagonal, and \mathbf{V} belongs to the Stiefel manifold $\mathcal{S}(n_r, t_c)$. Furthermore, we choose the auxiliary p.d.f. so that the matrices on the RHS of (58) are mutually independent, which leads to

$$r_{\mathbf{Y}}(\mathbf{Y}) = f_{\mathbf{U}}(\mathbf{U})f_{\Sigma}(\Sigma)f_{\mathbf{V}}(\mathbf{V})J(\Sigma) \quad (59)$$

where $f_{\mathbf{U}}(\cdot)$, $f_{\Sigma}(\cdot)$ and $f_{\mathbf{V}}(\cdot)$ are the p.d.f. of \mathbf{U} , Σ , and \mathbf{V} , respectively and

$$J(\Sigma) = \frac{1}{\mathcal{V}(\Sigma^2)^2 \det(\Sigma)^{2(t_c - n_r) + 1}} \quad (60)$$

is the Jacobian of the SVD [7, App. A]. Moreover, we assume that \mathbf{U} and \mathbf{V} are uniformly distributed (with respect to the Haar measure) over their respective domains, which implies that

$$f_{\mathbf{U}_Y}(\mathbf{U}) = \frac{\mathcal{P}(n_r)}{\pi^{\frac{n_r(n_r-1)}{2}}} \quad (61)$$

$$f_{\mathbf{V}_Y}(\mathbf{V}) = \frac{\mathcal{P}(t_c)}{\mathcal{P}(t_c - n_r) 2^{n_r} \pi^{n_r t_c - \frac{n_r(n_r-1)}{2}}}. \quad (62)$$

The p.d.f. of Σ is chosen as in [8, Sec. V]. Specifically, we take the first l singular values of \mathbf{Y} to be distributed as the

⁵This second requirement ensures that the SVD in (58) is a one-to-one map.

singular values of the noiseless channel output $\mathbf{S}\mathbf{X}$, with \mathbf{X} USTM distributed. We take the remaining $n_r - l$ singular values of \mathbf{Y} to be independent of the first l singular values and to be distributed as the singular values of an $(n_r - l) \times (t_c - l)$ matrix with i.i.d. $\mathcal{CN}(0, 1)$ entries. Let Σ_1 be an $l \times l$ diagonal matrix containing the first l singular values of \mathbf{Y} and let Σ_2 be an $(n_r - l) \times (n_r - l)$ matrix containing the remaining singular values. Then we have

$$f_{\Sigma}(\Sigma_1, \Sigma_2) = \frac{f_{\Sigma_1}(\Sigma_1)f_{\Sigma_2}(\Sigma_2)}{\kappa(\rho)}. \quad (63)$$

Here,

$$f_{\Sigma_1}(\Sigma_1) = \frac{2^l \mathcal{P}(p-l)}{\mathcal{P}(l)\mathcal{P}(p)\beta^{pl}} \exp\left(-\frac{1}{\beta} \text{Tr}\{\Sigma_1^2\}\right) \cdot \det(\Sigma_1)^{2(p-l)+1} \mathcal{V}(\Sigma_1^2)^2, \quad \Sigma_1 \in \mathbb{D}^{l \times l} \quad (64)$$

where $\beta = t_c \rho / n_t$ and

$$f_{\Sigma_2}(\Sigma_2) = \frac{2^{n_r-l} \mathcal{P}(t_c - n_r)}{\mathcal{P}(n_r - l)\mathcal{P}(t_c - l)} e^{-\text{Tr}\{\Sigma_2^2\}} \det(\Sigma_2)^{2(t_c - n_r) + 1} \cdot \mathcal{V}(\Sigma_2^2)^2, \quad \Sigma_2 \in \mathbb{D}^{n_r-l \times n_r-l}. \quad (65)$$

Furthermore, $\kappa(\rho)$ is a normalization constant that ensures that the singular values are ordered. Specifically, $\kappa(\rho)$ is the probability that the smallest singular value in Σ_1 is larger than the largest singular values in Σ_2 . Using $r_{\mathbf{Y}}(\cdot)$ in the duality bound [15, Eq. (186)] we obtain

$$\mathcal{C}(\rho) \leq \frac{1}{t_c} \sup_{\mathcal{J}(\rho)} \{-\mathbb{E}[\log(r_{\mathbf{Y}}(\mathbf{Y}))] - h(\mathbf{Y}|\mathbf{X})\}. \quad (66)$$

Note that the expectation in (66) is not with respect to $r_{\mathbf{Y}}(\cdot)$ but with respect to the probability distribution on \mathbf{Y} induced by the input distribution on \mathbf{X} through the channel (10). Fix $\lambda > 0$. We can further upper-bound the RHS of (66) by using (56) as follows:

$$\mathcal{C}(\rho) \leq \frac{1}{t_c} \inf_{\lambda > 0} \sup_{\mathcal{J}(\rho)} \left\{ -\mathbb{E}[\log(r_{\mathbf{Y}}(\mathbf{Y}))] - h(\mathbf{Y}|\mathbf{X}) + \lambda(t_c \rho - \mathbb{E}[\text{Tr}\{\mathbf{X}\mathbf{X}^\dagger\}]) \right\}. \quad (67)$$

Substituting (61), (64), (65), (62), and (60) into (67) and using [6, Thm. 2] we can rewrite the upper bound (67) as

$$\begin{aligned} \mathcal{C}(\rho) \leq & \frac{1}{t_c} \inf_{\lambda > 0} \sup_{\mathcal{J}(\rho)} \left\{ -n_r t_c + n_r n_t \log(\beta) + \log(\kappa(\rho)) \right. \\ & + \log\left(\frac{\mathcal{P}(n_r - l)\mathcal{P}(t_c - l)\mathcal{P}(n_t)}{\mathcal{P}(t_c)\mathcal{P}(p-l)}\right) \\ & - n_r \mathbb{E}[\log \det(\mathbf{I}_{n_t} + \mathbf{D}^2)] + \lambda(t_c \rho - \mathbb{E}[\text{Tr}\{\mathbf{D}^2\}]) \\ & + \frac{1}{\beta} \mathbb{E}[\text{Tr}\{\Sigma_1^2\}] + \mathbb{E}[\text{Tr}\{\Sigma_2^2\}] \\ & + (l + t_c - p - n_r) \mathbb{E}[\log \det(\Sigma_1^2)] \\ & \left. + \mathbb{E}\left[\log\left(\frac{\mathcal{V}(\Sigma^2)^2}{\mathcal{V}(\Sigma_1^2)^2 \mathcal{V}(\Sigma_2^2)^2}\right)\right]\right\}. \quad (68) \end{aligned}$$

We proceed now by upper-bounding some of the terms on the RHS of (68). First we bound $\mathbb{E}[\text{Tr}\{\Sigma_1^2\}]$:

$$\mathbb{E}[\text{Tr}\{\Sigma_1^2\}] \leq \mathbb{E}[\text{Tr}\{\mathbf{Y}\mathbf{Y}^\dagger\}] = n_r t_c + n_r \mathbb{E}[\text{Tr}\{\mathbf{D}^2\}]. \quad (69)$$

Let $\sigma_1^2 > \sigma_2^2 > \dots > \sigma_{n_t}^2$ denote the eigenvalues of $\mathbf{Y}\mathbf{Y}^\dagger$. To bound $\mathbb{E}[\text{Tr}\{\Sigma_2^2\}]$, we use the argument given in [7, p. 377], which yields

$$\mathbb{E}[\text{Tr}\{\Sigma_2^2\}] = \mathbb{E}\left[\sum_{i=l+1}^{n_r} \sigma_i^2\right] \leq (n_r - l)(t_c - l). \quad (70)$$

We also have that

$$\mathbb{E}\left[\log\left(\frac{\mathcal{V}(\Sigma^2)^2}{\mathcal{V}(\Sigma_1^2)^2\mathcal{V}(\Sigma_2^2)^2}\right)\right] = \mathbb{E}\left[\sum_{i=1}^l \sum_{j=l+1}^{n_r} 2\log(\sigma_i^2 - \sigma_j^2)\right] \leq 2(n_r - l)\mathbb{E}[\log \det(\Sigma_1^2)]. \quad (71)$$

Finally, to bound the term $\mathbb{E}[\log \det(\Sigma_1^2)]$, we start by noting that, given $\mathbf{X} = X = DQ$, the rows of the random matrix \mathbf{Y} are i.i.d. complex Gaussian random vectors with zero mean and covariance matrix $I_{t_c} + X^\dagger X$. This means that, given $\mathbf{X} = X = DQ$, the matrix $\mathbf{Y}^\dagger \mathbf{Y}$ has the same distribution as $\mathbf{G}(I_{n_t} + D^2)\mathbf{G}^\dagger + \mathbf{H}\mathbf{H}^\dagger$ where $\mathbf{G} \in \mathbb{C}^{n_t \times n_t}$ and $\mathbf{H} \in \mathbb{C}^{n_r \times (t_c - n_t)}$ are independent matrices with i.i.d. $\mathcal{CN}(0, 1)$ entries. Let $\lambda_1, \lambda_2, \dots, \lambda_l$ denote the l eigenvalues of $\mathbf{G}(I_{n_t} + D^2)\mathbf{G}^\dagger$ (note that this matrix has rank l). Furthermore, let μ_1 denote the largest eigenvalue of $\mathbf{H}\mathbf{H}^\dagger$, and let $\bar{\mu} = \mathbb{E}[\mu_1]$. We have

$$\mathbb{E}[\log \det(\Sigma_1^2)] = \mathbb{E}\left[\sum_{i=1}^l \log(\sigma_i^2)\right] \quad (72)$$

$$= \mathbb{E}_{\mathbf{D}, \mathbf{Q}} \left[\mathbb{E}_{\mathbf{Y} | \mathbf{D}, \mathbf{Q}} \left[\sum_{i=1}^l \log(\sigma_i^2) \middle| \mathbf{D}, \mathbf{Q} \right] \right] \quad (73)$$

$$\leq \mathbb{E}_{\mathbf{D}} \left[\mathbb{E}_{\mathbf{G}, \mathbf{H}} \left[\sum_{i=1}^l \log(\lambda_i + \mu_1) \middle| \mathbf{D} \right] \right] \quad (74)$$

$$\leq \mathbb{E}_{\mathbf{D}} \left[\mathbb{E}_{\mathbf{G}} \left[\sum_{i=1}^l \log(\lambda_i + \bar{\mu}) \middle| \mathbf{D} \right] \right] \quad (75)$$

$$= \mathbb{E}[\log \det(\mathbf{G}(I_{n_t} + \mathbf{D}^2)\mathbf{G}^\dagger + \bar{\mu}I_{n_r})] - (n_r - l)\log(\bar{\mu}). \quad (76)$$

Here, (74) follows from Weyl's theorem [27, Thm. 4.3.1] and in (75) we used Jensen's inequality. Substituting (69), (70), (76) into (68) we obtain

$$\mathcal{C}(\rho) \leq u(\rho) + \frac{1}{t_c} \inf_{\lambda \geq 0} \sup_{\mathcal{J}(\rho)} \mathbb{E}[g(\mathbf{D}, \lambda)] \quad (77)$$

$$\leq u(\rho) + \frac{1}{t_c} \inf_{\lambda \geq 0} \sup_{D \in \mathbb{D}^{n_t \times n_t}} g(D, \lambda) \quad (78)$$

where $u(\cdot)$ and $g(\cdot, \cdot)$ are given in (18) and (19), respectively, and where the last step follows by upper-bounding the supremum over the probability distribution on \mathbf{D} with the supremum over the set of deterministic diagonal matrices D .

APPENDIX C PROOF OF COROLLARY 2

When $n_t = n_r = n$, the upper bound (68) becomes

$$\mathcal{C}(\rho) \leq \frac{1}{t_c} \inf_{\lambda > 0} \sup_{\mathcal{J}(\rho)} \left\{ n^2 \log(\beta) + \log\left(\frac{\mathcal{P}(n)\mathcal{P}(t_c - n)}{\mathcal{P}(t_c)}\right) - nt_c + (t_c - n)\mathbb{E}[\log \det(\mathbf{Y}\mathbf{Y}^\dagger)] \right\}$$

$$- n\mathbb{E}[\log \det(I_n + \mathbf{D}^2)] + \frac{1}{\beta} \mathbb{E}[\text{Tr}\{\mathbf{Y}\mathbf{Y}^\dagger\}] + \lambda(t_c \rho - \mathbb{E}[\text{Tr}\{\mathbf{D}^2\}]) \}. \quad (79)$$

Differently from the general case treated in Appendix B, in the square case some of the terms on the RHS of (79) can be computed in closed form. Specifically,

$$\mathbb{E}[\text{Tr}\{\mathbf{Y}\mathbf{Y}^\dagger\}] = nt_c + n\mathbb{E}[\text{Tr}\{\mathbf{D}^2\}]. \quad (80)$$

To evaluate $\mathbb{E}[\log \det(\mathbf{Y}\mathbf{Y}^\dagger)]$ we use that, given $\mathbf{X} = X = DQ$, the rows of \mathbf{Y} are i.i.d. complex Gaussian random vectors with zero mean and covariance matrix $I_{t_c} + X^\dagger X$. This means that given $\mathbf{X} = X = DQ$, the matrix $\mathbf{Y}\mathbf{Y}^\dagger$ has the same probability distribution as $\mathbf{Z}(I_{t_c} + X^\dagger X)\mathbf{Z}^\dagger$, where $\mathbf{Z} \in \mathbb{C}^{n \times t_c}$ has i.i.d. $\mathcal{CN}(0, 1)$ entries. Hence, we conclude that

$$\mathbb{E}[\log \det(\mathbf{Y}\mathbf{Y}^\dagger)] = \mathbb{E}_{\mathbf{X}} [\mathbb{E}_{\mathbf{Z}} [\log \det(\mathbf{Z}(I_{t_c} + \mathbf{X}^\dagger \mathbf{X})\mathbf{Z}^\dagger)]] \quad (81)$$

$$= \mathbb{E} \left[\frac{\det(I_n + \mathbf{D}^2)^{t_c - n - 1}}{\mathcal{K}(\mathbf{D}^2, t_c - n)} \sum_{k=1}^n \det(R_k(I_n + \mathbf{D}^2)) \right]. \quad (82)$$

Here, the last step follows from Lemma 7 in Appendix A. Substituting (80) and (82) into (79) we obtain

$$\mathcal{C}(\rho) \leq u^*(\rho) + \frac{1}{t_c} \inf_{\lambda \geq 0} \sup_{\mathcal{J}(\rho)} \mathbb{E}[g^*(\mathbf{D}, \lambda)] \quad (83)$$

where $u^*(\cdot)$ and $g^*(\cdot, \cdot)$ are given in (21) and (22) respectively. We conclude the proof by noting that

$$\mathbb{E}[g^*(\mathbf{D}, \lambda)] \leq \sup_{D \in \mathbb{D}^{n \times n}} g^*(D, \lambda) \quad (84)$$

which yields (20).

APPENDIX D PROOF OF THEOREM 3

We lower-bound $\mathcal{C}(\rho)$ by evaluating the mutual information on the RHS of (15) for the MAC-USTM distribution (28), which yields

$$\mathcal{C}(\rho) \geq \mathcal{I}(\mathbf{X}; \mathbf{Y}). \quad (85)$$

Let $\mathbf{X} = \mathbf{U}\mathbf{D}\mathbf{Q}$ denote the singular value decomposition of \mathbf{X} , where \mathbf{U} belongs to the unitary group $\mathcal{U}(n_t)$, the diagonal matrix $\mathbf{D} \in \mathbb{D}^{n_t \times n_t}$ contains the ordered singular values of \mathbf{X} on its main diagonal, and \mathbf{Q} belongs to the Stiefel manifold $\mathcal{S}(n_t, t_c)$. We next decompose the mutual information as

$$\mathcal{I}(\mathbf{X}; \mathbf{Y}) = h(\mathbf{Y}) - h(\mathbf{Y}|\mathbf{X}) \quad (86)$$

and evaluate the two terms on the RHS of (86) separately. Since \mathbf{Y} is conditionally Gaussian given \mathbf{X} , the second term in (86) can be simplified as

$$h(\mathbf{Y}|\mathbf{X}) = n_r \mathbb{E}[\log \det(I_{n_t} + \mathbf{D}^2)] + n_r t_c \log(\pi e). \quad (87)$$

We next evaluate $h(\mathbf{Y})$. Since $\mathbf{X}\mathbf{V} \sim \mathbf{X}$ for every deterministic unitary matrix $\mathbf{V} \in \mathcal{U}(t_c)$, we conclude that \mathbf{Q} is uniformly

distributed over $\mathcal{S}(n_t, t_c)$ and independent of \mathbf{D} . We also have that

$$\mathbf{Y} = \mathbf{S}\mathbf{X} + \mathbf{W} \quad (88)$$

$$= \mathbf{S}\mathbf{U}\mathbf{D}\mathbf{Q} + \mathbf{W} \quad (89)$$

$$\sim \mathbf{S}\mathbf{D}\mathbf{Q} + \mathbf{W}. \quad (90)$$

Here, (89) follows from the singular value decomposition of \mathbf{X} , and (90) holds because \mathbf{S} is isotropically distributed, and, hence, $\mathbf{S}\mathbf{U} \sim \mathbf{S}$, which also implies that $\mathbf{S}\mathbf{U}$ is independent of \mathbf{D} . The fact that \mathbf{Q} is uniformly distributed over $\mathcal{S}(n_t, t_c)$ and independent of \mathbf{D} allows us to invoke [11, Prop. 6] and obtain a closed form expression for the conditional p.d.f. $f_{\mathbf{Y}|\mathbf{D}}$ of \mathbf{Y} given \mathbf{D} . Substituting this closed form expression into

$$h(\mathbf{Y}) = -\mathbb{E}[\log(f_{\mathbf{Y}}(\mathbf{Y}))] \quad (91)$$

$$= -\mathbb{E}[\log(\mathbb{E}_{\mathbf{D}}[f_{\mathbf{Y}|\mathbf{D}}(\mathbf{Y}|\mathbf{D})])] \quad (92)$$

and then substituting (92) and (87) into (86), we obtain (29).

APPENDIX E PROOF OF COROLLARY 4

We derive in closed form the p.d.f. of the eigenvalues of the 2×2 matrix $\mathbf{X}\mathbf{X}^\dagger$, where the two rows $\mathbf{X}_1 \in \mathbb{C}^{1 \times t_c}$ and $\mathbf{X}_2 \in \mathbb{C}^{1 \times t_c}$ of the $2 \times t_c$ dimensional matrix \mathbf{X} are given by

$$\mathbf{X}_i = \sqrt{\frac{t_c \rho}{2}} \mathbf{V}_i, \quad i = 1, 2 \quad (93)$$

with $\mathbf{V}_1 \in \mathbb{C}^{1 \times t_c}$ and $\mathbf{V}_2 \in \mathbb{C}^{1 \times t_c}$ i.i.d. and uniformly distributed on $\mathcal{S}(1, t_c)$ (with respect to the Haar measure). To compute the eigenvalues of $\mathbf{X}\mathbf{X}^\dagger$, we express it as follows

$$\mathbf{X}\mathbf{X}^\dagger = \frac{t_c \rho}{2} \begin{bmatrix} \mathbf{V}_1 \\ \mathbf{V}_2 \end{bmatrix} \begin{bmatrix} \mathbf{V}_1^\dagger & \mathbf{V}_2^\dagger \end{bmatrix} \quad (94)$$

$$= \frac{t_c \rho}{2} \begin{bmatrix} 1 & \mathbf{V}_1 \mathbf{V}_2^\dagger \\ \mathbf{V}_2 \mathbf{V}_1^\dagger & 1 \end{bmatrix}. \quad (95)$$

Let α denote the absolute value of the scalar $\mathbf{V}_1 \mathbf{V}_2^\dagger$. It follows from (95) that the eigenvalues of $\mathbf{X}\mathbf{X}^\dagger$ are $t_c \rho (1 \pm \alpha)/2$. Hence, in order to determine the p.d.f. of the eigenvalues of $\mathbf{X}\mathbf{X}^\dagger$, it is sufficient to obtain the p.d.f. of α . Since

$$\det(\mathbf{X}\mathbf{X}^\dagger) = \left(\frac{t_c \rho}{2}\right)^2 (1 - \alpha^2) \quad (96)$$

Bartlett's decomposition (see [18, Prop. 2.1]) implies that $\alpha^2 \sim \text{Beta}(1, t_c - 1)$. After some mathematical manipulations, we obtain that

$$f_\alpha(x) = (t_c - 1)2x(1 - x^2)^{t_c - 2}, \quad 0 \leq x \leq 1. \quad (97)$$

Using (97) in (29) together with the expression for the eigenvalues of $\mathbf{X}\mathbf{X}^\dagger$, we obtain (31).

REFERENCES

[1] A. Lozano and N. Jindal, "Transmit diversity vs. spatial multiplexing in modern MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 9, no. 1, pp. 186–197, Jan. 2010.
[2] A. Lapidath, "On the asymptotic capacity of stationary Gaussian fading channels," *IEEE Trans. Inf. Theory*, vol. 51, no. 2, pp. 437–446, Feb. 2005.

[3] S. Moser, "The fading number of multiple-input multiple-output fading channels with memory," *IEEE Trans. Inf. Theory*, vol. 55, no. 6, pp. 2716–2755, Jun. 2009.
[4] W. Yang, G. Durisi, T. Koch, and Y. Polyanskiy, "Diversity versus channel knowledge at finite block-length," in *Proc. IEEE Inf. Theory Workshop (ITW)*, Lausanne, Switzerland, Sep. 2012, pp. 577–581.
[5] B. Hochwald and T. Marzetta, "Unitary space-time modulation for multiple-antenna communications in Rayleigh flat fading," *IEEE Trans. Inf. Theory*, vol. 46, no. 2, pp. 543–564, Mar. 2000.
[6] T. L. Marzetta and B. M. Hochwald, "Capacity of a mobile multiple-antenna communication link in Rayleigh flat fading," *IEEE Trans. Inf. Theory*, vol. 45, no. 1, pp. 139–157, Jan. 1999.
[7] L. Zheng and D. N. C. Tse, "Communication on the Grassmann manifold: A geometric approach to the noncoherent multiple-antenna channel," *IEEE Trans. Inf. Theory*, vol. 48, no. 2, pp. 359–383, Feb. 2002.
[8] W. Yang, G. Durisi, and E. Riegler, "On the capacity of large-MIMO block-fading channels," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 117–132, Feb. 2013.
[9] F. Rusek, A. Lozano, and N. Jindal, "Mutual information of IID complex Gaussian signals on block Rayleigh-faded channels," *IEEE Trans. Inf. Theory*, vol. 58, no. 1, pp. 331–340, Jan. 2012.
[10] B. Hassibi and T. L. Marzetta, "Multiple-antennas and isotropically random unitary inputs: the received signal density in closed form," *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1473–1484, Jun. 2002.
[11] G. Alfano, C. Chiasserini, A. Nordin, and S. Zhou, "Closed-form output statistics of MIMO block-fading channels," *IEEE Trans. Inf. Theory*, vol. 60, no. 12, pp. 7782–7797, Dec. 2014.
[12] C. Itzykson and J. Zuber, "The planar approximation. II," *J. Math. Phys.*, vol. 21, no. 3, pp. 411–421, 1980.
[13] S. Shamai and T. Marzetta, "Multiuser capacity in block fading with no channel state information," *IEEE Trans. Inf. Theory*, vol. 48, no. 4, pp. 938–942, Apr. 2002.
[14] G.-R. Lin and S. Moser, "The fading number of a multiple-access Rician fading channel," *IEEE Trans. Inf. Theory*, vol. 57, no. 8, pp. 4983–4991, Aug. 2011.
[15] A. Lapidath and S. M. Moser, "Capacity bounds via duality with applications to multiple-antenna systems on flat-fading channels," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2426–2467, Oct. 2003.
[16] U. Mendlovic and M. Feder, "Collaboration gain in MAC is limited," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Honolulu, HI, Jun. 2014, pp. 3122–3126.
[17] M. Abramowitz and I. A. Stegun, Eds., *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, 10th ed. New York: Dover: Government Printing Office, 1972.
[18] A. Rouault, "Asymptotic behavior of random determinants in the Laguerre, Gram and Jacobi ensembles," *ALEA Lat. Am. J. Probab. Math. Stat.*, vol. 3, pp. 181–230, 2007.
[19] M. Chiani, M. Z. Win, and A. Zanella, "On the capacity of spatially correlated MIMO Rayleigh-fading channels," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2363–2371, Oct. 2003.
[20] E. Telatar, "Capacity of multi-antenna Gaussian channels," *Eur. Trans. Telecommun.*, vol. 10, no. 6, pp. 585–595, Nov. 1999.
[21] H. Shin, M. Win, J. H. Lee, and M. Chiani, "On the capacity of doubly correlated MIMO channels," *IEEE Trans. Wireless Commun.*, vol. 5, no. 8, pp. 2253–2265, Aug. 2006.
[22] A. Ghaderipoor, C. Tellambura, and A. Paulraj, "On the application of character expansions for MIMO capacity analysis," *IEEE Trans. Inf. Theory*, vol. 58, no. 5, pp. 2950–2962, May 2012.
[23] A. Lozano, A. Tulino, and S. Verdú, "High-SNR power offset in multi-antenna communication," *IEEE Trans. Inf. Theory*, vol. 51, no. 12, pp. 4134–4151, Dec. 2005.
[24] H. Gao and P. J. Smith, "A determinant representation for the distribution of quadratic forms in complex normal vectors," *J. Multivariate Analysis*, vol. 73, no. 2, pp. 155–165, 2000.
[25] C. Khatri, "On certain distribution problems based on positive definite quadratic functions in normal vectors," *Ann. Math. Statist.*, pp. 468–479, 1966.
[26] A. Edelman, "Eigenvalues and condition numbers of random matrices," Ph.D. dissertation, Massachusetts Institute of Technology, Cambridge, MA, May 1989.
[27] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 1985.