Quasi-Static Multiple-Antenna Fading Channels at Finite Blocklength

Wei Yang, *Student Member, IEEE*, Giuseppe Durisi, *Senior Member, IEEE*,
Tobias Koch, *Member, IEEE*, and Yury Polyanskiy, *Member, IEEE*

Abstract—This paper investigates the maximal achievable rate for a given blocklength and error probability over quasi-static multiple-input multiple-output (MIMO) fading channels, with and without channel state information (CSI) at the transmitter and/or the receiver. The principal finding is that outage capacity, despite being an asymptotic quantity, is a sharp proxy for the finite-blocklength fundamental limits of slow-fading channels. Specifically, the channel dispersion is shown to be zero regardless of whether the fading realizations are available at both transmitter and receiver, at only one of them, or at neither of them. These results follow from analytically tractable converse and achievability bounds. Numerical evaluation of these bounds verifies that zero dispersion may indeed imply fast convergence to the outage capacity as the blocklength increases. In the example of a particular 1×2 single-input multiple-output (SIMO) Rician fading channel, the blocklength required to achieve 90% of capacity is about an order of magnitude smaller compared to the blocklength required for an AWGN channel with the same capacity. For this specific scenario, the coding/decoding schemes adopted in the LTE-Advanced standard are benchmarked against the finite-blocklength achievability and converse bounds.

I. INTRODUCTION

Consider a delay-constrained communication system operating over a slowly-varying fading channel. In such a scenario, it is plausible to assume that the duration of each of the transmitted codewords is smaller than the coherence time of the channel, so the random fading coefficients stay constant over the duration of each codeword [1, p. 2631], [2, Sec. 5.4.1]. We shall refer to this channel model as *quasi-static fading channel*.[1]

When communicating over quasi-static fading channels at a given rate R, the realization of the random fading coefficient may be very small, in which case the block (frame) error probability ϵ is bounded away from zero even if the blocklength n tends

W. Yang and G. Durisi are with the Department of Signals and Systems, Chalmers University of Technology, 41296, Gothenburg, Sweden (e-mail: {ywei, durisi}@chalmers.se).

T. Koch is with the Signal Theory and Communications Department, Universidad Carlos III de Madrid, 28911, Leganés, Spain (e-mail: koch@tsc.uc3m.es).

Y. Polyanskiy is with the Department of Electrical Engineering and Computer Science, MIT, Cambridge, MA, 02139 USA (e-mail: yp@mit.edu).

[1]The term "quasi-static" is widely used in the communication literature (see, e.g., [2, Sec. 5.4.1], [3]). The quasi-static channel model belongs to the general class of *composite channels* [1, p. 2631], [4] (also known as *mixed channels* [5, Sec. 3.3]).

to infinity. In this case, the channel is said to be in *outage*. For fading distributions for which the fading coefficient can be arbitrarily small (such as for Rayleigh, Rician, or Nakagami fading), the probability of an outage is positive. Hence, the overall block error probability ϵ is bounded away from zero for every positive rate $R > 0$, in which case the Shannon capacity is zero. More generally, the Shannon capacity depends on the fading probability density function (pdf) only through its support [6], [7].

For applications in which a positive block error probability $\epsilon > 0$ is acceptable, the maximal achievable rate as a function of the outage probability (also known as *capacity versus outage*) [1, p. 2631], [8], may be a more relevant performance metric than Shannon capacity. The capacity versus outage coincides with the ϵ-capacity C_ϵ (which is the largest achievable rate under the assumption that the block error probability is less than $\epsilon > 0$) at the points where C_ϵ is a continuous function of ϵ [7, Sec. IV].

For the sake of simplicity, let us consider for a moment a single-antenna communication system operating over a quasi-static flat-fading channel. The outage probability as a function of the rate R is defined by

$$F(R) = \mathbb{P}\big[\log(1 + |H|^2 \rho) < R\big]. \tag{1}$$

Here, H denotes the random channel gain and ρ is the signal-to-noise ratio (SNR). For a given $\epsilon > 0$, the outage capacity (or ϵ-capacity) C_ϵ is the supremum of all rates R satisfying $F(R) \leq \epsilon$. The rationale behind this definition is that, for every realization of the fading coefficient $H = h$, the quasi-static fading channel can be viewed as an AWGN channel with channel gain $|h|^2$, for which communication with arbitrarily small block error probability is feasible if and only if $R < \log(1 + |h|^2 \rho)$, provided that the blocklength n is sufficiently large. Thus, the outage probability can be interpreted as the probability that the channel gain H is too small to allow for communication with arbitrarily small block error probability.

A major criticism of this definition is that it is somewhat contradictory to the underlying motivation of the channel model. Indeed, while $\log(1 + |h|^2 \rho)$ is meaningful only for codewords of sufficiently large blocklength, the assumption that the fading coefficient is constant during the transmission of the codeword is only reasonable if the blocklength is smaller than the coherence time of the channel. In other words, it is *prima facie* unclear whether for those blocklengths for which the quasi-static channel model is reasonable, the outage capacity is a meaningful performance metric.

In order to shed light on this issue, we study the maximal achievable rate $R^*(n, \epsilon)$ for a given blocklength n and block

error probability ϵ over a quasi-static multiple-input multiple-output (MIMO) fading channel, subject to a per-codeword power constraint.

Previous results: Building upon Dobrushin's and Strassen's asymptotic results, Polyanskiy, Poor, and Verdú recently showed that for various channels with positive Shannon capacity C, the maximal achievable rate can be tightly approximated by [9]

$$R^*(n, \epsilon) = C - \sqrt{\frac{V}{n}} Q^{-1}(\epsilon) + \mathcal{O}\left(\frac{\log n}{n}\right). \qquad (2)$$

Here, $Q^{-1}(\cdot)$ denotes the inverse of the Gaussian Q-function

$$Q(x) \triangleq \int_x^\infty \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt \qquad (3)$$

and V is the *channel dispersion* [9, Def. 1]. The approximation (2) implies that to sustain the desired error probability ϵ at a finite blocklength n, one pays a penalty on the rate (compared to the channel capacity) that is proportional to $1/\sqrt{n}$.

Recent works have extended (2) to some ergodic fading channels. Specifically, the dispersion of single-input single-output (SISO) stationary fading channels for the case when channel state information (CSI) is available at the receiver was derived in [10]. This result was extended to block-memoryless fading channels in [11]. Upper and lower bounds on the second-order coding rate of quasi-static MIMO Rayleigh-fading channels have been reported in [12] for the asymptotically ergodic setup when the number of antennas grows linearly with the blocklength. A lower bound on $R^*(n, \epsilon)$ for the imperfect CSI case has been developed in [13]. The second-order coding rate of single-antenna quasi-static fading channels for the case of perfect CSI and long-term power constraint has been derived in [14].

Contributions: We provide achievability and converse bounds on $R^*(n, \epsilon)$ for quasi-static MIMO fading channels. We consider both the case when the transmitter has full transmit CSI (CSIT) and, hence, can perform spatial water-filling, and the case when no CSIT is available. Our converse results are obtained under the assumption of perfect receive CSI (CSIR), whereas the achievability results are derived under the assumption of no CSIR.

By analyzing the asymptotic behavior of our achievability and converse bounds, we show that under mild conditions on the fading distribution,[2]

$$R^*(n, \epsilon) = C_\epsilon + \mathcal{O}\left(\frac{\log n}{n}\right). \qquad (4)$$

This results holds both for the case of perfect CSIT and for the case of no CSIT, and independently on whether CSIR is available at the receiver or not. By comparing (2) with (4), we observe that for the quasi-static fading case, the $1/\sqrt{n}$ rate penalty is absent. In other words, the ϵ-dispersion (see [9, Def. 2] or (52) below) of quasi-static fading channels is *zero*. This suggests that the maximal achievable rate $R^*(n, \epsilon)$ converges quickly to C_ϵ as n tends to infinity, thereby indicating that the outage capacity is indeed a meaningful performance metric for delay-constrained communication over slowly-varying fading channels.

[2]These conditions are satisfied by the fading distributions commonly used in the wireless communication literature (e.g., Rayleigh, Rician, Nakagami).

Fast convergence to the outage capacity provides mathematical support to the observation reported by several researchers in the past that the outage probability describes accurately the performance over quasi-static fading channels of actual codes (see [15] and references therein).

The following example supports our claims: for a 1×2 single-input multiple-output (SIMO) Rician-fading channel with $C_\epsilon = 1$ bit/channel use and $\epsilon = 10^{-3}$, the blocklength required to achieve 90% of C_ϵ for the perfect CSIR case is between 120 and 320 (see Fig. 2 on p. 10), which is about an order of magnitude smaller compared to the blocklength required for an AWGN channel with the same capacity (see [9, Fig. 12]).

Fast convergence to the outage capacity further suggests that communication strategies that are optimal with respect to outage capacity may perform also well at finite blocklength. Note, however, that this need not be true for very small blocklengths, where the $\mathcal{O}(n^{-1} \log n)$ term in (4) may dominate. Thus, for small n the derived achievability and converse bounds on $R^*(n, \epsilon)$ may behave differently than the outage capacity. Table I summarizes how the outage capacity and the achievability/converse bounds on $R^*(n, \epsilon)$ derived in this paper depend on system parameters such as the availability of CSI and the number of antennas at the transmitter/receiver. These observations may be relevant for delay-constrained communication over slowly-varying fading channels.

Proof techniques: Our converse bounds on $R^*(n, \epsilon)$ are based on the meta-converse theorem [9, Th. 30]. Our achievability bounds on $R^*(n, \epsilon)$ are based on the $\kappa\beta$ bound [9, Th. 25] applied to a stochastically degraded channel, whose choice is motivated by geometric considerations. The main tools used to establish (4) are a Cramer-Esseen-type central-limit theorem [16, Th. VI.1] and a result on the speed of convergence of $\mathbb{P}[B > A/\sqrt{n}]$ to $\mathbb{P}[B > 0]$ for $n \to \infty$, where A and B are independent random variables.

Notation: Upper case letters such as X denote scalar random variables and their realizations are written in lower case, e.g., x. We use boldface upper case letters to denote random vectors, e.g., $\boldsymbol{X}$, and boldface lower case letters for their realizations, e.g., $\boldsymbol{x}$. Upper case letters of two special fonts are used to denote deterministic matrices (e.g., Y) and random matrices (e.g., $\mathbb{Y}$). The superscripts $^\mathrm{T}$ and $^\mathrm{H}$ stand for transposition and Hermitian transposition, respectively. We use $\mathrm{tr}(\mathsf{A})$ and $\det(\mathsf{A})$ to denote the trace and determinant of the matrix A, respectively, and use $\mathrm{span}(\mathsf{A})$ to designate the subspace spanned by the column vectors of A. The Frobenius norm of a matrix A is denoted by $\|\mathsf{A}\|_\mathsf{F} \triangleq \sqrt{\mathrm{tr}(\mathsf{A}\mathsf{A}^\mathrm{H})}$. The notation $\mathsf{A} \succeq \boldsymbol{0}$ means that the matrix A is positive semi-definite. The function resulting from the composition of two functions f and g is denoted by $g \circ f$, i.e., $(g \circ f)(x) = g(f(x))$. For two functions $f(x)$ and $g(x)$, the notation $f(x) = \mathcal{O}(g(x))$, $x \to \infty$, means that $\limsup_{x\to\infty} |f(x)/g(x)| < \infty$, and $f(x) = o(g(x))$, $x \to \infty$, means that $\lim_{x\to\infty} |f(x)/g(x)| = 0$. We use I_a to denote the identity matrix of size $a \times a$, and designate by $\mathsf{I}_{a,b}$ $(a > b)$ the $a \times b$ matrix containing the first b columns of I_a. The distribution of a circularly-symmetric complex Gaussian random vector with covariance matrix A is denoted by $\mathcal{CN}(\boldsymbol{0}, \mathsf{A})$, the Wishart distribution [18, Def. 2.3] with n degrees of freedom and covariance matrix A defined on matrices of size $m \times m$ is denoted

TABLE I
OUTAGE CAPACITY VS. FINITE BLOCKLENGTH WISDOM; t IS THE NUMBER OF TRANSMIT ANTENNAS.

Wisdom	C_ϵ	Bounds on $R^*(n, \epsilon)$
CSIT is beneficial	only if $t > 1$	only if $t > 1$
CSIR is beneficial	no [1, p. 2632]	yes
With CSIT, waterfilling is optimal	yes [17]	no
With CSIT, the channel is reciprocal[3]	yes [17]	only with CSIR

by $\mathcal{W}_m(n, \mathsf{A})$, and the Beta distribution [19, Ch. 25] is denoted by $\mathrm{Beta}(\cdot, \cdot)$. The symbol $\mathbb{R}_+$ stands for the nonnegative real line, $\mathbb{R}_+^m \subset \mathbb{R}^m$ is the nonnegative orthant of the m-dimensional real Euclidean spaces, and $\mathbb{R}_\geq^m \subset \mathbb{R}_+^m$ is defined by

$$\mathbb{R}_\geq^m \triangleq \{\boldsymbol{x} \in \mathbb{R}_+^m : x_1 \geq \cdots \geq x_m\}. \tag{5}$$

The indicator function is denoted by $\mathbb{1}\{\cdot\}$, and $[\cdot]^+ \triangleq \max\{\cdot, 0\}$. Finally, $\log(\cdot)$ is the natural logarithm.

Given two distributions P and Q on a common measurable space $\mathcal{W}$, we define a randomized test between P and Q as a random transformation $P_{Z\,|\,W} : \mathcal{W} \mapsto \{0, 1\}$ where 0 indicates that the test chooses Q. We shall need the following performance metric for the test between P and Q:

$$\beta_\alpha(P, Q) \triangleq \min \int P_{Z\,|\,W}(1\,|\,w) Q(dw) \tag{6}$$

where the minimum is over all probability distributions $P_{Z\,|\,W}$ satisfying

$$\int P_{Z\,|\,W}(1\,|\,w) P(dw) \geq \alpha. \tag{7}$$

II. SYSTEM MODEL

We consider a quasi-static MIMO fading channel with t transmit and r receive antennas. Throughout this paper, we denote the minimum number of transmit and receive antennas by m, i.e., $m \triangleq \min\{t, r\}$. The channel input-output relation is given by

$$\mathbb{Y} = \mathbb{X}\mathbb{H} + \mathbb{W}. \tag{8}$$

Here, $\mathbb{X} \in \mathbb{C}^{n \times t}$ is the signal transmitted over n channel uses; $\mathbb{Y} \in \mathbb{C}^{n \times r}$ is the corresponding received signal; the matrix $\mathbb{H} \in \mathbb{C}^{t \times r}$ contains the complex fading coefficients, which are random but remain constant over the n channel uses; $\mathbb{W} \in \mathbb{C}^{n \times r}$ denotes the additive noise at the receiver, which is independent of $\mathbb{H}$ and has independent and identically distributed (i.i.d.) $\mathcal{CN}(0, 1)$ entries.

We consider the following four scenarios:

1) no-CSI: neither the transmitter nor the receiver is aware of the realizations of the fading matrix $\mathbb{H}$;
2) CSIT: the transmitter knows $\mathbb{H}$;
3) CSIR: the receiver knows $\mathbb{H}$;
4) CSIRT: both the transmitter and the receiver know $\mathbb{H}$.

To keep the notation compact, we shall abbreviate in mathematical formulas the acronyms no-CSI, CSIT, CSIR, and CSIRT as

no, tx, rx, and rt, respectively. Next, we introduce the notion of a channel code for each of these four settings.

Definition 1 (no-CSI): An $(n, M, \epsilon)_{\mathrm{no}}$ code consists of:

i) an encoder $f_{\mathrm{no}}: \{1, \ldots, M\} \mapsto \mathbb{C}^{n \times t}$ that maps the message $J \in \{1, \ldots, M\}$ to a codeword $\mathsf{X} \in \{\mathsf{C}_1, \ldots, \mathsf{C}_M\}$. The codewords satisfy the power constraint

$$\|\mathsf{C}_i\|_\mathsf{F}^2 \leq n\rho, \quad i = 1, \ldots, M. \tag{9}$$

ii) A decoder $g_{\mathrm{no}}: \mathbb{C}^{n \times r} \mapsto \{1, \ldots, M\}$ satisfying a *maximum probability of error* constraint

$$\max_{1 \leq j \leq M} \mathbb{P}[g_{\mathrm{no}}(\mathbb{Y}) \neq J \,|\, J = j] \leq \epsilon \tag{10}$$

where $\mathbb{Y}$ is the channel output induced by the transmitted codeword $\mathsf{X} = f_{\mathrm{no}}(j)$ according to (8).

Definition 2 (CSIR): An $(n, M, \epsilon)_{\mathrm{rx}}$ code consists of:

i) an encoder $f_{\mathrm{no}}: \{1, \ldots, M\} \mapsto \mathbb{C}^{n \times t}$ that maps the message $J \in \{1, \ldots, M\}$ to a codeword $\mathsf{X} \in \{\mathsf{C}_1, \ldots, \mathsf{C}_M\}$. The codewords satisfy the power constraint (9).

ii) A decoder $g_{\mathrm{rx}}: \mathbb{C}^{n \times r} \times \mathbb{C}^{t \times r} \mapsto \{1, \ldots, M\}$ satisfying

$$\max_{1 \leq j \leq M} \mathbb{P}[g_{\mathrm{rx}}(\mathbb{Y}, \mathbb{H}) \neq J \,|\, J = j] \leq \epsilon. \tag{11}$$

Definition 3 (CSIT): An $(n, M, \epsilon)_{\mathrm{tx}}$ code consists of:

i) an encoder $f_{\mathrm{tx}}: \{1, \ldots, M\} \times \mathbb{C}^{t \times r} \mapsto \mathbb{C}^{n \times t}$ that maps the message $j \in \{1, \ldots, M\}$ and the channel H to a codeword $\mathsf{X} = f_{\mathrm{tx}}(j, \mathsf{H})$ satisfying

$$\|\mathsf{X}\|_\mathsf{F}^2 = \|f_{\mathrm{tx}}(j, \mathsf{H})\|_\mathsf{F}^2 \leq n\rho,$$
$$\forall j = 1, \ldots, M, \forall \mathsf{H} \in \mathbb{C}^{t \times r}. \tag{12}$$

ii) A decoder $g_{\mathrm{no}}: \mathbb{C}^{n \times r} \mapsto \{1, \ldots, M\}$ satisfying (10).

Definition 4 (CSIRT): An $(n, M, \epsilon)_{\mathrm{rt}}$ code consists of:

i) an encoder $f_{\mathrm{tx}}: \{1, \ldots, M\} \times \mathbb{C}^{t \times r} \mapsto \mathbb{C}^{n \times t}$ that maps the message $j \in \{1, \ldots, M\}$ and the channel H to a codeword $\mathsf{X} = f_{\mathrm{tx}}(j, \mathsf{H})$ satisfying (12).

ii) A decoder $g_{\mathrm{rx}}: \mathbb{C}^{n \times r} \times \mathbb{C}^{t \times r} \mapsto \{1, \ldots, M\}$ satisfying (11).

The maximal achievable rate for the four cases listed above is defined as follows:

$$R_l^*(n, \epsilon) \triangleq \sup\left\{\frac{\log M}{n} : \exists (n, M, \epsilon)_l \text{ code}\right\},$$
$$l \in \{\mathrm{no}, \mathrm{rx}, \mathrm{tx}, \mathrm{rt}\}. \tag{13}$$

From Definitions 1–4, it follows that

$$R_{\mathrm{no}}^*(n, \epsilon) \leq R_{\mathrm{tx}}^*(n, \epsilon) \leq R_{\mathrm{rt}}^*(n, \epsilon) \tag{14}$$

$$R_{\mathrm{no}}^*(n, \epsilon) \leq R_{\mathrm{rx}}^*(n, \epsilon) \leq R_{\mathrm{rt}}^*(n, \epsilon). \tag{15}$$

[3]A channel is reciprocal for a given performance metric (e.g., outage capacity) if substituting $\mathbb{H}$ with $\mathbb{H}^\mathsf{H}$ does not change the metric.

III. ASYMPTOTIC RESULTS AND PREVIEW

It was noted in [1, p. 2632] that the ϵ-capacity of quasi-static MIMO fading channel does not depend on whether CSI is available at the receiver. Intuitively, this is true because the channel stays constant during the transmission of a codeword, so it can be accurately estimated at the receiver through the transmission of $\sqrt{n}$ pilot symbols with no rate penalty as $n \to \infty$. A rigorous proof of this statement follows by our zero-dispersion results (Theorems 3 and 9). In contrast, if CSIT is available and $t > 1$, then water-filling over space yields a larger ϵ-capacity [15]. We next define C_ϵ for both the CSIT and the no-CSIT case.

Let $\mathcal{U}_t$ be the set of $t \times t$ positive semidefinite matrices whose trace is upper-bounded by ρ, i.e.,

$$\mathcal{U}_t \triangleq \{\mathsf{A} \in \mathbb{C}^{t \times t} : \mathsf{A} \succeq \mathbf{0}, \operatorname{tr}(\mathsf{A}) \leq \rho\}. \tag{16}$$

When CSI is available at the transmitter, the ϵ-capacity C_ϵ^{tx} is given by [15, Prop. 2][4]

$$C_\epsilon^{\mathrm{tx}} = \lim_{n \to \infty} R_{\mathrm{tx}}^*(n, \epsilon) \tag{17}$$

$$= \lim_{n \to \infty} R_{\mathrm{rt}}^*(n, \epsilon) \tag{18}$$

$$= \sup\{R : F_{\mathrm{tx}}(R) \leq \epsilon\} \tag{19}$$

where

$$F_{\mathrm{tx}}(R) \triangleq \mathbb{P}\left[\max_{\mathsf{Q} \in \mathcal{U}_t} \log \det\left(\mathsf{I}_r + \mathbb{H}^{\mathrm{H}} \mathsf{Q} \mathbb{H}\right) < R\right] \tag{20}$$

denotes the outage probability. Given $\mathbb{H} = \mathsf{H}$, the function $\log \det\left(\mathsf{I}_r + \mathsf{H}^{\mathrm{H}} \mathsf{Q} \mathsf{H}\right)$ in (20) is maximized by the well-known water-filling power-allocation strategy (see, e.g., [17]), which results in

$$\max_{\mathsf{Q} \in \mathcal{U}_t} \log \det\left(\mathsf{I}_r + \mathsf{H}^{\mathrm{H}} \mathsf{Q} \mathsf{H}\right) = \sum_{j=1}^m [\log(\bar{\gamma}\lambda_j)]^+ \tag{21}$$

where the scalars $\lambda_1 \geq \cdots \geq \lambda_m$ denote the m largest eigenvalues of $\mathsf{H}^{\mathrm{H}}\mathsf{H}$, and $\bar{\gamma}$ is the solution of

$$\sum_{j=1}^m [\bar{\gamma} - 1/\lambda_j]^+ = \rho. \tag{22}$$

In Section IV, we study quasi-static MIMO channels with CSIT at finite blocklength. We present an achievability (lower) bound on $R_{\mathrm{tx}}^*(n, \epsilon)$ (Section IV-A, Theorem 1) and a converse (upper) bound on $R_{\mathrm{rt}}^*(n, \epsilon)$ (Section IV-B, Theorem 2). We show in Section IV-C (Theorem 3) that, under mild conditions on the fading distribution, the two bounds match asymptotically up to a $\mathcal{O}(\log(n)/n)$ term. This allows us to establish the zero-dispersion result (4) for the CSIT case.

When CSI is not available at the transmitter, the ϵ-capacity C_ϵ^{no} is given by [17], [6]

$$C_\epsilon^{\mathrm{no}} = \lim_{n \to \infty} R_{\mathrm{no}}^*(n, \epsilon) \tag{23}$$

$$= \lim_{n \to \infty} R_{\mathrm{rx}}^*(n, \epsilon) \tag{24}$$

$$= \sup\{R : F_{\mathrm{no}}(R) \leq \epsilon\} \tag{25}$$

where

$$F_{\mathrm{no}}(R) \triangleq \inf_{\mathsf{Q} \in \mathcal{U}_t} \mathbb{P}\left[\log \det\left(\mathsf{I}_r + \mathbb{H}^{\mathrm{H}} \mathsf{Q} \mathbb{H}\right) < R\right] \tag{26}$$

is the outage probability for the no-CSIT case. The matrix Q that minimizes the right-hand side (RHS) of (26) is in general not known, making this case more difficult to analyze and our nonasymptotic results less sharp and more difficult to evaluate numerically. The minimization in (26) can be restricted to all Q on the boundary of $\mathcal{U}_t$, i.e.,

$$F_{\mathrm{no}}(R) = \inf_{\mathsf{Q} \in \mathcal{U}_t^{\mathrm{e}}} \mathbb{P}\left[\log \det\left(\mathsf{I}_r + \mathbb{H}^{\mathrm{H}} \mathsf{Q} \mathbb{H}\right) < R\right] \tag{27}$$

where

$$\mathcal{U}_t^{\mathrm{e}} \triangleq \{\mathsf{A} \in \mathbb{C}^{t \times t} : \mathsf{A} \succeq \mathbf{0}, \operatorname{tr}(\mathsf{A}) = \rho\}. \tag{28}$$

We lower-bound $R_{\mathrm{no}}^*(n, \epsilon)$ in Section V-A (Theorem 4), and upper-bound $R_{\mathrm{rx}}^*(n, \epsilon)$ in Section V-B (Theorem 6). The asymptotic analysis of the bounds provided in Section V-C (Theorem 9) allows us to establish (4), although under slightly more stringent assumptions on the fading probability distribution than for the CSIT case.

For the i.i.d. Rayleigh-fading model (without CSIT), Telatar [17] conjectured that the optimal Q is of the form[5]

$$\frac{\rho}{t^*} \operatorname{diag}\{\underbrace{1, \ldots, 1}_{t^*}, \underbrace{0, \ldots, 0}_{t-t^*}\}, \quad 1 \leq t^* \leq t \tag{29}$$

and that for small ϵ values or for high SNR values, all available transmit antennas should be used, i.e., $t^* = t$. We define the ϵ-rate $C_\epsilon^{\mathrm{iso}}$ resulting from the choice $\mathsf{Q} = (\rho/t)\mathsf{I}_t$ as

$$C_\epsilon^{\mathrm{iso}} \triangleq \sup\{R : F_{\mathrm{iso}}(R) \leq \epsilon\} \tag{30}$$

where

$$F_{\mathrm{iso}}(R) \triangleq \mathbb{P}\left[\log \det\left(\mathsf{I}_r + \frac{\rho}{t} \mathbb{H}^{\mathrm{H}} \mathbb{H}\right) < R\right]. \tag{31}$$

The ϵ-rate $C_\epsilon^{\mathrm{iso}}$ is often taken as an accurate lower bound on the actual ϵ-capacity for the case of i.i.d Rayleigh fading and no CSIT. Motivated by this fact, we consider in Section V codes with isotropic codewords, i.e., chosen from the set

$$\mathcal{F}_{\mathrm{iso}} \triangleq \left\{\mathsf{X} \in \mathbb{C}^{n \times t} : \frac{1}{n} \mathsf{X}^{\mathrm{H}} \mathsf{X} = \frac{\rho}{t} \mathsf{I}_t\right\}. \tag{32}$$

We indicate by $(n, M, \epsilon)_{\mathrm{iso}}$ a code with M codewords chosen from $\mathcal{F}_{\mathrm{iso}}$ and with a maximal error probability smaller than ϵ. For this special class of codes, the maximal achievable rate $R_{\mathrm{no,iso}}^*(n, \epsilon)$ for the no-CSI case and $R_{\mathrm{rx,iso}}^*(n, \epsilon)$ for the CSIR case can be characterized more accurately at finite blocklength (Theorem 8) than for the general no-CSI case. Furthermore, we show in Section V-C (Theorem 11) that under mild conditions on the fading distributions (weaker than the ones required for the general no-CSI case)

$$\{R_{\mathrm{no,iso}}^*(n, \epsilon), R_{\mathrm{rx,iso}}^*(n, \epsilon)\} = C_\epsilon^{\mathrm{iso}} + \mathcal{O}\left(\frac{\log n}{n}\right). \tag{33}$$

A final remark on notation. For the single-transmit-antenna case (i.e., $t = 1$), the ϵ-capacity does not depend on whether CSIT is available or not [15, Prop. 3]. Hence, we shall denote the ϵ-capacity for this case simply as C_ϵ.

---

[4]More precisely, (19) and (25) hold provided that C_ϵ^{tx} and C_ϵ^{no} are continuous functions of ϵ [7, Th. 6].

[5]This conjecture has recently been proved for the multiple-input single-output case [20].

IV. CSI AVAILABLE AT THE TRANSMITTER

A. Achievability

In this section, we consider the case where CSI is available at the transmitter but not at the receiver. Before establishing our achievability bound in Section IV-A2, we provide some geometric intuition that will guide us in the choice of the decoder g_{no} (see Definition 3).

1) Geometric Intuition: Consider for simplicity a real-valued quasi-static SISO channel ($t = r = 1$), i.e., a channel with input-output relation

$$\boldsymbol{Y} = H\boldsymbol{x} + \boldsymbol{W} \tag{34}$$

where $\boldsymbol{Y}$, $\boldsymbol{x}$, and $\boldsymbol{W}$ are n-dimensional vectors, and H is a (real-valued) scalar. As reviewed in Section I, the typical error event for the quasi-static fading channel (in the large blocklength regime) is that the instantaneous channel gain H^2 is not large enough to support the desired rate R, i.e., $\frac{1}{2}\log(1 + \rho H^2) < R$ (outage event). For the channel in (34), the ϵ-capacity C_ϵ, i.e., the largest rate R for which the probability that the channel is in outage is less than ϵ, is given by

$$C_\epsilon = \sup\left\{R : \mathbb{P}\left[\frac{1}{2}\log(1 + \rho H^2) < R\right] \le \epsilon\right\}. \tag{35}$$

Roughly speaking, the decoder of a C_ϵ-achieving code may commit an error only when the channel is in outage. Pick now an arbitrary codeword $\boldsymbol{x}_1$ from the hypersphere $\{\boldsymbol{x} \in \mathbb{R}^n : \|\boldsymbol{x}\|^2 = n\rho\}$, and let $\boldsymbol{Y}$ be the received signal corresponding to $\boldsymbol{x}_1$. Following [21], we analyze the angle $\theta(\boldsymbol{x}_1, \boldsymbol{Y})$ between $\boldsymbol{x}_1$ and $\boldsymbol{Y}$ as follows. By the law of large numbers, the noise vector $\boldsymbol{W}$ is approximately orthogonal to $\boldsymbol{x}_1$ if n is large, i.e.,

$$\frac{\langle \boldsymbol{x}_1, \boldsymbol{W}\rangle}{\|\boldsymbol{x}_1\|\|\boldsymbol{W}\|} \to 0, \quad n \to \infty. \tag{36}$$

Also by the law of large numbers, $\|\boldsymbol{W}\|^2/n \to 1$ as $n \to \infty$. Hence, for a given H and for large n, the angle $\theta(\boldsymbol{x}_1, \boldsymbol{Y})$ can be approximated as

$$\theta(\boldsymbol{x}_1, \boldsymbol{Y}) \approx \arcsin\frac{\|\boldsymbol{W}\|}{\sqrt{H^2\|\boldsymbol{x}_1\|^2 + \|\boldsymbol{W}\|^2}} \tag{37}$$

$$\approx \arcsin\frac{1}{\sqrt{\rho H^2 + 1}} \tag{38}$$

where the first approximation follows by (36) and the second approximation follows because $\|\boldsymbol{W}\|^2 \approx n$. It follows from (35) and (38) that $\theta(\boldsymbol{x}_1, \boldsymbol{Y})$ is larger than $\theta_\epsilon \triangleq \arcsin(e^{-C_\epsilon})$ in the outage case, and smaller than θ_ϵ otherwise (see Fig. 1).

This geometric argument suggests the use of a threshold decoder that, for a given received signal $\boldsymbol{Y}$, declares $\boldsymbol{x}_i$ to be the transmitted codeword if $\boldsymbol{x}_i$ is the only codeword for which $\theta(\boldsymbol{x}_i, \boldsymbol{Y}) \le \theta_\epsilon$. If no codewords or more than one codeword meet this condition, the decoder declares an error. Thresholding angles instead of log-likelihood ratios (cf., [9, Th. 17 and Th. 25]) appears to be a natural approach when CSIR is unavailable. Note that the proposed threshold decoder does neither require CSIR nor knowledge of the fading distribution. As we shall see, it achieves (4) and yields a tight achievability bound at
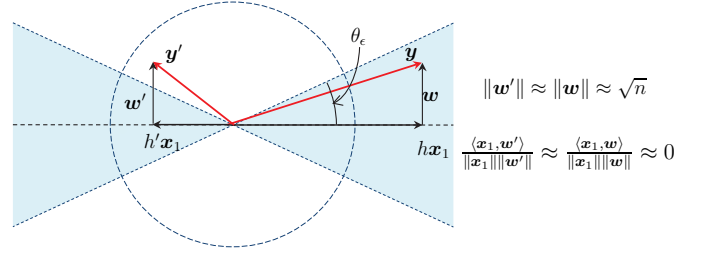

Fig. 1. A geometric illustration of the outage event for large blocklength n. In the example, the fading realization h' triggers an outage event, h does not.

finite blocklength, provided that the threshold θ_ϵ is chosen appropriately.

In the following, we generalize the aforementioned threshold decoder to the MIMO case and present our achievability results.

2) The Achievability Bound: To state our achievability (lower) bound on $R_{\mathrm{tx}}^*(n, \epsilon)$, we will need the following definition, which extends the notion of angle between real vectors to complex subspaces.

Definition 5: Let $\mathcal{A}$ and $\mathcal{B}$ be subspaces in $\mathbb{C}^n$ with $a = \dim(\mathcal{A}) \le \dim(\mathcal{B}) = b$. The *principal angles* $0 \le \theta_1 \le \cdots \le \theta_a \le \pi/2$ between $\mathcal{A}$ and $\mathcal{B}$ are defined recursively by

$$\cos\theta_k \triangleq \max_{\substack{\boldsymbol{a} \in \mathcal{A}, \boldsymbol{b} \in \mathcal{B}: \|\boldsymbol{a}\| = \|\boldsymbol{b}\| = 1, \\ \langle \boldsymbol{a}, \boldsymbol{a}_i\rangle = \langle \boldsymbol{b}, \boldsymbol{b}_i\rangle = 0, i = 1, \ldots, k-1}} |\langle \boldsymbol{a}, \boldsymbol{b}\rangle|,$$
$$k = 1, \ldots, a. \tag{39}$$

Here, $\boldsymbol{a}_k$ and $\boldsymbol{b}_k$, $k = 1, \ldots, a$, are the vectors that achieve the maximum in (39) at the kth recursion. The angle between the subspaces $\mathcal{A}$ and $\mathcal{B}$ is defined by

$$\sin\{\mathcal{A}, \mathcal{B}\} \triangleq \prod_{k=1}^{a} \sin\theta_k. \tag{40}$$

With a slight abuse of notation, for two matrices $\mathsf{A} \in \mathbb{C}^{n \times a}$ and $\mathsf{B} \in \mathbb{C}^{n \times b}$, we abbreviate $\sin\{\mathrm{span}(\mathsf{A}), \mathrm{span}(\mathsf{B})\}$ with $\sin\{\mathsf{A}, \mathsf{B}\}$. When the columns of A and B are orthonormal bases for $\mathrm{span}(\mathsf{A})$ and $\mathrm{span}(\mathsf{B})$, respectively, we have (see, e.g., [22, Sec. I])

$$\sin^2\{\mathsf{A}, \mathsf{B}\} = \det\left(\mathsf{I} - \mathsf{A}^{\mathrm{H}}\mathsf{B}\mathsf{B}^{\mathrm{H}}\mathsf{A}\right) \tag{41}$$

$$= \det\left(\mathsf{I} - \mathsf{B}^{\mathrm{H}}\mathsf{A}\mathsf{A}^{\mathrm{H}}\mathsf{B}\right). \tag{42}$$

Some additional properties of the operator $\sin\{\cdot, \cdot\}$ are listed in Appendix I.

We are now ready to state our achievability bound.

Theorem 1: Let $\Lambda_1 \ge \cdots \ge \Lambda_m$ be the m largest eigenvalues of $\mathbb{H}\mathbb{H}^{\mathrm{H}}$. For every $0 < \epsilon < 1$ and every $0 < \tau < \epsilon$, there exists an $(n, M, \epsilon)_{\mathrm{tx}}$ code for the channel (8) that satisfies

$$\frac{\log M}{n} \ge \frac{1}{n}\log\frac{\tau}{\mathbb{P}\left[\prod_{j=1}^{r} B_j \le \gamma_n\right]}. \tag{43}$$

Here, $B_j \sim \mathrm{Beta}(n-t-j+1, t)$, $j = 1, \ldots, r$, are independent Beta-distributed random variables, and $\gamma_n \in [0, 1]$ is chosen so

that

$$\mathbb{P}\Bigg[\sin^2\Big\{\mathsf{I}_{n,t}, \sqrt{n}\mathsf{I}_{n,t}\mathrm{diag}\Big\{\sqrt{v_1^*\Lambda_1}, \dots,$$
$$\sqrt{v_m^*\Lambda_m}, \underbrace{0, \dots, 0}_{t-m}\Big\} + \mathbb{W}\Big\} \leq \gamma_n\Bigg] \geq 1 - \epsilon + \tau \quad (44)$$

where

$$v_j^* = [\bar{\gamma} - 1/\Lambda_j]^+, \quad j = 1, \dots, r \quad (45)$$

are the water-filling power gains and $\bar{\gamma}$ is defined in (22).

Proof: The achievability bound is based on a decoder that operates as follows: it first computes the sine of the angle between the subspace spanned by the received matrix $\mathbb{Y}$ and the subspace spanned by each codeword; then, it chooses the first codeword for which the squared sine of the angle is below γ_n. To analyze the performance of this decoder, we apply the $\kappa\beta$ bound [9, Th. 25] to a physically degraded channel whose output is $\mathrm{span}(\mathbb{Y})$. See Appendix II for the complete proof. ∎

B. Converse

In this section, we shall assume both CSIR and CSIT. Our converse bound is based on the meta-converse theorem [9, Th. 30]. Since CSI is available at both the transmitter and the receiver, the MIMO channel (8) can be transformed into a set of parallel quasi-static channels. The proof of Theorem 2 below builds on [23, Sec. 4.5], which characterizes the nonasymptotic coding rate of parallel AWGN channels.

Theorem 2: Let $\Lambda_1 \geq \cdots \geq \Lambda_m$ be the m largest eigenvalues of $\mathbb{H}\mathbb{H}^{\mathrm{H}}$, and let $\mathbf{\Lambda} \triangleq [\Lambda_1, \dots, \Lambda_m]^{\mathrm{T}}$. Consider an arbitrary power-allocation function $\mathbf{v} : \mathbb{R}_+^m \mapsto \mathcal{V}_m$, where

$$\mathcal{V}_m \triangleq \Big\{[p_1, \dots, p_m] \in \mathbb{R}_+^m : \sum_{j=1}^m p_j \leq \rho\Big\}. \quad (46)$$

Let

$$L_n^{\mathrm{rt}}(\mathbf{v}, \mathbf{\Lambda}) \triangleq \sum_{i=1}^n \sum_{j=1}^m \Bigg(\log\big(1 + \Lambda_j v_j(\mathbf{\Lambda})\big) + 1$$
$$- \Big|\sqrt{\Lambda_j v_j(\mathbf{\Lambda})}Z_{i,j} - \sqrt{1 + \Lambda_j v_j(\mathbf{\Lambda})}\Big|^2\Bigg) \quad (47)$$

and

$$S_n^{\mathrm{rt}}(\mathbf{v}, \mathbf{\Lambda}) \triangleq \sum_{i=1}^n \sum_{j=1}^m \Bigg(\log\big(1 + \Lambda_j v_j(\mathbf{\Lambda})\big) + 1$$
$$- \frac{\big|\sqrt{\Lambda_j v_j(\mathbf{\Lambda})}Z_{ij} - 1\big|^2}{1 + \Lambda_j v_j(\mathbf{\Lambda})}\Bigg) \quad (48)$$

where $v_j(\cdot)$ is the jth coordinate of $\mathbf{v}(\cdot)$, and Z_{ij}, $i = 1, \dots, n$, $j = 1, \dots, m$, are i.i.d. $\mathcal{CN}(0, 1)$ distributed random variables. For every n and every $0 < \epsilon < 1$, the maximal achievable rate on the channel (8) with CSIRT is upper-bounded by

$$R_{\mathrm{rt}}^*(n, \epsilon) \leq \frac{1}{n}\log\frac{c_{\mathrm{rt}}(n)}{\inf_{\mathbf{v}(\cdot)} \mathbb{P}[L_n^{\mathrm{rt}}(\mathbf{v}, \mathbf{\Lambda}) \geq n\gamma_n(\mathbf{v})]} \quad (49)$$

where

$$c_{\mathrm{rt}}(n) \triangleq \Bigg(\frac{(n-1)^n e^{-(n-1)}}{\Gamma(n)} + \frac{\Gamma(n, n-1)}{\Gamma(n)}\Bigg)^m$$
$$\times \mathbb{E}_{\mathbb{H}}\big[\det(\mathsf{I}_t + \rho\mathbb{H}\mathbb{H}^{\mathrm{H}})\big] \quad (50)$$

and the scalar $\gamma_n(\mathbf{v})$ is the solution of

$$\mathbb{P}[S_n^{\mathrm{rt}}(\mathbf{v}, \mathbf{\Lambda}) \leq n\gamma_n(\mathbf{v})] = \epsilon. \quad (51)$$

The infimum on the RHS of (49) is taken over all power allocation functions $\mathbf{v} : \mathbb{R}_+^m \mapsto \mathcal{V}_m$.

Proof: See Appendix III. ∎

Remark 1: The infimum on the RHS of (49) makes the converse bound in Theorem 2 difficult to evaluate numerically. We can further upper-bound the RHS of (49) by lower-bounding $\mathbb{P}[L_n^{\mathrm{rt}}(\mathbf{v}, \mathbf{\Lambda}) \geq n\gamma_n(\mathbf{v})]$ for each $\mathbf{v}(\cdot)$ using [9, Eq. (102)] and the Chernoff bound. After doing so, the infimum can be computed analytically and the resulting upper bound on $R_{\mathrm{rt}}^*(n, \epsilon)$ allows for numerical evaluations. Unfortunately, this bound is in general loose.

Remark 2: As we shall discuss in Section V-B, the bound (49) can be tightened and evaluated numerically in the SIMO case or when the codewords are isotropic, i.e., are chosen from the set $\mathcal{F}_{\mathrm{iso}}$ in (32). Note that in both scenarios CSIT is not beneficial.

C. Asymptotic Analysis

Following [9, Def. 2], we define the ϵ-dispersion of the channel (8) with CSIT via $R_{\mathrm{tx}}^*(n, \epsilon)$ (resp. $R_{\mathrm{rt}}^*(n, \epsilon)$) as

$$V_\epsilon^l \triangleq \limsup_{n \to \infty} n\left(\frac{C_\epsilon^{\mathrm{tx}} - R_l^*(n, \epsilon)}{Q^{-1}(\epsilon)}\right)^2,$$
$$\epsilon \in (0, 1)\backslash\{1/2\}, \ l = \{\mathrm{tx}, \mathrm{rt}\}. \quad (52)$$

Theorem 3 below characterizes the ϵ-dispersion of the quasi-static fading channel (8) with CSIT.

Theorem 3: Assume that the fading channel $\mathbb{H}$ satisfies the following conditions:

1) the expectation $\mathbb{E}_{\mathbb{H}}\big[\det(\mathsf{I}_t + \rho\mathbb{H}\mathbb{H}^{\mathrm{H}})\big]$ is finite;
2) the joint pdf of the ordered nonzero eigenvalues of $\mathbb{H}^{\mathrm{H}}\mathbb{H}$ exists and is continuously differentiable;
3) C_ϵ^{tx} is a point of growth of the outage probability function (20), i.e.,[6]

$$F_{\mathrm{tx}}'\big(C_\epsilon^{\mathrm{tx}}\big) > 0. \quad (53)$$

Then

$$\{R_{\mathrm{tx}}^*(n, \epsilon), R_{\mathrm{rt}}^*(n, \epsilon)\} = C_\epsilon^{\mathrm{tx}} + \mathcal{O}\left(\frac{\log n}{n}\right). \quad (54)$$

Hence, the ϵ-dispersion is zero for both the CSIRT and the CSIT case:

$$V_\epsilon^{\mathrm{tx}} = V_\epsilon^{\mathrm{rt}} = 0, \quad \epsilon \in (0, 1)\backslash\{1/2\}. \quad (55)$$

Proof: To prove (54), we first establish in Appendix IV the converse result

$$R_{\mathrm{rt}}^*(n, \epsilon) \leq C_\epsilon^{\mathrm{tx}} + \mathcal{O}\left(\frac{\log n}{n}\right) \quad (56)$$

---

[6]Note that this condition implies that C_ϵ^{tx} is a continuous function of ϵ (see Section III).

by analyzing the upper bound (49) in the limit $n \to \infty$. We next prove in Appendix V the achievability result

$$R_{\text{tx}}^*(n, \epsilon) \geq C_\epsilon^{\text{tx}} + \mathcal{O}\left(\frac{\log n}{n}\right) \quad (57)$$

by expanding (43) for $n \to \infty$. The desired result then follows by (14). ∎

Remark 3: As mentioned in Section I, the quasi-static fading channel considered in this paper belongs to the general class of composite or mixed channels, whose ϵ-dispersion is known in some special cases. Specifically, the dispersion of a mixed channel with two states was derived in [24, Th. 7]. This result was extended to channels with finitely many states in [25, Th. 4]. In both cases, the rate of convergence to the ϵ-capacity is $\mathcal{O}(1/\sqrt{n})$ (positive dispersion), as opposed to $\mathcal{O}(\log(n)/n)$ in Theorem 3. Our result shows that moving from finitely many to uncountably many states (as in the quasi-static fading case) yields a drastic change in the value of the channel dispersion. For this reason, our result is not derivable from [24] or [25].

Remark 4: It can be shown that the assumptions on the fading matrix in Theorem 3 are satisfied by most probability distributions used to model MIMO fading channels, such as i.i.d. or correlated Rayleigh, Rician, and Nakagami. However, the (nonfading) AWGN MIMO channel, which can be seen as a quasi-static fading channel with fading distribution equal to a step function, does not meet these assumptions and has, in fact, positive dispersion [23, Th. 78].

While zero dispersion indeed may imply fast convergence to ϵ-capacity, this is not true anymore when the probability distribution of the fading matrix approaches a step function, in which case the higher-order terms in the expansion (54) become more dominant. Consider for example a SISO Rician fading channel with Rician factor K. For $\epsilon < 1/2$, one can refine (54) and show that [26]

$$C_\epsilon - \frac{\log n}{n} - \frac{c_1\sqrt{K} + c_2}{n} + o\left(\frac{1}{n}\right) \leq R_{\text{tx}}^*(n, \epsilon)$$

$$\leq R_{\text{rt}}^*(n, \epsilon) \leq C_\epsilon + \frac{\log n}{n} - \frac{\tilde{c}_1\sqrt{K} + \tilde{c}_2}{n} + o\left(\frac{1}{n}\right) \quad (58)$$

where c_1, c_2, $\tilde{c}_1$ and $\tilde{c}_2$ are finite constants with $c_1 > 0$ and $\tilde{c}_1 > 0$. As we let the Rician factor K become large, the fading distribution converges to a step function and the third term in both the left-most lower bound and the right-most upper bound becomes increasingly large in absolute value.

D. Normal Approximation

We define the *normal approximation* $R_{\text{rt}}^\mathcal{N}(n, \epsilon)$ of $R_{\text{rt}}^*(n, \epsilon)$ as the solution of

$$\epsilon = \mathbb{E}\left[Q\left(\frac{C(\mathbb{H}) - R_{\text{rt}}^\mathcal{N}(n, \epsilon)}{\sqrt{V(\mathbb{H})/n}}\right)\right]. \quad (59)$$

Here,

$$C(\mathsf{H}) = \sum_{j=1}^m \log(1 + v_j^*\lambda_j) \quad (60)$$

is the capacity of the channel (8) when $\mathbb{H} = \mathsf{H}$ (the water-filling power allocation values $\{v_j^*\}$ in (60) are given in (45) and $\{\lambda_j\}$ are the eigenvalues of $\mathsf{H}^\mathsf{H}\mathsf{H}$), and

$$V(\mathsf{H}) = m - \sum_{j=1}^m \frac{1}{(1 + v_j^*\lambda_j)^2} \quad (61)$$

is the dispersion of the channel (8) when $\mathbb{H} = \mathsf{H}$ [23, Th. 78]. Theorem 3 and the expansion

$$R_{\text{rt}}^\mathcal{N}(n, \epsilon) = C_\epsilon^{\text{tx}} + \mathcal{O}\left(\frac{1}{n}\right) \quad (62)$$

(which follows from Lemma 17 in Appendix IV-C and Taylor's theorem) suggest that this approximation is accurate, as confirmed by the numerical results reported in Section VI-A. Note that the same approximation has been concurrently proposed in [27]; see also [24, Def. 2] and [25, Sec. 4] for similar approximations for mixed channels with finitely many states.

V. CSI Not Available at the Transmitter

A. Achievability

In this section, we shall assume that neither the transmitter nor the receiver have *a priori* CSI. Using the decoder described in IV-A, we obtain the following achievability bound.

Theorem 4: Assume that for a given $0 < \epsilon < 1$ there exists a $\mathsf{Q}^* \in \mathcal{U}_t$ such that

$$F_{\text{no}}(C_\epsilon^{\text{no}}) = \inf_{\mathsf{Q} \in \mathcal{U}_t} \mathbb{P}\left[\log\det\left(\mathsf{I}_r + \mathbb{H}^\mathsf{H}\mathsf{Q}\mathbb{H}\right) \leq C_\epsilon^{\text{no}}\right] \quad (63)$$

$$= \mathbb{P}\left[\log\det\left(\mathsf{I}_r + \mathbb{H}^\mathsf{H}\mathsf{Q}^*\mathbb{H}\right) \leq C_\epsilon^{\text{no}}\right] \quad (64)$$

i.e., the infimum in (63) is a minimum. Then, for every $0 < \tau < \epsilon$ there exists an $(n, M, \epsilon)_{\text{no}}$ code for the channel (8) that satisfies

$$\frac{\log M}{n} \geq \frac{1}{n}\log\frac{\tau}{\mathbb{P}\left[\prod_{j=1}^r B_j \leq \gamma_n\right]}. \quad (65)$$

Here, $B_j \sim \text{Beta}(n - t^* - j + 1, t^*)$, $j = 1, \ldots, r$, are independent Beta-distributed random variables, $t^* \triangleq \text{rank}(\mathsf{Q}^*)$, and $\gamma_n \in [0, 1]$ is chosen so that

$$\mathbb{P}\left[\sin^2\{\mathsf{I}_{n,t^*}, \sqrt{n}\mathsf{I}_{n,t^*}\mathsf{U}\mathbb{H} + \mathbb{W}\} \leq \gamma_n\right] \geq 1 - \epsilon + \tau \quad (66)$$

with $\mathsf{U} \in \mathbb{C}^{t^* \times t}$ satisfying $\mathsf{U}^\mathsf{H}\mathsf{U} = \mathsf{Q}^*$.

Proof: The proof is identical to the proof of Theorem 1, with the only difference that the precoding matrix $\mathsf{P}(\mathbb{H})$ (defined in (108)) is replaced by $\sqrt{n}\mathsf{I}_{n,t^*}\mathsf{U}$. ∎

The assumption in (64) that the ϵ-capacity-achieving input covariance matrix of the channel (8) exists is mild. A sufficient condition for the existence of Q^* is given in the following proposition.

Proposition 5: Assume that $\mathbb{E}\left[\|\mathbb{H}\|_\mathsf{F}^2\right] < \infty$ and that the distribution of $\mathbb{H}$ is absolutely continuous with respect to the Lebesgue measure on $\mathbb{C}^{t \times r}$. Then, for every $R \in \mathbb{R}_+$, the infimum in (26) is a minimum.

Proof: See Appendix VI. ∎

For the SIMO case, the RHS of (43) and the RHS of (65) coincide, i.e.,

$$\left\{R_{\text{tx}}(n, \epsilon), R_{\text{no}}(n, \epsilon)\right\} \geq \frac{1}{n}\log\frac{\tau}{\mathbb{P}[B \leq \gamma_n]} \quad (67)$$

where $B \sim \text{Beta}(n-r, r)$, and $\gamma_n \in [0, 1]$ is chosen so that

$$\mathbb{P}[\sin^2\{\boldsymbol{e}_1, \sqrt{n\rho}\boldsymbol{e}_1\boldsymbol{H}^{\mathrm{T}} + \mathbb{W}\} \leq \gamma_n] \geq 1 - \epsilon + \tau. \quad (68)$$

Here, $\boldsymbol{e}_1$ stands for the first column of the identity matrix I_n. The achievability bound (67) follows from (43) and (65) by noting that the random variable B on the RHS of (67) has the same distribution as $\prod_{i=1}^{r} B_i$, where $B_i \sim \text{Beta}(n-i, 1)$, $i = 1, \ldots, r$.

B. Converse

For the converse, we shall assume CSIR but not CSIT. The counterpart of Theorem 2 is the following result.

Theorem 6: Let $\mathcal{U}_t^{\mathrm{e}}$ be as in (28). For an arbitrary $\mathsf{Q} \in \mathcal{U}_t^{\mathrm{e}}$, let $\Lambda_1 \geq \cdots \geq \Lambda_m$ be the ordered eigenvalues of $\mathbb{H}^{\mathrm{H}}\mathsf{Q}\mathbb{H}$. Let

$$L_n^{\mathrm{rx}}(\mathsf{Q}) \triangleq \sum_{i=1}^{n} \sum_{j=1}^{m} \left(\log(1+\Lambda_j) + 1 - \left| \sqrt{\Lambda_j} Z_{ij} - \sqrt{1+\Lambda_j} \right|^2 \right)$$
$$(69)$$

and

$$S_n^{\mathrm{rx}}(\mathsf{Q}) \triangleq \sum_{i=1}^{n} \sum_{j=1}^{m} \left(\log(1+\Lambda_j) + 1 - \frac{\left| \sqrt{\Lambda_j} Z_{ij} - 1 \right|^2}{1 + \Lambda_j} \right) (70)$$

where Z_{ij}, $i = 1, \ldots, n$, $j = 1, \ldots, m$, are i.i.d. $\mathcal{CN}(0, 1)$ distributed. Then, for every $n \geq r$ and every $0 < \epsilon < 1$, the maximal achievable rate on the quasi-static MIMO fading channel (8) with CSIR is upper-bounded by

$$R_{\mathrm{rx}}^*(n-1, \epsilon) \leq \frac{1}{n-1} \log \frac{c_{\mathrm{rx}}(n)}{\inf_{\mathsf{Q} \in \mathcal{U}_t^{\mathrm{e}}} \mathbb{P}[L_n^{\mathrm{rx}}(\mathsf{Q}) \geq n\gamma_n(\mathsf{Q})]} . \quad (71)$$

Here,

$$c_{\mathrm{rx}}(n) \triangleq \frac{\pi^{r(r-1)}}{\Gamma_r(n)\Gamma_r(r)} \mathbb{E}\left[\left(1 + \rho \|\mathbb{H}\|_{\mathsf{F}}^2 \right)^{\lfloor (r+1)^2/4 \rfloor} \right]$$
$$\times \prod_{i=1}^{r} \left[(n+r-2i)^{n+r-2i+1} e^{-(n+r-2i)} \right.$$
$$\left. + \Gamma(n+r-2i+1, n+r-2i) \right] \quad (72)$$

with $\Gamma_{(\cdot)}(\cdot)$ denoting the *complex* multivariate Gamma function [28, Eq. (83)], and $\gamma_n(\mathsf{Q})$ is the solution of

$$\mathbb{P}[S_n^{\mathrm{rx}}(\mathsf{Q}) \leq n\gamma_n(\mathsf{Q})] = \epsilon. \quad (73)$$

Proof: See Appendix VII. ∎

The infimum in (71) makes the upper bound more difficult to evaluate numerically and to analyze asymptotically up to $\mathcal{O}(\log(n)/n)$ terms than the upper bound (49) that we established for the CSIT case. In fact, even the simpler problem of finding the matrix Q that minimizes $\lim_{n\to\infty} \mathbb{P}[L_n^{\mathrm{rx}}(\mathsf{Q}) \geq n\gamma_n]$ is open. Next, we consider two special cases for which the bound (71) can be tightened and evaluated numerically: the SIMO case and the case where all codewords are chosen from the set $\mathcal{F}_{\mathrm{iso}}$.

1) SIMO case: For the SIMO case, CSIT is not beneficial [26] and the bounds (49) and (71) can be tightened as follows.

Theorem 7: Let

$$L_n \triangleq n \log(1 + \rho G) + \sum_{i=1}^{n} \left(1 - \left| \sqrt{\rho G} Z_i - \sqrt{1+\rho G} \right|^2 \right)$$
$$(74)$$

and

$$S_n \triangleq n \log(1 + \rho G) + \sum_{i=1}^{n} \left(1 - \frac{\left| \sqrt{\rho G} Z_i - 1 \right|^2}{1 + \rho G} \right) \quad (75)$$

with $G \triangleq \|\boldsymbol{H}\|^2$ and Z_i, $i = 1, \ldots, n$, i.i.d. $\mathcal{CN}(0, 1)$ distributed. For every n and every $0 < \epsilon < 1$, the maximal achievable rate on the quasi-static fading channel (8) with one transmit antenna and with CSIR (with or without CSIT) is upper-bounded by

$$R_{\mathrm{rx}}^*(n-1, \epsilon) \leq R_{\mathrm{rt}}^*(n-1, \epsilon) \leq \frac{1}{n-1} \log \frac{1}{\mathbb{P}[L_n \geq n\gamma_n]} (76)$$

where γ_n is the solution of

$$\mathbb{P}[S_n \leq n\gamma_n] = \epsilon. \quad (77)$$

Proof: See [26, Th. 1]. The main difference between the proof of Theorem 7 and the proof of Theorem 2 and Theorem 6 is that the simple bound $\epsilon' \geq 1 - 1/M$ on the maximal error probability of the auxiliary channel in the meta-converse theorem [9, Th. 30] suffices to establish the desired result. The more sophisticated bounds reported in Lemma 14 (Appendix III) and Lemma 19 (Appendix VII) are not needed. ∎

2) Converse for $(n, M, \epsilon)_{\mathrm{iso}}$ codes: In Theorem 8 below, we establish a converse bound on the maximal achievable rate of $(n, M, \epsilon)_{\mathrm{iso}}$ codes introduced in Section III. As such codes consist of isotropic codewords chosen from the set $\mathcal{F}_{\mathrm{iso}}$ in (32), CSIT is not beneficial also in this scenario.

Theorem 8: Let $L_n^{\mathrm{rx}}(\cdot)$ and $S_n^{\mathrm{rx}}(\cdot)$ be as in (69) and (70), respectively. Then, for every n and every $0 < \epsilon < 1$, the maximal achievable rate $R_{\mathrm{rx,iso}}^*(n, \epsilon)$ of $(n, M, \epsilon)_{\mathrm{iso}}$ codes over the quasi-static MIMO fading channel (8) with CSIR is upper-bounded by

$$R_{\mathrm{rx,iso}}^*(n, \epsilon) \leq R_{\mathrm{rt,iso}}^*(n, \epsilon) \leq \frac{1}{n} \log \frac{1}{\mathbb{P}[L_n^{\mathrm{rx}}((\rho/t)\mathsf{I}_t) \geq n\gamma_n]}$$
$$(78)$$

where γ_n is the solution of

$$\mathbb{P}[S_n^{\mathrm{rx}}((\rho/t)\mathsf{I}_t) \leq n\gamma_n] = \epsilon. \quad (79)$$

Proof: The proof follows closely the proof of Theorem 6. As in the SIMO case, the main difference is that the simple bound $\epsilon' \geq 1 - 1/M$ on the maximal error probability of the auxiliary channel in the meta-converse theorem [9, Th. 30] suffices to establish (79). ∎

C. Asymptotic Analysis

To state our dispersion result, we will need the following definition of the gradient ∇g of a differentiable function $g : \mathbb{C}^{t\times r} \mapsto \mathbb{R}$. Let $\mathsf{L} \in \mathbb{C}^{t\times r}$, then we shall write $\nabla g(\mathsf{H}) = \mathsf{L}$ if

$$\frac{d}{dt}g(\mathsf{H} + t\mathsf{A})\Big|_{t=0} = \text{Re}\{\text{tr}(\mathsf{A}^{\mathrm{H}}\mathsf{L})\}, \quad \forall \mathsf{A} \in \mathbb{C}^{t\times r}. \quad (80)$$

Theorem 9 below establishes the zero-dispersion result for the case of no CSIT. Because of the analytical intractability of the minimization in the converse bound (71), Theorem 9 requires more stringent conditions on the fading distribution compared to the CSIT case (cf., Theorem 3), and its proof is more involved.

Theorem 9: Let $f_{\mathbb{H}}$ be the pdf of the fading matrix $\mathbb{H}$. Assume that $\mathbb{H}$ satisfies the following conditions:

1) $f_{\mathbb{H}}$ is a smooth function, i.e., it has derivatives of all orders.
2) There exists a positive constant a such that

$$f_{\mathbb{H}}(\mathsf{H}) \leq a \, \|\mathsf{H}\|_{\mathsf{F}}^{-2tr - \lfloor (r+1)^2/2 \rfloor - 1} \qquad (81)$$

$$\|\nabla f_{\mathbb{H}}(\mathsf{H})\|_{\mathsf{F}} \leq a \, \|\mathsf{H}\|_{\mathsf{F}}^{-2tr-5} . \qquad (82)$$

3) The function $F_{\mathrm{no}}(\cdot)$ satisfies

$$\liminf_{\delta \to 0} \frac{F_{\mathrm{no}}(C_{\epsilon}^{\mathrm{no}} + \delta) - F_{\mathrm{no}}(C_{\epsilon}^{\mathrm{no}})}{\delta} > 0. \qquad (83)$$

Then,

$$\{R_{\mathrm{no}}^*(n, \epsilon), R_{\mathrm{rx}}^*(n, \epsilon)\} = C_{\epsilon}^{\mathrm{no}} + \mathcal{O}\left(\frac{\log n}{n}\right). \qquad (84)$$

Hence, the ϵ-dispersion is zero for both the CSIR and the no-CSI case:

$$V_{\epsilon}^{\mathrm{no}} = V_{\epsilon}^{\mathrm{rx}} = 0, \qquad \epsilon \in (0, 1) \backslash \{1/2\}. \qquad (85)$$

Proof: See Appendices VIII and IX. ∎

Remark 5: It can be shown that Conditions 1–3 in Theorem 9 are satisfied by the probability distributions commonly used to model MIMO fading channels, such as Rayleigh, Rician, and Nakagami. Condition 2 requires simply that $f_{\mathbb{H}}$ has a polynomially decaying tail. Condition 3 plays the same role as (53) in the CSIT case. The exact counterpart of (53) for the no-CSIT case would be

$$F_{\mathrm{no}}'(C_{\epsilon}^{\mathrm{no}}) > 0. \qquad (86)$$

However, different from (53), the inequality (86) does not necessarily hold for the commonly used fading distributions. Indeed, consider a MISO i.i.d. Rayleigh-fading channel. As proven in [20], the ϵ-capacity-achieving covariance matrix for this case is given by (29). The resulting outage probability function $F_{\mathrm{no}}(\cdot)$ may not be differentiable at the rates R for which the infimum in (27) is achieved by two input covariance matrices with different number of nonzero entries t^* on the main diagonal.

Next, we briefly sketch how to prove that Condition 3 holds for Rayleigh, Rician, and Nakagami distributions. Let

$$F_{\mathsf{Q}}(R) \triangleq \mathbb{P}[\log\det(\mathsf{I}_r + \mathbb{H}^{\mathsf{H}} \mathsf{Q} \mathbb{H}) < R]. \qquad (87)$$

Let $\mathcal{Q}_{\epsilon}$ be the set of all ϵ-capacity-achieving covariance matrices, i.e.,

$$\mathcal{Q}_{\epsilon} \triangleq \{\mathsf{Q} \in \mathcal{U}_t^e : F_{\mathsf{Q}}(C_{\epsilon}^{\mathrm{no}}) = F_{\mathrm{no}}(C_{\epsilon}^{\mathrm{no}})\}. \qquad (88)$$

By Proposition 5, the set $\mathcal{Q}_{\epsilon}$ is non-empty for the considered fading distributions. It follows from algebraic manipulations that

$$\liminf_{\delta \to 0} \frac{F_{\mathrm{no}}(C_{\epsilon}^{\mathrm{no}} + \delta) - F_{\mathrm{no}}(C_{\epsilon}^{\mathrm{no}})}{\delta} = \inf_{\mathsf{Q} \in \mathcal{Q}_{\epsilon}} F_{\mathsf{Q}}'(C_{\epsilon}^{\mathrm{no}}). \qquad (89)$$

To show that the RHS of (89) is positive, one needs to perform two steps. First, one shows that the set $\mathcal{Q}_{\epsilon}$ is compact with respect to the metric $d(\mathsf{A}, \mathsf{B}) = \|\mathsf{A} - \mathsf{B}\|_{\mathsf{F}}$ and that under

Conditions 1 and 2 of Theorem 9, the function $\mathsf{Q} \mapsto F_{\mathsf{Q}}'(C_{\epsilon}^{\mathrm{no}})$ is continuous with respect to the same metric. By the extreme value theorem [29, p. 34], these two properties imply that the infimum on the RHS of (89) is a minimum. Then, one shows that for Rayleigh, Rician, and Nakagami distributions

$$F_{\mathsf{Q}}'(C_{\epsilon}^{\mathrm{no}}) > 0, \qquad \forall \mathsf{Q} \in \mathcal{Q}_{\epsilon}. \qquad (90)$$

One way to prove (90) is to write $F_{\mathsf{Q}}'(C_{\epsilon}^{\mathrm{no}})$ in integral form using Lemma 22 in Appendix VIII-A1 and to show that the resulting integral is positive.

For the SIMO case, the conditions on the fading distribution can be relaxed and the following result holds.

Theorem 10: Assume that the pdf of $\|\boldsymbol{H}\|^2$ is continuously differentiable and that the ϵ-capacity C_{ϵ} is a point of growth for the outage probability function

$$F(R) = \mathbb{P}[\log(1 + \|\boldsymbol{H}\|^2 \rho) < R] \qquad (91)$$

i.e., $F'(C_{\epsilon}) > 0$. Then,

$$\{R_{\mathrm{no}}^*(n, \epsilon), R_{\mathrm{rx}}^*(n, \epsilon)\} = C_{\epsilon} + \mathcal{O}\left(\frac{\log n}{n}\right). \qquad (92)$$

Proof: In the SIMO case, CSIT is not beneficial [26, Th. 5]. Hence, the result follows directly from Theorem 3 and Proposition 23 in Appendix IX. ∎

Similarly, for the case of codes consisting of isotropic codewords, milder conditions on the fading distribution are sufficient to establish zero dispersion, as illustrated in the following theorem.

Theorem 11: Assume that the joint pdf of the nonzero eigenvalues of $\mathbb{H}^{\mathsf{H}}\mathbb{H}$ is continuously differentiable and that

$$F_{\mathrm{iso}}'(C_{\epsilon}^{\mathrm{iso}}) > 0 \qquad (93)$$

where F_{iso} is the outage probability function given in (31). Then, we have

$$\{R_{\mathrm{no,iso}}^*(n, \epsilon), R_{\mathrm{rx,iso}}^*(n, \epsilon)\} = C_{\epsilon}^{\mathrm{iso}} + \mathcal{O}\left(\frac{\log n}{n}\right). \qquad (94)$$

Proof: See Appendix X. ∎

D. Normal Approximation

For the general no-CSIT MIMO case, the unavailability of a closed-form expression for the ϵ-capacity $C_{\epsilon}^{\mathrm{no}}$ in (25) prevents us from obtaining a normal approximation for the maximum coding rate at finite blocklength. However, such an approximation can be obtained for the SIMO case and for the case of isotropic codewords. In both cases, CSIT is not beneficial and the outage capacity can be characterized in closed form.

For the SIMO case, the normal approximation follows directly from (59)–(61) by setting $m = 1$, $v_1^* = \rho$ and noting that $\lambda_1 = \|\boldsymbol{h}\|^2$.

For $(n, M, \epsilon)_{\mathrm{iso}}$ codes, the normal approximation $R_{\mathrm{rx,iso}}^{\mathcal{N}}(n, \epsilon)$ to the maximal achievable rate $R_{\mathrm{rx,iso}}^*(n, \epsilon)$ is obtained as the solution of

$$\epsilon = \mathbb{E}\left[Q\left(\frac{C_{\mathrm{iso}}(\mathbb{H}) - R_{\mathrm{rx,iso}}^{\mathcal{N}}(n, \epsilon)}{\sqrt{V_{\mathrm{iso}}(\mathbb{H})/n}}\right)\right]. \qquad (95)$$
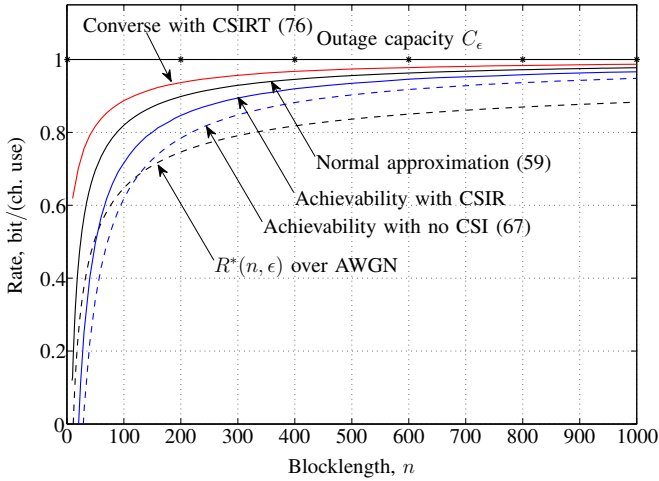
Fig. 2. Achievability and converse bounds for a quasi-static SIMO Rician-fading channel with K-factor equal to 20 dB, two receive antennas, SNR $= -1.55$ dB, and $\epsilon = 10^{-3}$. Note that in the SIMO case $C_\epsilon^{\text{tx}} = C_\epsilon^{\text{no}} = C_\epsilon$.
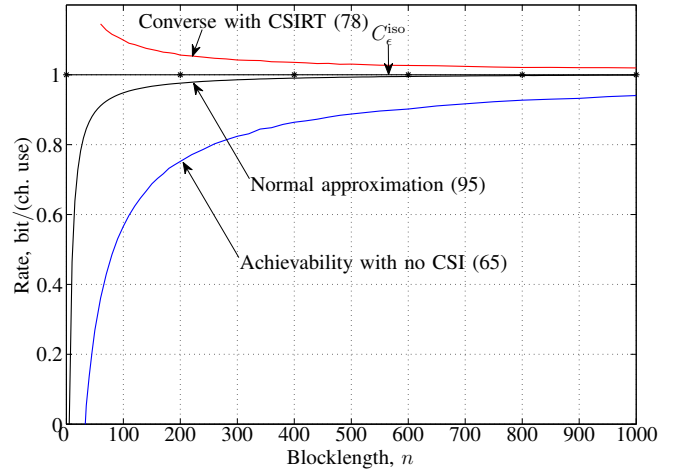

Fig. 3. Achievability and converse bounds for $(n, M, \epsilon)_{\text{iso}}$ codes over a quasi-static MIMO Rayleigh-fading channel with two transmit and three receive antennas, SNR $= 2.12$ dB, and $\epsilon = 10^{-3}$.

Here,

$$C_{\text{iso}}(\mathsf{H}) = \sum_{j=1}^{m} \log(1 + \rho \lambda_j / t) \tag{96}$$

and

$$V_{\text{iso}}(\mathsf{H}) = m - \sum_{j=1}^{m} \frac{1}{(1 + \rho \lambda_j / t)^2} \tag{97}$$

where $\{\lambda_j\}$ are the eigenvalues of $\mathsf{H}^{\mathsf{H}}\mathsf{H}$. A comparison between $R_{\text{rx,iso}}^{\mathcal{N}}(n, \epsilon)$ and the bounds (65) and (78) is provided in the next section.

VI. NUMERICAL RESULTS

A. Numerical Results

In this section, we compute the bounds reported in Sections IV and V. Fig. 2 compares $R_{\text{rt}}^{\mathcal{N}}(n, \epsilon)$ with the achievability bound (67) and the converse bound (76) for a quasi-static SIMO fading channel with two receive antennas. The channels between the transmit antenna and each of the two receive antennas are Rician-distributed with K-factor equal to 20 dB. The two channels are assumed to be independent. We set $\epsilon = 10^{-3}$ and choose $\rho = -1.55$ dB so that $C_\epsilon = 1$ bit/(ch. use). We also plot a lower bound on $R_{\text{rt}}^*(n, \epsilon)$ obtained by using the $\kappa\beta$ bound [9, Th. 25] and assuming CSIR.[7] For reference, Fig. 2 shows also the approximation (2) for $R^*(n, \epsilon)$ corresponding to an AWGN channel with $C = 1$ bit/(ch. use), replacing the term $\mathcal{O}(\log(n)/n)$ in (2) with $\log(n)/(2n)$ [9, Eq. (296)] [30].[8] The blocklength required to achieve 90% of the ϵ-capacity of the quasi-static fading channel is in the range $[120, 320]$ for the

CSIRT case and in the range $[120, 480]$ for the no-CSI case. For the AWGN channel, this number is approximately 1420. Hence, for the parameters chosen in Fig. 2, the prediction (based on zero dispersion) of fast convergence to capacity is validated. The gap between the normal approximation $R_{\text{rt}}^{\mathcal{N}}(n, \epsilon)$ defined implicitly in (59) and both the achievability (CSIR) and the converse bounds is less than 0.02 bit/(ch. use) for blocklengths larger than 400.

Note that although the AWGN curve in Fig. 2 lies below the achievability bound for the quasi-static fading channel, this does not mean that "fading helps". In Fig. 2, we chose the SNRs so that both channels have the same ϵ-capacity. This results in the received power for the quasi-static case being 1.45 dB larger than that for the AWGN case.

In Fig. 3, we compare the normal approximation $R_{\text{rx,iso}}^{\mathcal{N}}(n, \epsilon)$ defined (implicitly) in (95) with the achievability bound (65) and the converse bound (78) on the maximal achievable rate with $(n, M, \epsilon)_{\text{iso}}$ codes over a quasi-static MIMO fading channel with $t = 2$ transmit and $r = 3$ receive antennas. The channel between each transmit-receive antenna pair is Rayleigh-distributed, and the channels between different transmit-receive antenna pairs are assumed to be independent. We set $\epsilon = 10^{-3}$ and choose $\rho = 2.12$ dB so that $C_\epsilon^{\text{iso}} = 1$ bit/(ch. use). For this scenario, the blocklength required to achieve 90% of C_ϵ^{iso} is less than 500, which again demonstrates fast convergence to C_ϵ^{iso}.

B. Comparison with coding schemes in LTE-Advanced

The bounds reported in Sections IV and V can be used to benchmark the coding schemes adopted in current standards. In Fig. 4, we compare the performance of the coding schemes used in LTE-Advanced [31, Sec. 5.1.3.2] against the achievability and converse bounds for the same scenario as in Fig. 2. Specifically, Fig. 4 illustrates the performance of the family of turbo codes chosen in LTE-Advanced for the case of QPSK modulation. The decoder employs a max-log-MAP decoding algorithm [32] with 10 iterations. We further assume that the decoder has perfect CSI. For the AWGN case, it was observed in [9, Fig. 12] that about

---

[7] Specifically, we took $\mathcal{F} = \{\boldsymbol{x} \in \mathbb{C}^n : \|\boldsymbol{x}\|^2 = n\rho\}$, and $Q_{\mathbb{Y}\boldsymbol{H}} = P_{\boldsymbol{H}} \prod_{j=1}^{n} Q_{\boldsymbol{Y}_j \mid \boldsymbol{H}}$ where $Q_{\boldsymbol{Y}_j \mid \boldsymbol{H} = \boldsymbol{h}} = \mathcal{CN}(\boldsymbol{0}, \mathsf{I}_r + \rho \boldsymbol{h} \boldsymbol{h}^{\mathsf{H}})$.

[8] The approximation reported in [9, Eq. (296)], [30] holds for a real AWGN channel. Since a complex AWGN channel with blocklength n can be identified as a real AWGN channel with the same SNR and blocklength $2n$, the approximation [9, Eq. (296)], [30] with $C = \log(1 + \rho)$ and $V = \frac{\rho^2 + 2\rho}{(1+\rho)^2}$ is accurate for the complex case.

Fig. 4. Comparison between achievability and converse bounds and the rate achievable with the coding schemes in LTE-Advanced. We consider a quasi-static SIMO Rician-fading channel with K-factor equal to 20 dB, two receive antennas, SNR $= -1.55$ dB, $\epsilon = 10^{-3}$, and CSIR. The star-shaped markers indicate the rates achievable by the turbo codes in LTE-Advanced (QPSK modulation and 10 iterations of a max-log-MAP decoder [32]).

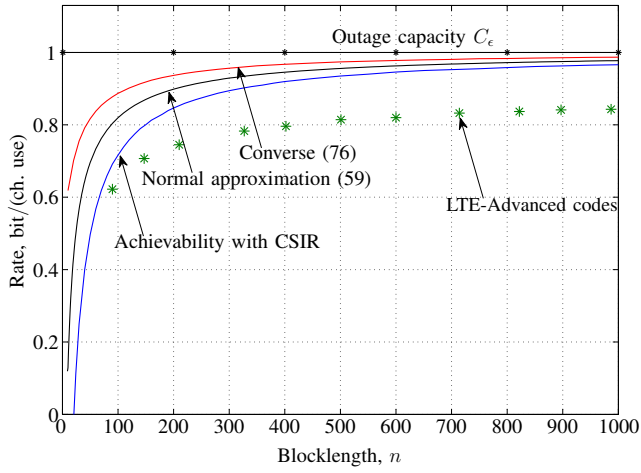

Fig. 5. Comparison between achievability and converse bounds and rate achievable with the coding schemes in LTE-Advanced. We consider a quasi-static SIMO Rayleigh-fading channel with two receive antennas, SNR $= 2.74$ dB, $\epsilon = 0.1$, and CSIR. The star-shaped markers indicate the rates achievable by the turbo codes in LTE-Advanced (QPSK modulation and 10 iterations of a max-log-MAP decoder [32]).
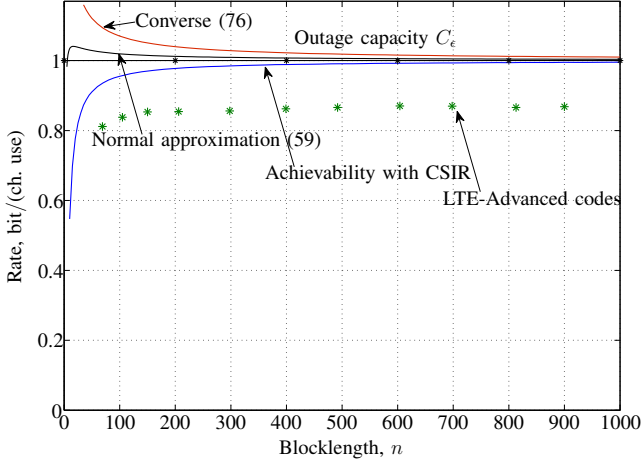
half of the gap between the rate achieved by the best available channel codes[9] and capacity is due to the $1/\sqrt{n}$ penalty in (2); the other half is due to the suboptimality of the codes. From Fig. 4, we conclude that for quasi-static fading channels the finite-blocklength penalty is significantly reduced because of the zero-dispersion effect. However, the penalty due to the code suboptimality remains. In fact, we see that the gap between the rate achieved by the LTE-Advanced turbo codes and the normal approximation $R_{\mathrm{rt}}^{\mathcal{N}}(n, \epsilon)$ is approximately constant up to a blocklength of 1000.

---

[9]The codes used in [9, Fig. 12] are a certain family of multiedge low-density parity-check (LDPC) codes.

LTE-Advanced uses hybrid automatic repeat request (HARQ) to compensate for packets loss due to outage events. When HARQ is used, the block error rate that maximizes the average throughput is about 10^{-1} [33, p. 218]. The performance of LTE-Advanced codes for $\epsilon = 10^{-1}$ is analyzed in Fig. 5. We set $\rho = 2.74$ dB and consider Rayleigh fading (the other parameters are as in Fig. 4). Again, we observe that there is a constant gap between the rate achieved by LTE-Advanced turbo codes and $R_{\mathrm{rt}}^{\mathcal{N}}(n, \epsilon)$.

VII. CONCLUSION

In this paper, we established achievability and converse bounds on the maximal achievable rate $R^*(n, \epsilon)$ for a given blocklength n and error probability ϵ over quasi-static MIMO fading channels. We proved that (under some mild conditions on the fading distribution) the channel dispersion is zero for all four cases of CSI availability. The bounds are easy to evaluate when CSIT is available, when the number of transmit antennas is one, or when the code has isotropic codewords. In all these cases the outage-capacity-achieving distribution is known.

The numerical results reported in Section VI-A demonstrate that, in some scenarios, zero dispersion implies fast convergence to C_ϵ as the blocklength increases. This suggests that the outage capacity is a valid performance metric for communication systems with stringent latency constraints operating over quasi-static fading channels. We developed an easy-to-evaluate approximation of $R^*(n, \epsilon)$ and demonstrated its accuracy by comparison to our achievability and converse bounds. Finally, we used our bounds to benchmark the performance of the coding schemes adopted in the LTE-Advanced standard. Specifically, we showed that for a blocklength between 500 and 1000 LTE-Advanced codes achieve about 85% of the maximal coding rate.

APPENDIX I
AUXILIARY LEMMAS CONCERNING THE PRODUCT OF SINES OF PRINCIPAL ANGLES

In this appendix, we state two properties of the product of principal sines defined in (40), which will be used in the proof of Theorem 3 and of Proposition 23. The first property, which is referred to in [34] as "equalized Hadamard inequality", is stated in Lemma 12 below.

Lemma 12: Let $\mathsf{A} = [\mathsf{A}_1, \mathsf{A}_2] \in \mathbb{C}^{n \times (a_1 + a_2)}$, where $\mathsf{A}_1 \in \mathbb{C}^{n \times a_1}$ and $\mathsf{A}_2 \in \mathbb{C}^{n \times a_2}$. If $\mathrm{rank}(\mathsf{A}_1) = a_1$ and $\mathrm{rank}(\mathsf{A}_2) = a_2$, then

$$\det(\mathsf{A}^{\mathsf{H}}\mathsf{A}) = \det(\mathsf{A}_1^{\mathsf{H}}\mathsf{A}_1) \det(\mathsf{A}_2^{\mathsf{H}}\mathsf{A}_2) \sin^2\{\mathsf{A}_1, \mathsf{A}_2\}. \quad (98)$$

Proof: The proof follows by extending [35, Th. 3.3] to the complex case. ∎

The second property provides an upper bound on $\sin\{\mathcal{A}, \mathcal{B}\}$ that depends on the angles between the basis vectors of the two subspaces.

Lemma 13: Let $\mathcal{A}$ and $\mathcal{B}$ be subspaces of $\mathbb{C}^n$ with $\dim(\mathcal{A}) = a$ and $\dim(\mathcal{B}) = b$. Let $\{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_a\}$ be an orthonormal basis for $\mathcal{A}$, and let $\{\boldsymbol{b}_1, \ldots, \boldsymbol{b}_b\}$ be an arbitrary basis (not necessarily orthonormal) for $\mathcal{B}$. Then

$$\sin\{\mathcal{A}, \mathcal{B}\} \leq \prod_{j=1}^{\min\{a,b\}} \sin\{\boldsymbol{a}_j, \boldsymbol{b}_j\}. \quad (99)$$

Proof: To keep notation simple, we define the following function, which maps a complex matrix X of arbitrary size to its volume:

$$\mathrm{vol}(\mathsf{X}) \triangleq \sqrt{\det(\mathsf{X}^{\mathsf{H}}\mathsf{X})}. \tag{100}$$

Let $\mathsf{A} = [\boldsymbol{a}_1, \ldots, \boldsymbol{a}_a] \in \mathbb{C}^{n \times a}$ and $\mathsf{B} = [\boldsymbol{b}_1, \ldots, \boldsymbol{b}_b] \in \mathbb{C}^{n \times b}$. If the vectors $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_a, \boldsymbol{b}_1, \ldots, \boldsymbol{b}_b$ are linearly dependent, then the LHS of (99) vanishes, in which case (99) holds trivially. In the following, we therefore assume that the vectors $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_a, \boldsymbol{b}_1, \ldots, \boldsymbol{b}_b$ form a linearly independent set. Below, we prove Lemma 13 for the case $a \leq b$. The proof for the case $a > b$ follows from similar steps.

Using Lemma 12, we get the following chain of (in)equalities:

$$\sin\{\mathsf{A}, \mathsf{B}\}$$
$$= \frac{\mathrm{vol}([\mathsf{A}, \mathsf{B}])}{\mathrm{vol}(\mathsf{A})\mathrm{vol}(\mathsf{B})} \tag{101}$$
$$= \frac{\mathrm{vol}([\mathsf{A}, \mathsf{B}])}{\mathrm{vol}(\mathsf{B})} \tag{102}$$
$$= \frac{1}{\mathrm{vol}(\mathsf{B})} \underbrace{\|\boldsymbol{a}_1\|}_{=1} \mathrm{vol}([\boldsymbol{a}_2, \ldots, \boldsymbol{a}_a, \mathsf{B}])$$
$$\cdot \sin\{\boldsymbol{a}_1, [\boldsymbol{a}_2, \ldots, \boldsymbol{a}_a, \mathsf{B}]\} \tag{103}$$
$$\vdots$$
$$= \frac{1}{\mathrm{vol}(\mathsf{B})} \left(\prod_{i=1}^{a} \sin\{\boldsymbol{a}_i, [\boldsymbol{a}_{i+1}, \ldots, \boldsymbol{a}_a, \mathsf{B}]\} \right) \mathrm{vol}(\mathsf{B}) \tag{104}$$
$$\leq \prod_{i=1}^{a} \sin\{\boldsymbol{a}_i, \boldsymbol{b}_i\}. \tag{105}$$

Here, (102) holds because the columns of A are orthonormal and, hence, $\det(\mathsf{A}^{\mathsf{H}}\mathsf{A}) = 1$; (103) and (104) follow from Lemma 12; (105) follows because

$$\sin\{\boldsymbol{a}_i, [\boldsymbol{a}_{i+1}, \ldots, \boldsymbol{a}_a, \mathsf{B}]\} \leq \sin\{\boldsymbol{a}_i, \boldsymbol{b}_i\}. \tag{106}$$

∎

APPENDIX II
PROOF OF THEOREM 1 (CSIT ACHIEVABILITY BOUND)

Given $\mathbb{H} = \mathsf{H}$, we perform a singular value decomposition (SVD) of H to obtain

$$\mathsf{H} = \mathsf{L}\Sigma\mathsf{V}^{\mathsf{H}} \tag{107}$$

where $\mathsf{L} \in \mathbb{C}^{t \times t}$ and $\mathsf{V} \in \mathbb{C}^{r \times r}$ are unitary matrices, and $\Sigma \in \mathbb{C}^{t \times r}$ is a (truncated) diagonal matrix of dimension $t \times r$, whose diagonal elements $\sqrt{\lambda_1}, \ldots, \sqrt{\lambda_m}$, are the ordered singular values of H. It will be convenient to define the following $t \times t$ precoding matrix for each H:

$$\mathsf{P}(\mathsf{H}) \triangleq \mathrm{diag}\{\sqrt{nv_1^*}, \ldots, \sqrt{nv_m^*}, \underbrace{0, \ldots, 0}_{t-m}\}\mathsf{L}^{\mathsf{H}}. \tag{108}$$

We consider a code whose codewords $\mathsf{X}_j(\mathbb{H})$, $j = 1, \ldots, M$, have the following structure

$$\mathsf{X}_j(\mathbb{H}) = \Phi_j \mathsf{P}(\mathbb{H}), \quad \Phi_j \in \mathcal{S}_{n,t} \tag{109}$$

where $\mathcal{S}_{n,t} \triangleq \{\mathsf{A} \in \mathbb{C}^{n \times t} : \mathsf{A}^{\mathsf{H}}\mathsf{A} = \mathsf{I}_t\}$ denotes the set of all $n \times t$ unitary matrices, (i.e., the complex *Stiefel manifold*). As

$\{\Phi_j\}$ are unitary, the codewords satisfy the power constraint (12). Motivated by the geometric considerations reported in Section IV-A1, we consider for a given input $\mathsf{X}(\mathbb{H}) = \Phi\mathsf{P}(\mathbb{H})$ a physically degraded version of the channel (8), whose output is given by

$$\Omega_{\mathbb{Y}} = \mathrm{span}(\Phi\mathsf{P}(\mathbb{H})\mathbb{H} + \mathbb{W}). \tag{110}$$

Note that the subspace $\Omega_{\mathbb{Y}}$ belongs with probability one to the *Grassmannian manifold* $\mathcal{G}_{n,r}$, i.e., the set of all r dimensional subspaces in $\mathbb{C}^n$. Because (110) is a physically degraded version of (8), the rate achievable on (110) is a lower bound on the rate achievable on (8).

To prove the theorem, we apply the $\kappa\beta$ bound [9, Th. 25] to the channel (110). Following [9, Eq. (107)], we define the following measure of performance for the composite hypothesis test between an *auxiliary* output distribution $Q_{\Omega_{\mathbb{Y}}}$ defined on the subspace $\Omega_{\mathbb{Y}}$ and the collection of channel-output distributions $\{P_{\Omega_{\mathbb{Y}} \mid \Phi=\Phi}\}_{\Phi \in \mathcal{S}_{n,t}}$:

$$\kappa_{\tau}(\mathcal{S}_{n,t}, Q_{\Omega_{\mathbb{Y}}}) \triangleq \inf \int P_{Z \mid \Omega_{\mathbb{Y}}}(1 \mid \Omega_{\mathsf{Y}}) Q_{\Omega_{\mathbb{Y}}}(d\Omega_{\mathsf{Y}}) \tag{111}$$

where the infimum is over all probability distributions $P_{Z \mid \Omega_{\mathbb{Y}}} : \mathcal{G}_{n,t} \mapsto \{0, 1\}$ satisfying

$$\int P_{Z \mid \Omega_{\mathbb{Y}}}(1 \mid \Omega_{\mathsf{Y}}) P_{\Omega_{\mathbb{Y}} \mid \Phi=\Phi}(d\Omega_{\mathsf{Y}}) \geq \tau, \quad \forall \Phi \in \mathcal{S}_{n,t}. \tag{112}$$

By [9, Th. 25], we have that for every auxiliary distribution $Q_{\Omega_{\mathbb{Y}}}$

$$M \geq \frac{\kappa_{\tau}(\mathcal{S}_{n,t}, Q_{\Omega_{\mathbb{Y}}})}{\sup_{\Phi \in \mathcal{S}_{n,t}} \beta_{1-\epsilon+\tau}(P_{\Omega_{\mathbb{Y}} \mid \Phi=\Phi}, Q_{\Omega_{\mathbb{Y}}})} \tag{113}$$

where $\beta_{(\cdot)}(\cdot, \cdot)$ is defined in (6). We next lower-bound the RHS of (113) to obtain an expression that can be evaluated numerically. Fix a $\Phi \in \mathcal{S}_{n,t}$ and let

$$Z_{\Phi}(\Omega_{\mathbb{Y}}) = \mathbb{1}\{\sin^2\{\mathrm{span}(\Phi), \Omega_{\mathbb{Y}}\} \leq \gamma_n\} \tag{114}$$

where $\gamma_n \in [0, 1]$ is chosen so that

$$P_{\Omega_{\mathbb{Y}}|\Phi=\Phi}[Z_{\Phi}(\Omega_{\mathbb{Y}}) = 1] \geq 1 - \epsilon + \tau. \tag{115}$$

Since the noise matrix $\mathbb{W}$ is isotropically distributed, the probability distribution of the random variable $\sin^2\{\mathrm{span}(\Phi), \Omega_{\mathbb{Y}}\}$ (where $\Omega_{\mathbb{Y}} \sim P_{\Omega_{\mathbb{Y}}|\Phi=\Phi}$) does not depend on Φ. Hence, the chosen γ_n satisfies (115) for all $\Phi \in \mathcal{S}_{n,t}$. Furthermore, $Z_{\Phi}(\Omega_{\mathbb{Y}})$ can be viewed as a hypothesis test between $P_{\Omega_{\mathbb{Y}} \mid \Phi=\Phi}$ and $Q_{\Omega_{\mathbb{Y}}}$. Hence, by definition

$$\beta_{1-\epsilon+\tau}(P_{\Omega_{\mathbb{Y}} \mid \Phi=\Phi}, Q_{\Omega_{\mathbb{Y}}}) \leq Q_{\Omega_{\mathbb{Y}}}[Z_{\Phi}(\Omega_{\mathbb{Y}}) = 1] \tag{116}$$

for every $\Phi \in \mathcal{S}_{n,t}$.

We next evaluate the RHS of (116), taking as the auxiliary output distribution the uniform distribution on $\mathcal{G}_{n,r}$, which we denote by $Q_{\Omega_{\mathbb{Y}}}^{\mathrm{u}}$. With this choice, $Q_{\Omega_{\mathbb{Y}}}^{\mathrm{u}}[\sin^2\{\mathrm{span}(\Phi), \Omega_{\mathbb{Y}}\} \leq \gamma_n]$ does not depend on $\Phi \in \mathcal{S}_{n,t}$. To simplify calculations, we can therefore set $\Phi = \mathsf{I}_{n,t}$. Observe that under $Q_{\Omega_{\mathbb{Y}}}^{\mathrm{u}}$, the squares of the sines of the principle angles between $\mathrm{span}(\mathsf{I}_{n,t})$ and $\Omega_{\mathbb{Y}}$ have the same distribution as the eigenvalues of a *complex* multivariate Beta-distributed matrix $\mathbb{B} \sim \mathrm{Beta}_r(n-t, t)$ [36, Sec. 2]. By [37, Cor. 1], the distribution of $\det \mathbb{B}$ coincides with the distribution of $\prod_{i=1}^{r} B_i$, where $\{B_i\}$, $i = 1, \ldots, r$, are independent with

$B_i \sim \text{Beta}(n - t - i + 1, t)$. Using this result to compute the RHS of (116) we obtain

$$\sup_{\mathbf{\Phi} \in \mathcal{S}_{n,t}} \beta_{1-\epsilon+\tau}(P_{\Omega_{\mathbb{Y}} \mid \mathbf{\Phi}=\mathbf{\Phi}}, Q_{\Omega_{\mathbb{Y}}}) \leq \mathbb{P}\left[\prod_{j=1}^{r} B_j \leq \gamma_n\right] \quad (117)$$

where γ_n satisfies

$$\mathbb{P}\left[\sin^2\{\mathsf{I}_{n,t}, \mathsf{I}_{n,t}\mathsf{P}(\mathbb{H})\mathbb{H} + \mathbb{W}\} \leq \gamma_n\right] \geq 1 - \epsilon + \tau. \quad (118)$$

Note that (118) is equivalent to (44). Indeed

$$\mathbb{P}\left[\sin^2\{\mathsf{I}_{n,t}, \sqrt{n}\mathsf{I}_{n,t}\mathsf{P}(\mathbb{H})\mathbb{H} + \mathbb{W}\} \leq \gamma_n\right]$$
$$= \mathbb{P}\left[\sin^2\{\mathsf{I}_{n,t}, \sqrt{n}\mathsf{I}_{n,t}\text{diag}\{\sqrt{v_1^*\Lambda_1}, \ldots, \sqrt{v_m^*\Lambda_m},\right.$$
$$\left.\underbrace{0, \ldots, 0}_{t-m}\}\mathbb{V}^{\mathrm{H}} + \mathbb{W}\} \leq \gamma_n\right] \quad (119)$$
$$= \mathbb{P}\left[\sin^2\{\mathsf{I}_{n,t}, \sqrt{n}\mathsf{I}_{n,t}\text{diag}\{\sqrt{v_1^*\Lambda_1}, \ldots, \sqrt{v_m^*\Lambda_m},\right.$$
$$\left.\underbrace{0, \ldots, 0}_{t-m}\} + \mathbb{W}\mathbb{V}\} \leq \gamma_n\right] \quad (120)$$
$$= \mathbb{P}\left[\sin^2\{\mathsf{I}_{n,t}, \sqrt{n}\mathsf{I}_{n,t}\text{diag}\{\sqrt{v_1^*\Lambda_1}, \ldots, \sqrt{v_m^*\Lambda_m},\right.$$
$$\left.\underbrace{0, \ldots, 0}_{t-m}\} + \mathbb{W}\} \leq \gamma_n\right] \quad (121)$$

where $\mathbb{V}$ contains the right singular vectors of $\mathbb{H}$ (see (107)). Here, (119) follows from (108); (120) follows because right-multiplying a matrix A by a unitary matrix does not change the subspace spanned by the columns of A and hence, it does not change $\sin\{\cdot, \cdot\}$; (121) follows because $\mathbb{W}$ is isotropically distributed and hence $\mathbb{W}\mathbb{V}$ has the same distribution as $\mathbb{W}$.

To conclude the proof, it remains to show that

$$\kappa_\tau(\mathcal{S}_{n,t}, Q_{\Omega_{\mathbb{Y}}}^{\mathrm{u}}) \geq \tau. \quad (122)$$

Once this is done, the desired lower bound (43) follows by using the inequality (117) and (122) in (113), by taking the logarithm of both sides of (113), and by dividing by the blocklength n.

To prove (122), we replace (112) with the less stringent constraint that

$$\mathbb{E}_{P_{\mathbf{\Phi}}^{\mathrm{u}}}\left[\int P_{Z \mid \Omega_{\mathbb{Y}}}(1 \mid \Omega_{\mathbb{Y}})P_{\Omega_{\mathbb{Y}} \mid \mathbf{\Phi}}(d\Omega_{\mathbb{Y}})\right] \geq \tau \quad (123)$$

where $P_{\mathbf{\Phi}}^{\mathrm{u}}$ is the uniform input distribution on $\mathcal{S}_{n,t}$. Since replacing (112) by (123) enlarges the feasible region of the minimization problem (111), we obtain an infimum in (111) (denoted by $\kappa_\tau^{\mathrm{u}}(\mathcal{S}_{n,t}, Q_{\Omega_{\mathbb{Y}}}^{\mathrm{u}})$) that is no larger than $\kappa_\tau(\mathcal{S}_{n,t}, Q_{\Omega_{\mathbb{Y}}}^{\mathrm{u}})$. The key observation is that the uniform distribution $P_{\mathbf{\Phi}}^{\mathrm{u}}$ induces an isotropic distribution on $\mathbb{Y}$. This implies that the induced distribution on $\Omega_{\mathbb{Y}}$ is the uniform distribution on $\mathcal{G}_{n,r}$, i.e., $Q_{\Omega_{\mathbb{Y}}}^{\mathrm{u}}$. Therefore, it follows that

$$\int P_{Z \mid \Omega_{\mathbb{Y}}}(1 \mid \Omega_{\mathbb{Y}})Q_{\Omega_{\mathbb{Y}}}^{\mathrm{u}}(d\Omega_{\mathbb{Y}})$$
$$= \mathbb{E}_{P_{\mathbf{\Phi}}^{\mathrm{u}}}\left[\int P_{Z \mid \Omega_{\mathbb{Y}}}(1 \mid \Omega_{\mathbb{Y}})P_{\Omega_{\mathbb{Y}} \mid \mathbf{\Phi}}(d\Omega_{\mathbb{Y}})\right] \quad (124)$$
$$\geq \tau \quad (125)$$

for all distributions $P_{Z \mid \Omega_{\mathbb{Y}}}$ that satisfy (123). This proves (122), since

$$\kappa_\tau(\mathcal{S}_{n,t}, Q_{\Omega_{\mathbb{Y}}}^{\mathrm{u}}) \geq \kappa_\tau^{\mathrm{u}}(\mathcal{S}_{n,t}, Q_{\Omega_{\mathbb{Y}}}^{\mathrm{u}}) \geq \tau. \quad (126)$$

APPENDIX III
PROOF OF THEOREM 2 (CSIRT CONVERSE BOUND)

When CSI is available at both the transmitter and the receiver, the MIMO channel (8) can be transformed into the following set of m parallel quasi-static channels

$$\mathbf{Y}_i = \boldsymbol{x}_i\sqrt{\Lambda_i} + \mathbf{W}_i, \qquad i = 1, \ldots, m \quad (127)$$

by performing a singular value decomposition [17, Sec. 3.1]. Here, $\Lambda_1 \geq \cdots \geq \Lambda_m$ denote the m largest eigenvalues of $\mathbb{H}\mathbb{H}^{\mathrm{H}}$, and $\mathbf{W}_i \sim \mathcal{CN}(\mathbf{0}, \mathsf{I}_n)$, $i = 1, \ldots, m$, are independent noise vectors.

Next, we establish a converse bound for the channel (127). Let $\mathsf{X} = [\boldsymbol{x}_1 \cdots \boldsymbol{x}_m]$ and fix an $(n, M, \epsilon)_{\mathrm{rt}}$ code. Note that (12) implies

$$\sum_{i=1}^{m} \|\boldsymbol{x}_i\|^2 \leq n\rho. \quad (128)$$

To simplify the presentation, we assume that the encoder f_{tx} is deterministic. Nevertheless, the theorem holds also if we allow for randomized encoders. We further assume that the encoder f_{tx} acts on the pairs $(j, \boldsymbol{\lambda})$ instead of (j, H) (cf., Definition 3). The channel (127) and the encoder f_{tx} define a random transformation $P_{\mathbb{Y},\boldsymbol{\Lambda} \mid J}$ from the message set $\{1, \ldots, M\}$ to the space $\mathbb{C}^{n \times m} \times \mathbb{R}_+^m$:

$$P_{\mathbb{Y},\boldsymbol{\Lambda} \mid J} = P_{\boldsymbol{\Lambda}}P_{\mathbb{Y} \mid \boldsymbol{\Lambda},J} \quad (129)$$

where $\mathbb{Y} = [\mathbf{Y}_1, \ldots, \mathbf{Y}_m]$ and

$$P_{\mathbb{Y} \mid \boldsymbol{\Lambda}=\boldsymbol{\lambda}, J=j} \triangleq P_{\mathbb{Y} \mid \boldsymbol{\Lambda}=\boldsymbol{\lambda}, \mathsf{X}=f_{\mathrm{tx}}(j,\boldsymbol{\lambda})}. \quad (130)$$

We can think of $P_{\mathbb{Y},\boldsymbol{\Lambda} \mid J}$ as the channel law associated with

$$J \longrightarrow \mathbb{Y}, \boldsymbol{\Lambda}. \quad (131)$$

To upper-bound $R_{\mathrm{rt}}^*(n, \epsilon)$, we use the meta-converse theorem [9, Th. 30] on the channel (131). We start by associating to each codeword X a power-allocation vector $\tilde{\boldsymbol{v}}(\mathsf{X})$ whose entries $\tilde{v}_i(\mathsf{X})$ are

$$\tilde{v}_i(\mathsf{X}) \triangleq \frac{1}{n}\|\boldsymbol{x}_i\|^2, \quad i = 1, \ldots, m. \quad (132)$$

We take as auxiliary channel $Q_{\mathbb{Y},\boldsymbol{\Lambda} \mid J} = P_{\boldsymbol{\Lambda}}Q_{\mathbb{Y} \mid \boldsymbol{\Lambda},J}$, where

$$Q_{\mathbb{Y} \mid \boldsymbol{\Lambda}=\boldsymbol{\lambda}, J=j} = \prod_{i=1}^{m} Q_{\mathbf{Y}_i \mid \boldsymbol{\Lambda}=\boldsymbol{\lambda}, J=j} \quad (133)$$

and

$$Q_{\mathbf{Y}_i \mid \boldsymbol{\Lambda}=\boldsymbol{\lambda}, J=j} = \mathcal{CN}\left(\mathbf{0}, \left[1 + (\tilde{v}_i \circ f_{\mathrm{tx}}(j,\boldsymbol{\lambda}))\lambda_i\right]\mathsf{I}_n\right). \quad (134)$$

By [9, Th. 30], we obtain

$$\min_{j \in \{1,\ldots,M\}} \beta_{1-\epsilon}(P_{\mathbb{Y}\boldsymbol{\Lambda} \mid J=j}, Q_{\mathbb{Y}\boldsymbol{\Lambda} \mid J=j}) \leq 1 - \epsilon' \quad (135)$$

where ϵ' is the maximal probability of error over $Q_{\mathbb{Y},\boldsymbol{\Lambda} \mid J}$. We shall prove Theorem 2 in the following two steps: in Appendix III-1, we evaluate $\beta_{1-\epsilon}(P_{\mathbb{Y}\boldsymbol{\Lambda} \mid J=j}, Q_{\mathbb{Y}\boldsymbol{\Lambda} \mid J=j})$; in Appendix III-2, we relate ϵ' to $R_{\mathrm{rt}}^*(n, \epsilon)$ by establishing a converse bound on the auxiliary channel $Q_{\mathbb{Y},\boldsymbol{\Lambda} \mid J}$.

1) Evaluation of $\beta_{1-\epsilon}$: Let j^* be the message that achieves the minimum in (135), let $f^*_{\text{tx}}(\boldsymbol{\lambda}) \triangleq f_{\text{tx}}(j^*, \boldsymbol{\lambda})$, and let

$$\beta_{1-\epsilon}(f^*_{\text{tx}}) \triangleq \beta_{1-\epsilon}(P_{\mathbb{Y}, \boldsymbol{\Lambda} \mid J=j^*}, Q_{\mathbb{Y}, \boldsymbol{\Lambda} \mid J=j^*}). \quad (136)$$

Using (136), we can rewrite (135) as

$$\beta_{1-\epsilon}(f^*_{\text{tx}}) \leq 1 - \epsilon'. \quad (137)$$

Let now

$$r(f^*_{\text{tx}}; \mathbb{Y}, \boldsymbol{\Lambda}) \triangleq \log \frac{dP_{\mathbb{Y}, \boldsymbol{\Lambda} \mid J=j^*}}{dQ_{\mathbb{Y}, \boldsymbol{\Lambda} \mid J=j^*}}. \quad (138)$$

Note that, under both $P_{\mathbb{Y}, \boldsymbol{\Lambda} \mid J=j^*}$ and $Q_{\mathbb{Y}, \boldsymbol{\Lambda} \mid J=j^*}$, the random variable $r(f^*_{\text{tx}}; \mathbb{Y}, \boldsymbol{\Lambda})$ has absolutely continuous cumulative distribution function (cdf) with respect to the Lebesgue measure. By the Neyman-Pearson lemma [38, p. 300]

$$\beta_{1-\epsilon}(f^*_{\text{tx}}) = Q_{\mathbb{Y}, \boldsymbol{\Lambda} \mid J=j^*}[r(f^*_{\text{tx}}; \mathbb{Y}, \boldsymbol{\Lambda}) \geq n\gamma_n(f^*_{\text{tx}})] \quad (139)$$

where $\gamma_n(f^*_{\text{tx}})$ is the solution of

$$P_{\mathbb{Y}, \boldsymbol{\Lambda} \mid J=j^*}[r(f^*_{\text{tx}}; \mathbb{Y}, \boldsymbol{\Lambda}) \leq n\gamma_n(f^*_{\text{tx}})] = \epsilon. \quad (140)$$

Let now $\boldsymbol{v} \triangleq \tilde{\boldsymbol{v}} \circ f^*_{\text{tx}}$. Because of the power constraint (128), $\boldsymbol{v}$ is a mapping from $\{1, \ldots, M\}$ to the set $\mathcal{V}_m$ defined in (46). Furthermore, under $Q_{\mathbb{Y}, \boldsymbol{\Lambda} \mid J=j^*}$, the random variable $r(f^*_{\text{tx}}; \mathbb{Y}, \boldsymbol{\Lambda})$ has the same distribution as $L^{\text{rt}}_n(\boldsymbol{v}, \boldsymbol{\Lambda})$ in (47), and under $P_{\mathbb{Y}, \boldsymbol{\Lambda} \mid J=j^*}$, it has the same distribution as $S^{\text{rt}}_n(\boldsymbol{v}, \boldsymbol{\Lambda})$ in (48). Thus, (137) is equivalent to

$$\mathbb{P}[L^{\text{rt}}_n(\boldsymbol{v}, \boldsymbol{\Lambda}) \geq n\gamma_n(\boldsymbol{v})] \leq 1 - \epsilon' \quad (141)$$

where $\gamma_n(\boldsymbol{v})$ is the solution of (51). Note that this upper bound depends on the chosen code only through the induced power allocation function $\boldsymbol{v}$. To conclude, we take the infimum of the LHS of (141) over all power allocation functions $\boldsymbol{v}$ to obtain a bound that holds for all $(n, M, \epsilon)_{\text{rt}}$ codes.

2) Converse on the auxiliary channel: We next relate ϵ' to $R^*_{\text{rt}}(n, \epsilon)$. The following lemma, whose proof can be found at the end of this appendix, serves this purpose.

Lemma 14: For every code with M codewords and block-length n, the maximum probability of error ϵ' over the channel $Q_{\mathbb{Y}, \boldsymbol{\Lambda} \mid J}$ satisfies

$$1 - \epsilon' \leq \frac{c_{\text{rt}}(n)}{M} \quad (142)$$

where $c_{\text{rt}}(n)$ is given in (50).

Using Lemma 14, we obtain

$$\inf_{\boldsymbol{v}(\cdot)} \mathbb{P}[L^{\text{rt}}_n(\boldsymbol{v}, \boldsymbol{\Lambda}) \geq n\gamma_n(\boldsymbol{v})] \leq \frac{c_{\text{rt}}(n)}{M}. \quad (143)$$

The desired lower bound (49) follows by taking the logarithm on both sides of (143) and dividing by n.

Proof of Lemma 14: By (133), given $\boldsymbol{\Lambda} = \boldsymbol{\lambda}$, the output of the channel $Q_{\mathbb{Y}, \boldsymbol{\Lambda} \mid J}$ depends on the input J only through $\boldsymbol{S} \triangleq \tilde{\boldsymbol{v}} \circ f_{\text{tx}}(J, \boldsymbol{\lambda})$, i.e., through the norm of each column of the codeword matrix $f_{\text{tx}}(J, \boldsymbol{\lambda})$. Let $\boldsymbol{U} \triangleq \tilde{\boldsymbol{v}}(\mathbb{Y})$. In words, the entries of $\boldsymbol{U}$ are the square of the norm of the columns of $\mathbb{Y}$ normalized by the blocklength n. It follows that $(\boldsymbol{U}, \boldsymbol{\Lambda})$ is a sufficient statistic for the detection of J from $(\mathbb{Y}, \boldsymbol{\Lambda})$. Hence, to lower-bound ϵ' and establish (142), it suffices to lower-bound

the maximal error probability over the channel $Q_{\boldsymbol{U}, \boldsymbol{\Lambda} \mid \boldsymbol{S}}$ defined by

$$U_i = \frac{1 + S_i \Lambda_i}{n} \sum_{l=1}^{n} |W_{i,l}|^2, \quad i = 1, \ldots, m. \quad (144)$$

Here, U_i denotes the ith entry of $\boldsymbol{U}$, the random variables $\{W_{i,l}\}$ are i.i.d. $\mathcal{CN}(0, 1)$-distributed, and the input $\boldsymbol{S} = [S_1 \ldots S_m]$ has nonnegative entries whose sum does not exceed ρ, i.e., $\boldsymbol{S} \in \mathcal{V}_m$. Note that, given S_i and Λ_i, the random variable U_i in (144) is Gamma-distributed, i.e., its pdf $q_{U_i \mid S_i, \Lambda_i}$ is given by

$$q_{U_i \mid S_i, \Lambda_i}(u_i \mid s_i, \lambda_i)$$
$$= \frac{n^n}{(1 + s_i \lambda_i)^n \Gamma(n)} u_i^{n-1} \exp\left(-\frac{n u_i}{1 + s_i \lambda_i}\right). \quad (145)$$

Furthermore, the random variables $U_1, \ldots, U_m$ are conditionally independent given $\boldsymbol{S}$ and $\boldsymbol{\Lambda}$.

We shall use that $q_{U_i \mid S_i, \Lambda_i}$ can be upper-bounded as

$$q_{U_i \mid S_i, \Lambda_i}(u_i \mid s_i, \lambda_i)$$
$$\leq g_i(u_i, \lambda_i) \quad (146)$$
$$\triangleq \begin{cases} \dfrac{n(n-1)^{n-1}}{\Gamma(n)} e^{-(n-1)}, & \text{if } u_i \leq \frac{n-1}{n}(1 + \rho\lambda_i) \\ \dfrac{n^n u_i^{n-1} e^{-n u_i/(1+\rho\lambda_i)}}{\Gamma(n)(1+\rho\lambda_i)^{n-1}}, & \text{if } u_i > \frac{n-1}{n}(1 + \rho\lambda_i) \end{cases} \quad (147)$$

which follows because $1 + s_i \lambda_i \leq 1 + \rho\lambda_i$, and because $q_{U_i \mid S_i, \Lambda_i}$ is a unimodal function with maximum at

$$u_i = \frac{n-1}{n}(1 + s_i \lambda_i). \quad (148)$$

The bound in (147) is useful because it is integrable and does not depend on the input s_i.

Consider now an arbitrary code $\{\boldsymbol{c}_1(\boldsymbol{\Lambda}), \ldots, \boldsymbol{c}_M(\boldsymbol{\Lambda})\} \subset \mathcal{V}_m$ for the channel $Q_{\boldsymbol{U}, \boldsymbol{\Lambda} \mid \boldsymbol{S}}$. Let $\mathcal{D}_j(\boldsymbol{\Lambda})$, $j = 1, \ldots, M$, be the (disjoint) *decoding sets* corresponding to the M codewords $\{\boldsymbol{c}_j(\boldsymbol{\Lambda})\}$. Let ϵ'_{avg} be the *average* probability of error over the channel $Q_{\boldsymbol{U}, \boldsymbol{\Lambda} \mid \boldsymbol{S}}$. We have

$$1 - \epsilon' \leq 1 - \epsilon'_{\text{avg}} \quad (149)$$
$$= \frac{1}{M} \mathbb{E}_{\boldsymbol{\Lambda}}\left[\sum_{j=1}^{M} \int_{\mathcal{D}_j(\boldsymbol{\Lambda})} q_{\boldsymbol{U} \mid \boldsymbol{S}, \boldsymbol{\Lambda}}(\boldsymbol{u} \mid \boldsymbol{c}_j(\boldsymbol{\Lambda}), \boldsymbol{\Lambda}) d\boldsymbol{u}\right] \quad (150)$$
$$\leq \frac{1}{M} \mathbb{E}_{\boldsymbol{\Lambda}}\left[\sum_{j=1}^{M} \int_{\mathcal{D}_j(\boldsymbol{\Lambda})} \left(\prod_{i=1}^{m} g_i(u_i, \Lambda_i)\right) d\boldsymbol{u}\right] \quad (151)$$
$$= \frac{1}{M} \mathbb{E}_{\boldsymbol{\Lambda}}\left[\int_{\mathbb{R}_+^m} \left(\prod_{i=1}^{m} g_i(u_i, \Lambda_i)\right) d\boldsymbol{u}\right] \quad (152)$$
$$= \frac{1}{M} \mathbb{E}_{\boldsymbol{\Lambda}}\left[\prod_{i=1}^{m} \int_0^{+\infty} g_i(u_i, \Lambda_i) du_i\right] \quad (153)$$

where (151) follows from (147), and where (152) follows because $g_i(u_i, \Lambda_i)$ is independent of the message j and because $\bigcup_{j=1}^{M} \mathcal{D}_j(\boldsymbol{\Lambda}) = \mathbb{R}_+^m$. After algebraic manipulations, we obtain

$$\int_0^\infty g_i(u_i, \lambda_i) du_i$$
$$= \frac{(1 + \rho\lambda_i)}{\Gamma(n)}\left[(n-1)^n e^{-(n-1)} + \Gamma(n, n-1)\right]. \quad (154)$$

Here, $\Gamma(\cdot,\cdot)$ denotes the *(upper) incomplete Gamma function* [39, Sec. 6.5]. Substituting (154) into (153), we finally obtain that for every code $\{c_1(\mathbf{\Lambda}), \ldots, c_M(\mathbf{\Lambda})\} \subset \mathcal{V}_m$,

$$1 - \epsilon' \leq \frac{1}{M} \left(\frac{(n-1)^n e^{-(n-1)}}{\Gamma(n)} + \frac{\Gamma(n, n-1)}{\Gamma(n)} \right)^m$$
$$\times \mathbb{E}\left[\prod_{i=1}^m (1 + \rho\Lambda_i) \right] \qquad (155)$$
$$= \frac{c_{\mathrm{rt}}(n)}{M}. \qquad (156)$$

This proves Lemma 14.

APPENDIX IV
PROOF OF THE CONVERSE PART OF THEOREM 3

As a first step towards establishing (56), we relax the upper bound (49) by lower-bounding its denominator. Recall that by definition (see Appendix III-1)

$$\mathbb{P}[L_n^{\mathrm{rt}}(\boldsymbol{v}, \mathbf{\Lambda}) \geq n\gamma_n(\boldsymbol{v})] = \beta_{1-\epsilon}(P_{\mathbb{Y}, \mathbf{\Lambda} \mid J=j^*}, Q_{\mathbb{Y}, \mathbf{\Lambda} \mid J=j^*}). \qquad (157)$$

We shall use the following inequality: for every $\eta > 0$ [9, Eq. (102)]

$$\beta_{1-\epsilon}(P, Q) \geq \frac{1}{\eta} \left(1 - P\left[\frac{dP}{dQ} \geq \eta \right] - \epsilon \right). \qquad (158)$$

Using (158) with $P = P_{\mathbb{Y}, \mathbf{\Lambda} \mid J=j^*}$, $Q = Q_{\mathbb{Y}, \mathbf{\Lambda} \mid J=j^*}$, $\eta = e^{n\gamma}$, and recalling that (see Appendix III-1)

$$1 - P_{\mathbb{Y}, \mathbf{\Lambda} \mid J=j^*}\left[\frac{dP_{\mathbb{Y}, \mathbf{\Lambda} \mid J=j^*}}{dQ_{\mathbb{Y}, \mathbf{\Lambda} \mid J=j^*}} \geq e^{n\gamma} \right] = \mathbb{P}[S_n^{\mathrm{rt}}(\boldsymbol{v}, \mathbf{\Lambda}) \leq n\gamma] \qquad (159)$$

we obtain that for every $\gamma > 0$

$$\beta_{1-\epsilon}\left(P_{\mathbb{Y}, \mathbf{\Lambda} \mid J=j^*}, Q_{\mathbb{Y}, \mathbf{\Lambda} \mid J=j^*} \right)$$
$$\geq e^{-n\gamma} \left(\mathbb{P}[S_n^{\mathrm{rt}}(\boldsymbol{v}, \mathbf{\Lambda}) \leq n\gamma] - \epsilon \right). \qquad (160)$$

Using (160) and the estimate

$$\log c_{\mathrm{rt}}(n) = \frac{m}{2} \log n + \mathcal{O}(1) \qquad (161)$$

(which follows from (50), Assumption 1 in Theorem 3, and from algebraic manipulations), we upper-bound the RHS of (49) as

$$R_{\mathrm{rt}}^*(n, \epsilon) \leq \gamma - \frac{1}{n} \log\left(\inf_{\boldsymbol{v}(\cdot)} \mathbb{P}[S_n^{\mathrm{rt}}(\boldsymbol{v}, \mathbf{\Lambda}) \leq n\gamma] - \epsilon \right)$$
$$+ \frac{m}{2} \frac{\log n}{n} + \mathcal{O}\left(\frac{1}{n} \right). \qquad (162)$$

To conclude the proof we show that for every γ in a certain neighborhood of C_ϵ^{tx} (recall that γ is a free optimization parameter),

$$\inf_{\boldsymbol{v}(\cdot)} \mathbb{P}[S_n^{\mathrm{rt}}(\boldsymbol{v}, \mathbf{\Lambda}) \leq n\gamma] \geq F_{\mathrm{tx}}(\gamma) + \mathcal{O}\left(\frac{1}{n} \right) \qquad (163)$$

where $F_{\mathrm{tx}}(\cdot)$ is the outage probability defined in (20) and the $\mathcal{O}(1/n)$ term is uniform in γ. The desired result (56) follows then by substituting (163) into (162), setting γ as the solution of

$$F_{\mathrm{tx}}(\gamma) - \epsilon + \mathcal{O}(1/n) = 1/n \qquad (164)$$

and by noting that this γ satisfies

$$\gamma = C_\epsilon^{\mathrm{tx}} + \mathcal{O}(1/n) \qquad (165)$$

i.e., it belongs to the desired neighborhood of C_ϵ^{tx} for sufficiently large n. Here, (165) follows by a Taylor series expansion [40, Th. 5.15] of $F_{\mathrm{tx}}(\gamma)$ around C_ϵ^{tx}, and because $F_{\mathrm{tx}}(C_\epsilon^{\mathrm{tx}}) = \epsilon$ and $F_{\mathrm{tx}}'(C_\epsilon^{\mathrm{tx}}) > 0$ by assumption.

In the reminder of this appendix, we will prove (163). Our proof consists of the three steps sketched below.

Step 1: Given $\boldsymbol{v}$ and $\mathbf{\Lambda}$, the random variable $S_n^{\mathrm{rt}}(\boldsymbol{v}, \mathbf{\Lambda})$ (see (48) for its definition) is the sum of n i.i.d. random variables with mean

$$\mu(\boldsymbol{v}, \mathbf{\Lambda}) \triangleq \sum_{j=1}^m \log\left(1 + \Lambda_j v_j(\mathbf{\Lambda}) \right) \qquad (166)$$

and variance

$$\sigma^2(\boldsymbol{v}, \mathbf{\Lambda}) \triangleq m - \sum_{j=1}^m \frac{1}{\left(1 + \Lambda_j v_j(\mathbf{\Lambda}) \right)^2}. \qquad (167)$$

Fix an arbitrary power allocation function $\boldsymbol{v}(\cdot)$, and assume that $\mathbf{\Lambda} = \boldsymbol{\lambda}$. Let

$$u(\boldsymbol{v}, \boldsymbol{\lambda}) \triangleq \frac{\gamma - \mu(\boldsymbol{v}, \boldsymbol{\lambda})}{\sigma(\boldsymbol{v}, \boldsymbol{\lambda})}. \qquad (168)$$

Using the Cramer-Esseen theorem (see Theorem 15 below), we show in Appendix IV-A that

$$\mathbb{P}[S_n^{\mathrm{rt}}(\boldsymbol{v}, \mathbf{\Lambda}) \leq n\gamma \mid \mathbf{\Lambda} = \boldsymbol{\lambda}] \geq q_n(u(\boldsymbol{v}, \boldsymbol{\lambda})) + \frac{k_3}{n} \quad (169)$$

where

$$q_n(x) \triangleq Q(-\sqrt{n}x) - \frac{[1 - nx^2]^+ e^{-nx^2/2}}{6\sqrt{n}} \qquad (170)$$

and k_3 is a finite constant independent of $\boldsymbol{\lambda}$, $\boldsymbol{v}$ and γ.

Step 2: We make the RHS of (169) independent of $\boldsymbol{v}$ by minimizing $q_n(u(\boldsymbol{v}, \boldsymbol{\lambda}))$ over $\boldsymbol{v}$. Specifically, we establish in Appendix IV-B the following result: for every γ in a certain neighborhood of C_ϵ^{tx}, we have that

$$\mathbb{P}[S_n^{\mathrm{rt}}(\boldsymbol{v}, \mathbf{\Lambda}) \leq n\gamma \mid \mathbf{\Lambda} = \boldsymbol{\lambda}] \geq q_n(\hat{u}(\boldsymbol{\lambda})) + \frac{k_3}{n} \qquad (171)$$

where $\hat{u}(\boldsymbol{\lambda})$ is defined in (187).

Step 3: We average (171) over $\mathbf{\Lambda}$ and establish in Appendix IV-C the bound (163). This concludes the proof.

A. Proof of (169)

We need the following version of the Cramer-Esseen Theorem.[10]

Theorem 15: Let $X_1, \ldots, X_n$ be a sequence of i.i.d. real random variables having zero mean and unit variance. Furthermore, let

$$\varphi(t) \triangleq \mathbb{E}\left[e^{itX_1} \right] \quad \text{and} \quad F_n(\xi) \triangleq \mathbb{P}\left[\frac{1}{\sqrt{n}} \sum_{j=1}^n X_j \leq \xi \right]. \quad (172)$$

---

[10]The Berry-Esseen Theorem used in [9] to prove (2) yields an asymptotic expansion in (163) up to a $\mathcal{O}(1/\sqrt{n})$ term. This is not sufficient here, since we need an expansion up to a $\mathcal{O}(1/n)$ term (see (163)).

If $\mathbb{E}\big[|X_1|^4\big] < \infty$ and if $\sup_{|t|\geq\zeta} |\varphi(t)| \leq k_0$ for some $k_0 < 1$, where $\zeta \triangleq 1/(12\mathbb{E}\big[|X_1|^3\big])$, then for every ξ and n

$$\left| F_n(\xi) - Q(-\xi) - k_1(1-\xi^2)e^{-\xi^2/2}\frac{1}{\sqrt{n}} \right|$$
$$\leq k_2 \left\{ \frac{\mathbb{E}\big[|X_1|^4\big]}{n} + n^6\left(k_0 + \frac{1}{2n}\right)^n \right\}. \quad (173)$$

Here, $k_1 \triangleq \mathbb{E}\big[X_1^3\big]/(6\sqrt{2\pi})$, and k_2 is a positive constant independent of $\{X_i\}$ and ξ.

Proof: The inequality (173) is a consequence of the tighter inequality reported in [16, Th. VI.1]. ∎

Let

$$T_l(\boldsymbol{v}, \boldsymbol{\Lambda}) \triangleq \frac{1}{\sigma(\boldsymbol{v}, \boldsymbol{\Lambda})} \sum_{j=1}^{m} \left(1 - \frac{\big|\sqrt{\Lambda_j v_j(\boldsymbol{\Lambda})} Z_{l,j} - 1\big|^2}{1 + \Lambda_j v_j(\boldsymbol{\Lambda})} \right) \quad (174)$$

where $Z_{l,j}$, $l = 1, \ldots, n$ and $j = 1, \ldots, m$, are i.i.d. $\mathcal{CN}(0,1)$ distributed. The random variables $T_1, \ldots, T_n$ have zero mean and unit variance, and are conditionally i.i.d. given $\boldsymbol{\Lambda}$. Furthermore, by construction

$$\mathbb{P}\big[S_n^{\mathrm{rt}}(\boldsymbol{v}, \boldsymbol{\Lambda}) \leq n\gamma\big] = \mathbb{P}\left[\frac{1}{\sqrt{n}} \sum_{l=1}^{n} T_l(\boldsymbol{v}, \boldsymbol{\Lambda}) \leq \sqrt{n}u(\boldsymbol{v}, \boldsymbol{\Lambda}) \right] \quad (175)$$

where $u(\boldsymbol{v}, \boldsymbol{\Lambda})$ was defined in (168). We next show that the conditions under which Theorem 15 holds are satisfied by the random variables $\{T_l\}$.

We start by noting that if $\lambda_j v_j(\boldsymbol{\lambda})$, $j = 1, \ldots, m$, are identically zero, then $S_n^{\mathrm{rt}}(\boldsymbol{v}, \boldsymbol{\Lambda}) = 0$, so (169) holds trivially. Hence, we will focus on the case where $\{\lambda_j v_j(\boldsymbol{\lambda})\}$ are not all identically zero. Let

$$\varphi_{T_l}(t) \triangleq \mathbb{E}\big[e^{itT_l}\big|\boldsymbol{\Lambda} = \boldsymbol{\lambda}\big] \quad (176)$$

and

$$\zeta \triangleq \frac{1}{12\mathbb{E}\big[|T_l|^3\big|\boldsymbol{\Lambda} = \boldsymbol{\lambda}\big]}. \quad (177)$$

We next show that there exists a $k_0 < 1$ such that $\sup_{|t|>\zeta} |\varphi_{T_l}(t)| \leq k_0$ for every $\boldsymbol{\lambda} \in \mathbb{R}_+^m$ and every function $\boldsymbol{v}(\cdot)$. We start by evaluating ζ. For every $\boldsymbol{\lambda} \in \mathbb{R}_+^m$ and every $\boldsymbol{v}(\cdot)$ such that $\lambda_j v_j(\boldsymbol{\lambda})$, $1 \leq j \leq m$, are not identically zero, it can be shown through algebraic manipulations that

$$\mathbb{E}\big[|T_l|^4\big|\boldsymbol{\Lambda} = \boldsymbol{\lambda}\big] \leq 9. \quad (178)$$

By Lyapunov's inequality [16, p. 18], this implies that

$$\mathbb{E}\big[|T_l|^3\big|\boldsymbol{\Lambda} = \boldsymbol{\lambda}\big] \leq \left(\mathbb{E}\big[|T_l|^4\big|\boldsymbol{\Lambda} = \boldsymbol{\lambda}\big]\right)^{3/4} \leq 9^{3/4}. \quad (179)$$

Hence,

$$\zeta = \frac{1}{12\mathbb{E}\big[|T_l|^3\big|\boldsymbol{\Lambda} = \boldsymbol{\lambda}\big]} \geq \frac{1}{12 \times 9^{3/4}} \triangleq \zeta_0. \quad (180)$$

By (180), we have

$$\sup_{|t|>\zeta} \big|\varphi_{T_l}(t)\big| \leq \sup_{|t|>\zeta_0} \big|\varphi_{T_l}(t)\big| \quad (181)$$

where ζ_0 does not depend on $\boldsymbol{\lambda}$ and $\boldsymbol{v}$. Through algebraic manipulations, we can further show that the RHS of (181) is upper-bounded by

$$\sup_{|t|>\zeta_0} \big|\varphi_{T_l}(t)\big| \leq \frac{1}{\sqrt{1 + \zeta_0^2/m}} \triangleq k_0 < 1. \quad (182)$$

The inequalities (178) and (182) imply that the conditions in Theorem 15 are met. Hence, we conclude that, by Theorem 15, for every n, $\boldsymbol{\lambda}$, and $\boldsymbol{v}(\cdot)$,

$$\mathbb{P}\left[\frac{1}{\sqrt{n}} \sum_{l=1}^{n} T_l \leq \sqrt{n}u(\boldsymbol{v}, \boldsymbol{\lambda}) \bigg| \boldsymbol{\Lambda} = \boldsymbol{\lambda} \right] - Q\big(-\sqrt{n}u(\boldsymbol{v}, \boldsymbol{\lambda})\big)$$
$$\geq \frac{\mathbb{E}\big[T_l^3\big|\boldsymbol{\Lambda} = \boldsymbol{\lambda}\big]}{6\sqrt{2\pi}\sqrt{n}}(1 - nu(\boldsymbol{v}, \boldsymbol{\lambda})^2)e^{-nu(\boldsymbol{v}, \boldsymbol{\lambda})^2/2} - \frac{9k_2}{n}$$
$$- k_2 n^6 \left(k_0 + \frac{1}{2n}\right)^n \quad (183)$$

where $u(\boldsymbol{v}, \boldsymbol{\lambda})$ was defined in (168). The inequality (169) follows then by noting that

$$0 \geq \mathbb{E}\big[T_l^3\big|\boldsymbol{\Lambda} = \boldsymbol{\lambda}\big] \geq -\sqrt{2\pi} \quad (184)$$

and that

$$\sup_{n\geq 1} n\left(k_2 n^6 \left(k_0 + \frac{1}{2n}\right)^n\right) < \infty. \quad (185)$$

B. Proof of (171)

For every fixed $\boldsymbol{\lambda}$, we minimize $q_n(u(\boldsymbol{v}, \boldsymbol{\lambda}))$ on the RHS of (169) over all power allocation functions $\boldsymbol{v}(\cdot)$. With a slight abuse of notation, we use $\boldsymbol{v} \in \mathcal{V}_m$ (where $\mathcal{V}_m$ was defined in (46)) to denote the vector $\boldsymbol{v}(\boldsymbol{\lambda})$ whenever no ambiguity arises. Since the function $q_n(x)$ in (170) is monotonically increasing in x, the vector $\boldsymbol{v} \in \mathcal{V}_m$ that minimizes $q_n(u(\boldsymbol{v}, \boldsymbol{\lambda}))$ is the solution of

$$\min_{\boldsymbol{v} \in \mathcal{V}_m} u(\boldsymbol{v}, \boldsymbol{\lambda}). \quad (186)$$

The minimization in (186) is difficult to solve since $u(\boldsymbol{v}, \boldsymbol{\lambda})$ is neither convex nor concave in $\boldsymbol{v}$. For our purposes, it suffices to obtain a lower bound on (186), which is given in Lemma 16 below. Together with (187) and the monotonicity of $q_n(\cdot)$, this then yields (171).

Lemma 16: Let $\boldsymbol{v}^*$, $\mu(\boldsymbol{v}, \boldsymbol{\lambda})$, $\sigma(\boldsymbol{v}, \boldsymbol{\lambda})$, and $u(\boldsymbol{v}, \boldsymbol{\lambda})$ be as in (45), (166), (167), and (168), respectively. Moreover, let $\mu^*(\boldsymbol{\lambda}) \triangleq \mu(\boldsymbol{v}^*, \boldsymbol{\lambda})$ and $\sigma^*(\boldsymbol{\lambda}) \triangleq \sigma(\boldsymbol{v}^*, \boldsymbol{\lambda})$. Then, there exist $\delta > 0$, $\tilde{\delta} > 0$ and $k < \infty$ such that for every $\gamma \in (C_\epsilon^{\mathrm{tx}} - \tilde{\delta}, C_\epsilon^{\mathrm{tx}} + \tilde{\delta})$

$$\min_{\boldsymbol{v} \in \mathcal{V}_m} u(\boldsymbol{v}, \boldsymbol{\lambda})$$
$$\geq \hat{u}(\boldsymbol{\lambda}) \triangleq \begin{cases} \delta/\sqrt{m}, & \text{if } \mu^*(\boldsymbol{\lambda}) \leq \gamma - \delta \\ \dfrac{\gamma - \mu^*(\boldsymbol{\lambda})}{\sigma^*(\boldsymbol{\lambda}) + k(\gamma - \mu^*(\boldsymbol{\lambda}))}, & \text{if } |\gamma - \mu^*(\boldsymbol{\lambda})| < \delta \\ -\infty, & \text{if } \mu^*(\boldsymbol{\lambda}) \geq \gamma + \delta. \end{cases}$$
$$\quad (187)$$

Proof: See Appendix IV-D. ∎

C. Proof of (163)

We shall need the following lemma, which concerns the speed of convergence of $\mathbb{P}[B > A/\sqrt{n}]$ to $\mathbb{P}[B > 0]$ as $n \to \infty$ for two independent random variables A and B.

Lemma 17: Let A be a real random variable with zero mean and unit variance. Let B be a real random variable independent of A with continuously differentiable pdf f_B. Then

$$\left|\mathbb{P}\left[B \geq \frac{A}{\sqrt{n}}\right] - \mathbb{P}[B \geq 0]\right| \leq \frac{1}{n}\left(\frac{2}{\delta^2} + \frac{k_1}{\delta} + \frac{k_1}{2}\right) \quad (188)$$

where

$$k_1 \triangleq \sup_{t \in (-\delta, \delta)} \max\{|f_B(t)|, |f'_B(t)|\} \quad (189)$$

and $\delta > 0$ is chosen so that k_1 is finite.

Proof: See Appendix IV-E. $\blacksquare$

To establish (163), we lower-bound $\mathbb{E}[q_n(\hat{u}(\mathbf{\Lambda}))]$ on the RHS of (171) using Lemma 17. This entails technical difficulties since the pdf of $\hat{u}(\mathbf{\Lambda})$ is not continuously differentiable due to the fact that the water-filling solution (45) may give rise to different numbers of active eigenmodes for different values of $\mathbf{\lambda}$. To circumvent this problem, we partition $\mathbb{R}^m_{\geq}$ into m non-intersecting subregions $\mathcal{W}_j$, $j = 1, \ldots, m$ [15, Eq. (24)]

$$\mathcal{W}_j \triangleq \left\{\boldsymbol{x} \in \mathbb{R}^m_{\geq} : \frac{1}{x_{j+1}} > \frac{1}{j}\sum_{l=1}^{j}\frac{1}{x_l} + \frac{\rho}{j} \geq \frac{1}{x_j}\right\},$$
$$j = 1, \ldots, m-1 \quad (190)$$

and

$$\mathcal{W}_m \triangleq \left\{\boldsymbol{x} \in \mathbb{R}^m_{\geq} : \frac{1}{m}\sum_{l=1}^{m}\frac{1}{x_l} + \frac{\rho}{m} \geq \frac{1}{x_m}\right\}. \quad (191)$$

In the interior of $\mathcal{W}_j$, $j = 1, \ldots, m$, the pdf of $\hat{u}(\mathbf{\Lambda})$ is continuously differentiable. Note that $\bigcup_{j=1}^{m}\mathcal{W}_j = \mathbb{R}^m_{\geq}$. For every $\mathbf{\lambda} \in \mathcal{W}_j$, the water-filling solution gives exactly j active eigenmodes, i.e.,

$$v_1^*(\mathbf{\lambda}) \geq \cdots \geq v_j^*(\mathbf{\lambda}) > v_{j+1}^*(\mathbf{\lambda}) = \cdots = v_m^*(\mathbf{\lambda}) = 0. \quad (192)$$

Let

$$\mathcal{K}_\delta \triangleq \left\{\mathbf{\lambda} \in \mathbb{R}^m_{\geq} : |\gamma - \mu^*(\mathbf{\lambda})| < \delta\right\}. \quad (193)$$

Using (193) and the sets $\{\mathcal{W}_j\}$, we express $\mathbb{E}[q_n(\hat{u}(\mathbf{\Lambda}))]$ as

$$\mathbb{E}[q_n(\hat{u}(\mathbf{\Lambda}))]$$
$$= \mathbb{E}[q_n(\hat{u}(\mathbf{\Lambda}))\mathbb{1}\{\mathbf{\Lambda} \notin \mathcal{K}_\delta\}]$$
$$+ \sum_{j=1}^{m}\mathbb{E}[q_n(\hat{u}(\mathbf{\Lambda}))\mathbb{1}\{\mathbf{\Lambda} \in \mathcal{K}_\delta \cap \text{Int}(\mathcal{W}_j)\}] \quad (194)$$

where $\text{Int}(\cdot)$ denotes the interior of a given set. To obtain (194), we used that $\mathbf{\Lambda}$ lies in $\bigcup_{j=1}^{m}\text{Int}(\mathcal{W}_j)$ almost surely, which holds because the joint pdf of $\{\Lambda_j\}_{j=1}^{m}$ exists by assumption and because the boundary of $\mathcal{W}_j$ has zero Lebesgue measure.

We next lower-bound the two terms on the RHS of (194) separately. We first consider the first term. When $\mu^*(\mathbf{\lambda}) \geq \gamma + \delta$, we have $\hat{u}(\mathbf{\lambda}) = -\infty$ and $q_n(u_1(\mathbf{\lambda})) = 0$; when $\mu^*(\mathbf{\lambda}) \leq \gamma - \delta$, we have $\hat{u}(\mathbf{\lambda}) = \delta/\sqrt{m}$ and

$$q_n(\hat{u}(\mathbf{\lambda})) = Q\left(-\sqrt{n}\frac{\delta}{\sqrt{m}}\right) - \frac{[1 - n\delta^2/m]^+ e^{-n\delta^2/(2m)}}{6\sqrt{n}}. \quad (195)$$

Assume without loss of generality that $n \geq m/\delta^2$ (recall that we are interested in the asymptotic regime $n \to \infty$). In this case, the second term on the RHS of (195) is zero. Hence,

$$\mathbb{E}[q_n(\hat{u}(\mathbf{\Lambda}))\mathbb{1}\{\mathbf{\Lambda} \notin \mathcal{K}_\delta\}]$$
$$= Q\left(-\sqrt{n}\frac{\delta}{\sqrt{m}}\right)\mathbb{P}[\mu^*(\mathbf{\Lambda}) \leq \gamma - \delta] \quad (196)$$
$$\geq \mathbb{P}[\mu^*(\mathbf{\Lambda}) \leq \gamma - \delta] - e^{-n\delta^2/(2m)}. \quad (197)$$

Here, (197) follows because $Q(-t) \geq 1 - e^{-t^2/2}$ for all $t \geq 0$ and because $\mathbb{P}[\mu^*(\mathbf{\Lambda}) \leq \gamma - \delta] \leq 1$.

We next lower-bound the second term on the RHS of (194). If $\mathbb{P}[\mathbf{\Lambda} \in \mathcal{K}_\delta \cap \text{Int}(\mathcal{W}_j)] = 0$, we have

$$\mathbb{E}[q_n(\hat{u}(\mathbf{\Lambda}))\mathbb{1}\{\mathbf{\Lambda} \in \mathcal{K}_\delta \cap \text{Int}(\mathcal{W}_j)\}] = 0 \quad (198)$$

since $q_n(\cdot)$ is bounded. We thus assume in the following that $\mathbb{P}[\mathbf{\Lambda} \in \mathcal{K}_\delta \cap \text{Int}(\mathcal{W}_j)] > 0$. Let $\hat{U}$ denote the random variable $\hat{u}(\mathbf{\Lambda})$. To emphasize that $\hat{U}$ depends on γ (see (187)), we write $\hat{U}(\gamma)$ in place of $\hat{U}$ whenever necessary. Using this definition and (170), we obtain

$$\mathbb{E}\left[q_n(\hat{U})\mathbb{1}\{\mathbf{\Lambda} \in \mathcal{K}_\delta \cap \text{Int}(\mathcal{W}_j)\}\right]$$
$$= \left(\mathbb{E}\left[Q(-\sqrt{n}\hat{U})\,|\,\mathbf{\Lambda} \in \mathcal{K}_\delta \cap \text{Int}(\mathcal{W}_j)\right]\right.$$
$$\left. - \frac{1}{6\sqrt{n}}\mathbb{E}\left[[1 - n\hat{U}^2]^+ e^{-n\hat{U}^2/2}\Big|\mathbf{\Lambda} \in \mathcal{K}_\delta \cap \text{Int}(\mathcal{W}_j)\right]\right)$$
$$\times \mathbb{P}[\mathbf{\Lambda} \in \mathcal{K}_\delta \cap \text{Int}(\mathcal{W}_j)]. \quad (199)$$

Observe that the transformation

$$(\lambda_1, \ldots, \lambda_j, \gamma) \mapsto (\hat{u}(\mathbf{\lambda}), \lambda_2, \ldots, \lambda_j, \gamma) \quad (200)$$

is one-to-one and twice continuously differentiable with nonsingular Jacobian for $\mathbf{\lambda} \in \mathcal{K}_\delta \cap \text{Int}(\mathcal{W}_j)$, i.e., it is a diffeomorphism of class C^2 [29, p. 147]. Consequently, the conditional pdf $f_{\hat{U}(\gamma)\,|\,\mathbf{\Lambda}\in\mathcal{K}_\delta\cap\text{Int}(\mathcal{W}_j)}(t)$ of $\hat{U}(\gamma)$ given $\mathbf{\Lambda} \in \mathcal{K}_\delta \cap \text{Int}(\mathcal{W}_j)$ as well as its first derivative are jointly continuous functions of γ and t. Hence, they are bounded on bounded sets. It thus follows that for every $j \in \{1, \ldots, m\}$, every $\gamma \in (C_\epsilon^{\text{tx}} - \tilde{\delta}, C_\epsilon^{\text{tx}} + \tilde{\delta})$ (where $\tilde{\delta}$ is given by Lemma 16), and every $\tilde{\delta}_1 > 0$, there exists a $k_2 < \infty$ such that the conditional pdf $f_{\hat{U}(\gamma)\,|\,\mathbf{\Lambda}\in\mathcal{K}_\delta\cap\text{Int}(\mathcal{W}_j)}$ and its derivative satisfy

$$\sup_{t \in [-\tilde{\delta}_1, \tilde{\delta}_1]} \sup_{\gamma \in (C_\epsilon^{\text{tx}}-\tilde{\delta}, C_\epsilon^{\text{tx}}+\tilde{\delta})} \left|f_{\hat{U}(\gamma)\,|\,\mathbf{\Lambda}\in\mathcal{K}_\delta\cap\text{Int}(\mathcal{W}_j)}(t)\right| \leq k_2 \quad (201)$$

$$\sup_{t \in [-\tilde{\delta}_1, \tilde{\delta}_1]} \sup_{\gamma \in (C_\epsilon^{\text{tx}}-\tilde{\delta}, C_\epsilon^{\text{tx}}+\tilde{\delta})} \left|f'_{\hat{U}(\gamma)\,|\,\mathbf{\Lambda}\in\mathcal{K}_\delta\cap\text{Int}(\mathcal{W}_j)}(t)\right| \leq k_2. \quad (202)$$

We next apply Lemma 17 with A being a standard normal random variable and B being the random variable $\hat{U}$ conditioned on $\mathbf{\Lambda} \in \mathcal{K}_\delta \cap \text{Int}(\mathcal{W}_j)$. This implies that there exists a finite constant k_3 independent of γ and n such that the first term on the RHS of (199) satisfies

$$\mathbb{E}\left[Q(-\sqrt{n}\hat{U}(\gamma))|\mathbf{\Lambda} \in \mathcal{K}_\delta \cap \text{Int}(\mathcal{W}_j)\right]$$
$$\geq \mathbb{P}[\mu^*(\mathbf{\Lambda}) \leq \gamma\,|\,\mathbf{\Lambda} \in \mathcal{K}_\delta \cap \text{Int}(\mathcal{W}_j)] + \frac{k_3}{n}. \quad (203)$$

We next bound the second term on the RHS of (199) for $n \geq \tilde{\delta}_1^{-2}$ as

$$\frac{1}{6\sqrt{n}}\mathbb{E}\left[\left[1 - n\hat{U}^2\right]^+ e^{-n\hat{U}^2/2}\Big|\boldsymbol{\Lambda} \in \mathcal{K}_\delta \cap \text{Int}(\mathcal{W}_j)\right]$$

$$\leq \frac{k_2}{6\sqrt{n}}\int_{-1/\sqrt{n}}^{1/\sqrt{n}}(1 - nt^2)e^{-nt^2/2}dt \qquad (204)$$

$$= \frac{k_2}{3\sqrt{en}} \qquad (205)$$

where (204) follows from (201). Substituting (203) and (205) into (199) we obtain

$$\mathbb{E}\left[q_n(\hat{U})\mathbb{1}\{\boldsymbol{\Lambda} \in \mathcal{K}_\delta \cap \text{Int}(\mathcal{W}_j)\}\right]$$

$$\geq \mathbb{P}\left[\mu^*(\boldsymbol{\Lambda}) \leq \gamma, \boldsymbol{\Lambda} \in \mathcal{K}_\delta \cap \text{Int}(\mathcal{W}_j)\right] + \frac{k_4}{n} \qquad (206)$$

for some finite k_4 independent of γ and n. Using (197), (198) and (206) in (194), and substituting (194) into (171), we conclude that

$$\mathbb{P}[S_n^{\text{rt}}(\boldsymbol{v}, \boldsymbol{\Lambda}) \leq n\gamma] \geq \mathbb{P}[\mu^*(\boldsymbol{\Lambda}) \leq \gamma] + \mathcal{O}\left(\frac{1}{n}\right) \qquad (207)$$

$$= F_{\text{tx}}(\gamma) + \mathcal{O}\left(\frac{1}{n}\right) \qquad (208)$$

where the $\mathcal{O}(1/n)$ term is uniform in $\gamma \in (C_\epsilon^{\text{tx}} - \tilde{\delta}, C_\epsilon^{\text{tx}} + \tilde{\delta})$. Here, the last step follows from (166) and (20).

D. Proof of Lemma 16

For an arbitrary $\boldsymbol{\lambda} \in \mathbb{R}_{\geq}^m$, the function $\mu(\boldsymbol{v}, \boldsymbol{\lambda})$ in the numerator of (168) is maximized by the (unique) water-filling power allocation $v_j = v_j^*$ defined in (45):

$$\mu^*(\boldsymbol{\lambda}) = \max_{\boldsymbol{v} \in \mathcal{V}_m} \mu(\boldsymbol{v}, \boldsymbol{\lambda}) = \mu(\boldsymbol{v}^*, \boldsymbol{\lambda}). \qquad (209)$$

The function $\sigma(\boldsymbol{v}, \boldsymbol{\lambda})$ on the denominator of (168) can be bounded as

$$0 \leq \sigma(\boldsymbol{v}, \boldsymbol{\lambda}) \leq \sqrt{m}. \qquad (210)$$

Using (209) and (210) we obtain that for an arbitrary $\delta > 0$

$$\min_{\boldsymbol{v} \in \mathcal{V}_m} u(\boldsymbol{v}, \boldsymbol{\lambda}) \geq \begin{cases} \delta/\sqrt{m}, & \mu^*(\boldsymbol{\lambda}) \leq \gamma - \delta \\ -\infty, & \mu^*(\boldsymbol{\lambda}) \geq \gamma + \delta. \end{cases} \qquad (211)$$

Let $\boldsymbol{v}_{\min}$ be the minimizer of $u(\boldsymbol{v}, \boldsymbol{\lambda})$ for a given $\boldsymbol{\lambda}$. To prove Lemma 16, it remains to show that there exist $\delta > 0$, $\tilde{\delta} > 0$ and $k < \infty$ such that for every $\gamma \in (C_\epsilon^{\text{tx}} - \tilde{\delta}, C_\epsilon^{\text{tx}} + \tilde{\delta})$ and every $\boldsymbol{\lambda} \in \mathbb{R}_{\geq}^m$ satisfying $|\mu^*(\boldsymbol{\lambda}) - \gamma| < \delta$,

$$\min_{\boldsymbol{v} \in \mathcal{V}_m} u(\boldsymbol{v}, \boldsymbol{\lambda}) = u(\boldsymbol{v}_{\min}, \boldsymbol{\lambda}) \qquad (212)$$

$$\geq \frac{\gamma - \mu^*(\boldsymbol{\lambda})}{\sigma^*(\boldsymbol{\lambda}) + k(\gamma - \mu^*(\boldsymbol{\lambda}))}. \qquad (213)$$

Since

$$u(\boldsymbol{v}_{\min}, \boldsymbol{\lambda}) = \frac{\gamma - \mu(\boldsymbol{v}_{\min}, \boldsymbol{\lambda})}{\sigma(\boldsymbol{v}_{\min}, \boldsymbol{\lambda})} \geq \frac{\gamma - \mu^*(\boldsymbol{\lambda})}{\sigma(\boldsymbol{v}_{\min}, \boldsymbol{\lambda})} \qquad (214)$$

it suffices to show that for every $\gamma \in (C_\epsilon^{\text{tx}} - \tilde{\delta}, C_\epsilon^{\text{tx}} + \tilde{\delta})$ and every $\boldsymbol{\lambda} \in \mathbb{R}_{\geq}^m$ satisfying $|\mu^*(\boldsymbol{\lambda}) - \gamma| < \delta$, we have

$$|\sigma(\boldsymbol{v}_{\min}, \boldsymbol{\lambda}) - \sigma^*(\boldsymbol{\lambda})| \leq k|\gamma - \mu^*(\boldsymbol{\lambda})| \qquad (215)$$

and that

$$\sigma^*(\boldsymbol{\lambda}) - k|\gamma - \mu^*(\boldsymbol{\lambda})| > 0. \qquad (216)$$

The desired bound (213) follows then by lower-bounding $\sigma(\boldsymbol{v}_{\min}, \boldsymbol{\lambda})$ in (214) by $\sigma^*(\boldsymbol{\lambda}) - k|\gamma - \mu^*(\boldsymbol{\lambda})|$ when $\mu^*(\boldsymbol{\lambda}) \geq \gamma$ and by upper-bounding $\sigma(\boldsymbol{v}_{\min}, \boldsymbol{\lambda})$ by $\sigma^*(\boldsymbol{\lambda}) + k|\gamma - \mu^*(\boldsymbol{\lambda})|$ when $\mu^*(\boldsymbol{\lambda}) < \gamma$.

We first establish (215). By the mean value theorem, there exist v_j' between v_j^* and $v_{\min,j}$, $j = 1, \ldots, m$, such that

$$\left|\sigma(\boldsymbol{v}_{\min}, \boldsymbol{\lambda}) - \sigma^*(\boldsymbol{\lambda})\right|$$

$$= \left|\sum_{j=1}^m \frac{2\lambda_j}{(1 + \lambda_j v_j')^3}(v_{\min,j} - v_j^*)\right| \qquad (217)$$

$$\leq \sum_{j=1}^m \frac{2\lambda_j}{(1 + \lambda_j v_j')^3}\left|v_{\min,j} - v_j^*\right| \qquad (218)$$

$$\leq 2\lambda_1 \sum_{j=1}^m \left|v_{\min,j} - v_j^*\right| \qquad (219)$$

$$\leq 2\lambda_1\sqrt{m}\|\boldsymbol{v}_{\min} - \boldsymbol{v}^*\|. \qquad (220)$$

Here, the last step follows because for every $\boldsymbol{a} = [a_1, \ldots, a_m] \in \mathbb{R}^m$, we have $\sum_{j=1}^m |a_j| \leq \sqrt{m}\|\boldsymbol{a}\|$.

Next, we upper-bound λ_1 and $\|\boldsymbol{v}_{\min} - \boldsymbol{v}^*\|$ separately. The variable λ_1 can be bounded as follows. Because the water-filling power levels $\{v_l^*\}$ in (45) are nonincreasing, we have that

$$\frac{\rho}{m} \leq v_1^* \leq \rho. \qquad (221)$$

Choose $\delta_1 > 0$ and $\tilde{\delta} > 0$ such that $\delta_1 + \tilde{\delta} < C_\epsilon^{\text{tx}}$. Using (221) together with

$$\log(1 + \lambda_1 v_1^*) \leq \mu^*(\boldsymbol{\lambda}) \leq m\log(1 + \lambda_1 v_1^*) \qquad (222)$$

and the assumption that $\gamma \in (C_\epsilon^{\text{tx}} - \tilde{\delta}, C_\epsilon^{\text{tx}} + \tilde{\delta})$, we obtain that whenever $|\mu^*(\boldsymbol{\lambda}) - \gamma| < \delta_1$

$$k_0 \triangleq \frac{1}{\rho}\left(e^{(C_\epsilon^{\text{tx}} - \delta_1 - \tilde{\delta})/m} - 1\right)$$

$$\leq \lambda_1 \leq \frac{m}{\rho}\left(e^{C_\epsilon^{\text{tx}} + \delta_1 + \tilde{\delta}} - 1\right) \triangleq k_1. \qquad (223)$$

The term $\|\boldsymbol{v}_{\min} - \boldsymbol{v}^*\|$ can be upper-bounded as follows. Since $\boldsymbol{v}_{\min}$ is the minimizer of $u(\boldsymbol{v}, \boldsymbol{\lambda})$, it must satisfy the Karush–Kuhn–Tucker (KKT) conditions [41, Sec. 5.5.3]:

$$-\frac{\partial u(\boldsymbol{v}, \boldsymbol{\lambda})}{\partial v_l}\bigg|_{v_l = v_{\min,l}} = \eta, \quad \forall l \text{ for which } v_{\min,l} > 0 \qquad (224)$$

$$-\frac{\partial u(\boldsymbol{v}, \boldsymbol{\lambda})}{\partial v_l}\bigg|_{v_l = v_{\min,l}} \leq \eta, \quad \forall l \text{ for which } v_{\min,l} = 0 \qquad (225)$$

for some η. The derivatives in (224) and (225) are given by

$$-\frac{\partial u(\boldsymbol{v}, \boldsymbol{\lambda})}{\partial v_l}\bigg|_{v_l = v_{\min,l}} = \left(1 + \frac{\gamma - \mu(\boldsymbol{v}_{\min}, \boldsymbol{\lambda})}{(1 + \lambda_l v_{\min,l})^2\sigma^2(\boldsymbol{v}_{\min}, \boldsymbol{\lambda})}\right)$$

$$\times \frac{1}{(v_{\min,l} + 1/\lambda_l)\sigma(\boldsymbol{v}_{\min}, \boldsymbol{\lambda})}. \qquad (226)$$

Let $\tilde{\eta} \triangleq 1/(\sigma(\boldsymbol{v}_{\min}, \boldsymbol{\lambda})\eta)$. Then, (224) and (225) can be rewritten as

$$v_{\min,l} = \left[\tilde{\eta}\left(1 + \frac{\gamma - \mu(\boldsymbol{v}_{\min}, \boldsymbol{\lambda})}{(1 + \lambda_l v_{\min,l})^2\sigma^2(\boldsymbol{v}_{\min}, \boldsymbol{\lambda})}\right) - \frac{1}{\lambda_l}\right]^+ \qquad (227)$$

where $\tilde{\eta}$ satisfies

$$\sum_{l=1}^{m}\left[\tilde{\eta}\left(1+\frac{\gamma-\mu(\boldsymbol{v}_{\min},\boldsymbol{\lambda})}{(1+\lambda_l v_{\min,l})^2\sigma^2(\boldsymbol{v}_{\min},\boldsymbol{\lambda})}\right)-\frac{1}{\lambda_l}\right]^+=\rho.\quad(228)$$

Here, the equality in (228) follows because $u(\boldsymbol{v},\boldsymbol{\lambda})$ is monotonically decreasing in v_j, which implies that the minimizer $\boldsymbol{v}_{\min}$ of $u(\boldsymbol{v},\boldsymbol{\lambda})$ must satisfy $\sum_{l=1}^{m}v_{\min,l}=\rho$. Comparing (227) and (228) with (45) and (22), we obtain, after algebraic manipulations

$$\|\boldsymbol{v}_{\min}-\boldsymbol{v}^*\|\le k_2|\gamma-\mu(\boldsymbol{v}_{\min},\boldsymbol{\lambda})|\qquad(229)$$

for some $k_2<\infty$ that does not depend on $\boldsymbol{\lambda}$, $\boldsymbol{v}_{\min}$, $\boldsymbol{v}^*$ and γ.

To further upper-bound the RHS of (229), recall that $\boldsymbol{v}_{\min}$ minimizes $u(\boldsymbol{v},\boldsymbol{\lambda})=(\gamma-\mu(\boldsymbol{v},\boldsymbol{\lambda}))/\sigma(\boldsymbol{v},\boldsymbol{\lambda})$ for a given $\boldsymbol{\lambda}$ and that $\mu^*(\boldsymbol{\lambda})=\max_{\boldsymbol{v}\in\mathcal{V}_m}\mu(\boldsymbol{v},\boldsymbol{\lambda})$. Thus, if $\mu^*(\boldsymbol{\lambda})\ge\gamma$ then we must have $u(\boldsymbol{v}_{\min},\boldsymbol{\lambda})\le u(\boldsymbol{v}^*,\boldsymbol{\lambda})\le 0$, which implies that

$$0\le\mu(\boldsymbol{v}_{\min},\boldsymbol{\lambda})-\gamma\le\mu^*(\boldsymbol{\lambda})-\gamma.\qquad(230)$$

If $\mu^*(\boldsymbol{\lambda})<\gamma$ then

$$0\le\frac{\gamma-\mu(\boldsymbol{v}_{\min},\boldsymbol{\lambda})}{\sqrt{m}}\le u(\boldsymbol{v}_{\min},\boldsymbol{\lambda})\le\frac{\gamma-\mu^*(\boldsymbol{\lambda})}{\sigma^*(\boldsymbol{\lambda})}\qquad(231)$$

where in the second inequality we used that $\sigma(\boldsymbol{v}_{\min},\boldsymbol{\lambda})\le\sqrt{m}$ (see (210)). Using (221) and (223), we can lower-bound $\sigma^*(\boldsymbol{\lambda})$ as

$$\sigma^*(\boldsymbol{\lambda})\ge\sqrt{1-\frac{1}{(1+\lambda_1 v_1^*)^2}}\qquad(232)$$

$$\ge\sqrt{1-\frac{1}{(1+\rho k_0/m)^2}}\triangleq k_3.\qquad(233)$$

Substituting (233) into (231), we obtain

$$0\le\gamma-\mu(\boldsymbol{v}_{\min},\boldsymbol{\lambda})\le\frac{\sqrt{m}}{k_3}\big[\gamma-\mu^*(\boldsymbol{\lambda})\big].\qquad(234)$$

Combining (234) with (230) and using that $\sqrt{m}/k_3>1$, we get

$$\big|\gamma-\mu(\boldsymbol{v}_{\min},\boldsymbol{\lambda})\big|\le\frac{\sqrt{m}}{k_3}\big|\gamma-\mu^*(\boldsymbol{\lambda})\big|.\qquad(235)$$

Finally, substituting (235) into (229), then (229) and (223) into (220), and writing $k\triangleq k_1 k_2\sqrt{m}/k_3$, we conclude that (215) holds for every $\gamma\in(C_\epsilon^{\mathrm{tx}}-\tilde{\delta},C_\epsilon^{\mathrm{tx}}+\tilde{\delta})$ and every $\boldsymbol{\lambda}$ satisfying $|\mu^*(\boldsymbol{\lambda})-\gamma|<\delta_1$.

To prove (216), we choose $0<\delta<\min\{\delta_1,k_3/k\}$. It then follows that for every $\boldsymbol{\lambda}$ satisfying $|\mu^*(\boldsymbol{\lambda})-\gamma|<\delta$ we have

$$\sigma^*(\boldsymbol{\lambda})-k|\gamma-\mu^*(\boldsymbol{\lambda})|\ge k_3-k\delta>0.\qquad(236)$$

Here, in (236) we used the bound (233). This concludes the proof.

E. Proof of Lemma 17

By assumption, there exist $\delta>0$ and $k_1<\infty$ such that

$$\sup_{t\in(-\delta,\delta)}\max\{|f_B(t)|,|f_B'(t)|\}\le k_1.\qquad(237)$$

Let F_A and F_B be the cdfs of A and B, respectively. We rewrite $\mathbb{P}[B\ge A/\sqrt{n}]$ as follows:

$$\mathbb{P}[B\ge A/\sqrt{n}]=\underbrace{\int_{|a|\ge\delta\sqrt{n}}\mathbb{P}[B\ge a/\sqrt{n}]dF_A(a)}_{\triangleq c_0(n)}$$
$$+\int_{|a|<\delta\sqrt{n}}\underbrace{\mathbb{P}[B\ge a/\sqrt{n}]}_{=1-F_B(a/\sqrt{n})}dF_A(a).\quad(238)$$

We next expand the argument of the second integral in (238) by applying Taylor's theorem [40, Th. 5.15] on $F_B(a/\sqrt{n})$ as follows: for all $a\in(-\delta\sqrt{n},\delta\sqrt{n})$

$$1-F_B(a/\sqrt{n})=1-F_B(0)-f_B(0)\frac{a}{\sqrt{n}}-\frac{f_B'(a_0)}{2}\frac{a^2}{n}\quad(239)$$

for some $a_0\in(0,a/\sqrt{n})$. Averaging over A, we get

$$\int_{|a|<\delta\sqrt{n}}1-F_B(a/\sqrt{n})dF_A(a)$$
$$=\underbrace{(1-F_B(0))}_{=\mathbb{P}[B\ge 0]}\mathbb{P}[|A|<\delta\sqrt{n}]$$
$$-\frac{f_B(0)}{\sqrt{n}}\underbrace{\mathbb{E}\big[A\cdot\mathbb{1}\{|A|<\delta\sqrt{n}\}\big]}_{\triangleq c_1(n)}$$
$$-\underbrace{\mathbb{E}\left[\frac{A^2 f_B'(A_0)}{2n}\cdot\mathbb{1}\{|A|<\delta\sqrt{n}\}\right]}_{\triangleq c_2(n)}.\qquad(240)$$

Hence,

$$\big|\mathbb{P}[B\ge A/\sqrt{n}]-\mathbb{P}[B\ge 0]\big|\qquad(241)$$
$$=\bigg|c_0(n)-\mathbb{P}[B\ge 0]\cdot\mathbb{P}[|A|\ge\delta\sqrt{n}]$$
$$-\frac{f_B(0)}{\sqrt{n}}c_1(n)-c_2(n)\bigg|\qquad(242)$$
$$\le c_0(n)+\mathbb{P}[|A|\ge\delta\sqrt{n}]+\frac{k_1}{\sqrt{n}}|c_1(n)|+|c_2(n)|\quad(243)$$
$$\le 2\mathbb{P}[|A|\ge\delta\sqrt{n}]+\frac{k_1}{\sqrt{n}}|c_1(n)|+|c_2(n)|\qquad(244)$$
$$\le\frac{2}{\delta^2 n}+\frac{k_1}{\sqrt{n}}|c_1(n)|+|c_2(n)|.\qquad(245)$$

Here, in (243) we used the triangle inequality together with (237) and the trivial bound $\mathbb{P}[B\ge 0]\le 1$; (244) follows because $c_0(n)\le\mathbb{P}[|A|\ge\delta\sqrt{n}]$; (245) follows from Chebyshev's inequality and because $\mathbb{E}\big[A^2\big]=1$ by assumption.

To conclude the proof, we next upper-bound $|c_1(n)|$, and $|c_2(n)|$. The term $|c_1(n)|$ can be bounded as

$$|c_1(n)|=\big|\mathbb{E}\big[A\cdot\mathbb{1}\{|A|\ge\delta\sqrt{n}\}\big]\big|\qquad(246)$$
$$\le\frac{1}{\delta\sqrt{n}}\mathbb{E}\big[\delta\sqrt{n}|A|\cdot\mathbb{1}\{|A|\ge\delta\sqrt{n}\}\big]\quad(247)$$
$$\le\frac{1}{\delta\sqrt{n}}\mathbb{E}\big[A^2\cdot\mathbb{1}\{|A|\ge\delta\sqrt{n}\}\big]\qquad(248)$$
$$\le\frac{1}{\delta\sqrt{n}}\qquad(249)$$

where (246) follows because $\mathbb{E}[A]=0$ by assumption.

Finally, $|c_2(n)|$ can be bounded as

$$|c_2(n)| \leq \mathbb{E}\left[\frac{A^2 |f'_B(A_0)|}{2n} \cdot \mathbb{1}\{|A| < \delta\sqrt{n}\}\right] \qquad (250)$$

$$\leq \mathbb{E}\left[A^2 \cdot \mathbb{1}\{|A| < \delta\sqrt{n}\}\right]\frac{k_1}{2n} \qquad (251)$$

$$\leq \frac{k_1}{2n}. \qquad (252)$$

Here, (251) follows because the support of A_0 is contained in $(0, \delta)$ and from (237). Substituting (249) and (252) into (245), we obtain the desired inequality (188).

APPENDIX V
PROOF OF THE ACHIEVABILITY PART OF THEOREM 3

In order to prove (57), we study the achievability bound (43) in the large-n limit. We start by analyzing the denominator on the RHS of (43). Let $\alpha = n - t - r > 0$. Then,

$$\mathbb{P}\left[\prod_{i=1}^{r} B_i \leq \gamma_n\right] = \mathbb{P}\left[\prod_{i=1}^{r} B_i^{-\alpha} \geq \gamma_n^{-\alpha}\right] \qquad (253)$$

$$\leq \frac{\mathbb{E}\left[\prod_{i=1}^{r} B_i^{-\alpha}\right]}{\gamma_n^{-\alpha}} \qquad (254)$$

$$= \gamma_n^{n-t-r} \prod_{i=1}^{r} \mathbb{E}\left[B_i^{-(n-t-r)}\right] \qquad (255)$$

where (254) follows from Markov's inequality, and (255) follows because the $B_1, \ldots, B_r$ are independent. Recalling that $B_i \sim \text{Beta}(n - t - i + 1, t)$, we obtain that for every $i \in \{1, \ldots, r\}$

$$\mathbb{E}\left[B_i^{-(n-t-r)}\right]$$

$$= \frac{\Gamma(n-i+1)}{\Gamma(n-t-i+1)\Gamma(t)}\int_0^1 s^{r-i}(1-s)^{t-1}ds \qquad (256)$$

$$\leq \frac{\Gamma(n-i+1)}{\Gamma(n-t-i+1)\Gamma(t)} \qquad (257)$$

$$\leq n^t. \qquad (258)$$

Substituting (258) into (255), we get

$$\mathbb{P}\left[\prod_{i=1}^{r} B_i \leq \gamma_n\right] \leq n^{rt}\gamma_n^{n-t-r}. \qquad (259)$$

Setting $\tau = 1/n$ and $\gamma_n = \exp(-C_\epsilon^{\text{tx}} + \mathcal{O}(1/n))$ in (43), and using (259), we obtain

$$\frac{\log M}{n} \geq C_\epsilon^{\text{tx}} - (1 + rt)\frac{\log n}{n} + \mathcal{O}\left(\frac{1}{n}\right). \qquad (260)$$

To conclude the proof, it remains to show that there exists a $\gamma_n = \exp(-C_\epsilon^{\text{tx}} + \mathcal{O}(1/n))$ satisfying (44). To this end, we note that

$$\mathbb{P}\left[\sin^2\left\{\mathsf{I}_{n,t}, \sqrt{n}\mathsf{I}_{n,t}\text{diag}\left\{\sqrt{v_1^*\Lambda_1}, \ldots, \sqrt{v_m^*\Lambda_m}, \right.\right.\right.$$
$$\left.\left.\left. \underbrace{0, \ldots, 0}_{t-m}\right\} + \mathbb{W}\right\} \leq \gamma_n\right]$$

$$\geq \mathbb{P}\left[\prod_{j=1}^{m}\sin^2\left\{\boldsymbol{e}_j, \sqrt{nv_j^*\Lambda_j}\boldsymbol{e}_j + \boldsymbol{W}_j\right\} \leq \gamma_n\right] \qquad (261)$$

$$= \mathbb{P}\left[\prod_{j=1}^{m}\sin^2\left\{\boldsymbol{e}_1, \sqrt{nv_j^*\Lambda_j}\boldsymbol{e}_1 + \boldsymbol{W}_j\right\} \leq \gamma_n\right]. \qquad (262)$$

Here, (261) follows from Lemma 13 (Appendix I) by letting $\boldsymbol{e}_j$ and $\boldsymbol{W}_j$ stand for the jth column of $\mathsf{I}_{n,t}$ and $\mathbb{W}$, respectively; (262) follows by symmetry. We next note that by (98), the random variable $\sin^2\{\boldsymbol{e}_1, \sqrt{nv_j^*\Lambda_j}\boldsymbol{e}_1 + \boldsymbol{W}_j\}$ has the same distribution as

$$T_j \triangleq \frac{\sum_{i=2}^{n}|W_{i,j}|^2}{|\sqrt{nv_j^*\Lambda_j} + W_{1,j}|^2 + \sum_{i=2}^{n}|W_{i,j}|^2}. \qquad (263)$$

Thus,

$$\mathbb{P}\left[\prod_{j=1}^{m}\sin^2\left\{\boldsymbol{e}_1, \sqrt{nv_j^*\Lambda_j}\boldsymbol{e}_1 + \boldsymbol{W}_j\right\} \leq \gamma_n\right] = \mathbb{P}\left[\prod_{j=1}^{m}T_j \leq \gamma_n\right]. \qquad (264)$$

To evaluate the RHS of (264), we observe that by the law of large numbers, the noise term $\frac{1}{n}\sum_{i=2}^{n}|W_{i,j}|^2$ in (263) concentrates around 1 as $n \to \infty$. Hence, we expect that for all $\gamma > 0$

$$\mathbb{P}\left[\prod_{j=1}^{m}T_j \leq \gamma\right] \to \mathbb{P}\left[\prod_{j=1}^{m}\frac{1}{v_j^*\Lambda_j + 1} \leq \gamma\right] \quad \text{as } n \to \infty. \qquad (265)$$

We shall next make this statement rigorous by showing that, for all γ in a certain neighborhood of $e^{-C_\epsilon^{\text{tx}}}$,

$$\mathbb{P}\left[\prod_{j=1}^{m}T_j \leq \gamma\right] \geq \mathbb{P}\left[\prod_{j=1}^{m}\frac{1}{v_j^*\Lambda_j + 1} \leq \gamma\right] + \mathcal{O}\left(\frac{1}{n}\right) \qquad (266)$$

where the term $\mathcal{O}(1/n)$ is uniform in γ. To this end, we build on Lemma 17 in Appendix IV-C. The technical difficulty is that the joint pdf of $\Lambda_1 v_1^*, \ldots, \Lambda_m v_m^*$ is not continuously differentiable because the functions $\{v_j^*(\cdot)\}$ are not differentiable on the boundary of the nonintersecting regions $\mathcal{W}_1, \ldots, \mathcal{W}_m$ defined in (190) and (191). To circumvent this problem, we study the asymptotic behavior of $\{T_j\}$ conditioned on $\boldsymbol{\Lambda} \in \text{Int}(\mathcal{W}_u)$, in which case the joint pdf of $\Lambda_j v_j^*(\boldsymbol{\Lambda})$, $j = 1, \ldots, m$, is continuously differentiable. This comes without loss of generality since $\boldsymbol{\Lambda}$ lies in $\bigcup_{u=1}^{m}\text{Int}(\mathcal{W}_u)$ almost surely (see also Appendix IV-C).

To simplify notation, we use $T_j^{(u)}$ to denote the random variable T_j conditioned on the event $\boldsymbol{\Lambda} \in \text{Int}(\mathcal{W}_u)$, $u = 1, \ldots, m$. We further denote by $\boldsymbol{\Lambda}^{(u)}$ and $\widetilde{\boldsymbol{\Lambda}}^{(u)}$ the random vectors $[\Lambda_1, \ldots, \Lambda_u]^{\text{T}}$ and $[\Lambda_1 v_1^*(\boldsymbol{\Lambda}), \ldots, \Lambda_u v_u^*(\boldsymbol{\Lambda})]^{\text{T}}$ conditioned on the event $\boldsymbol{\Lambda} \in \text{Int}(\mathcal{W}_u)$, respectively. Using these definitions, the LHS of (266) can be rewritten as

$$\mathbb{P}\left[\prod_{j=1}^{m}T_j \leq \gamma\right]$$

$$= \sum_{u=1}^{m}\left\{\mathbb{P}\left[\prod_{j=1}^{m}T_j \leq \gamma\Big|\boldsymbol{\Lambda} \in \text{Int}(\mathcal{W}_u)\right]\mathbb{P}[\boldsymbol{\Lambda} \in \text{Int}(\mathcal{W}_u)]\right\} \qquad (267)$$

$$= \sum_{u=1}^{m}\left\{\mathbb{P}\left[\left(\prod_{j=1}^{u}T_j^{(u)}\right)\cdot\underbrace{\left(\prod_{j=u+1}^{m}\frac{\sum_{i=2}^{n}|W_{i,j}|^2}{\sum_{i=1}^{n}|W_{i,j}|^2}\right)}_{\leq 1} \leq \gamma\right]\right.$$

$$\left. \times \mathbb{P}[\boldsymbol{\Lambda} \in \text{Int}(\mathcal{W}_u)]\right\} \qquad (268)$$

$$\geq \sum_{u=1}^{m} \left\{ \mathbb{P}\left[\prod_{j=1}^{u} T_j^{(u)} \leq \gamma \right] \mathbb{P}[\mathbf{\Lambda} \in \mathrm{Int}(\mathcal{W}_u)] \right\}. \qquad (269)$$

Here, (268) follows because, by (192), $T_j = (\sum_{i=2}^{n} |W_{i,j}|^2)$ $/(\sum_{i=1}^{n} |W_{i,j}|^2)$ for $j = u+1, \dots, m$.

The following lemma, built upon Lemma 17, allows us to establish (266).

Lemma 18: Let $\mathbf{G} = [G_1, \dots, G_u]^{\mathrm{T}} \in \mathbb{R}^{u}_{\geq}$ be a random vector with continuously differentiable joint pdf. Let

$$D_j \triangleq \frac{\sum_{i=2}^{n} |W_{i,j}|^2}{|\sqrt{nG_j} + W_{1,j}|^2 + \sum_{i=2}^{n} |W_{i,j}|^2}, \; j = 1, \dots, u \quad (270)$$

where $W_{i,j}$, $i = 1, \dots, n$, $j = 1, \dots, u$, are i.i.d. $\mathcal{CN}(0,1)$-distributed. Fix an arbitrary $\xi_0 \in (0,1)$. Then, there exist a $\delta > 0$ and a finite constant k such that

$$\inf_{\xi \in (\xi_0 - \delta, \xi_0 + \delta)} \left(\mathbb{P}\left[\prod_{j=1}^{u} D_j \leq \xi \right] - \mathbb{P}\left[\prod_{j=1}^{u} \frac{1}{1 + G_j} \leq \xi \right] \right) > \frac{k}{n}. \qquad (271)$$

Proof: See Appendix V-A. ∎

Using Lemma 18 with $G_j = \widetilde{\Lambda}_j^{(u)}$, it follows that there exist $\delta_u > 0$ and $0 \leq k_u < \infty$, such that for every $\gamma \in \left(e^{-C_\epsilon^{\mathrm{tx}} - \delta_u}, e^{-C_\epsilon^{\mathrm{tx}} + \delta_u} \right)$

$$\mathbb{P}\left[\prod_{j=1}^{u} T_j^{(u)} \leq \gamma \right] \geq \mathbb{P}\left[\prod_{j=1}^{u} \frac{1}{1 + \widetilde{\Lambda}_j^{(u)}} \leq \gamma \right] + \mathcal{O}\left(\frac{1}{n} \right). \quad (272)$$

To show that $\widetilde{\Lambda}_j^{(u)}$, $j = 1, \dots, u$, indeed satisfy the conditions in Lemma (18), we use (192), (45), and (22), to obtain

$$\widetilde{\Lambda}_j^{(u)} = \frac{\Lambda_j^{(u)}}{u} \left(\rho + \sum_{l=1}^{u} \frac{1}{\Lambda_l^{(u)}} \right) - 1, \quad j = 1, \dots, u. \quad (273)$$

Since the joint pdf of $\mathbf{\Lambda}$ is continuously differentiable by assumption, the joint pdf of $\mathbf{\Lambda}^{(u)}$ is also continuously differentiable. Moreover, it can be shown that the transformation $\mathbf{\Lambda}^{(u)} \mapsto \widetilde{\mathbf{\Lambda}}^{(u)}$ defined by (273) is a diffeomorphism of class C^2 [29, p. 147]. Therefore, the joint pdf of $\widetilde{\mathbf{\Lambda}}^{(u)}$ is continuously differentiable.

We next use (272) in (269) to conclude that for every $\gamma \in \left(e^{-C_\epsilon^{\mathrm{tx}} - \delta_{\mathrm{a}}}, e^{-C_\epsilon^{\mathrm{tx}} + \delta_{\mathrm{a}}} \right)$ (where $\delta_{\mathrm{a}} \triangleq \min\{\delta_1, \dots, \delta_m\}$)

$$\mathbb{P}\left[\prod_{u=1}^{m} T_j \leq \gamma \right]$$

$$\geq \sum_{u=1}^{m} \left\{ \mathbb{P}\left[\prod_{j=1}^{u} \frac{1}{1 + \widetilde{\Lambda}_j^{(u)}} \leq \gamma \right] \mathbb{P}[\mathbf{\Lambda} \in \mathrm{Int}(\mathcal{W}_u)] \right\} + \mathcal{O}\left(\frac{1}{n} \right) \qquad (274)$$

$$= \mathbb{P}\left[\prod_{j=1}^{m} \frac{1}{1 + \Lambda_j v_j^*(\mathbf{\Lambda})} \leq \gamma \right] + \mathcal{O}\left(\frac{1}{n} \right) \qquad (275)$$

$$= 1 - \mathbb{P}\left[\sum_{j=1}^{m} \log(1 + \Lambda_j v_j^*(\mathbf{\Lambda})) \leq -\log\gamma \right] + \mathcal{O}\left(\frac{1}{n} \right) (276)$$

$$= 1 - F_{\mathrm{tx}}(-\log\gamma) + \mathcal{O}\left(\frac{1}{n} \right) \qquad (277)$$

where $F_{\mathrm{tx}}(\cdot)$ is given in (20).

We next choose γ_n as the solution of

$$1 - F_{\mathrm{tx}}(-\log\gamma_n) + \mathcal{O}\left(\frac{1}{n} \right) = 1 - \epsilon + \frac{1}{n}. \qquad (278)$$

Since $F_{\mathrm{tx}}(C_\epsilon^{\mathrm{tx}}) = \epsilon$ and $F_{\mathrm{tx}}'(C_\epsilon^{\mathrm{tx}}) > 0$, it follow from Taylor's theorem that

$$-\log\gamma_n = C_\epsilon^{\mathrm{tx}} + \mathcal{O}\left(\frac{1}{n} \right). \qquad (279)$$

So, for sufficiently large n, γ_n in (279) belongs to the interval $\left(e^{-C_\epsilon^{\mathrm{tx}} - \delta_{\mathrm{a}}}, e^{-C_\epsilon^{\mathrm{tx}} + \delta_{\mathrm{a}}} \right)$. Hence, by (264), (277), and (278), this γ_n satisfies (44). This concludes the proof.

A. Proof of Lemma 18

Choose $\delta > 0$ such that $\delta \leq \xi_0/2$. Throughout this appendix, we shall use const to indicate a finite constant that does neither depend on $\xi \in (\xi_0 - \delta, \xi_0 + \delta)$ nor on n; its magnitude and sign may change at each occurrence.

Let $g_{\mathrm{th}} \triangleq 2/\xi_0 - 1$ and let

$$p_1 \triangleq \mathbb{P}\left[\prod_{j=1}^{u} D_j \leq \xi \middle| G_1 \geq g_{\mathrm{th}} \right] \qquad (280)$$

$$p_2 \triangleq \mathbb{P}\left[\prod_{j=1}^{u} D_j \leq \xi \middle| G_1 < g_{\mathrm{th}} \right]. \qquad (281)$$

To prove Lemma 18, we decompose $\mathbb{P}\left[\prod_{j=1}^{u} D_j \leq \xi \right]$ as

$$\mathbb{P}\left[\prod_{j=1}^{u} D_j \leq \xi \right] = p_1 \mathbb{P}[G_1 \geq g_{\mathrm{th}}] + p_2 \mathbb{P}[G_1 < g_{\mathrm{th}}]. \quad (282)$$

The proof consists of the following steps:

1) We show in Section V-A1 that for every $\xi \in (\xi_0 - \delta, \xi_0 + \delta)$, the term p_1 in (282) can be lower-bounded as

$$p_1 \geq 1 - \frac{\mathrm{const}}{n}. \qquad (283)$$

2) Using Lemma 17 in Appendix IV-C, we show in Section V-A2 that p_2 can be lower-bounded as

$$p_2 \geq \mathbb{P}\left[\frac{1}{1 + G_1} \prod_{j=2}^{u} D_j \leq \xi \middle| G_1 < g_{\mathrm{th}} \right] - \frac{\mathrm{const}}{n}. \quad (284)$$

3) Reiterating Step 2 for $D_2, \dots, D_u$, we conclude that (284) can be further lower-bounded as

$$p_2 \geq \mathbb{P}\left[\prod_{j=1}^{u} \frac{1}{1 + G_j} \leq \xi \middle| G_1 < g_{\mathrm{th}} \right] - \frac{\mathrm{const}}{n}. \quad (285)$$

4) Finally, using (283) and (285) in (282), we show in Section V-A3 that

$$\mathbb{P}\left[\prod_{j=1}^{u} D_j \leq \xi \right] \geq \mathbb{P}\left[\prod_{j=1}^{u} \frac{1}{1 + G_j} \leq \xi \right] - \frac{\mathrm{const}}{n}. \quad (286)$$

This proves Lemma 18.

1) Proof of (283)*:* Let δ_1 be an arbitrary real number in $(1/(\xi_0 - \delta), 2/\xi_0)$ and let $\delta_2 \triangleq \sqrt{g_{\text{th}}} - \sqrt{\delta_1 - 1} > 0$. Let $W_{n+1,1} \sim \mathcal{CN}(0,1)$ be independent of all other random variables appearing in the definition of the $\{D_j\}$ in (270). Finally, let W_{re} denote the real part of $W_{1,1}$. For every $\xi \in (\xi_0 - \delta, \xi_0 + \delta)$

$$p_1 \geq \mathbb{P}\big[D_1 \leq \xi \big| G_1 \geq g_{\text{th}}\big] \tag{287}$$

$$\geq \mathbb{P}\Bigg[\big|\sqrt{nG_1} + W_{1,1}\big|^2 \geq \frac{1-\xi}{\xi}\sum_{i=2}^{n}|W_{i,1}|^2,$$
$$W_{\text{re}} \geq -\sqrt{n}\delta_2 \bigg| G_1 \geq g_{\text{th}}\Bigg] \tag{288}$$

$$\geq \mathbb{P}\Bigg[n(\sqrt{G_1} - \delta_2)^2 \geq \frac{1-\xi}{\xi}\sum_{i=2}^{n}|W_{i,1}|^2 \bigg| G_1 \geq g_{\text{th}}\Bigg]$$
$$\times \mathbb{P}\big[W_{\text{re}} \geq -\sqrt{n}\delta_2\big] \tag{289}$$

$$\geq \mathbb{P}\Bigg[n(\delta_1 - 1) \geq \frac{1-\xi}{\xi}\sum_{i=2}^{n}|W_{i,1}|^2\Bigg] \mathbb{P}\big[W_{\text{re}} \geq -\sqrt{n}\delta_2\big] \tag{290}$$

$$\geq \mathbb{P}\Bigg[n(\delta_1 - 1) \geq \big(1/(\xi_0 - \delta) - 1\big)\sum_{i=2}^{n+1}|W_{i,1}|^2\Bigg]$$
$$\times \mathbb{P}\big[|W_{\text{re}}| \leq \sqrt{n}\delta_2\big] \tag{291}$$

$$\geq \left(1 - \frac{1}{n}\left(\frac{\delta_1(\xi_0 - \delta) - 1}{1 - (\xi_0 - \delta)}\right)^2\right)\left(1 - \frac{1}{2n\delta_2^2}\right) \tag{292}$$

$$\geq 1 - \frac{\text{const}}{n}. \tag{293}$$

Here, (287) follows because $D_i \leq 1$, $i = 2, \ldots, u$, with probability one (see (270)); (290) follows because $\delta_1 - 1 = (\sqrt{g_{\text{th}}} - \delta_2)^2$; (291) follows because $\xi > \xi_0 - \delta$ and because $\sum_{i=2}^{n+1}|W_{i,1}|^2$ is stochastically larger than $\sum_{i=2}^{n}|W_{i,1}|^2$; (292) follows from Chebyshev's inequality applied to both probabilities in (291). This proves (283).

Before proceeding to the next step, we first argue that, if $\mathbb{P}[G_1 \geq g_{\text{th}}] = 1$, then (271) follows directly from (293). Indeed, in this case we obtain from (293) and (282) that

$$\mathbb{P}\Bigg[\prod_{j=1}^{u} D_j \leq \xi\Bigg] = p_1 \geq 1 - \frac{\text{const}}{n}. \tag{294}$$

We further have, with probability one,

$$\prod_{j=1}^{u}\frac{1}{1+G_j} \leq \frac{1}{1+G_1} \leq \frac{1}{1+g_{\text{th}}} = \frac{\xi_0}{2} \leq \xi_0 - \delta < \xi \tag{295}$$

which gives

$$\mathbb{P}\Bigg[\prod_{j=1}^{u}\frac{1}{1+G_j} \leq \xi\Bigg] = 1. \tag{296}$$

Subtracting (294) from (296) yields (271). In the following, we shall focus exclusively on the case $\mathbb{P}[G_1 \geq g_{\text{th}}] < 1$.

2) Proof of (284)*:* To evaluate p_2 in (282), we proceed as follows. Defining $Z \triangleq \xi/\prod_{j=2}^{u} D_j$, we obtain

$$p_2 = \mathbb{P}\Bigg[\prod_{j=1}^{u} D_j \leq \xi \bigg| G_1 < g_{\text{th}}\Bigg] \tag{297}$$

$$= \mathbb{P}\big[D_1 \leq Z | G_1 < g_{\text{th}}\big] \tag{298}$$

$$= \mathbb{P}\big[D_1 \leq Z, Z \geq 1 \big| G_1 < g_{\text{th}}\big]$$
$$+ \mathbb{P}\big[D_1 \leq Z, Z < 1 \big| G_1 < g_{\text{th}}\big] \tag{299}$$

$$= \mathbb{P}\big[Z \geq 1 \big| G_1 < g_{\text{th}}\big] + \mathbb{P}\big[D_1 \leq Z, Z < 1 \big| G_1 < g_{\text{th}}\big] \tag{300}$$

where (300) follows because

$$\mathbb{P}\big[D_1 \leq Z | Z \geq 1, G_1 < g_{\text{th}}\big] = 1. \tag{301}$$

The second term on the RHS of (300) can be rewritten as

$$\mathbb{P}\big[D_1 \leq Z, Z < 1 \big| G_1 < g_{\text{th}}\big]$$
$$= \mathbb{E}_{Z, G_2, \ldots, G_u \,|\, G_1 < g_{\text{th}}}\Big[\mathbb{1}\{Z < 1\}$$
$$\times \mathbb{P}\big[D_1 \leq Z | Z, G_2, \ldots, G_u, G_1 < g_{\text{th}}\big]\Big]. \tag{302}$$

Since events of measure zero do not affect (302), we can assume without loss of generality that the conditional joint pdf of $Z, G_2, \ldots, G_u$ given $G_1 < g_{\text{th}}$ is strictly positive. To lower-bound (302), we first bound the conditional probability $\mathbb{P}\big[D_1 \leq Z | Z, G_2, \ldots, G_u, G_1 < g_{\text{th}}\big]$. Again, let W_{re} denote the real part of $W_{1,1}$, and let $W_{n+1,1} \sim \mathcal{CN}(0,1)$ be independent of all other random variables appearing in the definition of the $\{D_j\}$ in (270). Following similar steps as in (287)–(293), we obtain for $Z < 1$

$$\mathbb{P}[D_1 \leq Z \,|\, Z, G_2, \ldots, G_u, G_1 < g_{\text{th}}]$$
$$= \mathbb{P}\Bigg[\frac{\sum_{i=2}^{n}|W_{i,1}|^2}{\big|\sqrt{nG_1} + W_{1,1}\big|^2 + \sum_{i=2}^{n}|W_{i,1}|^2} \leq Z \bigg|$$
$$Z, G_2, \ldots, G_u, G_1 < g_{\text{th}}\Bigg] \tag{303}$$

$$= \mathbb{P}\Bigg[\big|\sqrt{nG_1} + W_{1,1}\big|^2 \geq \big(Z^{-1} - 1\big)\sum_{i=2}^{n}|W_{i,1}|^2 \bigg|$$
$$Z, G_2, \ldots, G_u, G_1 < g_{\text{th}}\Bigg] \tag{304}$$

$$\geq \mathbb{P}\Bigg[\big|\sqrt{nG_1} + W_{\text{re}}\big|^2 \geq \big(Z^{-1} - 1\big)\sum_{i=2}^{n+1}|W_{i,1}|^2 \bigg|$$
$$Z, G_2, \ldots, G_u, G_1 < g_{\text{th}}\Bigg] \tag{305}$$

$$\geq \mathbb{P}\Bigg[\sqrt{nG_1} \geq -W_{\text{re}} + \sqrt{Z^{-1} - 1}\sqrt{\sum_{i=2}^{n+1}|W_{i,1}|^2} \bigg|$$
$$Z, G_2, \ldots, G_u, G_1 < g_{\text{th}}\Bigg]. \tag{306}$$

Next, we lower-bound the RHS of (306) using Lemma 17 in Appendix IV-C. Let μ_W and σ_W^2 be the mean and the variance

of the random variable $\sqrt{\sum_{i=2}^{n+1} |W_{i,1}|^2}$. Let $Z_2 \triangleq \sqrt{Z^{-1} - 1}$. Furthermore, let

$$K_1 \triangleq \frac{1}{\sqrt{1/2 + Z_2^2 \sigma_W^2}} \left(-W_{\mathrm{re}} + Z_2 \sqrt{\sum_{i=2}^{n+1} |W_{i,1}|^2} - \mu_W Z_2 \right)$$
(307)

and

$$\overline{G}_1 \triangleq \frac{1}{\sqrt{1/2 + Z_2^2 \sigma_W^2}} \left(\sqrt{G_1} - \frac{\mu_W}{\sqrt{n}} Z_2 \right).$$
(308)

Note that K_1 is a zero-mean, unit-variance random variable that is conditionally independent of $\overline{G}_1$ given Z_2. Using these definitions, we can rewrite the RHS of (306) as

$$\mathbb{P}\left[\overline{G}_1 \geq K_1/\sqrt{n} \Big| Z_2, G_2, \ldots, G_u, G_1 < g_{\mathrm{th}} \right].$$
(309)

In order to use Lemma 17, we need to establish an upper bound on the conditional pdf of $\overline{G}_1$ given $Z_2, G_2, \ldots, G_u$ and $G_1 < g_{\mathrm{th}}$, which we denote by $\tilde{f}_{\overline{G}_1}$, and on its derivative. As $f_{G_1,\ldots,G_u}$ is continuously differentiable by assumption, $f_{G_1,\ldots,G_u}$ and its partial derivatives are bounded on bounded sets. Together with the assumption that $\mathbb{P}[G_1 \geq g_{\mathrm{th}}] < 1$, this implies that the conditional pdf $f_{G_1,\ldots,G_u \mid G_1 < g_{\mathrm{th}}}$ of $G_1, \ldots, G_u$ given $G_1 < g_{\mathrm{th}}$ and its partial derivatives are all bounded on $[0, g_{\mathrm{th}}]^u$. Namely, for every $\{x_1, \ldots, x_u\} \in [0, g_{\mathrm{th}}]^u$ and every $i \in \{1, \ldots, u\}$

$$f_{G_1,\ldots,G_u \mid G_1 < g_{\mathrm{th}}}(x_1, \ldots, x_u) \leq \mathrm{const}$$
(310)

$$\left| \frac{\partial f_{G_1,\ldots,G_u \mid G_1 < g_{\mathrm{th}}}(x_1, \ldots, x_u)}{\partial x_i} \right| \leq \mathrm{const}.$$
(311)

Let $\tilde{f}_{G_1}$ be the conditional pdf of G_1 given $G_2, \ldots, G_u$ and $G_1 < g_{\mathrm{th}}$, and let $f_{G_2,\ldots,G_u \mid G_1 < g_{\mathrm{th}}}$ be the conditional pdf of $G_2, \ldots, G_u$ given $G_1 < g_{\mathrm{th}}$. Then, $\tilde{f}_{\overline{G}_1}$ can be bounded as

$$\tilde{f}_{\overline{G}_1}(x \mid z_2, g_2 \ldots, g_u)$$
$$= 2\tilde{f}_{G_1}\left(\left(\sqrt{1/2 + z_2^2 \sigma_W^2}\, x + z_2 \mu_W/\sqrt{n} \right)^2 \Big| g_2, \ldots, g_u \right)$$
$$\times \sqrt{1/2 + z_2^2 \sigma_W^2} \left(\sqrt{1/2 + z_2^2 \sigma_W^2}\, x + z_2 \mu_W/\sqrt{n} \right)$$
(312)

$$\leq \frac{\mathrm{const} \cdot \sqrt{g_{\mathrm{th}}} \sqrt{1/2 + \sigma_W^2 z_2^2}}{f_{G_2,\ldots,G_u \mid G_1 < g_{\mathrm{th}}}(g_2, \ldots, g_u)}.$$
(313)

Here, (312) follows from (308), and (313) follows from (310) and because we condition on the event that $G_1 < g_{\mathrm{th}}$, so

$$\sqrt{1/2 + z_2^2 \sigma_W^2}\, x + z_2 \mu_W/\sqrt{n} \leq \sqrt{g_{\mathrm{th}}}.$$
(314)

To further upper-bound (313), we shall use that σ_W and Z_2 are bounded:

$$\sigma_W^2 = n - \left(\frac{\Gamma(n + 1/2)}{\Gamma(n)} \right)^2$$
(315)

$$\leq 1/4$$
(316)

and

$$Z_2^2 = Z^{-1} - 1$$
(317)

$$\leq 1/\xi - 1$$
(318)

$$\leq (\xi_0 - \delta)^{-1} - 1.$$
(319)

Here, (315) follows by using that $\sqrt{2 \sum_{i=2}^{n+1} |W_{i,1}|^2}$ is χ-distributed with $2n$ degrees of freedom and by using [42, Eq. (18.14)]; (316) follows from [43, Sec. 2.2]; (318) follows from the definition of Z and because $\prod_{j=2}^u D_j \leq 1$. Substituting (316) and (319) into (313), we obtain

$$\tilde{f}_{\overline{G}_1}(x \mid z_2, g_2 \ldots, g_u) \leq \frac{\mathrm{const}}{f_{G_2,\ldots,G_u \mid G_1 < g_{\mathrm{th}}}(g_2, \ldots, g_u)}.$$
(320)

Following similar steps, we can also establish that

$$\left| \tilde{f}'_{\overline{G}_1}(x \mid z_2, g_2 \ldots, g_u) \right| \leq \frac{\mathrm{const}}{f_{G_2,\ldots,G_u \mid G_1 < g_{\mathrm{th}}}(g_2, \ldots, g_u)}.$$
(321)

Using (320)–(321) and Lemma 17, we obtain that

$$\mathbb{P}\left[\overline{G}_1 \geq K_1/\sqrt{n} \Big| Z_2, G_2 = g_2, \ldots, G_u = g_u, G_1 < g_{\mathrm{th}} \right]$$
$$\geq \mathbb{P}\left[\overline{G}_1 \geq 0 \Big| Z_2, G_2 = g_2, \ldots, G_u = g_u, G_1 < g_{\mathrm{th}} \right]$$
$$- \frac{\mathrm{const}}{n} \left(1 + \frac{1}{f_{G_2,\ldots,G_u \mid G_1 < g_{\mathrm{th}}}(g_2, \ldots, g_u)} \right).$$
(322)

Returning to the analysis of (302), we combine (306), (309) and (322) to obtain

$$\mathbb{P}\left[D_1 \leq Z, Z < 1 \big| G_1 < g_{\mathrm{th}} \right]$$
$$\geq \mathbb{E}_{Z,G_2,\ldots,G_u \mid G_1 < g_{\mathrm{th}}} \Bigg[\mathbb{1}\{Z < 1\}$$
$$\times \Bigg(\mathbb{P}\left[\overline{G}_1 \geq 0 \Big| Z, G_2, \ldots, G_u, G_1 < g_{\mathrm{th}} \right]$$
$$- \frac{\mathrm{const}}{n} \left(1 + \frac{1}{f_{G_2,\ldots,G_u \mid G_1 < g_{\mathrm{th}}}(G_2, \ldots, G_u)} \right) \Bigg) \Bigg]$$
(323)

$$\geq \mathbb{P}\left[\frac{1}{1 + nG_1/\mu_W^2} \leq Z, Z < 1 \Big| G_1 < g_{\mathrm{th}} \right] - \frac{\mathrm{const}}{n}$$
$$\times \left(1 + \int_0^{g_{\mathrm{th}}} \cdots \int_0^{g_{\mathrm{th}}} \frac{f_{G_2,\ldots,G_u \mid G_1 < g_{\mathrm{th}}}(g_2, \ldots, g_u)}{f_{G_2,\ldots,G_u \mid G_1 < g_{\mathrm{th}}}(g_2, \ldots, g_u)} dg_2 \cdots dg_u \right)$$
(324)

$$\geq \mathbb{P}\left[\frac{1}{1 + G_1} \leq Z, Z < 1 \Big| G_1 < g_{\mathrm{th}} \right] - \frac{\mathrm{const}}{n}.$$
(325)

Here, in (324) we used (308), that $\mathbb{1}\{Z < 1\} \leq 1$, that $G_1, \ldots, G_u$ are nonincreasing, and that const in (323) is positive; (325) follows because [42, Eq. (18.14)]

$$\mu_W = \frac{\Gamma(n + 1/2)}{\Gamma(n)} \leq \sqrt{n}$$
(326)

and because the integral on the RHS of (324) is bounded. Substituting (325) into (300), we obtain

$$p_2 \geq \mathbb{P}\left[Z \geq 1 \big| G_1 < g_{\mathrm{th}} \right]$$
$$+ \mathbb{P}\left[\frac{1}{1 + G_1} \leq Z, Z < 1 \Big| G_1 < g_{\mathrm{th}} \right] - \frac{\mathrm{const}}{n}$$
(327)

$$= \mathbb{P}\left[\frac{1}{1 + G_1} \leq Z, Z \geq 1 \Big| G_1 < g_{\mathrm{th}} \right]$$
$$+ \mathbb{P}\left[\frac{1}{1 + G_1} \leq Z, Z < 1 \Big| G_1 < g_{\mathrm{th}} \right] - \frac{\mathrm{const}}{n}$$
(328)

$$= \mathbb{P}\left[\frac{1}{1+G_1} \leq Z \,\Big|\, G_1 < g_{\text{th}}\right] - \frac{\text{const}}{n} \tag{329}$$

$$= \mathbb{P}\left[\frac{1}{1+G_1}\prod_{j=2}^{u} D_j \leq \xi \,\Big|\, G_1 < g_{\text{th}}\right] - \frac{\text{const}}{n} \tag{330}$$

where (328) follows because $1/(1+G_1) \leq 1$ with probability one. This proves (284).

3) Proof of (286)*:* Set $p_0 \triangleq \mathbb{P}[G_1 \geq g_{\text{th}}]$. Substituting (283) and (285) into (282), we obtain

$$\mathbb{P}\left[\prod_{j=1}^{u} D_j \leq \xi\right]$$

$$\geq p_0 + (1-p_0)\mathbb{P}\left[\prod_{j=1}^{u}\frac{1}{1+G_j} \leq \xi \,\Big|\, G_1 < g_{\text{th}}\right] - \frac{\text{const}}{n} \tag{331}$$

$$= \underbrace{\mathbb{P}\left[\prod_{j=1}^{u}\frac{1}{1+G_j} \leq \xi \,\Big|\, G_1 \geq g_{\text{th}}\right]}_{=1} p_0$$

$$+ (1-p_0)\mathbb{P}\left[\prod_{j=1}^{u}\frac{1}{1+G_j} \leq \xi \,\Big|\, G_1 < g_{\text{th}}\right] - \frac{\text{const}}{n} \tag{332}$$

$$= \mathbb{P}\left[\prod_{j=1}^{u}\frac{1}{1+G_j} \leq \xi\right] - \frac{\text{const}}{n}. \tag{333}$$

The first factor in (332) is equal to one because of (295). This proves (286) and concludes the proof of Lemma 18.

APPENDIX VI
PROOF OF PROPOSITION 5 (EXISTENCE OF OPTIMAL COVARIANCE MATRIX)

Since the set $\mathcal{U}_t$ is compact, by the extreme value theorem [29, p. 34], it is sufficient to show that, under the assumptions in the proposition, the function $\mathsf{Q} \mapsto \mathbb{P}\left[\log\det\left(\mathsf{I}_r + \mathbb{H}^{\mathrm{H}}\mathsf{Q}\mathbb{H}\right) \leq \xi\right]$ is continuous in $\mathsf{Q} \in \mathcal{U}_t$ with respect to the metric $d(\mathsf{A},\mathsf{B}) = \|\mathsf{A} - \mathsf{B}\|_{\mathsf{F}}$.

Consider an arbitrary sequence $\{\mathsf{Q}_l\}$ in $\mathcal{U}_t$ that converges to Q. Then

$$\det(\mathsf{I}_r + \mathsf{H}^{\mathrm{H}}\mathsf{Q}_l\mathsf{H})$$
$$= \det(\mathsf{I}_r + \mathsf{H}^{\mathrm{H}}\mathsf{Q}\mathsf{H} + \mathsf{H}^{\mathrm{H}}(\mathsf{Q}_l - \mathsf{Q})\mathsf{H}) \tag{334}$$

$$= \det(\mathsf{I}_r + \mathsf{H}^{\mathrm{H}}\mathsf{Q}\mathsf{H})$$
$$\times \det\left(\mathsf{I}_r + \mathsf{H}^{\mathrm{H}}(\mathsf{Q}_l - \mathsf{Q})\mathsf{H}(\mathsf{I}_r + \mathsf{H}^{\mathrm{H}}\mathsf{Q}\mathsf{H})^{-1}\right) \tag{335}$$

$$\leq \det(\mathsf{I}_r + \mathsf{H}^{\mathrm{H}}\mathsf{Q}\mathsf{H})$$
$$\times \left(1 + \left\|\mathsf{H}^{\mathrm{H}}(\mathsf{Q}_l - \mathsf{Q})\mathsf{H}(\mathsf{I}_r + \mathsf{H}^{\mathrm{H}}\mathsf{Q}\mathsf{H})^{-1}\right\|_{\mathsf{F}}\right)^r \tag{336}$$

$$\leq \det(\mathsf{I}_r + \mathsf{H}^{\mathrm{H}}\mathsf{Q}\mathsf{H})$$
$$\times \left(1 + \|\mathsf{Q}_l - \mathsf{Q}\|_{\mathsf{F}}\|\mathsf{H}\|_{\mathsf{F}}^2\left\|(\mathsf{I}_r + \mathsf{H}^{\mathrm{H}}\mathsf{Q}\mathsf{H})^{-1}\right\|_{\mathsf{F}}\right)^r \tag{337}$$

$$\leq \det(\mathsf{I}_r + \mathsf{H}^{\mathrm{H}}\mathsf{Q}\mathsf{H})\left(1 + \|\mathsf{Q}_l - \mathsf{Q}\|_{\mathsf{F}}\|\mathsf{H}\|_{\mathsf{F}}^2\sqrt{r}\right)^r. \tag{338}$$

Here, (336) follows from Hadamard's inequality; (337) follows from the sub-multiplicative property of the Frobenius norm, namely, $\|\mathsf{AB}\|_{\mathsf{F}} \leq \|\mathsf{A}\|_{\mathsf{F}}\|\mathsf{B}\|_{\mathsf{F}}$; (338) follows because

$\left\|(\mathsf{I}_r + \mathsf{H}^{\mathrm{H}}\mathsf{Q}\mathsf{H})^{-1}\right\|_{\mathsf{F}} \leq \|\mathsf{I}_r\|_{\mathsf{F}} = \sqrt{r}$. Similarly, by replacing Q_l with Q in the above steps, we obtain

$$\det(\mathsf{I}_r + \mathsf{H}^{\mathrm{H}}\mathsf{Q}\mathsf{H})$$
$$\leq \det(\mathsf{I}_r + \mathsf{H}^{\mathrm{H}}\mathsf{Q}_l\mathsf{H})(1 + \|\mathsf{Q}_l - \mathsf{Q}\|_{\mathsf{F}}\|\mathsf{H}\|_{\mathsf{F}}^2\sqrt{r})^r. \tag{339}$$

The inequalities (338) and (339) imply that

$$\left|\log\det(\mathsf{I}_r + \mathsf{H}^{\mathrm{H}}\mathsf{Q}_l\mathsf{H}) - \log\det(\mathsf{I}_r + \mathsf{H}^{\mathrm{H}}\mathsf{Q}\mathsf{H})\right|$$
$$\leq r\log(1 + \|\mathsf{Q}_l - \mathsf{Q}\|_{\mathsf{F}}\|\mathsf{H}\|_{\mathsf{F}}^2\sqrt{r}) \tag{340}$$
$$\leq r^{3/2}\|\mathsf{Q}_l - \mathsf{Q}\|_{\mathsf{F}}\|\mathsf{H}\|_{\mathsf{F}}^2. \tag{341}$$

Hence, for every $c > 0$

$$\mathbb{P}\left[\left|\log\det(\mathsf{I}_r + \mathbb{H}^{\mathrm{H}}\mathsf{Q}_l\mathbb{H}) - \log\det(\mathsf{I}_r + \mathbb{H}^{\mathrm{H}}\mathsf{Q}\mathbb{H})\right| \geq c\right]$$

$$\leq \mathbb{P}\left[\|\mathbb{H}\|_{\mathsf{F}}^2 \geq \frac{c}{r^{3/2}}\frac{1}{\|\mathsf{Q}_l - \mathsf{Q}\|_{\mathsf{F}}}\right] \tag{342}$$

$$\leq \mathbb{E}\left[\|\mathbb{H}\|_{\mathsf{F}}^2\right] \cdot \|\mathsf{Q}_l - \mathsf{Q}\|_{\mathsf{F}}\frac{r^{3/2}}{c} \tag{343}$$

$$\to 0, \qquad \text{as } \mathsf{Q}_l \to \mathsf{Q} \tag{344}$$

where (343) follows from Markov's inequality and (344) follows because, by assumption, $\mathbb{E}\left[\|\mathbb{H}\|_{\mathsf{F}}^2\right] < \infty$. Thus, the sequence of random variables $\{\log\det(\mathsf{I}_r + \mathbb{H}^{\mathrm{H}}\mathsf{Q}_l\mathbb{H})\}$ converges in probability to $\log\det(\mathsf{I}_r + \mathbb{H}^{\mathrm{H}}\mathsf{Q}\mathbb{H})$. Since convergence in probability implies convergence in distribution, we conclude that

$$\mathbb{P}\left[\log\det\left(\mathsf{I}_r + \mathbb{H}^{\mathrm{H}}\mathsf{Q}_l\mathbb{H}\right) \leq \xi\right]$$
$$\to \mathbb{P}\left[\log\det\left(\mathsf{I}_r + \mathbb{H}^{\mathrm{H}}\mathsf{Q}\mathbb{H}\right) \leq \xi\right] \text{ as } \mathsf{Q}_l \to \mathsf{Q} \tag{345}$$

for every $\xi \in \mathbb{R}$ for which the cdf of $\log\det(\mathsf{I}_r + \mathbb{H}^{\mathrm{H}}\mathsf{Q}\mathbb{H})$ is continuous [44, p. 308]. However, the cdf of $\log\det(\mathsf{I}_r + \mathbb{H}^{\mathrm{H}}\mathsf{Q}\mathbb{H})$ is continuous for every $\xi \in \mathbb{R}$ since the distribution of $\mathbb{H}$ is, by assumption, absolutely continuous with respect to the Lebesgue measure and the function $\mathsf{H} \mapsto \log\det(\mathsf{I}_r + \mathsf{H}^{\mathrm{H}}\mathsf{Q}\mathsf{H})$ is continuous. Consequently, (345) holds for every $\xi \in \mathbb{R}$, thus proving Proposition 5.

APPENDIX VII
PROOF OF THEOREM 6 (CSIR CONVERSE BOUND)

For the CSIR case, the input of the channel (8) is X and the output is the pair $(\mathbb{Y}, \mathbb{H})$. An $(n, M, \epsilon)_{\mathrm{e}}$ code is defined in a similar way as the $(n, M, \epsilon)_{\mathrm{rx}}$ code in Definition 2, except that each codeword satisfies the power constraint (9) with equality, i.e., each codeword belongs to the set

$$\mathcal{F}_{n,t} \triangleq \{\mathsf{X} \in \mathbb{C}^{n \times t} : \|\mathsf{X}\|_{\mathsf{F}}^2 = n\rho\}. \tag{346}$$

Denote by $R_{\mathrm{e}}^*(n, \epsilon)$ the maximal achievable rate with an $(n, M, \epsilon)_{\mathrm{e}}$ code. Then by [21, Sec. XIII] (see also [9, Lem. 39],

$$R_{\mathrm{rx}}^*(n-1, \epsilon) \leq \frac{n}{n-1}R_{\mathrm{e}}^*(n, \epsilon). \tag{347}$$

We next establish an upper bound on $R_{\mathrm{e}}^*(n, \epsilon)$. Consider an arbitrary $(M, n, \epsilon)_{\mathrm{e}}$ code. To each codeword $\mathsf{X} \in \mathcal{F}_{n,t}$, we associate a matrix $\mathsf{U}(\mathsf{X}) \in \mathbb{C}^{t \times t}$:

$$\mathsf{U}(\mathsf{X}) \triangleq \frac{1}{n}\mathsf{X}^{\mathrm{H}}\mathsf{X}. \tag{348}$$

To upper-bound $R_e^*(n, \epsilon)$, we use the meta-converse theorem [9, Th. 30]. As *auxiliary* channel $Q_{\mathbb{YH}\,|\,\mathbb{X}}$, we take

$$Q_{\mathbb{YH}\,|\,\mathbb{X}} = P_{\mathbb{H}} \times Q_{\mathbb{Y}\,|\,\mathbb{XH}} \tag{349}$$

where

$$Q_{\mathbb{Y}\,|\,\mathbb{X}=\mathsf{X},\mathbb{H}=\mathsf{H}} = \prod_{i=1}^{n} Q_{\boldsymbol{Y}_i\,|\,\mathbb{X}=\mathsf{X},\mathbb{H}=\mathsf{H}} \tag{350}$$

with $\boldsymbol{Y}_i$, $i = 1, \ldots, n$ denoting the rows of $\mathbb{Y}$, and

$$Q_{\boldsymbol{Y}_i\,|\,\mathbb{X}=\mathsf{X},\mathbb{H}=\mathsf{H}} = \mathcal{CN}\big(\boldsymbol{0}, \mathsf{I}_r + \mathsf{H}^{\mathsf{H}}\mathsf{U}(\mathsf{X})\mathsf{H}\big). \tag{351}$$

By [9, Th. 30], we have

$$\inf_{\mathsf{X}\in\mathcal{F}_{n,t}} \beta_{1-\epsilon}\big(P_{\mathbb{YH}\,|\,\mathbb{X}=\mathsf{X}}, Q_{\mathbb{YH}\,|\,\mathbb{X}=\mathsf{X}}\big) \leq 1 - \epsilon' \tag{352}$$

where ϵ' is the *maximal probability of error* of the optimal code with M codewords over the auxiliary channel (349). To shorten notation, we define

$$\beta_{1-\epsilon}^n(\mathsf{X}) \triangleq \beta_{1-\epsilon}\big(P_{\mathbb{YH}\,|\,\mathbb{X}=\mathsf{X}}, Q_{\mathbb{YH}\,|\,\mathbb{X}=\mathsf{X}}\big). \tag{353}$$

To prove the theorem, we proceed as in Appendix III: we first evaluate $\beta_{1-\epsilon}^n(\mathsf{X})$, then we relate ϵ' to $R_e^*(n, \epsilon)$ by establishing a converse bound on the channel $Q_{\mathbb{YH}\,|\,\mathbb{X}}$.

1) Evaluation of $\beta_{1-\epsilon}(\mathsf{X})$: Let G be an arbitrary $n \times n$ unitary matrix. Let $g_i : \mathcal{F}_{n,t} \mapsto \mathcal{F}_{n,t}$ and $g_o : \mathbb{C}^{n \times r} \times \mathbb{C}^{t \times r} \mapsto \mathbb{C}^{n \times r} \times \mathbb{C}^{t \times r}$ be two mappings defined as

$$g_i(\mathsf{X}) \triangleq \mathsf{GX} \quad \text{and} \quad g_o(\mathsf{Y}, \mathsf{H}) \triangleq (\mathsf{GY}, \mathsf{H}). \tag{354}$$

Note that

$$P_{\mathbb{YH}\,|\,\mathbb{X}}(g_o^{-1}(\mathcal{E})\,|\,g_i(\mathsf{X})) = P_{\mathbb{YH}\,|\,\mathbb{X}}(\mathcal{E}\,|\,\mathsf{X}) \tag{355}$$

for all measurable sets $\mathcal{E} \subset \mathbb{C}^{n \times r} \times \mathbb{C}^{t \times r}$ and $\mathsf{X} \in \mathcal{F}_{n,t}$, i.e., the pair (g_i, g_o) is a symmetry [45, Def. 3] of $P_{\mathbb{YH}\,|\,\mathbb{X}}$. Furthermore, (350) and (351) imply that

$$Q_{\mathbb{YH}\,|\,\mathbb{X}=\mathsf{X}} = Q_{\mathbb{YH}\,|\,\mathbb{X}=g_i(\mathsf{X})} \tag{356}$$

and that $Q_{\mathbb{YH}\,|\,\mathbb{X}=\mathsf{X}}$ is invariant under g_o for all $\mathsf{X} \in \mathcal{F}$. Hence, by [45, Prop. 19], we have that

$$\beta_{1-\epsilon}^n(\mathsf{X}) = \beta_{1-\epsilon}^n(g_i(\mathsf{X})) = \beta_{1-\epsilon}^n(\mathsf{GX}). \tag{357}$$

Since G is arbitrary, this implies that $\beta_{1-\epsilon}^n(\mathsf{X})$ depends on X only through $\mathsf{U}(\mathsf{X})$. Consider the QR decomposition [46, p. 113] of X

$$\mathsf{X} = \mathsf{VX}_0 \tag{358}$$

where $\mathsf{V} \in \mathbb{C}^{n \times n}$ is unitary and $\mathsf{X}_0 \in \mathbb{C}^{n \times t}$ is upper triangular. By (357) and (358),

$$\beta_{1-\epsilon}^n(\mathsf{X}_0) = \beta_{1-\epsilon}^n(\mathsf{X}). \tag{359}$$

Let

$$r(\mathsf{X}_0; \mathbb{YH}) \triangleq \log \frac{dP_{\mathbb{YH}\,|\,\mathbb{X}=\mathsf{X}_0}}{dQ_{\mathbb{YH}\,|\,\mathbb{X}=\mathsf{X}_0}}. \tag{360}$$

Under both $P_{\mathbb{YH}\,|\,\mathbb{X}=\mathsf{X}_0}$ and $Q_{\mathbb{YH}\,|\,\mathbb{X}=\mathsf{X}_0}$, the random variable $r(\mathsf{X}_0; \mathbb{YH})$ has absolutely continuous cdf with respect to the Lebesgue measure. By the Neyman-Pearson lemma [38, p. 300]

$$\beta_{1-\epsilon}^n(\mathsf{X}_0) = Q_{\mathbb{YH}\,|\,\mathbb{X}=\mathsf{X}_0}\big[r(\mathsf{X}_0; \mathbb{YH}) \geq n\gamma_n(\mathsf{X}_0)\big] \tag{361}$$

where $\gamma_n(\mathsf{X}_0)$ is the solution of

$$P_{\mathbb{YH}\,|\,\mathbb{X}=\mathsf{X}_0}\big[r(\mathsf{X}_0; \mathbb{YH}) \leq n\gamma_n(\mathsf{X}_0)\big] = \epsilon. \tag{362}$$

It can be shown that under $P_{\mathbb{YH}\,|\,\mathbb{X}=\mathsf{X}_0}$, the random variable $r(\mathsf{X}_0; \mathbb{YH})$ has the same distribution as $S_n^{\mathrm{rx}}(\mathsf{U}(\mathsf{X}_0))$ in (70), and under $Q_{\mathbb{YH}\,|\,\mathbb{X}=\mathsf{X}_0}$, it has the same distribution as $L_n^{\mathrm{rx}}(\mathsf{U}(\mathsf{X}_0))$ in (69).

2) Converse on the auxiliary channel: To prove the theorem, it remains to lower-bound ϵ', which is the maximal probability of error over the auxiliary channel (349). The following lemma serves this purpose.

Lemma 19: For every code with M codewords and block-length $n \geq r$, the maximum probability of error ϵ' over the auxiliary channel (349) satisfies

$$1 - \epsilon' \leq \frac{c_{\mathrm{rx}}(n)}{M} \tag{363}$$

where $c_{\mathrm{rx}}(n)$ is given in (72).

Substituting (361) into (352) and using (363), we then obtain upon minimizing (361) over all matrices in $\mathcal{U}_t^e$

$$R_e^*(n, \epsilon) \leq \frac{1}{n} \frac{c_{\mathrm{rx}}(n)}{\inf\limits_{\mathsf{Q}\in\mathcal{U}_t^e} \mathbb{P}[L_n^{\mathrm{rx}}(\mathsf{Q}) \geq n\gamma_n]}. \tag{364}$$

The final bound (71) follows by combining (364) with (347) and by noting that the upper bound does not depend on the chosen code.

Proof of Lemma 19: According to (351), given $\mathbb{H} = \mathsf{H}$, the output of the auxiliary channel depends on X only through $\mathsf{U}(\mathsf{X})$. In the following, we shall omit the argument of $\mathsf{U}(\mathsf{X})$ where it is immaterial. Let $\mathbb{V} \triangleq \mathsf{U}(\mathbb{Y})$. Then, $(\mathbb{V}, \mathbb{H})$ is a sufficient statistic for the detection of X from $(\mathbb{Y}, \mathbb{H})$. Therefore, to establish (363), it is sufficient to lower-bound the maximal probability of error ϵ' over the equivalent auxiliary channel

$$Q_{\mathbb{VH}\,|\,\mathbb{U}} = P_{\mathbb{H}} \times Q_{\mathbb{V}\,|\,\mathbb{U},\mathbb{H}} \tag{365}$$

where $Q_{\mathbb{V}\,|\,\mathbb{U}=\mathsf{U},\mathbb{H}=\mathsf{H}}$ is the Wishart distribution [18, Def. 2.3]:

$$Q_{\mathbb{V}\,|\,\mathbb{U}=\mathsf{U},\mathbb{H}=\mathsf{H}} = \mathcal{W}_r\left(n, \frac{1}{n}(\mathsf{I}_r + \mathsf{H}^{\mathsf{H}}\mathsf{UH})\right). \tag{366}$$

Let $\mathsf{B} \triangleq \mathsf{I}_r + \mathsf{H}^{\mathsf{H}}\mathsf{UH}$, and let $q_{\mathbb{V}\,|\,\mathbb{B}}(\mathsf{V}\,|\,\mathsf{B})$ be the pdf associated with (366), i.e., [18, Def. 2.3]

$$q_{\mathbb{V}\,|\,\mathbb{B}}(\mathsf{V}\,|\,\mathsf{B}) = \frac{\det \mathsf{V}^{n-r}}{\Gamma_r(n)\det\big(\frac{1}{n}\mathsf{B}\big)^n} \exp\Big(-\mathrm{tr}\Big(\big(n^{-1}\mathsf{B}\big)^{-1}\mathsf{V}\Big)\Big). \tag{367}$$

It will be convenient to express $q_{\mathbb{V}\,|\,\mathbb{B}}(\mathsf{V}\,|\,\mathsf{B})$ in the coordinate system of the eigenvalue decomposition

$$\mathbb{V} = \mathbb{Q}\mathbb{D}\mathbb{Q}^{\mathsf{H}} \tag{368}$$

where $\mathbb{Q} \in \mathbb{C}^{r \times r}$ is unitary, and $\mathbb{D}$ is a diagonal matrix whose diagonal elements $D_1, \ldots, D_r$ are the eigenvalues of $\mathbb{V}$ in descending order. In order to make the eigenvalue decomposition (368) unique, we assume that the first row of $\mathbb{Q}$ is real

and non-negative. Thus, $\mathbb{Q}$ only lies in a *submanifold* $\widetilde{\mathcal{S}}_{r,r}$ of the Stiefel manifold $\mathcal{S}_{r,r}$. Using (368), we rewrite (367) as

$$q_{\mathbb{Q},\mathbb{D}\,|\,\mathbb{B}}(\mathsf{Q},\mathsf{D}\,|\,\mathsf{B}) = \frac{n^{rn}\exp\big(-n\cdot\mathrm{tr}(\mathsf{B}^{-1}\mathsf{QDQ}^{\mathrm{H}})\big)}{\Gamma_r(n)\det\mathsf{B}^n}$$
$$\times \det\mathsf{D}^{n-r}\prod_{i<j}^{r}(d_i-d_j)^2 \qquad (369)$$

where in (369) we used the fact that the Jacobian of the eigenvalue decomposition (368) is $\prod_{i<j}^{r}(d_i-d_j)^2$ (see [47, Th. 3.1]).

We next establish an upper bound on (369) that is integrable and does not depend on B. To this end, we will bound each of the factors on the RHS of (369). To bound the argument of the exponential function, we apply the trace inequality [48, Th. 20.A.4]

$$\mathrm{tr}(\mathsf{B}^{-1}\mathsf{QDQ}^{\mathrm{H}}) \geq \sum_{i=1}^{r}\frac{d_i}{b_i} \qquad (370)$$

for every unitary matrix Q, where $b_1 \geq \ldots \geq b_r$ are the ordered eigenvalues of B. Using (370) in (369) and further upper-bounding the terms $(d_i-d_j)^2$ in (369) with d_i^2, we obtain

$$q_{\mathbb{Q}\mathbb{D}\,|\,\mathbb{B}}(\mathsf{Q},\mathsf{D}\,|\,\mathsf{B}) \leq \frac{n^{rn}}{\Gamma_r(n)}\prod_{i=1}^{r}\left\{\frac{d_i^{n+r-2i}}{b_i^n}\exp\left(-n\frac{d_i}{b_i}\right)\right\}. \qquad (371)$$

Since $\mathsf{B} = \mathsf{I}_r + \mathsf{H}^{\mathrm{H}}\mathsf{UH}$, we have that

$$1 \leq b_i \leq 1 + \mathrm{tr}\big(\mathsf{H}^{\mathrm{H}}\mathsf{UH}\big) \qquad (372)$$
$$\leq 1 + \|\mathsf{H}\|_{\mathsf{F}}^2\,\mathrm{tr}\,(\mathsf{U}) \qquad (373)$$
$$= 1 + \|\mathsf{H}\|_{\mathsf{F}}^2\,\rho \triangleq b_0 \qquad (374)$$

where (373) follows from the Cauchy-Schwarz inequality and (374) follows because $\mathsf{U} \in \mathcal{U}_t^{\mathrm{e}}$. Using (374), we can upper-bound each factor on the RHS of (371) as follows:

$$\frac{d_i^{n+r-2i}}{b_i^n}\exp\left(-n\frac{d_i}{b_i}\right)$$
$$\leq g_i(d_i) \triangleq \begin{cases} \left(\dfrac{n+r-2i}{n}\right)^{n+r-2i} b_0^{[r-2i]^+}\,e^{-(n+r-2i)}, \\ \qquad\qquad\qquad\qquad \text{if } d_i \leq \dfrac{b_0(n+r-2i)}{n} \\ \left(\dfrac{d_i}{b_0}\right)^{n+r-2i} b_0^{[r-2i]^+}\,e^{-nd_i/b_0}, \\ \qquad\qquad\qquad\qquad \text{if } d_i > \dfrac{b_0(n+r-2i)}{n}. \end{cases} \qquad (375)$$

We are now ready to establish the desired converse result for the auxiliary channel Q. Consider an arbitrary code for the auxiliary channel Q with encoding function $f_0 : \{1,\ldots,M\} \mapsto \mathcal{U}_t^{\mathrm{e}}$. Let $\mathcal{D}_j(\mathsf{H})$ be the decoding set for the jth codeword $f_0(j)$ in the eigenvalue decomposition coordinate such that

$$\bigcup_{j=1}^{M}\mathcal{D}_j(\mathsf{H}) = \widetilde{\mathcal{S}}_{r,r} \times \mathbb{R}_{\geq}^r. \qquad (376)$$

Let ϵ'_{avg} denote the average probability of error over the auxiliary channel. Then,

$$1 - \epsilon'$$
$$\leq 1 - \epsilon'_{\mathrm{avg}} \qquad (377)$$
$$= \frac{1}{M}\mathbb{E}_{\mathbb{H}}\left[\sum_{j=1}^{M}\int_{\mathcal{D}_j(\mathbb{H})}q_{\mathbb{Q},\mathbb{D}\,|\,\mathbb{B}=\mathsf{I}_r+\mathbb{H}^{\mathrm{H}}f_0(j)\mathbb{H}}(\mathsf{Q},\mathsf{D})d\mathsf{Q}d\mathsf{D}\right] \qquad (378)$$
$$\leq \frac{n^{rn}}{\Gamma_r(n)M}\mathbb{E}_{\mathbb{H}}\left[\sum_{j=1}^{M}\int_{\mathcal{D}_j(\mathbb{H})}\prod_{i=1}^{r}g_i(d_i)d\mathsf{Q}d\mathsf{D}\right] \qquad (379)$$
$$= \frac{n^{rn}}{\Gamma_r(n)M}\mathbb{E}_{\mathbb{H}}\left[\int_{\widetilde{\mathcal{S}}_{r,r}\times\mathbb{R}_{\geq}^r}\prod_{i=1}^{r}g_i(d_i)d\mathsf{Q}d\mathsf{D}\right] \qquad (380)$$
$$\leq \frac{\pi^{r(r-1)}n^{rn}}{\Gamma_r(r)\Gamma_r(n)M}\mathbb{E}_{\mathbb{H}}\left[\prod_{i=1}^{r}\int_{\mathbb{R}_+}g_i(x_i)dx_i\right] \qquad (381)$$

where (379) follows from (371) and (375); (380) follows from (376); (381) holds because the integrand does not depend on Q, because $\mathbb{R}_{\geq}^r \subset \mathbb{R}_+^r$ and because the volume of $\widetilde{\mathcal{S}}_{r,r}$ (with respect to the Lebesgue measure on $\widetilde{\mathcal{S}}_{r,r}$) is $\pi^{r(r-1)}/\Gamma_r(r)$. After algebraic manipulations, we obtain

$$\int_{\mathbb{R}_+}g_i(x_i)dx_i = \frac{b_0^{[r-2i]^++1}}{n^{n+r-2i+1}}\bigg[\Gamma(n+r-2i+1, n+r-2i)$$
$$+ (n+r-2i)^{n+r-2i+1}\,e^{-(n+r-2i)}\bigg]. \qquad (382)$$

Substituting (382) into (381) and using (374), we obtain

$$1 - \epsilon' \leq \frac{c_{\mathrm{rx}}(n)}{M}. \qquad (383)$$

Note that the RHS of (383) is valid for every code.

APPENDIX VIII
PROOF OF THE CONVERSE PART OF THEOREM 9

In this appendix, we prove the converse asymptotic expansion for Theorem 9. More precisely, we show the following:

Proposition 20: Let the pdf of the fading matrix $\mathbb{H}$ satisfy the conditions in Theorem 9. Then

$$R_{\mathrm{rx}}^*(n,\epsilon) \leq C_\epsilon^{\mathrm{no}} + \mathcal{O}\left(\frac{\log n}{n}\right). \qquad (384)$$

Proof: Proceeding as in (158)–(162), we obtain from Theorem 6 that

$$(n-1)R_{\mathrm{rx}}^*(n-1,\epsilon)$$
$$\leq n\gamma - \log\left(\inf_{\mathsf{Q}\in\mathcal{U}_t^{\mathrm{e}}}\mathbb{P}[S_n^{\mathrm{rx}}(\mathsf{Q})\leq n\gamma]-\epsilon\right) + \log c_{\mathrm{rx}}(n) \qquad (385)$$

where $\gamma > 0$ is arbitrary. The third term on the RHS of (385) is upper-bounded by

$$\log c_{\mathrm{rx}}(n) \leq \frac{r^2}{2}\log n + \log\left(\mathbb{E}\left[\left(1+\rho\|\mathbb{H}\|_{\mathsf{F}}^2\right)^{\lfloor(r+1)^2/4\rfloor}\right]\right)$$
$$+ \mathcal{O}(1) \qquad (386)$$
$$= \frac{r^2}{2}\log n + \mathcal{O}(1). \qquad (387)$$

Here, (386) follows from algebraic manipulations, and (387) follows from the assumption (81), which ensures that the second term on the RHS of (386) is finite.

To evaluate $\mathbb{P}[S_n^{\mathrm{rx}}(\mathsf{Q}) \leq n\gamma]$ on the RHS of (385), we note that given $\mathbb{H} = \mathsf{H}$, the random variable $S_n^{\mathrm{rx}}(\mathsf{Q})$ is the sum of n i.i.d. random variables. Hence, using Theorem 15 (Appendix IV-A) and following similar steps as the ones reported in Appendix IV-A, we obtain

$$\mathbb{P}[S_n^{\mathrm{rx}}(\mathsf{Q}) \leq n\gamma \mid \mathbb{H} = \mathsf{H}] \geq q_n(\varphi_{\gamma,\mathsf{Q}}(\mathsf{H})) + \mathcal{O}\!\left(\frac{1}{n}\right) \quad (388)$$

where the function $\varphi_{\gamma,\mathsf{Q}} : \mathbb{C}^{t\times r} \mapsto \mathbb{R}$ is given by

$$\varphi_{\gamma,\mathsf{Q}}(\mathsf{H}) \triangleq \frac{\gamma - \log\det(\mathsf{I}_r + \mathsf{H}^{\mathrm{H}}\mathsf{Q}\mathsf{H})}{\sqrt{\mathrm{tr}(\mathsf{I}_r - (\mathsf{I}_r + \mathsf{H}^{\mathrm{H}}\mathsf{Q}\mathsf{H})^{-2})}} \quad (389)$$

the function $q_n(\cdot)$ was defined in (170), and the $\mathcal{O}(1/n)$ term is uniform in Q, γ and H. Let

$$U(\gamma,\mathsf{Q}) \triangleq \varphi_{\gamma,\mathsf{Q}}(\mathbb{H}). \quad (390)$$

Averaging (388) over $\mathbb{H}$, we obtain

$$\mathbb{P}[S_n^{\mathrm{rx}}(\mathsf{Q}) \leq n\gamma]$$
$$\geq \mathbb{E}\left[Q(-\sqrt{n}U(\gamma,\mathsf{Q}))\right]$$
$$- \mathbb{E}\left[\frac{[1 - nU^2(\gamma,\mathsf{Q})]^+ e^{-nU^2(\gamma,\mathsf{Q})/2}}{6\sqrt{n}}\right] + \mathcal{O}\!\left(\frac{1}{n}\right). \quad (391)$$

We proceed to lower-bound the first two terms on the RHS of (391). To this end, we show in Lemma 21 ahead that there exist $\delta_1 \in (0, C_\epsilon^{\mathrm{no}})$ and $\delta > 0$ such that $u \mapsto f_{U(\gamma,\mathsf{Q})}(u)$, where $f_{U(\gamma,\mathsf{Q})}$ denotes the pdf of $U(\gamma,\mathsf{Q})$, is continuously differentiable on $(-\delta,\delta)$, and that $f_{U(\gamma,\mathsf{Q})}(u)$ and $f'_{U(\gamma,\mathsf{Q})}(u)$ are uniformly bounded for every $\gamma \in (C_\epsilon^{\mathrm{no}} - \delta_1, C_\epsilon^{\mathrm{no}} + \delta_1)$, every $\mathsf{Q} \in \mathcal{U}_t^{\mathrm{e}}$, and every $u \in (-\delta,\delta)$. We then apply Lemma 17 in Appendix IV-C with A being a standard normal random variable and $B = U(\gamma,\mathsf{Q})$ to lower-bound the first term on the RHS of (391) for every $\delta > 0$ as

$$\mathbb{E}\left[Q(-\sqrt{n}U(\gamma,\mathsf{Q}))\right]$$
$$\geq \mathbb{P}\left[\log\det(\mathsf{I}_r + \mathbb{H}^{\mathrm{H}}\mathsf{Q}\mathbb{H}) \leq \gamma\right] - \frac{1}{n}\frac{2}{\delta^2}$$
$$- \frac{1}{n}\left(\frac{1}{\delta} + \frac{1}{2}\right)\sup_{u\in(-\delta,\delta)}\max\left\{f_{U(\gamma,\mathsf{Q})}(u), \left|f'_{U(\gamma,\mathsf{Q})}(u)\right|\right\}. \quad (392)$$

We upper-bound the second term on the RHS of (391) for $n > \delta^{-2}$ as

$$\mathbb{E}\left[\frac{\left|1 - nU^2(\gamma,\mathsf{Q})\right|^+ e^{-nU^2(\gamma,\mathsf{Q})/2}}{6\sqrt{n}}\right]$$
$$\leq \frac{1}{6\sqrt{n}}\sup_{u\in(-\delta,\delta)} f_{U(\gamma,\mathsf{Q})}(u) \int_{-1/\sqrt{n}}^{1/\sqrt{n}} \underbrace{(1 - nt^2)e^{-nt^2/2}}_{\leq 1}\, dt \quad (393)$$
$$\leq \frac{1}{3n}\sup_{u\in(-\delta,\delta)} f_{U(\gamma,\mathsf{Q})}(u). \quad (394)$$

The following lemma establishes that $f_{U(\gamma,\mathsf{Q})}$ and $f'_{U(\gamma,\mathsf{Q})}$ are indeed uniformly bounded.

Lemma 21: Let $\mathbb{H}$ have pdf $f_{\mathbb{H}}$ satisfying Conditions 1 and 2 in Theorem 9. Let $\varphi_{\gamma,\mathsf{Q}} : \mathbb{C}^{t\times r} \mapsto \mathbb{R}$ be defined as in (389) and let $U(\gamma,\mathsf{Q})$ with pdf $f_{U(\gamma,\mathsf{Q})}$ denote the random variable $\varphi_{\gamma,\mathsf{Q}}(\mathbb{H})$. Then, there exist $\delta_1 \in (0, C_\epsilon^{\mathrm{no}})$ and $\delta > 0$ such that $u \mapsto f_{U(\gamma,\mathsf{Q})}(u)$ is continuously differentiable on $(-\delta,\delta)$ and that

$$\sup_{\gamma\in(C_\epsilon^{\mathrm{no}}-\delta_1, C_\epsilon^{\mathrm{no}}+\delta_1)} \sup_{\mathsf{Q}\in\mathcal{U}_t^{\mathrm{e}}} \sup_{u\in(-\delta,\delta)} f_{U(\gamma,\mathsf{Q})}(u) < \infty \quad (395)$$

$$\sup_{\gamma\in(C_\epsilon^{\mathrm{no}}-\delta_1, C_\epsilon^{\mathrm{no}}+\delta_1)} \sup_{\mathsf{Q}\in\mathcal{U}_t^{\mathrm{e}}} \sup_{u\in(-\delta,\delta)} \left|f'_{U(\gamma,\mathsf{Q})}(u)\right| < \infty. \quad (396)$$

Proof: See Appendix VIII-A. ∎

Using (392), (394), and Lemma 21 in (391), and then (391) and (387) in (385), we obtain for every $\gamma \in (C_\epsilon^{\mathrm{no}} - \delta_1, C_\epsilon^{\mathrm{no}} + \delta_1)$ that

$$(n-1)R_{\mathrm{rx}}^*(n-1, \epsilon)$$
$$\leq n\gamma - \log\Big(\inf_{\mathsf{Q}\in\mathcal{U}_t^{\mathrm{e}}}\mathbb{P}[\log\det(\mathsf{I}_r + \mathsf{H}^{\mathrm{H}}\mathsf{Q}\mathsf{H}) \leq \gamma] - \epsilon$$
$$+ \mathcal{O}(1/n)\Big) + \mathcal{O}(\log n) \quad (397)$$
$$= n\gamma - \log\big(F_{\mathrm{no}}(\gamma) - \epsilon + \mathcal{O}(1/n)\big) + \mathcal{O}(\log n) \quad (398)$$

where (398) follows from (27).

We next set γ so that

$$F_{\mathrm{no}}(\gamma) - \epsilon + \mathcal{O}(1/n) = 1/n. \quad (399)$$

In words, we choose γ so that the argument of the logarithm in (398) is equal to $1/n$. Since the function $(\mathsf{Q}, R) \mapsto F_{\mathsf{Q}}(R)$ is continuous and $\mathcal{U}_t^{\mathrm{e}}$ is compact, by the maximum theorem [49, Sec. VI.3] the function $F_{\mathrm{no}}(R) = \inf_{\mathsf{Q}\in\mathcal{U}_t^{\mathrm{e}}} F_{\mathsf{Q}}(R)$ is continuous in R. This guarantees that such a γ indeed exists. We next show that, for sufficiently large n, this γ satisfies

$$|\gamma - C_\epsilon^{\mathrm{no}}| \leq \mathcal{O}(1/n). \quad (400)$$

This implies that, for sufficiently large n, γ belongs to the interval $(C_\epsilon^{\mathrm{no}} - \delta_1, C_\epsilon^{\mathrm{no}} + \delta_1)$. We then obtain (384) by combining (398) with (399) and (400), and dividing both sides of (398) by $n - 1$.

To prove (400), we note that by (83) and the definition of $\liminf$, there exists a $\delta_2 \in (0, \delta_1)$ such that

$$\inf_{\gamma\in(C_\epsilon^{\mathrm{no}}-\delta_2, C_\epsilon^{\mathrm{no}}+\delta_2)} \frac{F_{\mathrm{no}}(\gamma) - F_{\mathrm{no}}(C_\epsilon^{\mathrm{no}})}{\gamma - C_\epsilon^{\mathrm{no}}} > 0. \quad (401)$$

Substituting (401) into (399) and using that $F_{\mathrm{no}}(C_\epsilon^{\mathrm{no}}) = \epsilon$, we obtain (400). This concludes the proof of Proposition 20. ∎

A. Proof of Lemma 21

Throughout this section, we shall use const to indicate a finite constant that does not depend on any parameter of interest; its magnitude and sign may change at each occurrence. The proof of this lemma is technical and makes use of concepts from Riemannian geometry.

Denote by $\{\mathcal{M}_l\}$ the open subsets

$$\mathcal{M}_l \triangleq \{\mathsf{H} \in \mathbb{C}^{t\times r} : \|\mathsf{H}\|_{\mathsf{F}} < l\} \quad (402)$$

indexed by $l \in \mathbb{N}$. We shall use the following *flat Riemannian metric* [50, pp. 13 and 165] on $\mathcal{M}_l$

$$\langle \mathsf{H}_1, \mathsf{H}_2 \rangle \triangleq \mathrm{Re}\big\{\mathrm{tr}(\mathsf{H}_1^{\mathrm{H}}\mathsf{H}_2)\big\}. \quad (403)$$

Using this metric, we define the gradient ∇g of an arbitrary function $g : \mathcal{M}_l \mapsto \mathbb{R}$ as in (80). Note that the metric (403) induces a norm on the tangent space of $\mathcal{M}_l$, which can be identified with the Frobenius norm.

Our proof consists of two steps. Let $f_l(u)$ denote the pdf of the random variable $U(\gamma, \mathsf{Q})$ conditioned on $\mathbb{H} \in \mathcal{M}_l$. We first show that there exist $l_0 \in \mathbb{N}$, $\delta > 0$, and $\delta_1 \in (0, C_\epsilon^{\mathrm{no}})$ such that $f_l(u)$ and $f_l'(u)$ are uniformly bounded for every $\gamma \in (C_\epsilon^{\mathrm{no}} - \delta_1, C_\epsilon^{\mathrm{no}} + \delta_1)$, every $\mathsf{Q} \in \mathcal{U}_t^{\mathrm{e}}$, every $u \in [-\delta, \delta]$, and every $l \geq l_0$. We then show that $u \mapsto f_{U(\gamma, \mathsf{Q})}(u)$ is continuously differentiable on $(-\delta, \delta)$, and that for every $u \in (-\delta, \delta)$, the sequences $\{f_l(u)\}$ and $\{f_l'(u)\}$ converge uniformly to $f_{U(\gamma, \mathsf{Q})}(u)$ and $f_{U(\gamma, \mathsf{Q})}'(u)$, respectively, i.e.,

$$\lim_{l \to \infty} \sup_{u \in (-\delta, \delta)} \left| f_l(u) - f_{U(\gamma, \mathsf{Q})}(u) \right| = 0 \qquad (404)$$

$$\lim_{l \to \infty} \sup_{u \in (-\delta, \delta)} \left| f_l'(u) - f_{U(\gamma, \mathsf{Q})}'(u) \right| = 0 \qquad (405)$$

from which Lemma 21 follows.

1) Uniform Boundedness of $\{f_l\}$ and $\{f_l'\}$: To establish that $\{f_l\}$ and $\{f_l'\}$ are uniformly bounded, we shall need the following lemma.

Lemma 22: Let $\mathcal{M}$ be an oriented Riemannian manifold with Riemannian metric (403) and let $\varphi : \mathcal{M} \mapsto \mathbb{R}$ be a smooth function with $\|\nabla\varphi\|_{\mathsf{F}} \neq 0$ on $\mathcal{M}$. Let P be a random variable on $\mathcal{M}$ with smooth pdf f. Then,

1) the pdf f_* of $\varphi(P)$ at u is

$$f_*(u) = \int_{\varphi^{-1}(u)} f \frac{\mathrm{d}S}{\|\nabla\varphi\|_{\mathsf{F}}} \qquad (406)$$

where $\varphi^{-1}(u)$ denotes the preimage $\{x \in \mathcal{M} : \varphi(x) = u\}$ and $\mathrm{d}S$ denotes the surface area form on $\varphi^{-1}(u)$, chosen so that $\mathrm{d}S(\nabla\varphi) > 0$;

2) if the pdf f is compactly supported, then the derivative of f_* is

$$f_*'(u) = \int_{\varphi^{-1}(u)} \psi_1 \frac{\mathrm{d}S}{\|\nabla\varphi\|_{\mathsf{F}}} \qquad (407)$$

where ψ_1 is defined implicitly via

$$\psi_1 \mathrm{d}V = d\left(f \frac{\mathrm{d}S}{\|\nabla\varphi\|_{\mathsf{F}}} \right) \qquad (408)$$

with $\mathrm{d}V$ denoting the volume form on $\mathcal{M}$ and $d(\cdot)$ being exterior differentiation [29, p. 256].

Proof: To prove (406), we note that for arbitrary $a, b \in \mathbb{R}$

$$\int_a^b f_*(u) du = \int_{\varphi^{-1}((a,b))} f \mathrm{d}V \qquad (409)$$

$$= \int_a^b \left(\int_{\varphi^{-1}(u)} f \frac{\mathrm{d}S}{\|\nabla\varphi\|_{\mathsf{F}}} \right) du \qquad (410)$$

where (410) follows from the smooth coarea formula [51, p. 160]. This implies (406).

To prove (407), we shall use that for an arbitrary $\delta > 0$,

$$f_*(u + \delta) - f_*(u)$$
$$= \int_{\varphi^{-1}(u+\delta)} f \frac{\mathrm{d}S}{\|\nabla\varphi\|_{\mathsf{F}}} - \int_{\varphi^{-1}(u)} f \frac{\mathrm{d}S}{\|\nabla\varphi\|_{\mathsf{F}}} \qquad (411)$$

$$= \int_{\varphi^{-1}((u,u+\delta))} d\left(f \frac{\mathrm{d}S}{\|\nabla\varphi\|_{\mathsf{F}}} \right) \qquad (412)$$

$$= \int_{\varphi^{-1}((u,u+\delta))} \psi_1 \mathrm{d}V \qquad (413)$$

where in (412) we used Stoke's theorem [51, Th. III.7.2], that f is compactly supported, and that the restriction of the form $f \frac{\mathrm{d}S}{\|\nabla\varphi\|_{\mathsf{F}}}$ to $\varphi^{-1}((u, u + \delta))$ is also compactly supported; (413) follows from the definition of ψ_1 (see (408)). Equation (407) follows then from similar steps as in (409)–(410). ∎

Using Lemma 22, we obtain

$$f_l(u) = \int_{\varphi_{\gamma,\mathsf{Q}}^{-1}(u) \cap \mathcal{M}_l} \frac{f_{\mathbb{H}}}{\mathbb{P}[\mathbb{H} \in \mathcal{M}_l]} \frac{\mathrm{d}S}{\|\nabla\varphi_{\gamma,\mathsf{Q}}\|_{\mathsf{F}}} \qquad (414)$$

and

$$f_l'(u) = \int_{\varphi_{\gamma,\mathsf{Q}}^{-1}(u) \cap \mathcal{M}_l} \frac{\psi_1}{\mathbb{P}[\mathbb{H} \in \mathcal{M}_l]} \frac{\mathrm{d}S}{\|\nabla\varphi_{\gamma,\mathsf{Q}}\|_{\mathsf{F}}} \qquad (415)$$

where ψ_1 satisfies

$$\psi_1 \mathrm{d}V = d\left(f_{\mathbb{H}} \frac{\mathrm{d}S}{\|\nabla\varphi_{\gamma,\mathsf{Q}}\|_{\mathsf{F}}} \right). \qquad (416)$$

Since $\mathbb{P}[\mathbb{H} \in \mathcal{M}_l] \to 1$ as $l \to \infty$, there exists a l_0 such that $\mathbb{P}[\mathbb{H} \in \mathcal{M}_l] \geq 1/2$ for every $l \geq l_0$.

We next show that there exist $\delta > 0$, $0 < \delta_1 < C_\epsilon^{\mathrm{no}}$, such that for every $\gamma \in (C_\epsilon^{\mathrm{no}} - \delta_1, C_\epsilon^{\mathrm{no}} + \delta_1)$, every $u \in (-\delta, \delta)$, every $\mathsf{Q} \in \mathcal{U}_t^{\mathrm{e}}$, every $\mathsf{H} \in \varphi_{\gamma,\mathsf{Q}}^{-1}(u) \cap \mathcal{M}_l$, and every $l \geq l_0$

$$f_{\mathbb{H}}(\mathsf{H}) \leq \mathrm{const} \cdot \|\mathsf{H}\|_{\mathsf{F}}^{-2tr-3} \qquad (417)$$

$$|\psi_1(\mathsf{H})| \leq \mathrm{const} \cdot \|\mathsf{H}\|_{\mathsf{F}}^{-2tr-3} \qquad (418)$$

and

$$A_l(u) \triangleq \int_{\varphi_{\xi,\mathsf{Q}}^{-1}(u) \cap \mathcal{M}_l} \frac{\|\mathsf{H}\|_{\mathsf{F}}^{-2tr-3} \mathrm{d}S}{\|\nabla\varphi_{\gamma,\mathsf{Q}}\|_{\mathsf{F}}} \leq \mathrm{const}. \qquad (419)$$

The uniform boundedness of $\{f_l\}$ and $\{f_l'\}$ follows then by using the bounds (417)–(419) in (414) and (415).

Proof of (417): Since $f_{\mathbb{H}}(\mathsf{H})$ is continuous by assumption, it is uniformly bounded for every $\mathsf{H} \in \mathcal{M}_1$. Hence, (417) holds for every $\mathsf{H} \in \mathcal{M}_1$. For $\mathsf{H} \notin \mathcal{M}_1$, we have by (81)

$$f_{\mathbb{H}}(\mathsf{H}) \leq a \|\mathsf{H}\|_{\mathsf{F}}^{-2tr-\lfloor(1+r)^2/2\rfloor-1} \leq a \|\mathsf{H}\|_{\mathsf{F}}^{-2tr-3}. \qquad (420)$$

This proves (417).

Proof of (418): The surface area form $\mathrm{d}S$ on $\varphi_{\gamma,\mathsf{Q}}^{-1}(u) \cap \mathcal{M}_l$ is given by

$$\mathrm{d}S = \frac{\star d\varphi_{\gamma,\mathsf{Q}}}{\|\nabla\varphi_{\gamma,\mathsf{Q}}\|_{\mathsf{F}}} \qquad (421)$$

where $\star$ denotes the Hodge star operator [50, p. 103] induced by the metric (403). Using (421) and the definition of the Hodge star operator, the RHS of (416) becomes

$$d\left(\frac{f_{\mathbb{H}}}{\|\nabla\varphi_{\gamma,\mathsf{Q}}\|_{\mathsf{F}}^2}\right)\wedge\star d\varphi_{\gamma,\mathsf{Q}}+\frac{f_{\mathbb{H}}}{\|\nabla\varphi_{\gamma,\mathsf{Q}}\|_{\mathsf{F}}^2}\wedge d\star d\varphi_{\xi,\mathsf{Q}}$$
$$=\left(\frac{\langle\nabla f_{\mathbb{H}},\nabla\varphi_{\gamma,\mathsf{Q}}\rangle}{\|\nabla\varphi_{\gamma,\mathsf{Q}}\|_{\mathsf{F}}^2}-\frac{f_{\mathbb{H}}\langle\nabla\|\nabla\varphi_{\gamma,\mathsf{Q}}\|_{\mathsf{F}}^2,\nabla\varphi_{\gamma,\mathsf{Q}}\rangle}{\|\nabla\varphi_{\gamma,\mathsf{Q}}\|_{\mathsf{F}}^4}\right.$$
$$\left.-\frac{f_{\mathbb{H}}\cdot\Delta\varphi_{\gamma,\mathsf{Q}}}{\|\nabla\varphi_{\gamma,\mathsf{Q}}\|_{\mathsf{F}}^2}\right)\mathrm{d}V \tag{422}$$

where $\wedge$ denotes the wedge product [29, p. 237] and Δ denotes the Laplace operator [50, Eq. (3.1.6)].[11] From (422) we get

$$|\psi_1|=\left|\frac{\langle\nabla f_{\mathbb{H}},\nabla\varphi_{\gamma,\mathsf{Q}}\rangle}{\|\nabla\varphi_{\gamma,\mathsf{Q}}\|_{\mathsf{F}}^2}-\frac{f_{\mathbb{H}}\langle\nabla\|\nabla\varphi_{\gamma,\mathsf{Q}}\|_{\mathsf{F}}^2,\nabla\varphi_{\gamma,\mathsf{Q}}\rangle}{\|\nabla\varphi_{\gamma,\mathsf{Q}}\|_{\mathsf{F}}^4}\right.$$
$$\left.-\frac{f_{\mathbb{H}}\cdot\Delta\varphi_{\gamma,\mathsf{Q}}}{\|\nabla\varphi_{\gamma,\mathsf{Q}}\|_{\mathsf{F}}^2}\right| \tag{423}$$
$$\leq\frac{\|\nabla f_{\mathbb{H}}\|_{\mathsf{F}}}{\|\nabla\varphi_{\gamma,\mathsf{Q}}\|_{\mathsf{F}}}+\frac{f_{\mathbb{H}}\left\|\nabla\|\nabla\varphi_{\gamma,\mathsf{Q}}\|_{\mathsf{F}}^2\right\|_{\mathsf{F}}}{\|\nabla\varphi_{\gamma,\mathsf{Q}}\|_{\mathsf{F}}^3}+\frac{f_{\mathbb{H}}\cdot|\Delta\varphi_{\gamma,\mathsf{Q}}|}{\|\nabla\varphi_{\gamma,\mathsf{Q}}\|_{\mathsf{F}}^2} \tag{424}$$

where the last step follows from the triangle inequality and the Cauchy-Schwarz inequality.

We proceed to lower-bound $\|\nabla\varphi_{\gamma,\mathsf{Q}}\|_{\mathsf{F}}$. Using the definition of the gradient (80) together with the matrix identities [52, p. 29]

$$\det(\mathsf{I}+\varepsilon\mathsf{A})=1+\varepsilon\mathrm{tr}(\mathsf{A})+\mathcal{O}(\varepsilon^2),\quad\varepsilon\to0 \tag{425}$$
$$(\mathsf{I}+\varepsilon\mathsf{A})^{-1}=\mathsf{I}-\varepsilon\mathsf{A}+\mathcal{O}(\varepsilon^2),\qquad\varepsilon\to0 \tag{426}$$

for every bounded square matrix A, we obtain

$$\nabla\varphi_{\gamma,\mathsf{Q}}(\mathsf{H})=-\frac{2\mathsf{Q}\mathsf{H}\Phi^{-3}}{\left(\mathrm{tr}(\mathsf{I}_r-\Phi^{-2})\right)^{3/2}}$$
$$\times\left(\underbrace{\mathrm{tr}(\mathsf{I}_r-\Phi^{-2})\Phi^2+(\gamma-\log\det\Phi)\mathsf{I}_r}_{\triangleq\mathsf{T}}\right) \tag{427}$$

where $\Phi\triangleq\mathsf{I}_r+\mathsf{H}^{\mathrm{H}}\mathsf{Q}\mathsf{H}$.

Fix an arbitrary $\delta_1\in(0,C_\epsilon^{\mathrm{no}})$ and choose $\delta\in(0,(C_\epsilon^{\mathrm{no}}-\delta_1)/\sqrt{r})$. We first bound $\mathrm{tr}(\mathsf{I}_r-\Phi^{-2})$ as

$$r\geq\mathrm{tr}(\mathsf{I}_r-\Phi^{-2})\geq1-(1+\lambda_{\max}(\mathsf{H}^{\mathrm{H}}\mathsf{Q}\mathsf{H}))^{-2}. \tag{428}$$

It follows from the first inequality in (428) and from (389) that for every $u\in(-\delta,\delta)$

$$|\gamma-\log\det\Phi|=|u|\sqrt{\mathrm{tr}(\mathsf{I}_r-\Phi^{-2})}\leq\delta\sqrt{r}. \tag{429}$$

Using (429) and that the determinant is given by the product of the eigenvalues, we obtain that, for every $\gamma\in(C_\epsilon^{\mathrm{no}}-\delta_1,C_\epsilon^{\mathrm{no}}-\delta_1)$ and every $u\in(-\delta,\delta)$,

$$r\log(1+\lambda_{\max}(\mathsf{H}^{\mathrm{H}}\mathsf{Q}\mathsf{H}))\geq\log\det\Phi \tag{430}$$
$$\geq\gamma-\sqrt{r}\delta \tag{431}$$
$$\geq C_\epsilon^{\mathrm{no}}-\delta_1-\sqrt{r}\delta>0 \tag{432}$$

---

[11]The Laplace operator used here and in [50, Eq. (3.1.6)] differs from the usual one on $\mathbb{R}^n$, as defined in calculus, by a minus sign. See [50, Sec. 3.1] for a more detailed discussion.

which implies that

$$\lambda_{\max}(\mathsf{H}^{\mathrm{H}}\mathsf{Q}\mathsf{H})\geq e^{(C_\epsilon^{\mathrm{no}}-\delta_1-\sqrt{r}\delta)/r}-1>0. \tag{433}$$

Combing (433) with the second inequality in (428), we obtain

$$\mathrm{tr}(\mathsf{I}_r-\Phi^{-2})\geq1-e^{-2(C_\epsilon^{\mathrm{no}}-\delta_1-\sqrt{r}\delta)/r}. \tag{434}$$

We use (429) and (434) to lower-bound the smallest eigenvalue of the matrix T defined in (427) as

$$\lambda_{\min}(\mathsf{T})=\mathrm{tr}(\mathsf{I}_r-\Phi^{-2})\underbrace{\lambda_{\min}(\Phi^2)}_{\geq1}+(\gamma-\log\det\Phi) \tag{435}$$
$$\geq\mathrm{tr}(\mathsf{I}_r-\Phi^{-2})-\delta\sqrt{r} \tag{436}$$
$$\geq1-e^{-2(C_\epsilon^{\mathrm{no}}-\delta_1-\sqrt{r}\delta)/r}-\delta\sqrt{r}. \tag{437}$$

The RHS of (437) can be made positive if we choose δ sufficiently small, in which case T is invertible. We can theorefore lower-bound $\|\nabla\varphi_{\gamma,\mathsf{Q}}\|_{\mathsf{F}}$ as

$$\|\nabla\varphi_{\xi,\mathsf{Q}}\|_{\mathsf{F}}=\frac{2}{\left(\mathrm{tr}(\mathsf{I}_r-\Phi^{-2})\right)^{3/2}}\left\|\mathsf{Q}\mathsf{H}\Phi^{-3}\mathsf{T}\right\|_{\mathsf{F}} \tag{438}$$
$$\geq\frac{2}{r^{3/2}}\left\|\mathsf{Q}\mathsf{H}\Phi^{-3}\right\|_{\mathsf{F}}\cdot\frac{1}{\|\mathsf{T}^{-1}\|_{\mathsf{F}}} \tag{439}$$
$$\geq\frac{2}{r^{3/2}}\|\mathsf{Q}\mathsf{H}\|_{\mathsf{F}}\cdot\frac{1}{\|\Phi^3\|_{\mathsf{F}}}\cdot\frac{1}{\|\mathsf{T}^{-1}\|_{\mathsf{F}}}. \tag{440}$$

Here, we use the first inequality in (428) and the submultiplicativity of the Frobenius norm. The term $\|\mathsf{Q}\mathsf{H}\|_{\mathsf{F}}$ can be bounded as

$$\|\mathsf{Q}\mathsf{H}\|_{\mathsf{F}}\geq\frac{\|\mathsf{H}^{\mathrm{H}}\mathsf{Q}\mathsf{H}\|_{\mathsf{F}}}{\|\mathsf{H}\|_{\mathsf{F}}} \tag{441}$$
$$\geq\frac{\lambda_{\max}(\mathsf{H}^{\mathrm{H}}\mathsf{Q}\mathsf{H})}{\|\mathsf{H}\|_{\mathsf{F}}} \tag{442}$$
$$\geq\frac{e^{(C_\epsilon^{\mathrm{no}}-\delta_1-\sqrt{r}\delta)/r}-1}{\|\mathsf{H}\|_{\mathsf{F}}} \tag{443}$$

where (443) follows from (433).

The term $\|\Phi^3\|_{\mathsf{F}}$ in (440) can be upper-bounded as

$$\|\Phi^3\|_{\mathsf{F}}\leq\sqrt{r}(1+\lambda_{\max}(\mathsf{H}^{\mathrm{H}}\mathsf{Q}\mathsf{H}))^3 \tag{444}$$
$$\leq\sqrt{r}(1+\det\Phi)^3 \tag{445}$$
$$\leq\mathrm{const.} \tag{446}$$

Here, (446) follows from (429) and because $\gamma\leq C_\epsilon^{\mathrm{no}}+\delta$. Finally, $\|\mathsf{T}^{-1}\|_{\mathsf{F}}$ in (440) can be bounded as

$$\|\mathsf{T}^{-1}\|_{\mathsf{F}}\leq\sqrt{r}\lambda_{\max}(\mathsf{T}^{-1})=\frac{\sqrt{r}}{\lambda_{\min}(\mathsf{T})}. \tag{447}$$

The RHS of (447) is bounded because of (437). Substituting (443), (446) and (447) into (440), we conclude that

$$\|\nabla\varphi_{\gamma,\mathsf{Q}}\|^{-1}\leq\mathrm{const}\cdot\|\mathsf{H}\|_{\mathsf{F}}. \tag{448}$$

Following similar steps as the ones reported in (425)–(448), we can show that

$$\left\|\nabla\|\nabla\varphi_{\gamma,\mathsf{Q}}\|_{\mathsf{F}}^2\right\|_{\mathsf{F}}<\mathrm{const}\cdot\|\nabla\varphi_{\gamma,\mathsf{Q}}\|_{\mathsf{F}} \tag{449}$$

and

$$|\Delta\varphi_{\gamma,\mathsf{Q}}|<\mathrm{const.} \tag{450}$$

Substituting (448)–(450) into (424) and using the bounds (81) and (82), we obtain (418).

Proof of (419)*:* We begin by observing that for every $\mathsf{H} \in \varphi_{\gamma,\mathsf{Q}}^{-1}(u) \cap \mathcal{M}_l$, every $\gamma \in (C_\epsilon^{\mathrm{no}} - \delta_1, C_\epsilon^{\mathrm{no}} + \delta_1)$, every $u \in (-\delta, \delta)$ and every $\mathsf{Q} \in \mathcal{U}_t^{\mathrm{e}}$

$$\|\mathsf{H}\|_\mathsf{F}^2 \geq \frac{\mathrm{tr}(\mathsf{H}^\mathsf{H}\mathsf{Q}\mathsf{H})}{\mathrm{tr}(\mathsf{Q})} \tag{451}$$

$$\geq \frac{1}{\rho}\lambda_{\max}(\mathsf{H}^\mathsf{H}\mathsf{Q}\mathsf{H}) \tag{452}$$

$$\geq \frac{1}{\rho}\left(e^{(C_\epsilon^{\mathrm{no}} - \delta_1 - \sqrt{r}\delta)/r} - 1\right) \triangleq k_0. \tag{453}$$

Here, (451) follows from Cauchy-Schwarz inequality; (452) follows because $\mathrm{tr}(\mathsf{Q}) = \rho$ for every $\mathsf{Q} \in \mathcal{U}_t^{\mathrm{e}}$; (453) follows from (433). From (453) we conclude that

$$\left(\varphi_{\gamma,\mathsf{Q}}^{-1}((-\delta,\delta)) \cap \mathcal{M}_l\right) \subset \mathcal{M}' \triangleq \{\mathsf{H} \in \mathbb{C}^{t \times r} : \|\mathsf{H}\|_\mathsf{F} \geq \sqrt{k_0}\}. \tag{454}$$

To upper-bound $A_l(u)$, we note that

$$\int_{-\delta}^{\delta} A_l(u)du = \int_{\varphi_{\gamma,\mathsf{Q}}^{-1}((-\delta,\delta)) \cap \mathcal{M}_l} \|\mathsf{H}\|_\mathsf{F}^{-2tr-3}\,\mathrm{d}V \tag{455}$$

$$\leq \int_{\mathcal{M}'} \|\mathsf{H}\|_\mathsf{F}^{-2tr-3}\,\mathrm{d}V \tag{456}$$

$$= \mathrm{const} \cdot \int_{\sqrt{k_0}}^{\infty} x^{-4}dx \tag{457}$$

$$= \mathrm{const}. \tag{458}$$

Here, (455) follows from the smooth coarea formula [51, p. 160]; (456) follows from (454); (457) follows by writing the RHS of (456) in polar coordinates and by using that, by (433), $k_0 > 0$. By the mean value theorem, it follows from (458) that there exists a $u_0 \in (-\delta, \delta)$ satisfying

$$A_l(u_0) = \frac{\int_{-\delta}^{\delta} A_l(u)du}{2\delta} \leq \mathrm{const}. \tag{459}$$

Next, for every $u \in (u_0, \delta)$ we have that

$$A_l(u) - A_l(u_0) = \int_{\varphi_{\gamma,\mathsf{Q}}^{-1}(u) \cap \mathcal{M}_l} \frac{\|\mathsf{H}\|_\mathsf{F}^{-2tr-3}}{\|\nabla\varphi_{\gamma,\mathsf{Q}}\|_\mathsf{F}}\mathrm{d}S$$
$$- \int_{\varphi_{\gamma,\mathsf{Q}}^{-1}(u_0) \cap \mathcal{M}_l} \frac{\|\mathsf{H}\|_\mathsf{F}^{-2tr-3}}{\|\nabla\varphi_{\gamma,\mathsf{Q}}\|_\mathsf{F}}\mathrm{d}S \tag{460}$$

$$= \int_{\varphi_{\gamma,\mathsf{Q}}^{-1}((u_0,u)) \cap \mathcal{M}_l} d\left(\frac{\|\mathsf{H}\|_\mathsf{F}^{-2tr-3}}{\|\nabla\varphi_{\gamma,\mathsf{Q}}\|_\mathsf{F}}\mathrm{d}S\right) \tag{461}$$

$$= \int_{\varphi_{\gamma,\mathsf{Q}}^{-1}((u_0,u)) \cap \mathcal{M}_l} \psi_2\mathrm{d}V \tag{462}$$

where ψ_2 is defined implicitly via

$$\psi_2\mathrm{d}V = d\left(\frac{\|\mathsf{H}\|_\mathsf{F}^{-2tr-3}}{\|\nabla\varphi_{\gamma,\mathsf{Q}}\|_\mathsf{F}}\mathrm{d}S\right). \tag{463}$$

Here, (461) follows from Stokes' theorem. Following similar steps as the ones reported in (421)–(450), we obtain that

$$|\psi_2| \leq \mathrm{const} \cdot \|\mathsf{H}\|_\mathsf{F}^{-2tr-1}. \tag{464}$$

We can therefore upper-bound $A_l(u)$ as

$$A_l(u) = A_l(u_0) + \int_{\varphi_{\gamma,\mathsf{Q}}^{-1}((u_0,u)) \cap \mathcal{M}_l} \psi_2\mathrm{d}V \tag{465}$$

$$\leq \mathrm{const} + \int_{\mathcal{M}'} \mathrm{const} \cdot \|\mathsf{H}\|_\mathsf{F}^{-2tr-1}\,\mathrm{d}V \tag{466}$$

$$\leq \mathrm{const} + \int_{\sqrt{k_0}}^{\infty} \mathrm{const} \cdot x^{-2}dx \tag{467}$$

$$= \mathrm{const}. \tag{468}$$

Here, (465) follows from (462); (467) follows from (459), (464), and (454). Note that the bound (468) is uniform in γ, Q, u, and l. Following similar steps as the ones reported in (460)–(468), we obtain the same result for $u \in (-\delta, u_0)$. This proves (419).

2) Convergence of $f_l(u)$ and $f_l'(u)$: In this section, we will prove (404) and (405). By Lemma 22,

$$f_{U(\gamma,\mathsf{Q})}(u) = \int_{\varphi_{\gamma,\mathsf{Q}}^{-1}(u)} \frac{f_\mathbb{H}\mathrm{d}S}{\|\nabla\varphi_{\gamma,\mathsf{Q}}\|_\mathsf{F}}. \tag{469}$$

We have the following chain of inequalities

$$|f_l(u) - f_{U(\gamma,\mathsf{Q})}(u)|$$
$$\leq |\mathbb{P}[\mathbb{H} \in \mathcal{M}_l]f_l(u) - f_{U(\gamma,\mathsf{Q})}(u)|$$
$$\quad + |(1 - \mathbb{P}[\mathbb{H} \in \mathcal{M}_l])f_l(u)| \tag{470}$$

$$\leq \int_{\varphi_{\gamma,\mathsf{Q}}^{-1}(u) \cap (\mathbb{C}^{t \times r} \setminus \mathcal{M}_l)} \frac{f_\mathbb{H}\mathrm{d}S}{\|\nabla\varphi_{\gamma,\mathsf{Q}}\|_\mathsf{F}}$$
$$\quad + \mathrm{const} \cdot (1 - \mathbb{P}[\mathbb{H} \in \mathcal{M}_l]) \tag{471}$$

$$\leq \mathrm{const} \cdot \int_{\varphi_{\gamma,\mathsf{Q}}^{-1}(u) \cap (\mathbb{C}^{t \times r} \setminus \mathcal{M}_l)} \frac{\|\mathsf{H}\|_\mathsf{F}^{-2tr-3}\,\mathrm{d}S}{\|\nabla\varphi_{\gamma,\mathsf{Q}}\|_\mathsf{F}}$$
$$\quad + \mathrm{const} \cdot (1 - \mathbb{P}[\mathbb{H} \in \mathcal{M}_l]). \tag{472}$$

Here, (470) follows from the triangle inequality; (471) follows from (414) and because $\{f_l(u)\}$ is uniformly bounded; (472) follows from (417). Following similar steps as the ones reported in (455)–(468), we upper-bound the first term on the RHS of (472) as

$$\int_{\varphi_{\gamma,\mathsf{Q}}^{-1}(u) \cap (\mathbb{C}^{t \times r} \setminus \mathcal{M}_l)} \frac{\|\mathsf{H}\|_\mathsf{F}^{-2tr-3}\,\mathrm{d}S}{\|\nabla\varphi_{\gamma,\mathsf{Q}}\|_\mathsf{F}} \leq \frac{\mathrm{const}}{l}. \tag{473}$$

Substituting (473) into (472), and using that $\mathbb{P}[\mathbb{H} \in \mathcal{M}_l] \to 1$ as $l \to \infty$, we obtain that

$$\lim_{l \to \infty} \sup_{u \in (-\delta,\delta)} |f_l(u) - f_{U(\gamma,\mathsf{Q})}(u)| = 0. \tag{474}$$

This proves (404).

To prove (405), we proceed as follows. Let $C^1([-\delta, \delta])$ denote the set of continuously differentiable functions on the compact interval $[-\delta, \delta]$. The space $C^1([-\delta, \delta])$ is a Banach space (i.e., a complete normed vector space) when equipped with the C^1 norm [53, p. 92]

$$\|f\|_{C^1([-\delta,\delta])} = \sup_{x \in [-\delta,\delta]} (|f(x)| + |f'(x)|). \tag{475}$$

Following similar steps as in (460)–(468), we obtain that $\{f_l'\}$ is continuous on $[-\delta, \delta]$, i.e., the restriction of $\{f_l\}$ to $[-\delta, \delta]$

belongs to $C^1([-\delta,\delta])$. Moreover, following similar steps as in (470)–(474), we obtain that for all $m > l > 0$

$$\lim_{l\to\infty}\sup_{u\in[-\delta,\delta]}\Big(|f_m(u)-f_l(u)|+|f'_m(u)-f'_l(u)|\Big)=0. \quad (476)$$

This means that $\{f_l\}$ restricted to $[-\delta,\delta]$ is a Cauchy sequence, and, hence, converges in $C^1([-\delta,\delta])$ with respect to the C^1 norm (475). Moreover, by (474) the limit of $\{f_l\}$ is $f_{U(\gamma,\mathsf{Q})}$. Therefore, we have $f_{U(\gamma,\mathsf{Q})}\in C^1([-\delta,\delta])$, and $\{f'_l\}$ converges to $f'_{U(\gamma,\mathsf{Q})}$ with respect to the sup-norm $\|\cdot\|_\infty$. This proves (405).

APPENDIX IX
PROOF OF THE ACHIEVABILITY PART OF THEOREM 9

We prove the achievability asymptotic expansion for Theorem 9. More precisely, we prove the following:

Proposition 23: Assume that there exists a $\mathsf{Q}^* \in \mathcal{U}_t$ satisfying (64). Let $F_{\mathsf{Q}^*}(\cdot)$ be as in (87). Assume that the joint pdf of the nonzero eigenvalues of $\mathbb{H}^H\mathsf{Q}^*\mathbb{H}$ is continuously differentiable and that $F_{\mathsf{Q}^*}(\cdot)$ is differentiable and strictly increasing at C^{no}_ϵ, i.e.,

$$F'_{\mathsf{Q}^*}(C^{\mathrm{no}}_\epsilon) > 0. \quad (477)$$

Let $t^* = \mathrm{rank}(\mathsf{Q}^*)$. Then,

$$R^*_{\mathrm{no}}(n,\epsilon) \geq C^{\mathrm{no}}_\epsilon - (1+rt^*)\frac{\log n}{n} + \mathcal{O}\Big(\frac{1}{n}\Big). \quad (478)$$

Note that the conditions on the distribution of the fading matrix $\mathbb{H}$ under which Proposition 23 holds are less stringent than (and, because of Proposition 5 on p. 7 and Lemma 21 on p. 27, implied by) the conditions under which Proposition 20 (converse part of Theorem 9) holds.

Proof: The proof follows closely the proof of the achievability part of Theorem 3. Following similar steps as the ones reported in (253)–(259), we obtain

$$\mathbb{P}\Big[\prod_{i=1}^r B_i \leq \gamma_n\Big] \leq n^{rt^*}\gamma_n^{n-t^*-r}. \quad (479)$$

Setting $\tau = 1/n$ and $\gamma_n = \exp(-C^{\mathrm{no}}_\epsilon + \mathcal{O}(1/n))$ in Theorem 4, and using (479), we obtain

$$\frac{\log M}{n} \geq C^{\mathrm{no}}_\epsilon - (1+rt^*)\frac{\log n}{n} + \mathcal{O}\Big(\frac{1}{n}\Big). \quad (480)$$

To conclude the proof, we show that there exists indeed a $\gamma_n = \exp(-C^{\mathrm{no}}_\epsilon + \mathcal{O}(1/n))$ satisfying

$$\mathbb{P}\big[\sin^2\{\mathsf{I}_{n,t^*}, \sqrt{n}\mathsf{I}_{n,t^*}\mathsf{U}\mathbb{H}+\mathbb{W}\} \leq \gamma_n\big] \geq 1-\epsilon+1/n \quad (481)$$

where $\mathsf{U} \in \mathbb{C}^{t^*\times t}$ satisfies $\mathsf{U}^H\mathsf{U} = \mathsf{Q}^*$. Hereafter, we restrict ourselves to $\gamma_n \in \big(e^{-C^{\mathrm{no}}_\epsilon-\delta}, e^{-C^{\mathrm{no}}_\epsilon+\delta}\big)$ for some $\delta \in (0, C^{\mathrm{no}}_\epsilon)$. Let $m^* \triangleq \min\{t^*, r\}$. Consider the SVD of $\mathsf{U}\mathbb{H}$

$$\mathsf{U}\mathbb{H} = \mathsf{L}\underbrace{\begin{pmatrix} \Sigma_{m^*} & \mathbf{0}_{m^*\times(r-m^*)} \\ \mathbf{0}_{(t^*-m^*)\times m^*} & \mathbf{0}_{(t^*-m^*)\times(r-m^*)} \end{pmatrix}}_{\triangleq\Sigma}\mathsf{V}^H \quad (482)$$

where $\mathsf{L} \in \mathbb{C}^{t^*\times t^*}$ and $\mathsf{V} \in \mathbb{C}^{r\times r}$ are unitary matrices, $\Sigma_{m^*} = \mathrm{diag}\{\sqrt{\lambda_1}, \dots, \sqrt{\lambda_{m^*}}\}$ with $\lambda_1, \dots, \lambda_{m^*}$ being the m^* largest

eigenvalues of $\mathsf{H}^H\mathsf{Q}^*\mathsf{H}$, and $\mathbf{0}_{a,b}$ denotes the all zero matrix of size $a \times b$. Conditioned on $\mathbb{H} = \mathsf{H}$, we have

$$\sin^2\{\mathsf{I}_{n,t^*}, \sqrt{n}\mathsf{I}_{n,t^*}\mathsf{U}\mathsf{H}+\mathbb{W}\}$$
$$= \sin^2\{\mathsf{I}_{n,t^*}\mathsf{L}, (\sqrt{n}\mathsf{I}_{n,t^*}\mathsf{U}\mathsf{H}+\mathbb{W})\mathsf{V}\} \quad (483)$$
$$= \sin^2\{\widetilde{\mathsf{L}}\mathsf{I}_{n,t^*}\mathsf{L}, \widetilde{\mathsf{L}}(\sqrt{n}\mathsf{I}_{n,t^*}\mathsf{U}\mathsf{H}+\mathbb{W})\mathsf{V}\} \quad (484)$$
$$= \sin^2\{\mathsf{I}_{n,t^*}, \sqrt{n}\mathsf{I}_{n,t^*}\Sigma+\mathbb{W}\} \quad (485)$$

where

$$\widetilde{\mathsf{L}} \triangleq \begin{pmatrix} \mathsf{L}^H & \mathbf{0}_{(n-t^*)\times t^*} \\ \mathbf{0}_{t^*\times(n-t^*)} & \mathsf{I}_{n-t^*} \end{pmatrix} \quad (486)$$

is unitary. Here, (483) follows because $\mathrm{span}(\mathsf{A}) = \mathrm{span}(\mathsf{AB})$ for every invertible matrix B; (484) follows because the principal angles between two subspaces are invariant under simultaneous rotation of the two subspaces; (485) follows because $\mathbb{W}$ is isotropically distributed, which implies that $\widetilde{\mathsf{L}}\mathbb{W}\mathsf{V}$ has the same distribution as $\mathbb{W}$.

Let $\boldsymbol{e}_j$ and $\boldsymbol{W}_j$ be the jth column of I_{n,t^*} and $\mathbb{W}$, respectively. Then

$$\mathbb{P}\big[\sin^2\{\mathsf{I}_{n,t^*}, \sqrt{n}\mathsf{I}_{n,t^*}\Sigma+\mathbb{W}\} \leq \gamma_n\big]$$
$$\geq \mathbb{P}\Big[\prod_{j=1}^{m^*}\sin^2\{\boldsymbol{e}_j, \sqrt{n\Lambda_j}\boldsymbol{e}_j+\boldsymbol{W}_j\} \leq \gamma_n\Big] \quad (487)$$
$$= \mathbb{P}\Big[\prod_{j=1}^{m^*}\sin^2\{\boldsymbol{e}_1, \sqrt{n\Lambda_j}\boldsymbol{e}_1+\boldsymbol{W}_j\} \leq \gamma_n\Big]. \quad (488)$$

Here, (487) follows from Lemma 13 (Appendix I) and (488) follows by symmetry. By repeating the same steps as in (263)–(279), we obtain from (488) that there exists a $\gamma_n = \exp(-C^{\mathrm{no}}_\epsilon + \mathcal{O}(1/n))$ that satisfies (481). ∎

APPENDIX X
PROOF OF THEOREM 11 (DISPERSION OF CODES WITH ISOTROPIC CODEWORDS)

Using Proposition 23 with Q^* replaced by $(\rho/t)\mathsf{I}_t$, we obtain

$$R^*_{\mathrm{no,iso}}(n,\epsilon) \geq C^{\mathrm{iso}}_\epsilon + \mathcal{O}\Big(\frac{\log n}{n}\Big). \quad (489)$$

Since $R^*_{\mathrm{no,iso}}(n,\epsilon) \leq R^*_{\mathrm{rx,iso}}(n,\epsilon)$, the proof is completed by showing that

$$R^*_{\mathrm{rx,iso}}(n,\epsilon) \leq C^{\mathrm{iso}}_\epsilon + \mathcal{O}\Big(\frac{\log n}{n}\Big). \quad (490)$$

To prove (490), we evaluate the converse bound (78) in the large-n limit. This evaluation follows closely the proof of (56) in Appendix IV. Let $\Lambda_1 \geq \cdots \geq \Lambda_m$ be the ordered nonzero eigenvalues of $\mathbb{H}^H\mathbb{H}$. Following similar steps as in (158)–(162), we obtain that for every $\gamma > 0$

$$R^*_{\mathrm{rx,iso}}(n,\epsilon)$$
$$\leq \gamma - \frac{1}{n}\log\Big(\mathbb{P}[S^{\mathrm{rx}}_n((\rho/t)\mathsf{I}_t) \leq n\gamma] - \epsilon\Big) + \mathcal{O}\Big(\frac{1}{n}\Big) \quad (491)$$

with $S^{\mathrm{rx}}_n(\cdot)$ defined in (70). To evaluate the second term on the RHS of (491), we proceed as in Appendix IV-A to obtain

$$\mathbb{P}[S^{\mathrm{rx}}_n((\rho/t)\mathsf{I}_t) \leq n\gamma \mid \boldsymbol{\Lambda} = \boldsymbol{\lambda}] \geq q_n(\tilde{u}_\gamma(\boldsymbol{\lambda})) + \frac{k_1}{n} \quad (492)$$

for γ in a certain neighborhood of $C_\epsilon^{\mathrm{iso}}$. Here, the function $q_n(\cdot)$ is given in (170); the function $\tilde{u}_\gamma(\cdot) : \mathbb{R}_+^m \mapsto \mathbb{R}$ is defined as

$$\tilde{u}_\gamma(\boldsymbol{\lambda}) \triangleq \frac{\gamma - \sum_{j=1}^m \log(1 + \rho\lambda_j/t)}{\sqrt{m - \sum_{j=1}^m (1 + \rho\lambda_j/t)^{-2}}}; \qquad (493)$$

$\boldsymbol{\Lambda} = [\Lambda_1, \ldots, \Lambda_m]$; and k_1 is a finite constant independent of γ and $\boldsymbol{\lambda}$. A lower bound on $\mathbb{P}[S_n^{\mathrm{rx}}((\rho/t)\mathsf{I}_t) \le n\gamma]$ follows then by averaging both sides of (492) with respect to $\boldsymbol{\Lambda}$

$$\mathbb{P}[S_n^{\mathrm{rx}}((\rho/t)\mathsf{I}_t) \le n\gamma] \ge \mathbb{E}\big[q_n\big(\tilde{u}_\gamma(\boldsymbol{\Lambda})\big)\big] + \frac{k_1}{n}. \quad (494)$$

Proceeding as in (199)–(206) and using the assumption that the joint pdf of $\Lambda_1, \ldots, \Lambda_m$ is continuously differentiable, we obtain that for all $\gamma \in (C_\epsilon^{\mathrm{iso}} - \delta, C_\epsilon^{\mathrm{iso}} + \delta)$

$$\mathbb{E}\big[q_n\big(\tilde{u}_\gamma(\boldsymbol{\Lambda})\big)\big] \ge \mathbb{P}\left[\sum_{j=1}^m \log(1 + \rho\Lambda_j/t) \le \gamma\right] + \frac{k_2}{n} \quad (495)$$

for some $\delta > 0$ and $k_2 > -\infty$. Substituting (495) into (494), we see that for every $\gamma \in (C_\epsilon^{\mathrm{iso}} - \delta, C_\epsilon^{\mathrm{iso}} + \delta)$

$$\mathbb{P}[S_n^{\mathrm{rx}}((\rho/t)\mathsf{I}_t) \le n\gamma]$$
$$\ge \mathbb{P}\left[\sum_{j=1}^m \log(1 + \rho\Lambda_j/t) \le \gamma\right] + \frac{k_1 + k_2}{n} \quad (496)$$
$$= F_{\mathrm{iso}}(\gamma) + \frac{k_1 + k_2}{n}. \qquad (497)$$

The proof of (490) is concluded by repeating the same steps as in (164)–(165).

References

[1] E. Biglieri, J. Proakis, and S. Shamai (Shitz), "Fading channels: Information-theoretic and communications aspects," *IEEE Trans. Inf. Theory*, vol. 44, no. 6, pp. 2619–2692, Oct. 1998.

[2] D. N. C. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge, U.K.: Cambridge Univ. Press, 2005.

[3] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wirel. Personal Commun.*, vol. 6, pp. 311–335, 1998.

[4] M. Effros and A. Goldsmith, "Capacity definitions and coding strategies for general channels with receiver side information," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Cambridge MA, Aug. 1998, p. p. 39.

[5] T. S. Han, *Information-Spectrum Methods in Information Theory*. Berlin, Germany: Springer-Verlag, 2003.

[6] M. Effros, A. Goldsmith, and Y. Liang, "Generalizing capacity: New definitions and capacity theorems for composite channels," *IEEE Trans. Inf. Theory*, vol. 56, no. 7, pp. 3069–3087, Jul. 2010.

[7] S. Verdú and T. S. Han, "A general formula for channel capacity," *IEEE Trans. Inf. Theory*, vol. 40, no. 4, pp. 1147–1157, Jul. 1994.

[8] L. H. Ozarow, S. S. (Shitz), and A. D. Wyner, "Information theoretic considerations for cellular mobile radio," *IEEE Trans. Inf. Theory*, vol. 43, no. 2, pp. 359–378, May 1994.

[9] Y. Polyanskiy, H. V. Poor, and S. Verdú, "Channel coding rate in the finite blocklength regime," *IEEE Trans. Inf. Theory*, vol. 56, no. 5, pp. 2307–2359, May 2010.

[10] Y. Polyanskiy and S. Verdú, "Scalar coherent fading channel: dispersion analysis," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Saint Petersburg, Russia, Aug. 2011, pp. 2959–2963.

[11] 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.

[12] J. Hoydis, R. Couillet, and P. Piantanida, "The second-order coding rate of the MIMO Rayleigh block-fading channel," *IEEE Trans. Inf. Theory*, 2013, submitted. [Online]. Available: http://arxiv.org/abs/1303.3400

[13] C. Potter, K. Kosbar, and A. Panagos, "On achievable rates for MIMO systems with imperfect channel state information in the finite length regime," *IEEE Trans. Commun.*, vol. 61, no. 7, pp. 2772–2781, Jul. 2013.

[14] W. Yang, G. Caire, G. Durisi, and Y. Polyanskiy, "Finite blocklength channel coding rate under a long-term power constraint," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Honolulu, HI, USA, Jul. 2014.

[15] G. Caire, G. Taricco, and E. Biglieri, "Optimum power control over fading channels," *IEEE Trans. Inf. Theory*, vol. 45, no. 5, pp. 1468–1489, May 1999.

[16] V. V. Petrov, *Sums of Independent Random Variates*. Springer-Verlag, 1975, translated from the Russian by A. A. Brown.

[17] İ. E. Telatar, "Capacity of multi-antenna Gaussian channels," *Eur. Trans. Telecommun.*, vol. 10, pp. 585–595, Nov. 1999.

[18] A. M. Tulino and S. Verdú, "Random matrix theory and wireless communications," in *Foundations and Trends in Communications and Information Theory*. Delft, The Netherlands: now Publishers, 2004, vol. 1, no. 1, pp. 1–182.

[19] N. Johnson, S. Kotz, and N. Balakrishnan, *Continuous Univariate Distributions*, 2nd ed. New York: Wiley, 1995, vol. 2.

[20] E. Abbe, S.-L. Huang, and İ. E. Telatar, "Proof of the outage probability conjecture for MISO channels," *IEEE Trans. Inf. Theory*, vol. 59, no. 5, pp. 2596–2602, May 2013.

[21] C. E. Shannon, "Probability of error for optimal codes in a Gaussian channel," *Bell Syst. Tech. J.*, vol. 38, no. 3, pp. 611–656, May 1959.

[22] A. Barg and D. Y. Nogin, "Bounds on packings of spheres in the Grassmann manifold," *IEEE Trans. Inf. Theory*, vol. 48, no. 9, pp. 2450–2454, Sep. 2002.

[23] Y. Polyanskiy, "Channel coding: non-asymptotic fundamental limits," Ph.D. dissertation, Princeton University, 2010.

[24] Y. Polyanskiy, H. V. Poor, and S. Verdú, "Dispersion of the Gilbert-Elliott channel," *IEEE Trans. Inf. Theory*, vol. 57, no. 4, pp. 1829–1848, Apr. 2011.

[25] M. Tomamichel and V. Y. F. Tan, "ϵ-capacities and second-order coding rates for channels with general state," *IEEE Trans. Inf. Theory*, May 2013, submitted. [Online]. Available: http://arxiv.org/abs/1305.6789

[26] W. Yang, G. Durisi, T. Koch, and Y. Polyanskiy, "Quasi-static SIMO fading channels at finite blocklength," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT 2013)*, Istanbul, Turkey, Jul. 2013.

[27] E. MolavianJazi and J. N. Laneman, "On the second-order coding rate of block fading channels," in *Proc. Allerton Conf. Commun., Contr., Comput.*, Monticello, IL, USA, Oct. 2013, to appear.

[28] A. T. James, "Distribution of matrix variates and latent roots derived from normal samples," *Ann. Math. Statist.*, vol. 35, pp. 475–501, 1964.

[29] J. R. Munkres, *Analysis on Manifolds*. Redwood City, CA: Addison-Wesley, 1991.

[30] V. Y. F. Tan and M. Tomamichel, "The third-order term in the normal approximation for the AWGN channel," *IEEE Trans. Inf. Theory*, Dec. 2013, submitted. [Online]. Available: http://arxiv.org/abs/1311.2337

[31] 3GPP TS 36.212, "Technical specification group radio access network; evolved universal terrestrial radio access (E-UTRA); multiplexing and channel coding (release 10)," Dec. 2012.

[32] P. Robertson, E. Villebrun, and P. Hoeher, "A comparison of optimal and sub-optimal MAP decoding algorithms operating in the log domain," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Seattle, USA, Jun. 1995, pp. 1009–1013.

[33] S. Sesia, I. Toufik, and M. Baker, Eds., *LTE–The UMTS Long Term Evolution: From Theory to Practice*, 2nd ed. UK: Wiley, 2011.

[34] J. Miao and A. Ben-Israel, "On principal angles between subspaces in $\mathbb{R}^n$," *Linear Algebra Appl.*, vol. 171, pp. 81–98, 1992.

[35] S. Afriat, "Orthogonal and oblique projectors and the characteristics of pairs of vector spaces," *Proc. Cambridge Phil. Soc.*, vol. 53, no. 4, pp. 800–816, 1957.

[36] P. Absil, P. Koev, and A. Edelman, "On the largest principal angle between random subspaces," *Linear Algebra Appl.*, vol. 414, no. 1, pp. 288–294, Apr. 2006.

[37] J. C. Roh and B. D. Rao, "Design and analysis of MIMO spatial multiplexing systems with quantized feedback," *IEEE Trans. Signal Process.*, vol. 54, no. 8, pp. 2874–2886, Aug. 2006.

[38] J. Neyman and E. S. Pearson, "On the problem of the most efficient tests of statistical hypotheses," *Philosophical Trans. Royal Soc. A*, vol. 231, pp. 289–337, 1933.

[39] 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.

[40] W. Rudin, *Principles of Mathematical Analysis*, 3rd ed. Singapore: McGraw-Hill, 1976.

[41] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.

[42] N. Johnson, S. Kotz, and N. Balakrishnan, *Continuous Univariate Distributions*, 2nd ed. New York: Wiley, 1995, vol. 1.

[43] F. Qi and Q.-M. Luo, "Bounds for the ratio of two gamma functions—from Wendel's and related inequalities to logarithmically completely monotonic functions," *Banach J. Math. Anal.*, vol. 6, no. 2, pp. 132–158, 2012.

[44] G. Grimmett and D. Stirzaker, *Probability and Random Processes*, 3rd ed. New York, USA: Oxford University Press, 2001.

[45] Y. Polyanskiy, "Saddle point in the minimax converse for channel coding," *IEEE Trans. Inf. Theory*, vol. 59, no. 5, pp. 2576–2595, May 2013.

[46] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 1985.

[47] A. Edelman, "Eigenvalues and condition numbers of random matrices," Ph.D. dissertation, MIT, May 1989.

[48] A. W. Marshall and I. Olkin, *Inequalities: Theory of Majorization and its Application*. Orlando, FL: Academic Press, Inc., 1979.

[49] C. Berge, *Topological Spaces*. Edinburg, UK: Oliver and Boyd, 1963.

[50] J. Jost, *Riemannian Geometry and Geometric Analysis*, 6th ed. Berlin, Germany: Springer, 2011.

[51] I. Chavel, *Riemannian Geometry: A Modern Introduction*. Cambridge, UK: Cambridge Univ. Press, 2006.

[52] H. Lütkepohl, *Handbook of Matrices*. Chichester, England: John Wiley & Sons, 1996.

[53] J. K. Hunter and B. Nachtergaele, *Applied Analysis*. Singapore: World Scientific Publishing Co., 2001.

Giuseppe Durisi (S'02, M'06, SM'12) received the Laurea degree summa cum laude and the Doctor degree both from Politecnico di Torino, Italy, in 2001 and 2006, respectively. From 2002 to 2006, he was with Istituto Superiore Mario Boella, Torino, Italy. From 2006 to 2010 he was a postdoctoral researcher at ETH Zurich, Switzerland. Since 2010 he has been with Chalmers University of Technology, Gothenburg, Sweden, where is now associate professor. He held visiting researcher positions at IMST, Germany, University of Pisa, Italy, ETH Zurich, Switzerland, and Vienna University of Technology, Austria.

Dr. Durisi is a senior member of the IEEE. He is the recipient of the 2013 IEEE ComSoc Best Young Researcher Award for the Europe, Middle East, and Africa Region, and is co-author of a paper that won a "student paper award" at the 2012 International Symposium on Information Theory, and of a paper that won the 2013 IEEE Sweden VT-COM-IT joint chapter best student conference paper award. He served as TPC member in several IEEE conferences, and is currently publications editor of the IEEE Transactions on Information Theory. His research interests are in the areas of communication and information theory.

Tobias Koch (S'02, M'09) is a Visiting Professor with the Signal Theory and Communications Department of Universidad Carlos III de Madrid (UC3M), Spain. He received the M.Sc. degree in electrical engineering (with distinction) in 2004 and the Ph.D. degree in electrical engineering in 2009, both from ETH Zurich, Switzerland. From June 2010 until May 2012 he was a Marie Curie Intra-European Research Fellow with the University of Cambridge, UK. He was also a research intern at Bell Labs, Murray Hill, NJ in 2004 and at Universitat Pompeu Fabra (UPF), Spain, in 2007. He joined UC3M in 2012. His research interests include digital communication theory and information theory.

Dr. Koch is serving as Vice Chair of the Spain Chapter of the IEEE Information Theory Society in 2013–2014.

Wei Yang (S'09) received the B.E. degree in communication engineering and M.E. degree in communication and information systems from the Beijing University of Posts and Telecommunications, Beijing, China, in 2008 and 2011, respectively. He is currently pursuing a Ph.D. degree in electrical engineering at Chalmers University of Technology, Gothenburg, Sweden. From July to August 2012, he was a visiting student at the Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, Cambridge, MA.

Mr. Yang is the recipient of a Student Paper Award at the 2012 IEEE International Symposium on Information Theory (ISIT), Cambridge, MA, and the 2013 IEEE Sweden VT-COM-IT joint chapter best student conference paper award. His research interests are in the areas of information and communication theory.

Yury Polyanskiy (S'08-M'10) is an Assistant Professor of Electrical Engineering and Computer Science at MIT. He received the M.S. degree in applied mathematics and physics from the Moscow Institute of Physics and Technology, Moscow, Russia in 2005 and the Ph.D. degree in electrical engineering from Princeton University, Princeton, NJ in 2010. In 2000-2005 he lead the development of the embedded software in the Department of Surface Oilfield Equipment, Borets Company LLC (Moscow). Currently, his research focuses on basic questions in information theory, error-correcting codes, wireless communication and fault-tolerant circuits. Over the years Dr. Polyanskiy won the 2013 NSF CAREER award, the 2011 IEEE Information Theory Society Paper Award and Best Student Paper Awards at the 2008 and 2010 IEEE International Symposia on Information Theory (ISIT).