A High-SNR Normal Approximation for Single-Antenna Rayleigh Block-Fading Channels

Alejandro Lancho†‡, Tobias Koch†‡, and Giuseppe Durisi∗

†Signal Theory and Communications Department, Universidad Carlos III de Madrid, 28911, Leganés, Spain
‡Gregorio Marañón Health Research Institute, 28007, Madrid, Spain.
∗Chalmers University of Technology, Gothenburg, Sweden
Emails: {alancho, koch}@tsc.uc3m.es, durisi@chalmers.se

Abstract

This paper concerns the maximal achievable rate at which data can be transmitted over a non-coherent, single-antenna, Rayleigh block-fading channel using an error-correcting code of a given blocklength with a block-error probability not exceeding a given value. In particular, a high-SNR normal approximation of the maximal achievable rate is presented that becomes accurate as the signal-to-noise ratio (SNR) and the number of coherence intervals $L$ over which we code tend to infinity. Numerical analyses suggest that the approximation is accurate already at SNR values of $15\,\text{dB}$.

I. INTRODUCTION

There exists an increasing interest in the problem of transmitting short packets in wireless communications. For example, the vast majority of wireless connections in the fifth generation of cellular systems (5G) will most likely be originated by autonomous machines and devices, which predominantly exchange short packets. It is also expected that enhanced mobile-broadband services will be complemented by new services that target systems requiring reliable real-time communication with stringent requirements on latency and reliability. For more details see [1] and references therein. While in the absence of latency constraints, capacity and outage capacity provide accurate benchmarks for the throughput achievable in wireless communication systems, for low-latency wireless communications a more refined analysis of the maximal achievable rate as a function of the blocklength is needed. Such an analysis is provided in this paper.

Let $R^*(n, \epsilon)$ denote the maximal achievable rate at which data can be transmitted using an error-correcting code of a determined length $n$ with a block-error probability not larger than $\epsilon$. Building upon Dobrushin’s and Strassen’s asymptotic results, Polyanskiy, Poor and Verdú showed that for various channels with a positive capacity $C$, the maximal achievable rate can be tightly approximated by

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

where $V$ denotes the channel dispersion, $Q^{-1}(\epsilon)$ denotes the inverse $Q$-function, and $O\left(\log n/n\right)$ comprises terms that decay no slower than $\log n/n$. The approximation that follows from (1) by ignoring the $O\left(\log n/n\right)$ terms is sometimes referred to as normal approximation.

The work by Polyanskiy et al. [2] has been generalized to some wireless communication channels; see, e.g., [3]–[10]. However, the channel dispersion in the non-coherent setting—where neither the transmitter nor the receiver have a priori knowledge of the realizations of the fading coefficients—is only known in the quasi-static case [5], where it is zero. For general non-coherent block-fading channels, non-asymptotic bounds on the maximal achievable rate that can be evaluated numerically were presented, e.g., in [3], [7]. Obtaining an expression for the channel dispersion for non-coherent block-fading channels is difficult because for such channels the capacity-achieving input

A. Lancho and T. Koch have received funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (grant agreement number 714161), from the 7th European Union Framework Programme under Grant 333680, from the Spanish Ministerio de Economía y Competitividad under Grants TEC2013-41718-R, RYC-2014-16332 and TEC2016-78434-C3-3-R (AEI/FEDER, EU), from an FPU fellowship from the Spanish Ministerio de Educación, Cultura y Deporte under Grant FPU014/01274 and from the Comunidad de Madrid under Grant S2103/ICE-2845. G. Durisi has been supported by the Swedish Research Council under Grant 2012-4571 and 2016-03293.
distribution is in general unknown. Thus, the standard approach of obtaining expressions of the form (1), which consists of first evaluating non-asymptotic upper and lower bounds on $R^*(n, \epsilon)$ for the capacity-achieving input distribution and then analyzing these bounds in the limit as $n \to \infty$, cannot be followed.

In this paper, we present an expression similar to (1) of the maximal rate $R^*(L, \epsilon, \rho)$ achievable over non-coherent, single-antenna, Rayleigh block-fading channels using error-correcting codes that span $L$ coherence intervals, have a block-error probability not larger than $\epsilon$, and satisfy the power constraint $\rho$. By replacing the capacity and channel dispersion by asymptotically tight approximations, we obtain a high-SNR normal approximation of $R^*(L, \epsilon, \rho)$.

The obtained normal approximation is useful in two ways: On the one hand, it complements the non-asymptotic bounds provided in [3], [7], whose evaluation is computationally demanding. On the other hand, it allows for a mathematical analysis of $R^*(L, \epsilon, \rho)$.

The rest of this paper is organized as follows. Section II presents the system model. Section III introduces the most important quantities and defines the notation used in this paper. Section IV presents the main result of the presented results. Some of the proofs are deferred to the appendices.

II. SYSTEM MODEL

We consider a single-antenna Rayleigh block-fading channel with coherence interval $T$. For this channel model, the input-output relation within the $\ell$-th coherence interval is given by

$$\mathbf{Y}_\ell = \mathbf{H}_\ell \mathbf{X}_\ell + \mathbf{W}_\ell$$

(2)

where $\mathbf{X}_\ell$ and $\mathbf{Y}_\ell$ are $T$-dimensional, complex valued, random vectors containing the input and output signals, respectively; $\mathbf{W}_\ell$ is the additive noise, which is assumed to be a random vector with independent and identically distributed (i.i.d.) zero-mean, unit-variance, circularly-symmetric, complex Gaussian entries; and $\mathbf{H}_\ell$ is Rayleigh fading, i.e., it is a zero-mean, unit-variance, circularly-symmetric, complex Gaussian random variable. We assume that $\mathbf{H}_\ell$ and $\mathbf{W}_\ell$ are independent and take on independent realizations over successive coherence intervals. We further assume that the joint law of $(\mathbf{H}_\ell, \mathbf{W}_\ell)$ does not depend on the channel inputs. We consider a non-coherent setting where transmitter and receiver are aware of the distribution of $\mathbf{H}_\ell$ but not of its realization.

We next introduce the notion of a channel code. For simplicity, we shall restrict ourselves to codes whose blocklength $n$ satisfies $n = LT$, where $L$ denotes the number of blocks of length $T$ needed to transmit the whole code. An $(M, L, \epsilon, \rho)$ code for the channel (2) consists of the following:

1) An encoder $f: \{1, \ldots, M\} \to \mathbb{C}^{LT}$ that maps the message $A \in \{1, \ldots, M\}$ to a codeword $\mathbf{X}^L = [\mathbf{X}_1, \ldots, \mathbf{X}_L]$.

The codewords are assumed to satisfy the power constraint

$$||\mathbf{X}_\ell||^2 \leq TP, \quad \ell = 1, \ldots, L.$$  

(3)

Since the variance of $\mathbf{H}_\ell$ and of the entries of $\mathbf{W}_\ell$ are normalized to one, $\rho$ in (3) can be interpreted as the average SNR at the receiver.

2) A decoder $g: \mathbb{C}^{LT} \to \{1, \ldots, M\}$ satisfying a maximum error probability constraint

$$\max_{1 \leq a \leq M} \mathbb{P} \left[ g(\mathbf{Y}^L) \neq A | A = a \right] \leq \epsilon$$

(4)

where $\mathbf{Y}^L = [\mathbf{Y}_1, \ldots, \mathbf{Y}_L]$ is the channel output induced by the transmitted codeword $\mathbf{X}^L = f(a)$ according to (2).

The maximal coding rate is defined as

$$R^*(L, \epsilon, \rho) \triangleq \sup \left\{ \frac{\log M}{LT} : \exists (M, L, \epsilon, \rho) \text{ code} \right\}.$$  

(5)

1In contrast to [7], where the power constraint (3) is assumed to hold with equality, here we consider the more general case where the power constraint may also be satisfied with strict inequality.
where \( (\cdot)^H \) denotes Hermitian transposition. We shall refer to the distribution \( \mathbb{P}_{\mathbf{X}_\ell}^{(U)} \), according to which \( \mathbf{X}^L = \sqrt{T\rho} \mathbf{U}^L \) (where \( \mathbf{U}^L = [\mathbf{U}_1, \ldots, \mathbf{U}_L] \) are i.i.d. and uniformly distributed on the unit sphere in \( \mathbb{C}^T \) as unitary space-time modulation (USTM) [11]. This distribution is relevant because it gives rise to a lower bound on capacity that is asymptotically tight at high SNR [12], [13]. In fact, it can be shown that this lower-bound accurately approximates capacity already for intermediate SNR values. For example, [3, Fig. 1] illustrates that the lower-bound is indistinguishable from an upper bound on capacity given in [3, Eq. (17)] for \( \rho \geq 10 \text{ dB} \). The outputs \( \mathbf{Y}^L \) induced by the USTM input distribution have pdf \( q^\ell (y) = \prod_{i=1}^L q^\ell_i (y_i) \), where [3, Eq. (18)]

\[
q^\ell_i (y) = \frac{e^{-\|y\|^2/(1+T\rho)}\|y\|^{2(1-T)}\Gamma(T(T - 1, \frac{T\rho\|y\|^2}{1+T\rho})(1+\frac{1}{T\rho})^{T-1}}{\pi^T (1+T\rho)}, \quad y \in \mathbb{C}^T.
\]

Here, \( \gamma(\cdot, \cdot) \) denotes the regularized lower incomplete gamma function [14, Sec. 6.5] and \( \Gamma(\cdot) \) denotes the gamma function [14, Sec. 6.1.1].

Conditioned on \( \|\mathbf{X}_\ell\|^2 = T\alpha_\ell \), \( \alpha_\ell \in [0, \rho] \), the distributions of \( \|\mathbf{Y}^\ell_\ell \mathbf{X}_\ell\|^2 \) and \( \|\mathbf{Y}_\ell\|^2 \) are as follows:

\[
\|\mathbf{Y}^\ell_\ell \mathbf{X}_\ell\|^2 \overset{\triangle}{=} T\alpha_\ell(1+T\alpha_\ell)Z_{1,\ell}
\]

\[
\|\mathbf{Y}_\ell\|^2 \overset{\triangle}{=} (1+T\alpha_\ell)Z_{1,\ell} + Z_{2,\ell}
\]

where \( \overset{\triangle}{=} \) denotes equality in distribution, \( \{Z_{1,\ell}, \ell \in \mathbb{Z}\} \) is a sequence of i.i.d. Gamma(1,1)-distributed random variables, and \( \{Z_{2,\ell}, \ell \in \mathbb{Z}\} \) is a sequence of i.i.d. Gamma\((T-1, 1)\)-distributed random variables (with Gamma\((z, q)\) denoting the gamma distribution with parameters \( z \) and \( q \)). Here, \( \alpha_\ell \) can be thought of as the power allocated over the coherence interval \( \ell \). We will omit the subscript \( \ell \) when immaterial.

We next introduce some notation and preliminary results that will be helpful in the remainder of the paper. The \textit{information density} between \( \mathbf{X}^L \) and \( \mathbf{Y}^L \) is denoted by

\[
i(\mathbf{X}^L; \mathbf{Y}^L) \overset{\triangle}{=} \log \left( \frac{p_{\mathbf{Y}^L | \mathbf{X}^L}(\mathbf{Y}^L | \mathbf{X}^L)}{p_{\mathbf{Y}^L}(\mathbf{Y}^L)} \right)
\]

where \( p_{\mathbf{X}^L}(\mathbf{Y}^L) \) is the output distribution induced by the input distribution. When the input distribution is USTM, the conditional information density, conditioned on \( \|\mathbf{X}_\ell\|^2 = T\rho \), can be expressed as

\[
i(\mathbf{X}^L; \mathbf{Y}^L) = \sum_{\ell=1}^L i_\ell(\rho)
\]

where

\[
i_\ell(\rho) = (T - 1) \log(T\rho) - \log T(T) - \frac{T\rho Z_{2,\ell}}{1 + T\rho}
\]

\[+ (T - 1) \log \left( \frac{(1+T\rho)Z_{1,\ell} + Z_{2,\ell}}{1+T\rho} \right) - \log \tilde{\gamma} \left( T - 1, \frac{T\rho((1+T\rho)Z_{1,\ell} + Z_{2,\ell})}{1+T\rho} \right).
\]

Consider the following bounds on the \( \log \tilde{\gamma}(\cdot, \cdot) \) term:

\[
0 \leq - \log \tilde{\gamma} \left( T - 1, \frac{T\rho((1+\|\mathbf{X}_\ell\|^2)Z_{1,\ell} + Z_{2,\ell})}{1+\|\mathbf{X}_\ell\|^2} \right) \leq (T - 1) \log \left( 1 + \frac{\Gamma(T)\frac{T\rho}{1+\|\mathbf{X}_\ell\|^2}Z_{1,\ell} + Z_{2,\ell}}{T\rho} \right).
\]

where the right-most inequality follows by Alzer [15, Th. 1], and the left-most inequality follows because the normalized gamma function is between 0 and 1. Using the left-most inequality in [13], we can lower-bound [12] by

\[
i_\ell(\rho) \overset{\triangle}{=} (T - 1) \log(T\rho) - \log T(T) - \frac{T\rho Z_{2,\ell}}{1 + T\rho} + (T - 1) \log \left( \frac{(1+T\rho)Z_{1,\ell} + Z_{2,\ell}}{1+T\rho} \right).
\]
We next compute the first moment of (14), denoted by \( I(\rho) \triangleq E[J(\rho)] \), as
\[
I(\rho) = (T - 1) \log(T \rho) - \log \Gamma(T) - \frac{T \rho(T - 1)}{1 + T \rho} \\
- (T - 1) \log(1 + T \rho) + (T - 1) \log((1 + T \rho)Z_1 + Z_2)] \\
= (T - 1) \log(T \rho) - \log \Gamma(T) - (T - 1) \frac{T \rho}{1 + T \rho} \\
+ \psi(T - 1) - \log(1 + T \rho) + \frac{2 F_1(1, T - 1; T; \frac{T \rho}{1 + T \rho})}{T - 1}
\]  
(15)
where the expected value has been solved using \[16\] Eq. 4.337-1 to integrate with respect to \( Z_1 \), and \[16\] Eq. 4.352-1, \[16\] Eq. 3.381-4, and \[17\] Eq. 4.2.20 to integrate with respect to \( Z_2 \). Here, \( \psi(\cdot) \) denotes the digamma function \[14\] Sec. 6.3.2 and \( \psi(\cdot; \cdot; \cdot; \cdot) \) denotes the hypergeometric function \[16\] Sec. 9.1.

We define the mismatched information density as
\[
j(X^L; Y^L) \triangleq \log \left( \frac{p_{Y^L|X^L}(Y^L | X^L)}{q_{Y^L}(Y^L)} \right)
\]  
(16)
Using this definition together with (6) and (7), the mismatched information density \( J(X^L; Y^L) \) can be written as
\[
j(X^L; Y^L) = \sum_{\ell=1}^{L} j_\ell(X_\ell; Y_\ell)
\]  
(17)
where
\[
j_\ell(X_\ell; Y_\ell) = \log \left( \frac{1 + T \rho}{\Gamma(T)} \right) + \frac{|X^H_\ell X_\ell|^2}{1 + \|X_\ell\|^2} - T \rho \|Y_\ell\|^2 - (T - 1) \log \left( \frac{T \rho \|Y_\ell\|^2}{1 + T \rho} \right) \\
- \log(1 + \|X_\ell\|^2) - \log \gamma \left( T - 1, \frac{T \rho \|Y_\ell\|^2}{1 + T \rho} \right).
\]  
(18)
By (8) and (9), \( j_\ell(\alpha_\ell) \) depends on \( X_\ell \) only via \( \|X_\ell\|^2 = T \alpha_\ell \). We can thus express \( j(X_\ell; Y_\ell) \triangleq j_\ell(\alpha_\ell) \) as
\[
j_\ell(\alpha_\ell) = (T - 1) \log(T \rho) - \log \Gamma(T) - \frac{(T \rho - T \alpha_\ell)Z_{1,\ell}}{1 + T \rho} - T \rho Z_{2,\ell} + \log \left( \frac{1 + T \rho}{1 + T \alpha_\ell} \right) \\
+ (T - 1) \log \left( \frac{(1 + T \alpha_\ell)Z_{1,\ell} + Z_{2,\ell}}{1 + T \rho} \right) - \log \gamma \left( T - 1, \frac{T \rho((1 + T \alpha_\ell)Z_{1,\ell} + Z_{2,\ell})}{1 + T \rho} \right).
\]  
(19)
Note that \( j_\ell(\rho) = i_\ell(\rho) \) when the input distribution is USTM.

Let
\[
\bar{j}_\ell(\alpha_\ell) \triangleq (T - 1) \log(T \rho) - \log \Gamma(T) - \frac{(T \rho - T \alpha_\ell)Z_{1,\ell}}{1 + T \rho} - T \rho Z_{2,\ell} + \log \left( \frac{1 + T \rho}{1 + T \alpha_\ell} \right) \\
+ (T - 1) \log \left( \frac{(1 + T \alpha_\ell)Z_{1,\ell} + Z_{2,\ell}}{1 + T \rho} \right)
\]  
(20)
where \( \beta(T, \rho) \triangleq \Gamma(T) \frac{1 + T \rho}{1 + T \alpha_\ell} \). By (13), we have with probability 1
\[
j_\ell(\alpha_\ell) \leq \bar{j}_\ell(\alpha_\ell), \quad \alpha_\ell \in [0, \rho].
\]  
(21)
Next we compute the conditional expected value of (20) given \( \|X_\ell\|^2 = T \alpha_\ell \), denoted by \( \bar{J}(\alpha_\ell) \triangleq E[\bar{j}_\ell(\alpha_\ell)] \), as
\[
\bar{J}(\alpha_\ell) = (T - 1) \log(T \rho) - \log \Gamma(T) - \frac{T \rho - T \alpha_\ell}{1 + T \rho} - T \rho(T - 1) \\
+ \log \left( \frac{1 + T \rho}{1 + T \alpha_\ell} \right) - (T - 1) \log(1 + T \rho) + (T - 1) \log((1 + T \alpha_\ell)Z_1 + Z_2 + \beta(T, \rho)).
\]  
(22)
Let \( J(\alpha_t) = \mathbb{E}[j_t(\alpha_t)] \) and \( I(\rho) = \mathbb{E}[i_t(\rho)] \). Note that \( J(\cdot) \) and \( I(\rho) \) bound the capacity. Indeed, on the one hand we have

\[
C(\rho) = \lim_{L \to \infty} \sup_{p_x: \|X_1\|_p \leq T \rho} \frac{\mathbb{E}[i(X_1; Y_1)]}{LT} \leq \sup_{0 \leq \alpha \leq \rho} \frac{J(\alpha)}{T} \leq \sup_{0 \leq \alpha \leq \rho} \frac{\tilde{J}(\alpha)}{T}
\]

where the first inequality follows from [18, Th. 5.1], and the second inequality follows from (21). On the other hand,

\[
C(\rho) \geq \frac{I(\rho)}{T} \geq \frac{\tilde{I}(\rho)}{T}
\]

where the first inequality follows because USTM is a valid input distribution and the second inequality follows by (10), (12) and (13).

Let

\[
U(\rho) \triangleq \mathbb{E}\left[(i_t(\rho) - I(\rho))^2\right]
\]

\[
\tilde{V}_\rho(\alpha) \triangleq \mathbb{E}\left[(\tilde{j}_t(\alpha) - \tilde{J}(\alpha))^2\right]
\]

where the subscript \( \rho \) in \( \tilde{V}_\rho(\alpha) \) is introduced to highlight that \( \tilde{V}_\rho(\alpha) \) depends both on \( \alpha \) and \( \rho \), but it is omitted when \( \alpha = \rho \). In Lemma 3 (Appendix I) and Lemma 6 (Appendix IV), we show that \( I(\rho), U(\rho), \tilde{J}(\rho) \) and \( \tilde{V}_\rho(\alpha) \) can be approximated as

\[
I(\rho) = \bar{I}(\rho) + o_\rho(1)
\]

\[
U(\rho) = \bar{U} + o_\rho(1)
\]

\[
\tilde{J}(\rho) = \bar{J}(\rho) + o_\rho(1)
\]

\[
\tilde{V}(\rho) = \bar{\tilde{V}} + o_\rho(1)
\]

where \( o_\rho(1) \) comprises terms that are independent of \( L \) and that vanish as \( \rho \to \infty \). The closed form expression of \( \bar{I}(\rho) \) in (26a) is defined in (15). Moreover, \( \bar{U} \) in (26b) is given by

\[
\bar{U} \triangleq (T - 1)^2 \frac{\pi^2}{6} + (T - 1).
\]

IV. MAIN RESULT

The main result of this paper is a high-SNR normal approximation of \( R^*(L, \epsilon, \rho) \) presented in Section IV-A. A discussion of this approximation is provided in Section IV-B.

A. High-SNR normal approximation

Theorem 1: Assume that \( T > 2 \) and that \( 0 < \epsilon < 1/2 \). Then, \( R^*(L, \epsilon, \rho) \) can be approximated as

\[
R^*(L, \epsilon, \rho) = \frac{\bar{I}(\rho)}{T} + o_\rho(1) - \sqrt{\frac{\bar{U} + o_\rho(1)}{LT^2}} Q^{-1}(\epsilon) + \mathcal{O}_L\left(\frac{\log L}{L}\right)
\]

where \( \mathcal{O}_L(L/\log L) \) comprises terms that are independent of \( \rho \) and that decay in \( L \) no slower than \( (\log L)/L \). The quantities \( \bar{I}(\rho) \) and \( \bar{U} \) in (28) are defined in (15) and (27), respectively.

Proof: See Section IV-

Ignoring the \( \mathcal{O}_L(L/\log L) \) and the \( o_\rho(1) \) terms in (28), we obtain the following high-SNR normal approximation:

\[
R^*(L, \epsilon, \rho) \approx \frac{\bar{I}(\rho)}{T} - \sqrt{\frac{\bar{U}}{LT^2}} Q^{-1}(\epsilon).
\]

The quantity \( \frac{\bar{I}(\rho)}{T} \) is a high-SNR approximation of the information rate achievable with i.i.d. USTM inputs; cf. [19, Eq. (12)] (see also [3, Eq. (5)]). It is shown in [11, Th. 4] that \( \frac{\bar{I}(\rho)}{T} \) is an asymptotically-tight lower bound on the capacity \( C(\rho) \) in the sense that

\[
\lim_{\rho \to \infty}\left\{ C(\rho) - \frac{\bar{I}(\rho)}{T} \right\} = 0.
\]
According to Theorem 1, \( \tilde{I}/T^2 \) can be viewed as a high-SNR approximation of the dispersion.

Equation (15) provides a closed-form expression for \( I(\rho) \). However, it contains a hypergeometric function, which is difficult to analyze mathematically. We therefore present also a simplified expression that is less accurate than (15) but easier to analyze. Specifically, by Lemma 3 in Appendix I,

\[
I(\rho) = (T - 1) \log(T \rho) - \log \Gamma(T) - (T - 1)(1 + \gamma) + o(\rho(1))
\]  

(31)

where \( \gamma \) denotes Euler’s constant.

**B. Numerical Discussion**

We illustrate the accuracy of the approximation given in (29) by means of numerical examples, both for the case when we use the expression for \( I(\rho) \) given in (15), and for the case when we use the approximation (31).

In Figs. 1 and 2, we show the high-SNR normal approximation (29) as a function of \( L = n/T \) for a fixed coherence interval \( T \) and for different SNR values. We evaluate \( I(\rho) \) using both the exact expression (15) as well as the approximation (31). We further plot a non-asymptotic (in \( \rho \) and \( L \)) lower bound on \( R^*(L, \epsilon, \rho) \) that is based on the Dependence Testing (DT) lower bound [2, Th. 22] (see (32) below) as well as a non-asymptotic (in \( \rho \) and \( L \)) upper bound on \( R^*(L, \epsilon, \rho) \) that is based on the Meta Converse (MC) upper bound [2, Th. 31] (see (42) below). We finally plot \( I(\rho) \). Observe that the high-SNR approximation \( R^*(L, \epsilon, \rho) \) is accurate already for \( \rho = 15 \) dB when we use the exact expression (15) for \( I(\rho) \). For \( \rho = 25 \) dB, the expression is accurate even when we approximate \( I(\rho) \) using the simplified expression (31). Further observe that the high-SNR normal approximation is pessimistic for \( \rho = 15 \) dB and optimistic for \( \rho = 25 \) dB. This suggests that the term \( o(\rho(1)) \) is negative and that \( O_L(\log L/L) \) is positive.

In Fig. 3, we show the high-SNR normal approximation (29) as a function of the coherence interval \( T \) for a fixed blocklength \( n \) (hence \( L \) is inversely proportional to \( T \)). For comparison, we also show the DT bound, evaluated for an USTM input distribution and implemented in the communication toolbox SPECTRE [20, Sec. 3.1], and the MC bound given in (44) below. Finally, we present the normal approximation that was given in [5, Eq. (95)] for quasi-static multiple-input multiple-output (MIMO) block-fading channels. To adapt the quasi-static MIMO block-fading
Figure 2. Bounds on $R^*(L, \epsilon, \rho)$ for non-coherent Rayleigh block-fading channels for $\rho = 25$ dB, $T = 20$, $\epsilon = 10^{-3}$.

Figure 3. Bounds on $R^*(L, \epsilon, \rho)$ for non-coherent Rayleigh block-fading channels when $LT = 500$, $\epsilon = 10^{-3}$, $\rho = 25$ dB.
channel to our system model, we replace $H$ in [5] by an $L \times L$ diagonal matrix with diagonal entries $H_1, \ldots, H_L$. Observe that the high-SNR normal approximation (29) is accurate for $L \geq 10$ and then becomes less accurate as $L$ decreases. Further observe that the normal approximation for the quasi-static case [5, Eq. (95)], which is tailored towards the case where $L$ is small, becomes accurate only for $L \leq 2$.

V. PROOF OF THEOREM 1

In this section we provide the proof of Theorem 1. The proof is based on a lower bound on $R^*(L, \epsilon, \rho)$, given in Section V-A, and on an upper bound on $R^*(L, \epsilon, \rho)$, given in Section V-B. Since these bounds coincide up to error terms of the order $O_L(\log L/L)$ and $o_\rho(1)$ (compare (33) with (57) below), together with (26a) and (26b) they prove (28).

A. Dependence Testing (DT) Lower Bound

As the capacity-achieving input distribution is unknown for Rayleigh block-fading channels, to obtain a lower bound on $R^*(L, \rho, \epsilon)$, we evaluate the DT bound [2, Th. 22] for USTM inputs. Thus, assume that $X^L \sim P^{(U)}_{X^L}$, which implies $Y^L \sim q^{(U)}_{Y^L}$. Then, the cumulative distribution function $P[i(x^L; y^L) \leq \alpha]$ does not depend on $x^L$, see also [7, App. A]. A lower bound on $R^*(L, \epsilon, \rho)$ follows therefore from the DT bound (maximum probability of error) [2, Th. 22], which after a standard change of measure can be stated as follows: there exists a code with $M$ codewords and maximal probability of error $\rho$ not exceeding

$$
\epsilon \leq P\left[i(x^L; y^L) \leq \log(M - 1)\right] + (M - 1)E\left[e^{-\epsilon(x^L; y^L)}I[i(x^L; y^L) > \log(M - 1)]\right]
$$

(32)

where $I\{\cdot\}$ denotes the indicator function. To show that (32) yields the lower bound

$$
R^*(L, \rho, \epsilon) \geq \frac{I(\rho)}{T} - \sqrt{\frac{U(\rho)}{LT^2}}Q^{-1}(\epsilon) + O_L\left(\frac{1}{L}\right)
$$

(33)

(where $O_L(1/L)$ comprises terms that are independent of $\rho$ and that decay in $L$ no slower than $1/L$) we follow almost verbatim the steps (258)–(267) in [2] (with $\gamma$ in [2] replaced by $M - 1$). The main difference is that in our case $U(\rho)$ and $B(\rho)$ (cf. [2, Eq. (254)])

$$
B(\rho) \triangleq \frac{6E\left[|i_\rho(\rho) - I(\rho)|^3\right]}{U(\rho)^{3/2}}
$$

(34)

depend on $\rho$. To ensure that the term $O_L(1/L)$ is independent of $\rho$, we thus need that both $U(\rho)$ and $B(\rho)$ are bounded in $\rho$. We then apply the Berry-Esseen theorem to obtain [2, Eq. (259)] with $B(\rho)$ replaced by an upper bound $B(T)$, followed by [2, Eq. (261)–(266)], that yields

$$
R^*(L, \epsilon, \rho) \geq \frac{I(\rho)}{T} - \sqrt{\frac{U(\rho)}{LT^2}}Q^{-1}(\tau)
$$

(35)

where

$$
\tau \triangleq \epsilon - \left(\frac{2\log 2}{\sqrt{2\pi}} + 5B(T)\right)\frac{1}{\sqrt{L}}.
$$

(36)

Finally, by the differentiability of $Q^{-1}(\cdot)$ we obtain (33).

To show that $U(\rho)$ and $B(\rho)$ are bounded in $\rho$, we resort to Lemmas 4 (Appendix B), 5 (Appendix III), and 7 (Appendix V). Indeed, by Lemma 4, $\hat{V}_\rho(\alpha)$ satisfies for $\rho(1 - \delta) \leq \alpha \leq \rho$

$$
\hat{V}_\rho(\alpha) \geq \left(\frac{T\rho}{1 + T\rho}\right)^2(T - 1) + o_\rho(1) + o_\delta(1)
$$

(37)

where $o_\delta(1)$ comprises terms that are independent of $L$ and $\rho$ and that vanish as $\delta \downarrow 0$. sufficiently large $\rho > 0$ and sufficiently small $\delta > 0$. Together with (26d), this implies that for sufficiently large $\rho_0 > 0$

$$
U(\rho) \geq \left(\frac{T\rho}{1 + T\rho}\right)^2(T - 1) - \frac{1}{2}, \quad \rho \geq \rho_0.
$$

(38)
Furthermore, Lemma 5 shows that for every $\rho_0 > 0$ there exists a $U_{\text{UB}}(T, \rho_0)$ that is independent of $\rho$ and that satisfies

$$U(\rho) \leq U_{\text{UB}}(T, \rho_0), \quad \rho \geq \rho_0.$$  \hfill (39)

Finally, Lemma 7 shows that for every $\rho_0 > 0$ there exists a $S(T, \rho_0)$ that is independent of $\rho$ and that satisfies

$$E \left( |i_\ell(\rho) - I(\rho)|^3 \right) \leq S(T, \rho_0), \quad \rho \geq \rho_0.$$  \hfill (40)

Combining (38) and (40), it follows that for every $\rho \geq \rho_0$ there exists a $B(T, \rho_0)$ that is independent of $\rho$ and that satisfies

$$B(\rho) \leq \frac{6S(T, \rho_0)}{\left( \frac{T \rho}{1 + T \rho} \right)^3 (\frac{T - 1}{2})^{3/2}} \triangleq B(T, \rho_0), \quad \rho \geq \rho_0.$$  \hfill (41)

This concludes the proof of the lower bound (33).

### B. Meta Converse (MC) Upper Bound

An upper bound on $R^*(L, \epsilon, \rho)$ follows from the MC bound [2, Th. 31] computed for the auxiliary pdf $q^{(U)}_{Y^L}$

$$R^*(L, \epsilon, \rho) \leq \sup_{\alpha \in [0, \rho]^L} \log \left( \frac{1}{\beta(\alpha, q_{Y^L}^{(U)})} \right).$$  \hfill (42)

Here, $\alpha = (\alpha_1, \ldots, \alpha_L)$ denotes the vector of power allocations, and $\beta(\alpha, q_{Y^L}^{(U)})$ denotes the minimum probability of error under hypothesis $q_{Y^L}^{(U)}$ if the probability of error under hypothesis $p_{Y^L|X^L = \epsilon^L}$ does not exceed $\epsilon$ [2, Eq. (100)]. Note that $\beta(\alpha, q_{Y^L}^{(U)})$ depends on $x^L$ only via $\alpha$.

Fix an arbitrary $0 < \delta < 1$. The following lemma demonstrates that, without loss of optimality, in at least half of the coherence intervals $\alpha_\ell$ is between $\rho(1 - \delta)$ and $\rho$.

**Lemma 1**: For sufficiently large $L$ and $\rho$, the supremum in (42) can be replaced by a supremum over $\alpha \in A_{\rho, \delta}$, where

$$A_{\rho, \delta} \triangleq \{ \alpha \in [0, \rho]^L : L_{\alpha}(\delta) \geq L/2 \}$$  \hfill (43)

and $L_{\alpha}(\delta)$ denotes the number of $\alpha_\ell$’s in $\alpha$ that satisfy $\rho(1 - \delta) \leq \alpha_\ell \leq \rho$.

**Proof**: See Appendix VII.

In the following, we implicitly assume that $L \geq L_0$ and $\rho \geq \rho_0$ for sufficiently large $L_0$ and $\rho_0$. Applying Lemma 1 to (42), upper-bounding the right-hand side (RHS) of (42) using [2, Eq. (106)] and because by (21) $j_\ell(\alpha_\ell) \leq j(\alpha_\ell)$, we obtain

$$R^*(L, \epsilon, \rho) \leq \sup_{\alpha \in A_{\rho, \delta}} \left\{ \frac{\log \xi(\alpha)}{LT} \right\} - \frac{\log \left( 1 - \epsilon - \frac{\sum_{\ell=1}^L j_\ell(\alpha_\ell)}{LT} \right)}{LT}. \right\}$$  \hfill (44)

for an arbitrary $\xi : [0, \rho]^L \rightarrow (0, \infty)$.

Let

$$B(\alpha) \triangleq \frac{6}{\left( \sum_{\ell=1}^L V_\rho(\alpha_\ell) \right)^{3/2}}.$$  \hfill (45)

By Lemma 4 (Appendix V), the expectation $E \left[ |\tilde{j}_\ell(\alpha) - J(\alpha)|^3 \right]$ can be upper-bounded by a constant $\tilde{S}(T)$ that depends on $T$. Furthermore, by the nonnegativity of $V_\rho(\alpha_\ell)$,

$$\sum_{\ell=1}^L V_\rho(\alpha_\ell) \geq \sum_{\ell \in A_{\rho, \delta}} V_\rho(\alpha_\ell).$$  \hfill (46)
Lemma 4 (Appendix II) demonstrates that, for $\alpha \geq \rho(1 - \delta)$,

$$V_\rho(\alpha) \geq \left(\frac{T\rho}{1 + T\rho}\right)^2 (T - 1) + o_\rho(1) + o_\delta(1).$$

(47)

It thus follow that, for $\rho$ sufficiently large and $\delta$ sufficiently small,

$$\sum_{\ell=1}^L V_\rho(\alpha_\ell) \geq L_\alpha(\delta) \left(\frac{T\rho}{1 + T\rho}\right)^2 \frac{T - 1}{2}.$$  

(48)

Hence, for every $\alpha \in A_{\rho, \delta}$ and $\delta$ sufficiently small,

$$\bar{B}(\alpha) \leq \frac{6LS(T)}{(T-1/4)^{3/2}} \left(\frac{T\rho}{1 + T\rho}\right)^{3/2} \leq \frac{B(T)}{\sqrt{L}}.$$  

(49)

Let

$$\lambda = \frac{2\bar{B}(T)}{\sqrt{L}}$$  

(50)

and

$$\log \xi(\alpha) = \sum_{\ell=1}^L \bar{J}(\alpha_\ell) - \lambda \left(\sum_{\ell=1}^L V_\rho(\alpha_\ell)\right).$$  

(51)

With this choice, the Berry-Esseen theorem and (49) imply that, for every $\alpha \in A_{\rho, \delta}$,

$$\left| P \left[ \sum_{l=1}^L \bar{j}(\alpha_\ell) \leq \log \xi(\alpha) \right] - Q(\lambda) \right| \leq \bar{B}(\alpha) \leq \frac{B(T)}{\sqrt{L}}.$$  

(52)

It thus follows that

$$P \left[ \sum_{l=1}^L \bar{j}(\alpha_\ell) \leq \log \xi(\alpha) \right] \geq \epsilon + \frac{B(T)}{\sqrt{L}}.$$  

(53)

Introducing (53) into the upper bound (44), we obtain

$$R^*(L, \epsilon, \rho) \leq \sup_{\alpha \in [0, \rho]^L} \left\{ \sum_{\ell=1}^L \bar{J}(\alpha_\ell) \frac{L^2 T^2}{L^2 T^2} - \frac{\sum_{\ell=1}^L \bar{V}_\rho(\alpha_\ell)}{L^2 T^2} \right\} \log \frac{B(T)}{L^2 T^2} + \frac{1}{2} \log L.$$  

(54)

By the assumption $0 < \epsilon < \frac{1}{2}$, the inverse Q-function on the RHS of (54) is positive for sufficiently large $L$. It thus follows by the concavity of $x \mapsto \sqrt{x}$ and Jensen’s inequality that (54) can be further upper-bounded as

$$R^*(L, \epsilon, \rho) \leq \sup_{0 \leq \alpha \leq \rho} \left\{ \frac{\bar{J}(\alpha)}{T} - \frac{\bar{V}_\rho(\alpha)}{\sqrt{L}} \right\} \log \frac{B(T)}{L^2 T^2} + \frac{1}{2} \log L.$$  

(55)

where the second step follows because the channel is blockwise i.i.d., so $\bar{J}(\cdot)$ and $\bar{V}_\rho(\cdot)$ do not depend on $\ell$.

Applying a Taylor series expansion over $Q^{-1}(\epsilon + 2B(T)/\sqrt{L})$, using that by Lemma 5 (Appendix III) $\bar{V}_\rho(\cdot)$ is bounded in $\rho$ and $\alpha$, and collecting terms of order $\log L/L$, (55) can be written as

$$R^*(L, \epsilon, \rho) \leq \sup_{0 \leq \alpha \leq \rho} \left\{ \frac{\bar{J}(\alpha)}{T} - \frac{\bar{V}_\rho(\alpha)}{L^2 T^2} \right\} \log \frac{L}{L} + \mathcal{O}_L \left( \frac{\log L}{L} \right).$$  

(56)
We next show that (56) can be upper-bounded by
\[
R^*(L, \epsilon, \rho) \leq \frac{I(\rho) + o_\rho(1)}{T} - \sqrt{\frac{U + o_\rho(1)}{LT^2} Q^{-1}(\epsilon)} + O_L \left( \frac{\log L}{L} \right)
\]
(57)
which then concludes the proof of the converse bound. To this end, we show that
\[
\sup_{0 \leq \alpha \leq \rho} \left\{ \frac{\bar{J}^B(\alpha)}{T} - \sqrt{\frac{V_\rho^B(\alpha)}{LT^2} Q^{-1}(\epsilon)} \right\} = \bar{J}^B(\rho) - \sqrt{\frac{V(\rho)}{LT^2} Q^{-1}(\epsilon)} + O_L \left( \frac{1}{L} \right).
\]
(58)
The claim (57) follows then by (26c) and (26d).

To prove (58), we first present the following auxiliary result.

**Lemma 2 (Maximization over \(\alpha\)):**

1) Assume that \(T > 2\). For sufficiently large \(\rho\), we have
\[
\sup_{0 \leq \alpha \leq \rho} \bar{J}(\alpha) = \bar{J}(\rho).
\]
(59)
2) Assume that \(T > 2\) and that \(0 < \epsilon < \frac{1}{2}\). Consider the maximization problem on the left-hand side of (58).
For sufficiently large \(L\) and \(\rho\), we can assume that \(\alpha \in [\rho(1 - \delta_L), \rho]\), where
\[
0 \leq \delta_L \leq \frac{K}{L}
\]
(60)
for some nonnegative constant \(K\) that is independent of \((L, \rho, \alpha)\).

**Proof:** See Appendix VI.

We next set out to prove (58). By Part 2) of Lemma 2 we can assume that
\[
\alpha \geq \rho \left( 1 - \frac{K}{L} \right).
\]
(61)
Consider the following lower bound on \(\bar{V}_\rho(\alpha)\) given in Appendix VIII particularized for \(\delta = K/L\),
\[
\bar{V}_\rho(\alpha) \geq \bar{V}(\rho) - \Upsilon \frac{K}{L}, \quad \rho \left( 1 - \frac{K}{L} \right) \leq \alpha \leq \rho
\]
(62)
for some constant \(\Upsilon\). Combining this lower bound with Part 1) of Lemma 2 and using that, by assumption, \(Q^{-1}(\epsilon) > 0\), we obtain
\[
\sup_{0 \leq \alpha \leq \rho} \left\{ \frac{\bar{J}(\alpha)}{LT} - \sqrt{\frac{\bar{V}_\rho(\alpha)}{LT^2} Q^{-1}(\epsilon)} \right\} \leq \bar{J}(\rho) - \sqrt{\frac{\bar{V}(\rho) - \Upsilon K}{LT^2} Q^{-1}(\epsilon)}
\]
\[
= \bar{J}(\rho) - \sqrt{\frac{\bar{V}(\rho)}{LT^2} Q^{-1}(\epsilon)} + O_L \left( \frac{1}{L} \right).
\]
(63)
This proves (58) and concludes the proof of the upper bound.

**VI. CONCLUSION**

We have presented a high-SNR normal approximation for the maximal rate \(R^*(L, \epsilon, \rho)\) achievable over non-coherent, single-antenna, Rayleigh block-fading channels using an error-correcting code that spans \(L\) coherence intervals, has a block-error probability not larger than \(\epsilon\), and satisfies the power constraint \(\rho\). While the approximation was derived under the assumption that the number of coherence intervals \(L\) and the SNR \(\rho\) tend to infinity, numerical analyses suggest that it becomes accurate already for SNR values above 15dB and for 10 coherence intervals or more.

The obtained normal approximation is useful in two ways: Firstly, it complements the non-asymptotic bounds provided in [3], [7], whose evaluation is computationally demanding. Secondly, it lays the foundation for analytical studies that analyze the behavior of the maximum achievable rates as a function of system parameters such as SNR, number of coherence intervals, or blocklength.


**APPENDIX I**

**HIGH-SNR APPROXIMATIONS OF FIRST MOMENTS**

**Lemma 3:** The quantities $\bar{J}(\rho), I(\rho)$ and $\bar{I}(\rho)$ can be approximated as
\begin{align*}
\bar{J}(\rho) &= (T-1) \log(T\rho) - \log \Gamma(T) - (T-1)(1 + \gamma) + o_\rho(1) \\
I(\rho) &= (T-1) \log(T\rho) - \log \Gamma(T) - (T-1)(1 + \gamma) + o_\rho(1) \\
\bar{I}(\rho) &= (T-1) \log(T\rho) - \log \Gamma(T) - (T-1)(1 + \gamma) + o_\rho(1).
\end{align*}

**Proof:** We can express $\bar{J}(\rho), I(\rho)$ and $\bar{I}(\rho)$ as
\begin{align*}
\bar{J}(\rho) &= (T-1) \log(T\rho) - \log \Gamma(T) - (T-1) \frac{T\rho}{1+T\rho} \\
&+ (T-1) \mathbb{E} \left[ \log \left( Z_1 + \frac{Z_2}{(1+T\rho)} \right) \right] + (T-1) \mathbb{E} \left[ \log \left( 1 + \frac{\beta(T, \rho)}{(1+T\rho)Z_1 + Z_2} \right) \right] \\
I(\rho) &= (T-1) \log(T\rho) - \log \Gamma(T) - (T-1) \frac{T\rho}{1+T\rho} \\
&+ (T-1) \mathbb{E} \left[ \log \left( Z_1 + \frac{Z_2}{(1+T\rho)} \right) \right] - \mathbb{E} \left[ \log \tilde{\gamma} \left( T - 1, \frac{T\rho((1+T\rho)Z_1 + Z_2)}{1+T\rho} \right) \right] \\
\bar{I}(\rho) &= (T-1) \log(T\rho) - \log \Gamma(T) - (T-1) \frac{T\rho}{1+T\rho} + (T-1) \mathbb{E} \left[ \log \left( Z_1 + \frac{Z_2}{(1+T\rho)} \right) \right].
\end{align*}

Note that these expressions differ only in terms that vanish as $\rho \to \infty$. Indeed, we have
\begin{align*}
(T-1) \mathbb{E} \left[ \log \left( Z_1 + \frac{Z_2}{(1+T\rho)} \right) \right] &= -(T-1) \gamma + o_\rho(1) \\
(T-1) \mathbb{E} \left[ \log \left( 1 + \frac{\beta(T, \rho)}{(1+T\rho)Z_1 + Z_2} \right) \right] &= o_\rho(1) \\
\mathbb{E} \left[ \log \tilde{\gamma} \left( T - 1, \frac{T\rho((1+T\rho)Z_1 + Z_2)}{1+T\rho} \right) \right] &= o_\rho(1).
\end{align*}

Here, (66) follows by applying the Dominated Convergence Theorem [21, Section 1.26] to swap the limit as $\rho \to \infty$ and the expectation and because $\mathbb{E}[\log Z_1] = -\gamma$. The Dominated Convergence Theorem can be applied since
\begin{equation}
\left| \log \left( Z_1 + \frac{Z_2}{1+T\rho} \right) \right| \leq \left| \log(Z_1 + Z_2) \right| + \left| \log(Z_1) \right|
\end{equation}
and $\mathbb{E} \left[ \left| \log(Z_1 + Z_2) \right| + \left| \log(Z_1) \right| \right] < \infty$. Similarly, (67) and (68) follow by the Dominated Convergence Theorem and by noting that the terms inside the expected values on the LHS of (67) and (68) vanish as $\rho \to \infty$. The Dominated Convergence Theorem can be applied because for any arbitrary $\tilde{\rho}_* > 0$
\begin{align*}
\left| \log \tilde{\gamma} \left( T - 1, \frac{T\rho((1+T\rho)Z_1 + Z_2)}{1+T\rho} \right) \right| &\leq (T-1) \log \left( 1 - \frac{1}{1 - e^{-1/\beta(T, \rho)((1+T\rho)Z_1 + Z_2)}} \right) \\
&= (T-1) \log \left( 1 + \frac{1}{e^{1/\beta(T, \rho)((1+T\rho)Z_1 + Z_2)} - 1} \right) \\
&\leq (T-1) \log \left( 1 + \frac{\beta(T, \rho)}{(1+T\rho)Z_1 + Z_2} \right) \\
&\leq (T-1) \log \left( 1 + \frac{\beta(T, \tilde{\rho}_*)}{Z_1 + Z_2} \right), \quad \rho \geq \tilde{\rho}_*
\end{align*}
and because the expected value of the RHS of (70) is finite. Here, the first step follows by Alzer [15, Th. 1], the third step follows because $e^x \geq 1 + x$, and the last follows because $\rho \mapsto \beta(T, \rho)$ is monotonically decreasing in $\rho$.

Finally, note that $(T-1) \frac{T\rho}{1+T\rho}$ in (65a)–(65c) can be expressed as
\begin{equation}
(T-1) \frac{T\rho}{1+T\rho} = (T-1) + o_\rho(1).
\end{equation}
Hence (64a)–(64c) follow.
APPENDIX II

Lemma 4: For $\rho(1 - \delta) \leq \alpha \leq \rho$,

$$V_\rho(\alpha) \geq \left( \frac{T \rho}{1 + T \rho} \right)^2 (T - 1) + o_\rho(1) + o_\delta(1). \tag{72}$$

Proof: We have for every $\rho(1 - \delta) \leq \alpha \leq \rho$

$$V_\rho(\alpha) = E \left[ \left( - \frac{T \rho - T \alpha}{1 + T \rho} (Z_1 - 1) - \frac{T \rho}{1 + T \rho} (Z_2 - (T - 1)) \right) \right]$$

$$+ (T - 1) \log \left( \left( 1 + T \alpha \right) Z_1 + Z_2 + \beta(T, \rho) \right) - (T - 1) E \left[ \log \left( \left( 1 + T \alpha \right) Z_1 + Z_2 + \beta(T, \rho) \right) \right]$$

$$\geq E \left[ \left( - \frac{T \rho - T \alpha}{1 + T \rho} (Z_1 - 1) - \frac{T \rho}{1 + T \rho} (Z_2 - (T - 1)) \right)^2 \right]$$

$$+ 2 E \left[ \left( - \frac{T \rho - T \alpha}{1 + T \rho} (Z_1 - 1) - \frac{T \rho}{1 + T \rho} (Z_2 - (T - 1)) \right) \right]$$

$$\times \left( (T - 1) \log \left( \left( 1 + T \alpha \right) Z_1 + Z_2 + \beta(T, \rho) \right) - (T - 1) E \left[ \log \left( \left( 1 + T \alpha \right) Z_1 + Z_2 + \beta(T, \rho) \right) \right] \right)$$

$$\geq \left( \frac{T \rho - T \alpha}{1 + T \rho} \right)^2 + \left( \frac{T \rho}{1 + T \rho} \right)^2 (T - 1) - 2(T - 1) \frac{T \rho - T \alpha}{1 + T \rho} \left[ (Z_1 - 1) \log \left( \frac{Z_1 + Z_2}{1 + T \alpha} \right) \right]$$

$$- 2(T - 1) \frac{T \rho}{1 + T \rho} \left( E \left[ (Z_2 - (T - 1)) \log \left( \frac{Z_1 + Z_2}{1 + T \alpha} \right) \right] \right)$$

$$- E \left[ (Z_2 - (T - 1)) \log \left( \frac{1 + T \rho}{1 + T \alpha} \right) \right]. \tag{73}$$

Here, the second inequality follows because $E[Z_1 - 1] = E[Z_2 - (T - 1)] = 0$ and because

$$E \left[ (Z_1 - 1) \log \left( \frac{1 + T \rho}{1 + T \alpha} Z_1 + Z_2 \right) \right] \leq 0 \tag{74}$$

$$E \left[ (Z_2 - (T - 1)) \log \left( \frac{1 + T \rho}{1 + T \alpha} Z_1 + Z_2 \right) \right] \leq 0. \tag{75}$$

The latter claims (74) and (75) follow because

$$\left( \frac{Z_1 - 1}{Z_1 + Z_2} \right) \log \left( 1 + \frac{Z_1 + Z_2}{1 + T \alpha} \right) \leq \left( \frac{Z_1 - 1}{Z_1 + Z_2} \right) \log \left( 1 + \frac{Z_1 + Z_2}{1 + T \alpha} \right) \tag{76}$$

and

$$\left( \frac{Z_2 - (T - 1)}{Z_1 + Z_2} \right) \log \left( 1 + \frac{Z_1 + Z_2}{1 + T \alpha} \right) \leq \left( \frac{Z_2 - (T - 1)}{Z_1 + Z_2} \right) \log \left( 1 + \frac{Z_1 + Z_2}{1 + T \alpha} \right) \tag{77}$$

and

$$E \left[ (Z_1 - 1) \log \left( \frac{1 + T \rho}{1 + T \alpha} Z_1 + Z_2 \right) \right] = E \left[ (Z_2 - (T - 1)) \log \left( 1 + \frac{Z_1 + Z_2}{1 + T \alpha} \right) \right] = 0. \tag{78}$$

The first term on the RHS of (73) is nonnegative, so discarding it yields a lower bound. Furthermore, the third term in (73) can be bounded by upper-bounding for $\rho(1 - \delta) \leq \alpha \leq \rho$

$$2(T - 1) \frac{T \rho - T \alpha}{1 + T \rho} E \left[ (Z_1 - 1) \log \left( \frac{Z_1 + Z_2}{1 + T \alpha} \right) \right]$$
Finally, the remaining terms in (83) only depend on $\delta$

\[
\leq 2(T - 1) \frac{T\rho - T\alpha}{1 + T\rho} \sqrt{E[(Z_1 - 1)^2]E\left[\log^2\left(Z_1 + \frac{Z_2}{1 + T\alpha}\right)\right]}
\]

\[
\leq 2(T - 1)\delta \sqrt{\left(\frac{\pi^2}{6} + \gamma^2 + \psi^2(T) + \zeta(2,T)\right)}
\]  \hspace{1cm} (79)

where $\zeta(z,q)$ denotes Riemann's zeta function [16 Eq. (9.511)]. Here, the first inequality follows from the Cauchy-Schwarz inequality, and the last inequality follows by evaluating $E[(Z_1 - 1)^2] = 1$ and

\[
E\left[\log^2\left(Z_1 + \frac{Z_2}{1 + T\alpha}\right)\right] \leq E[\log^2(Z_1 + Z_2) + \log^2(Z_1)]
\]

\[
= \frac{\pi^2}{6} + \gamma^2 + \psi^2(T) + \zeta(2,T).
\]  \hspace{1cm} (80)

Finally, the fifth term on the RHS (72) can be bounded by upper-bounding for $\rho(1 - \delta) \leq \alpha \leq \rho$

\[
\left|E\left[(Z_2 - (T - 1)) \log\left(\frac{(1 + T\rho)Z_1 + Z_2}{(1 + T\alpha)Z_1 + Z_2}\right)\right]\right| \leq E\left[|Z_2 - (T - 1)| \log\left(\frac{(1 + T\rho)Z_1 + Z_2}{(1 + T\alpha)Z_1 + Z_2}\right)\right]
\]

\[
\leq E\left[|Z_2 - (T - 1)|\log\left(\frac{\rho}{\alpha}\right)\right]
\]

\[
\leq E\left[|Z_2 - (T - 1)|\log\left(\frac{1}{1 - \delta}\right)\right].
\]  \hspace{1cm} (81)

Combining (79)–(81) with (73), we obtain the lower bound

\[
V_\rho(\alpha) \geq \left(\frac{T\rho}{1 + T\rho}\right)^2 (T - 1) - 2(T - 1)\delta \sqrt{\left(\frac{\pi^2}{6} + \gamma^2 + \psi^2(T) + \zeta(2,T)\right)}
\]

\[
- 2(T - 1) \frac{T\rho}{1 + T\rho} \left\{E\left[(Z_2 - (T - 1)) \log\left(Z_1 + \frac{Z_2}{1 + T\rho}\right)\right] + E\left[|Z_2 - (T - 1)|\log\left(\frac{1}{1 - \delta}\right)\right]\right\}.
\]  \hspace{1cm} (82)

We conclude the proof of Lemma 4 by demonstrating that the RHS of (83) can be expressed as

\[
\left(\frac{T\rho}{1 + T\rho}\right)^2 (T - 1) + o_\rho(1) + o_\delta(1).
\]  \hspace{1cm} (84)

Indeed, by the Dominated Convergence Theorem, the third term on the RHS of (83) vanishes as $\rho \to \infty$, since

\[
\left|(Z_2 - (T - 1)) \log\left(Z_1 + \frac{Z_2}{1 + T\rho}\right)\right| \leq |Z_2 - (T - 1)| \log\left(Z_1 + \frac{Z_2}{1 + T\rho}\right)
\]

\[
\leq |(Z_2 - (T - 1))| \sqrt{\log^2(Z_1) + \log^2(Z_1 + Z_2)}
\]  \hspace{1cm} (85)

whose expected value is finite. Hence

\[
E\left[\left|(Z_2 - (T - 1)) \log\left(Z_1 + \frac{Z_2}{1 + T\rho}\right)\right|\right] = o_\rho(1).
\]  \hspace{1cm} (86)

Finally, the remaining terms in (83) only depend on $\delta$ and $T$ and vanish as $\delta \downarrow 0$, so they can be expressed as $o_\delta(1)$. The claim thus follows.

\section*{Appendix III}

$V_\rho(\alpha)$ and $U(\rho)$ are bounded

Lemma 5: For any arbitrary $\bar{\rho}_* > 0$, we have that

\[
\sup_{\alpha \geq 0} V_\rho(\alpha) < \infty
\]  \hspace{1cm} (87a)

\[
\sup_{\rho \geq \bar{\rho}_*} U(\rho) < \infty.
\]  \hspace{1cm} (87b)
Proof: We first show \((87a)\). Using the definitions of \(\tilde{j}(\alpha)\) and \(\tilde{J}(\alpha)\) in \((20)\) and \((22)\) together with the facts that \((a_1 + \ldots + a_\eta)^2 \leq ca_1^2 + \ldots + ca_\eta^2\) for some positive constant \(c\) and \(E[X^2] \geq E[X]^2\) for any random variable \(X\), we can upper-bound \(\tilde{V}_\rho(\alpha) \triangleq E[(\tilde{j}(\alpha) - \tilde{J}(\alpha))^2]\) as

\[
\tilde{V}_\rho(\alpha) = E \left[ \frac{T\rho - T\alpha}{1 + T\rho} (1 - Z_1) + \frac{T\rho}{1 + T\rho} (T - 1 - Z_2) \right]
+ (T - 1) \log \left( Z_1 + \frac{Z_2}{1 + T\alpha} \right) - (T - 1) E \left[ \log \left( Z_1 + \frac{Z_2}{1 + T\alpha} \right) \right]
+ (T - 1) \log \left( 1 + \frac{\beta(T, \rho)}{(1 + T\alpha) Z_1 + Z_2} \right) - (T - 1) E \left[ \log \left( 1 + \frac{\beta(T, \rho)}{(1 + T\alpha) Z_1 + Z_2} \right) \right]^2
\leq c \left( \frac{T\rho - T\alpha}{1 + T\rho} \right)^2 E \left[ (Z_1 - 1)^2 \right] + \left( \frac{T\rho}{1 + T\rho} \right)^2 E \left[ (Z_2 - T + 1)^2 \right]
+ 2(T - 1)^2 E \left[ \log^2 \left( Z_1 + \frac{Z_2}{1 + T\alpha} \right) \right] + 2(T - 1)^2 E \left[ \log^2 \left( 1 + \frac{\beta(T, \rho)}{(1 + T\alpha) Z_1 + Z_2} \right) \right]
\tag{88}
\]

for some positive constant \(c\).

We next show that each summand on the RHS of \((88)\) is bounded in \((\rho, \alpha)\). Indeed, we have \(E \left[ (Z_1 - 1)^2 \right] = 1\) and \(E \left[ (Z_2 - T + 1)^2 \right] = (T - 1)\), so noting that \((T\rho - T\alpha)/(1 + T\rho) \leq 1\) and \(T\rho/(1 + T\rho) \leq 1\), we obtain that the first two terms on the RHS of \((88)\) are bounded in \(\rho\) and \(\alpha\). The third term on the RHS of \((88)\) can be upper-bounded by (cf. \((80)\))

\[
(T - 1)^2 E \left[ \log^2 \left( Z_1 + \frac{Z_2}{1 + T\alpha} \right) \right] \leq (T - 1)^2 E \left[ \log^2 (Z_1 + Z_2) \right] + (T - 1)^2 E \left[ \log^2 (Z_1) \right] < \infty.
\tag{89}
\]

Finally, for any arbitrary \(\hat{\rho}_s > 0\) and \(\rho \geq \hat{\rho}_s\), the fourth term on the RHS of \((88)\) can be upper-bounded by

\[
E \left[ (T - 1)^2 \log^2 \left( 1 + \frac{\beta(T, \rho)}{(1 + T\alpha) Z_1 + Z_2} \right) \right] \leq (T - 1)^2 E \left[ \log^2 \left( 1 + \frac{\beta(T, \rho)}{Z_1 + Z_2} \right) \right]
\leq (T - 1)^2 E \left[ \log^2 \left( 1 + \frac{\beta(T, \hat{\rho}_s)}{Z_1 + Z_2} \right) \right] < \infty
\tag{90}
\]

where the last step follows because \(\rho \mapsto \beta(T, \rho)\) is monotonically decreasing in \(\rho\). Combining the above steps with \((88)\) proves \((87a)\).

To prove \((87b)\), we follow along similar lines. Indeed, using the definitions of \(i_i(\rho)\) and \(I(\rho)\) together with the facts that \((a_1 + \ldots + a_\eta)^2 \leq ca_1^2 + \ldots + ca_\eta^2\) for some positive constant \(c\) and \(E[X^2] \geq E[X]^2\) for any random variable \(X\), we can upper-bound \(U(\rho) \triangleq E[(i_i(\rho) - I(\rho))^2]\) as

\[
U(\rho) = E \left[ \left( \frac{T\rho}{1 + T\rho} (T - 1 - Z_2) + (T - 1) \log \left( Z_1 + \frac{Z_2}{1 + T\alpha} \right) - (T - 1) E \left[ \log \left( Z_1 + \frac{Z_2}{1 + T\alpha} \right) \right] \right]
- \log \tilde{\gamma} \left( T - 1, \frac{T\rho(1 + T\rho)Z_1 + Z_2}{1 + T\rho} \right) + E \left[ \log \tilde{\gamma} \left( T - 1, \frac{T\rho((1 + T\rho)Z_1 + Z_2)}{1 + T\rho} \right) \right]^2 \right]
\leq c \left( \frac{T\rho}{1 + T\rho} \right)^2 E \left[ (Z_2 - T + 1)^2 \right] + 2(T - 1)^2 E \left[ \log^2 \left( Z_1 + \frac{Z_2}{1 + T\alpha} \right) \right]
+ 2(T - 1)^2 E \left[ \log^2 \left( T - 1, \frac{T\rho((1 + T\rho)Z_1 + Z_2)}{1 + T\rho} \right) \right]
\tag{91}
\]

for some positive constant \(c\).
We next show that each summand is bounded in $\rho$. Indeed, as shown above, the first and second term on the RHS of (91) are bounded in $\rho$. Regarding the third term on the RHS of (91), we first lower-bound $\tilde{\gamma}(\cdot, \cdot)$ using the bound by Alzer [15, Th. 1]:

$$
E \left[ \log^2 \tilde{\gamma} \left( T - 1, \frac{T \rho((1 + T \rho)Z_1 + Z_2)}{1 + T \rho} \right) \right] \leq (T - 1)^2 E \left[ \log^2 \left( \frac{1}{1 - e^{-1/\beta(T, \rho)((1+T \rho)Z_1+Z_2)}} \right) \right]
$$

$$
= (T - 1)^2 E \left[ \log^2 \left( 1 + \frac{e^{1/\beta(T, \rho)((1+T \rho)Z_1+Z_2) - 1}}{1} \right) \right]. 
$$

(92)

Using that $e^x \geq 1 + x$, this can be further upper-bounded as

$$
E \left[ \log^2 \tilde{\gamma} \left( T - 1, \frac{T \rho((1 + T \rho)Z_1 + Z_2)}{1 + T \rho} \right) \right] \leq (T - 1)^2 E \left[ \log^2 \left( 1 + \frac{\beta(T, \rho)}{Z_1 + Z_2} \right) \right]. 
$$

(93)

By the monotonicity of $\rho \mapsto \beta(T, \rho)$, it thus follows that for any arbitrary $\tilde{\rho}_s > 0$ and $\rho \geq \tilde{\rho}_s$, the third term on the RHS of (91) is upper-bounded by

$$
E \left[ \log^2 \tilde{\gamma} \left( T - 1, \frac{T \rho((1 + T \rho)Z_1 + Z_2)}{1 + T \rho} \right) \right] \leq (T - 1)^2 E \left[ \log^2 \left( 1 + \frac{\beta(T, \tilde{\rho}_s)}{Z_1 + Z_2} \right) \right] < \infty. 
$$

(94)

Combining the above steps with (91) proves (87b).

\section*{APPENDIX IV
HIGH-SNR APPROXIMATIONS OF SECOND MOMENTS

\textbf{Lemma 6:} The quantities $\tilde{V}(\rho)$ and $U(\rho)$ can be approximated as

$$
\tilde{V}(\rho) = (T - 1)^2 \frac{\pi^2}{6} + (T - 1) + o_\rho(1) 
$$

(95a)

$$
U(\rho) = (T - 1)^2 \frac{\pi^2}{6} + (T - 1) + o_\rho(1) 
$$

(95b)

\textbf{Proof:} We first show (95a) by analyzing $\tilde{V}(\rho) \triangleq E[\log^2(\tilde{J}_t(\rho) - \tilde{J}(\rho))^2]$ in the limit as $\rho$ tends to infinity. To this end, we first note that

$$
(\tilde{J}_t(\rho) - \tilde{J}(\rho))^2 = \left( \frac{T \rho}{1 + T \rho}(T - 1 - Z_2) + (T - 1) \log \left( Z_1 + \frac{Z_2}{1 + T \rho} \right) - (T - 1) E \left[ \log \left( Z_1 + \frac{Z_2}{1 + T \rho} \right) \right] \right)
$$

$$
+ (T - 1) \log \left( 1 + \frac{\beta(T, \rho)}{(1 + T \rho)Z_1 + Z_2} \right) - (T - 1) E \left[ \log \left( \frac{\beta(T, \rho)}{(1 + T \rho)Z_1 + Z_2} \right) \right]
$$

(96)

tends to

$$
(T - 1 - Z_2 + (T - 1) \log(Z_1) - (T - 1) E[\log(Z_1)])^2 
$$

(97)
as $\rho$ tends to infinity. (To obtain $E[\log(Z_1)]$, we changed the order of taking limit and expectation, which can be justified by the Dominated Convergence Theorem; cf. [101].) Since $Z_1$ and $Z_2$ are independent, it follows that

$$
E \left[ (T - 1 - Z_2 + (T - 1) \log(Z_1) - (T - 1) E[\log(Z_1)])^2 \right] = E \left[ (T - 1 - Z_2)^2 \right] + (T - 1)^2 \left( E[\log^2(Z_1)] - E[\log(Z_1)]^2 \right)
$$

$$
= (T - 1) + (T - 1)^2 \frac{\pi^2}{6}. 
$$

(98)

It remains to show that the order of taking limit (as $\rho \to \infty$) and expectation can be swapped. To this end, we next argue that the Dominated Convergence Theorem applies. Indeed, proceeding similarly as in Appendix III we obtain for any arbitrary $\tilde{\rho}_s > 0$

$$
(\tilde{J}_t(\rho) - \tilde{J}(\rho))^2 \leq c \left( \frac{T \rho}{1 + T \rho} \right)^2 (Z_2 - T + 1)^2
$$


\[ + (T - 1)^2 \log^2 \left( Z_1 + \frac{Z_2}{1 + T \rho} \right) + (T - 1)^2 \mathbb{E} \left[ \log \left( Z_1 + \frac{Z_2}{1 + T \rho} \right) \right]^2 \]
\[ + (T - 1)^2 \log^2 \left( 1 + \frac{\beta(T, \rho)}{(1 + T \rho)Z_1 + Z_2} \right) + (T - 1)^2 \mathbb{E} \left[ \log \left( 1 + \frac{\beta(T, \rho)}{(1 + T \rho)Z_1 + Z_2} \right) \right]^2 \]
\[ \leq c \left( Z_2 - T + 1 \right)^2 + (T - 1)^2 \log^2 (Z_1 + Z_2) + (T - 1)^2 \log^2 (Z_1) \]
\[ + (T - 1)^2 \mathbb{E} \left[ \left| \log (Z_1 + Z_2) \right| + \left| \log (Z_1) \right| \right]^2 + (T - 1)^2 \log^2 \left( 1 + \frac{\beta(T, \tilde{\rho}_*)}{Z_1 + Z_2} \right) \]
\[ + (T - 1)^2 \mathbb{E} \left[ \log \left( 1 + \frac{\beta(T, \tilde{\rho}_*)}{Z_1 + Z_2} \right) \right]^2 \], \quad \rho \geq \tilde{\rho}_*, \tag{99} \]

for some positive constant \( c \). Here, we upper-bounded the second term using that (cf. (89), Appendix III)
\[ \log^2 \left( Z_1 + \frac{Z_2}{1 + T \rho} \right) \leq \log^2 (Z_1 + Z_2) + \log^2 (Z_1); \tag{100} \]
the third term using that (cf. (69), Appendix I)
\[ \left| \log \left( Z_1 + \frac{Z_2}{1 + T \rho} \right) \right| \leq \left| \log (Z_1 + Z_2) \right| + \left| \log (Z_1) \right|; \tag{101} \]
the fourth term using that, for any arbitrary \( \tilde{\rho}_* > 0 \), (cf. (90), Appendix III)
\[ \log^2 \left( 1 + \frac{\beta(T, \rho)}{(1 + T \rho)Z_1 + Z_2} \right) \leq \log^2 \left( 1 + \frac{\beta(T, \tilde{\rho}_*)}{Z_1 + Z_2} \right) \]
\[ \leq \log^2 \left( 1 + \frac{\beta(T, \tilde{\rho}_*)}{Z_1 + Z_2} \right), \quad \rho \geq \tilde{\rho}_*; \tag{102} \]
and the fifth term using that, for any arbitrary \( \tilde{\rho}_* > 0 \),
\[ \log \left( 1 + \frac{\beta(T, \rho)}{(1 + T \rho)Z_1 + Z_2} \right) \leq \log \left( 1 + \frac{\beta(T, \tilde{\rho}_*)}{Z_1 + Z_2} \right), \quad \rho \geq \tilde{\rho}_*. \tag{103} \]

Since the expected value on the RHS of (99) is finite, the Dominated Convergence Theorem applies, hence (95a) follows.

To prove (95b), we follow along the same lines. Indeed,
\[ \left( i_T(\rho) - I(\rho) \right)^2 = \left( \frac{T \rho}{1 + T \rho} (T - 1 - Z_2) + (T - 1) \log \left( Z_1 + \frac{Z_2}{1 + T \rho} \right) - (T - 1) \mathbb{E} \left[ \log \left( Z_1 + \frac{Z_2}{1 + T \rho} \right) \right] \right) \]
\[ - \log \tilde{\gamma} \left( T - 1, \frac{T \rho((1 + T \rho)Z_1 + Z_2)}{1 + T \rho} \right) + \mathbb{E} \left[ \log \tilde{\gamma} \left( T - 1, \frac{T \rho((1 + T \rho)Z_1 + Z_2)}{1 + T \rho} \right) \right]^2 \tag{104} \]
tends to (27) as \( \rho \) tends to infinity. It remains to show that the order of taking limit (as \( \rho \to \infty \)) and expectation can be swapped. We next argue that this can be justified by the Dominated Convergence Theorem. Indeed, we have, for any arbitrary \( \tilde{\rho}_* > 0 \),
\[ \left( i_T(\rho) - I(\rho) \right)^2 \leq c \left( \frac{T \rho}{1 + T \rho} \right)^2 (Z_2 - T + 1)^2 + (T - 1)^2 \log^2 \left( Z_1 + \frac{Z_2}{1 + T \rho} \right) \]
\[ + (T - 1)^2 \mathbb{E} \left[ \log \left( Z_1 + \frac{Z_2}{1 + T \rho} \right) \right]^2 + \log^2 \tilde{\gamma} \left( T - 1, \frac{T \rho((1 + T \rho)Z_1 + Z_2)}{1 + T \rho} \right) \]
\[ + \mathbb{E} \left[ \log \tilde{\gamma} \left( T - 1, \frac{T \rho((1 + T \rho)Z_1 + Z_2)}{1 + T \rho} \right) \right]^2 \]
\[ \leq c \left( Z_2 - T + 1 \right)^2 + (T - 1)^2 \log^2 (Z_1 + Z_2) + (T - 1)^2 \log^2 (Z_1) \]
\[ \begin{align*}
&+ (T - 1)^2 E \left[ |\log(Z_1 + Z_2)| + |\log(Z_1)| \right]^2 + (T - 1)^2 \log^2 \left( 1 + \frac{\beta(T, \bar{\rho}_*)}{Z_1 + Z_2} \right) \\
&+ (T - 1)^2 E \left[ \log \left( 1 + \frac{\beta(T, \bar{\rho}_*)}{Z_1 + Z_2} \right) \right]^2, \quad \rho \geq \bar{\rho}_* \tag{105}
\end{align*} \]

for some positive constant \(c\). Here, we upper-bounded the first three terms as in (99); the fourth term was upper-bounded by following the steps [97]-[99] in Appendix III

\[ \log^2 \tilde{\gamma} \left( T - 1, \frac{T \rho((1 + T \rho)Z_1 + Z_2)}{1 + T \rho} \right) \leq (T - 1)^2 \log^2 \left( 1 + \frac{\beta(T, \bar{\rho}_*)}{Z_1 + Z_2} \right); \tag{106} \]

and the last term is upper-bounded using (70) in Appendix I

\[ \left| \log \tilde{\gamma} \left( T - 1, \frac{T \rho((1 + T \rho)Z_1 + Z_2)}{1 + T \rho} \right) \right| \leq (T - 1) \log \left( 1 + \frac{\beta(T, \bar{\rho}_*)}{Z_1 + Z_2} \right). \tag{107} \]

Since the expected value on the RHS of (105) is finite, the Dominated Convergence Theorem applies, hence (95b) follows.

\[ \blacksquare \]

**APPENDIX V**

**BOUNDED THIRD MOMENT**

**Lemma 7:** For any arbitrary \(\bar{\rho}_* > 0\), we have that

\[ \sup_{\alpha \geq 0, \rho \geq \bar{\rho}_*} E \left[ |\tilde{j}_e(\alpha) - J(\alpha)|^3 \right] < \infty \tag{108a} \]

\[ \sup_{\rho \geq \bar{\rho}_*} E \left[ |i_e(\rho) - I(\rho)|^3 \right] < \infty \tag{108b} \]

**Proof:** We shall first prove (108a). Using the definitions of \(\tilde{j}_e(\alpha)\) and \(J(\alpha)\) in (20) and (22) together with the facts that \(|a_1 + \ldots + a_q|^3 \leq c |a_1|^3 + \ldots + |a_q|^3\) for some positive constant \(c\) and \(E[|X|^3] \geq |E[X]|^3\) for any random variable \(X\), we can upper-bound \(E \left[ |\tilde{j}_e(\alpha) - J(\alpha)|^3 \right] \) as

\[ E \left[ |\tilde{j}_e(\alpha) - J(\alpha)|^3 \right] = E \left[ \frac{T \rho - T \alpha}{1 + T \rho} (1 - Z_1) + \frac{T \rho}{1 + T \rho} (T - 1 - Z_2) \right. \]

\[ + (T - 1) \log \left( Z_1 + \frac{Z_2}{1 + T \alpha} \right) - (T - 1) E \left[ \log \left( Z_1 + \frac{Z_2}{1 + T \alpha} \right) \right] \]

\[ - \log \left( 1 + \frac{\beta(T, \rho)}{1 + T \alpha} \right) Z_1 + Z_2) + E \left[ \log \left( 1 + \frac{\beta(T, \rho)}{(1 + T \alpha)Z_1 + Z_2} \right) \right] \right]^3 \\
\leq c \left( \frac{T \rho - T \alpha}{1 + T \rho} \right)^3 \left( \frac{T \rho}{1 + T \rho} \right)^3 E \left[ |Z_1 - 1|^3 \right] + \left( \frac{T \rho}{1 + T \rho} \right)^3 E \left[ |Z_2 - T + 1|^3 \right] \\
+ 2(T - 1)^3 E \left[ \log \left( Z_1 + \frac{Z_2}{1 + T \alpha} \right) \right]^3 \]

\[ + 2(T - 1)^3 E \left[ \log \left( 1 + \frac{\beta(T, \rho)}{(1 + T \alpha)Z_1 + Z_2} \right) \right]^3 \tag{109} \]

for some positive constant \(c\).

We next show that each summand on the RHS of (109) is bounded in \(\rho\) and \(\alpha\). Indeed, the first two terms on the RHS of (109) are bounded because the third central moments of the Gamma-distributed random variables \(Z_1\) and \(Z_2\) are bounded, and because \(0 \leq (T \rho - T \alpha)/(1 + T \rho) \leq 1\) and \(0 \leq T \rho/(1 + T \rho) \leq 1\).
The third term on the RHS of (109) can be upper-bounded by using that
\[
\log\left(Z_1 + \frac{Z_2}{1 + T_\alpha}\right) \leq |\log Z_1| + |\log(Z_1 + Z_2)|
\]
hence
\[
\mathbb{E} \left[ \log\left(Z_1 + \frac{Z_2}{1 + T_\alpha}\right) \right]^3 \leq \tilde{c} \left( \mathbb{E} \left[ |\log Z_1|^3 \right] + \mathbb{E} \left[ |\log(Z_1 + Z_2)|^3 \right] \right) < \infty
\]
for some positive constant \(\tilde{c}\).

Finally, the fourth term on the RHS of (109) can be upper-bounded by
\[
(T-1)^3\mathbb{E} \left[ \left( \log\left(1 + \frac{\beta(T, \rho)}{(1 + T_\alpha)Z_1 + Z_2} \right) \right)^3 \right] \leq (T-1)^3\mathbb{E} \left[ \log\left(1 + \frac{\beta(T, \rho)}{Z_1 + Z_2} \right) \right]^3.
\]
By the monotonicity of \(\rho \mapsto \beta(T, \rho)\), it thus follows that for any arbitrary \(\bar{\rho}_* > 0\) and \(\rho \geq \bar{\rho}_*\),
\[
(T-1)^3\mathbb{E} \left[ \left( \log\left(1 + \frac{\beta(T, \rho)}{(1 + T_\alpha)Z_1 + Z_2} \right) \right)^3 \right] \leq (T-1)^3\mathbb{E} \left[ \left( \log\left(1 + \frac{\beta(T, \bar{\rho}_*)}{Z_1 + Z_2} \right) \right)^3 \right] < \infty.
\]
Combining the above steps with (109) proves (108a).

We prove (108b) along similar lines. Using the definitions of \(i_\ell(\rho)\) and \(I(\rho)\) together with the facts that \(|a_1 + \ldots + a_\ell| \leq c|a_1|^3 + \ldots + c|a_\ell|^3\) for some positive constant \(c\) and \(\mathbb{E}[|X|^3] \geq |\mathbb{E}[X]|^3\) for any random variable \(X\), we can upper-bound \(\mathbb{E}[|i_\ell(\rho) - I(\rho)|^3]\) as
\[
\mathbb{E}[|i_\ell(\rho) - I(\rho)|^3] = \mathbb{E}\left[ \frac{T \rho}{1 + T \rho} (T - 1 - Z_2) + (T - 1) \log\left(Z_1 + \frac{Z_2}{1 + T \rho}\right) - (T - 1) \mathbb{E}\left[ \log\left(Z_1 + \frac{Z_2}{1 + T \rho}\right) \right] \right.
\]
\[
\left. - \log\tilde{\gamma}\left(T - 1, \frac{T \rho((1 + T \rho)Z_1 + Z_2)}{1 + T \rho}\right) + \mathbb{E}\left[ \log\tilde{\gamma}\left(T - 1, \frac{T \rho((1 + T \rho)Z_1 + Z_2)}{1 + T \rho}\right) \right]^3 \right]
\]
\[
\leq \left( \frac{T \rho}{1 + T \rho} \right)^3 \mathbb{E}\left[ |Z_2 - T + 1|^3 + 2(T - 1)^3 \mathbb{E}\left[ |Z_1 - \frac{Z_2}{1 + T \rho}|^3 \right] \right]
\]
\[
+ 2(T - 1)^3 \mathbb{E}\left[ |Z_1 - \frac{Z_2}{1 + T \rho}|^3 \right] \right)
\]
for some positive constant \(\tilde{c}\).

We next show that each summand is bounded in \(\rho\). Indeed, as shown above, the first two terms on the RHS of (114) are bounded in \(\rho\). With respect to the third term on the RHS of (114), we first use the bound by Alzer [15, Th. 1]:
\[
\mathbb{E}\left[ \log\tilde{\gamma}\left(T - 1, \frac{T \rho((1 + T \rho)Z_1 + Z_2)}{1 + T \rho}\right) \right]^3 \leq \mathbb{E}\left[ \log^3\left( \frac{1}{1 - e^{-1/\beta(T, \rho)((1 + T \rho)Z_1 + Z_2)}} \right) \right]
\]
\[
= (T - 1)^3 \mathbb{E}\left[ \log^3\left( 1 + \frac{1}{e^{1/\beta(T, \rho)((1 + T \rho)Z_1 + Z_2)}} - 1 \right) \right].
\]
Using that \(e^x \geq 1 + x\), this can be further upper-bounded as
\[
\mathbb{E}\left[ \log\tilde{\gamma}\left(T - 1, \frac{T \rho((1 + T \rho)Z_1 + Z_2)}{1 + T \rho}\right) \right]^3 \leq (T - 1)^3 \mathbb{E}\left[ \log^3\left( 1 + \frac{\beta(T, \rho)}{(1 + T \rho)Z_1 + Z_2} \right) \right].
\]
By the monotonicity of \(\rho \mapsto \beta(T, \rho)\), it thus follows that for any arbitrary \(\bar{\rho}_* > 0\) and \(\rho \geq \bar{\rho}_*\), the third term on the RHS of (114) is upper-bounded by
\[
\mathbb{E}\left[ \log\tilde{\gamma}\left(T - 1, \frac{T \rho((1 + T \rho)Z_1 + Z_2)}{1 + T \rho}\right) \right]^3 \leq (T - 1)^3 \mathbb{E}\left[ \log^3\left( 1 + \frac{\beta(T, \bar{\rho}_*)}{(1 + T \rho)Z_1 + Z_2} \right) \right] < \infty.
\]
Combining the above steps with (114) proves (108b).
A. Proof of Lemma 7
Consider the upper bound (42), namely,

\[ R^*(L, \epsilon, \rho) \leq \sup_{\alpha \in [0, \rho]^L} \log \left( \frac{1}{\beta(\alpha, q_{Y_{L^L}})} \right), \]  

(118)

We next show that, without loss of optimality, we can assume that \( \alpha \in A_{\rho, \delta} \). To this end, we demonstrate that for all \( \alpha \notin A_{\rho, \delta} \), we can find a lower bound on \( R^*(L, \epsilon, \rho) \) that exceeds an upper bound on (118). Hence, such \( \alpha \) cannot be optimal.

A lower bound on \( R^*(L, \epsilon, \rho) \) follows by (33), and by bounding \( I(\rho) \geq I(\rho) \), and \( U(\rho) \leq U_{UB}(T, \rho_0) \) (cf. (39)):

\[ R^*(L, \epsilon, \rho) \geq \frac{I(\rho)}{T} - \sqrt{\frac{U_{UB}(T, \rho_0)}{LT^2}} Q^{-1}(\tau) \triangleq R_{LB}(\rho). \]  

(119)

(Note that, by assumption, \( 0 < \epsilon < \frac{1}{2} \), so \( Q^{-1}(\tau) > 0 \) for \( L \) sufficiently large.)

It follows from [2, Eq. (106)] and (21) that the RHS of (118) can be upper-bounded as

\[ \sup_{\alpha \in [0, \rho]^L} \log \left( \frac{1}{\beta(\alpha, q_{Y_{L^L}})} \right) \leq \sup_{\alpha \in [0, \rho]^L} \left\{ \log \xi(\alpha) \right\} \leq \sup_{\alpha \in [0, \rho]^L} \left\{ \frac{\log(1 - \epsilon - \frac{L V_{UB}(T, \rho_0)}{(1 - \epsilon) - \frac{1}{\sqrt{L}}})}{LT} \right\}. \]  

(120)

for an arbitrary \( \xi: [0, \rho]^L \rightarrow (0, \infty) \). By Lemma 5 for every \( \rho > 0 \) there exists a \( V_{UB}(T, \rho_0) \) that is independent of \( \alpha \) and \( \rho \) and that satisfies

\[ V_{UB}(T, \rho_0) \leq V_{UB}(T, \rho_0), \quad \alpha \geq 0, \ \rho \geq \rho_0. \]  

(121)

Let

\[ \log \xi(\alpha) = \sum_{\ell=1}^{L} J(\alpha_{\ell}) + \sqrt{\frac{L V_{UB}(T, \rho_0)}{(1 - \epsilon - \frac{1}{\sqrt{L}})}}. \]  

(122)

By Chebyshev’s inequality and (121), we thus obtain

\[ P \left[ \sum_{\ell=1}^{L} J(\alpha_{\ell}) \geq \log \xi(\alpha) \right] \leq \sum_{\ell=1}^{L} \frac{\bar{V}_{UB}(T, \rho_0)}{LT} \left( 1 - \epsilon - \frac{1}{\sqrt{L}} \right) \leq 1 - \frac{1}{\sqrt{L}}. \]  

(123)

Combining (120) with (123) gives

\[ R^*(L, \epsilon, \rho) \leq \sup_{\alpha \in [0, \rho]^L} \left\{ \frac{1}{L} \sum_{\ell=1}^{L} R_{UB}(\alpha_{\ell}) \right\}, \]  

(124)

The \( \alpha \)'s for which (124) is smaller than (119) can be discarded without loss of optimality, since the upper bound can never be smaller than the lower bound. We next use this argument to show that the fraction of \( \alpha_{\ell} \)'s in \( \alpha \) that satisfy \( \alpha_{\ell} \geq \rho(1 - \delta) \) tends to 1 as \( L \) and \( \rho \) tend to infinity. Specifically, we consider the difference

\[ \frac{1}{L} \sum_{\ell=1}^{L} [R_{LB}(\rho) - R_{UB}(\alpha_{\ell})] = \frac{1}{L} \sum_{\ell=1}^{L} \left[ \frac{T \rho - T \alpha_{\ell}}{1 + T \rho} + \log \frac{1 + T \alpha_{\ell}}{1 + T \rho} + (T - 1) \mathbb{E} \left[ \log \left( \frac{1 + T \rho}{1 + T \alpha_{\ell}} Z_{1,\ell} + Z_{2,\ell} + \beta(T, \rho) \right) \right. \
\left. - (T - 1) \mathbb{E} \left[ \log \left( \frac{1 + T \rho}{1 + T \rho} Z_{1,\ell} + Z_{2,\ell} + \beta(T, \rho) \right) \right] \right] \right. \]  

\[ - \left. \sqrt{\frac{V_{UB}(T)}{L}} Q^{-1}(\tau) - \sqrt{\frac{V_{UB}(T)}{L(1 - \epsilon) - \sqrt{L}}} \right) \]  

(125)
where we have evaluated $R_{\text{LB}}(\rho)$ and $R_{\text{UB}}(\alpha_\ell)$ using (15) and (22). Since $\rho \mapsto \beta(T, \rho)$ is decreasing in $\rho$, we can lower-bound the third-term on the RHS of (125) by replacing $\beta(T, \rho)$ by $\beta(T, \rho_0)$. Furthermore, using that, for $\alpha_\ell \leq \rho$, the first term on the RHS of (125) is non-negative, we obtain

$$
\frac{1}{L} \sum_{\ell=1}^{L} [R_{\text{LB}}(\rho) - R_{\text{UB}}(\alpha_\ell)] \geq \frac{1}{L} \sum_{\ell=1}^{L} \left[ \log \frac{1 + T \alpha_\ell}{1 + T \rho} + (T - 1) \mathbb{E} \left[ \log \frac{(1 + T \rho)Z_{1,\ell} + T - 1 + \beta(T, \rho_0)}{(1 + T \alpha_\ell)Z_{1,\ell} + T - 1 + \beta(T, \rho_0)} \right] 
- (T - 1) \mathbb{E} \left[ \log \left( 1 + \frac{\beta(T, \rho)}{(1 + T \rho)Z_{1,\ell} + T - 1 + \beta(T, \rho_0)} \right) \right] 
- \sqrt{\frac{\bar{V}_{\text{UB}}(T)}{L}} Q^{-1}(\tau) 
- \sqrt{\frac{\bar{V}_{\text{UB}}(T)}{L(1 - \rho) - \sqrt{L}}} \right]
$$

$$
\triangleq \frac{1}{L} \sum_{\ell=1}^{L} \Delta_{L,\rho}(\alpha_\ell), \quad \rho \geq \rho_0
$$

(126)

We next analyze the behaviour of $\alpha_\ell \mapsto \Delta_{L,\rho}(\alpha_\ell)$. Let

$$
g_{\rho}(\alpha_\ell) \triangleq \log \frac{1 + T \alpha_\ell}{1 + T \rho} + (T - 1) \mathbb{E} \left[ \log \frac{(1 + T \rho)Z_{1,\ell} + T - 1 + \beta(T, \rho_0)}{(1 + T \alpha_\ell)Z_{1,\ell} + T - 1 + \beta(T, \rho_0)} \right]
$$

(127)

and

$$
\omega_{L,\rho} \triangleq (T - 1) \mathbb{E} \left[ \log \left( 1 + \frac{\beta(T, \rho)}{(1 + T \rho)Z_{1,\ell} + T - 1 + \beta(T, \rho_0)} \right) \right] + \sqrt{\frac{\bar{V}_{\text{UB}}(T)}{L}} Q^{-1}(\tau) + \sqrt{\frac{\bar{V}_{\text{UB}}(T)}{L(1 - \rho) - \sqrt{L}}} + \log \frac{L}{2L}
$$

(128)

Thus, $\Delta_{L,\rho}(\alpha_\ell) = g_{\rho}(\alpha_\ell) - \omega_{L,\rho}$. Note that $\frac{\partial}{\partial \alpha_\ell} g_{\rho}(\alpha_\ell) = \frac{\partial}{\partial \alpha_\ell} \Delta_{L,\rho}(\alpha_\ell)$, since $\omega_{L,\rho}$ does not depend on $\alpha_\ell$. Further note that

$$
\lim_{L \to \infty} \lim_{\rho \to \infty} (T - 1) \mathbb{E} \left[ \log \left( 1 + \frac{\beta(T, \rho)}{(1 + T \rho)Z_{1,\ell} + T - 1 + \beta(T, \rho_0)} \right) \right] + \lim_{L \to \infty} \left( \sqrt{\frac{\bar{V}_{\text{UB}}(T)}{L}} Q^{-1}(\tau) + \sqrt{\frac{\bar{V}_{\text{UB}}(T)}{L(1 - \rho) - \sqrt{L}}} + \log \frac{L}{2L} \right) = 0
$$

(129)

where the first term in (129) vanishes by the Dominated Convergence Theorem. We next study $\alpha_\ell \mapsto g_{\rho}(\alpha_\ell)$.

**Lemma 8:** The function $\alpha \mapsto g_{\rho}(\alpha)$ has the following properties:

1. The derivative of $\alpha \mapsto g_{\rho}(\alpha)$ is either always positive, always negative, or changes its sign once from positive to negative. This implies that the function is either increasing, decreasing, or unimodal. Hence, it is minimized at an endpoint of $[0, \rho]$.
2. The derivative of $\alpha \mapsto g_{\rho}(\alpha)$ does not depend on $\rho$.
3. We have $g_{\rho}(\rho) = 0$. Furthermore, $\lim_{\rho \to \infty} g_{\rho}(0) = \infty$ for $T > 2$.

**Proof:** See Appendix [VII]

We next study those $\alpha$’s for which $\Delta_{L,\rho}(\alpha_\ell) \geq 0$, since they can be discarded without loss of optimality. Let $\mathcal{L}_\alpha(\delta) \triangleq \{ \ell = 1, \ldots, L : \alpha \geq \rho(1 - \delta) \}$ and $L_\alpha(\delta)$ denote the number of $\alpha_\ell$’s in $\alpha$ that satisfy $\rho(1 - \delta) \leq \alpha_\ell \leq \rho$. Further let

$$
\Delta_{\text{min}} \triangleq \inf_{0 \leq \alpha < \rho(1 - \delta)} \Delta_{L,\rho}(\alpha)
$$

(130)

We can express (126) as

$$
\sum_{\ell=1}^{L} \Delta_{L,\rho}(\alpha_\ell) = \sum_{\mathcal{L}_\alpha(\delta)} \Delta_{L,\rho}(\alpha_\ell) + \sum_{\mathcal{L}_\alpha^c(\delta)} \Delta_{L,\rho}(\alpha_\ell)
$$

(131)

where $\mathcal{L}_\alpha^c(\delta)$ denotes the complement of $\mathcal{L}_\alpha(\delta)$. By Lemma 8, we have that,

$$
\Delta_{L,\rho}(\alpha_\ell) \geq -\omega_{L,\rho}
$$

(132)
for \( \rho \) sufficiently large. Thus, lower-bounding the first sum on the RHS of (131) by \(-L_{\alpha}(\delta)\omega_{L,\rho}\), and the second sum on the RHS of (131) by \((L - L_{\alpha}(\delta))\Delta_{\min}\), we can discard without loss of optimality those \( \alpha \)'s for which

\[
L\Delta_{\min} \geq L_{\alpha}(\delta)(\omega_{L,\rho} + \Delta_{\min}).
\]  

(133)

This implies that an optimal power allocation \( \alpha \) must satisfy

\[
\frac{L_{\alpha}(\delta)}{L} > \frac{\Delta_{\min}}{\omega_{L,\rho} + \Delta_{\min}}.
\]  

(134)

As we shall show below, we have

\[
\lim_{L \to \infty, \rho \to \infty} \Delta_{\min} > 0.
\]  

(135)

Furthermore, \( \omega_{L,\rho} \) vanishes as \( \rho \) and \( L \) tend to infinity. It thus follows that \( L_{\alpha}(\delta)/L \) tends to one as \( \rho \) and \( L \) tend to infinity. This implies that there exist \( L_0 \) and \( \rho_0 \) sufficiently large such that, for \( L \geq L_0 \) and \( \rho \geq \rho_0 \), \( L_{\alpha}(\delta) \geq L/2 \), thereby proving Lemma 1.

It remains to show (135). Let \( \alpha_{\min} = \rho(1 - \delta) \). By Part 1) of Lemma 8 \( \alpha \mapsto g_\rho(\alpha) \) has exactly one maximizer, which we shall denote by \( \alpha^* \). Since \( \omega_{L,\rho} \) does not depend on \( \alpha \), it follows that \( \alpha^* \) also maximizes \( \alpha \mapsto \Delta_{L,\rho}(\alpha) \).

Furthermore, the infimum of \( \Delta_{L,\rho}(\alpha) \) over \( 0 \leq \alpha < \rho(1 - \delta) \) is either achieved at \( \alpha = 0 \) or at \( \alpha_{\min} \).

By Part 3) of Lemma 8 and (129), we have that

\[
\lim_{L \to \infty, \rho \to \infty} \Delta_{L,\rho}(0) > 0.
\]  

(136)

We next show that

\[
\lim_{L \to \infty, \rho \to \infty} \Delta_{L,\rho}(\alpha_{\min}) > 0.
\]  

(137)

If \( \alpha_{\min} \leq \alpha^* \), then this is automatically satisfied, since in this case \( \Delta_{L,\rho}(\alpha_{\min}) \geq \Delta_{L,\rho}(0) \). We thus focus on the case where \( \alpha_{\min} > \alpha^* \). Note that

\[
\Delta_{L,\rho}(\rho) - \Delta_{L,\rho}(\alpha_{\min}) = -\omega_{L,\rho} - \Delta_{L,\rho}(\alpha_{\min})
\]  

(138)

since \( g_\rho(\rho) = 0 \), so by the Mean Value Theorem there exists an \( x_0 \in [\alpha_{\min}, \rho] \) satisfying

\[
-\omega_{L,\rho} - \Delta_{L,\rho}(\alpha_{\min}) = \int_{\alpha_{\min}}^{\rho} \Delta'_{L,\rho}(\alpha) d\alpha = \rho \delta \Delta'_{L,\rho}(x_0)
\]  

(139)

where \( \Delta'_{L,\rho}(\cdot) \) denotes the derivative of \( \Delta_{L,\rho}(\cdot) \). We can thus lower-bound

\[
\Delta_{L,\rho}(\alpha_{\min}) \geq -\omega_{L,\rho} - \delta \sup_{x \in [\alpha_{\min}, \rho]} \rho \Delta'_{L,\rho}(x).
\]  

(140)

To prove (137), it remains to show that

\[
\lim_{L \to \infty, \rho \to \infty} \left\{-\delta \sup_{x \in [\alpha_{\min}, \rho]} \rho \Delta'_{L,\rho}(x)\right\} > 0.
\]  

(141)

Indeed, recall that \( \Delta'_{L,\rho}(x) = g''_\rho(x) \). Furthermore, (cf. (165), Appendix VII),

\[
\rho g'_\rho(\alpha) = \frac{T \rho}{1 + T \alpha} \left\{-\left( T - 2 \right) + \left( T - 1 \right) \frac{T - 1 + \beta(T, \rho_0)}{1 + T \alpha} e^{\frac{-\left( T - 1 + \beta(T, \rho_0) \right) \alpha}{1 + T \alpha}} \right\}
\]  

(142)

where \( E_1(\cdot) \) denotes the exponential integral function [14 Sec. 5.1.1]. The term inside the curly brackets is independent of \( L \) and \( \rho \) and strictly decreasing in \( \alpha \) (cf. Appendix VII), hence it is strictly negative for every \( \alpha > \alpha^* \). (This also implies that \( \alpha^* \) is independent of \( \rho \), whereas \( \alpha_{\min} \) grows to infinity as \( \rho \to \infty \)), hence, \( (\alpha_{\min} - \alpha^*) \to \infty \) as \( \rho \to \infty \).) Moreover, the term outside the curly brackets is bounded away from zero for any \( \alpha_{\min} \leq \alpha \leq \rho \). Thus, the claim (141) follows, which in turn proves (135).
B. Proof of Lemma 2

1) Part 1): The difference between $\bar{J}(\alpha)$ and $\bar{J}(\rho)$ can be lower-bounded by

$$J(\rho) - J(\alpha) \geq g_\rho(\alpha).$$

(143)

By Parts 1) and 3) of Lemma 8, $g_\rho(\cdot)$ is nonnegative for sufficiently large $\rho$. It thus follows that, for sufficiently large $\rho$,

$$\sup_{0 \leq \alpha \leq \rho} J(\alpha) = J(\rho).$$

(144)

This proves Part 1) of Lemma 2.

2) Part 2): To study the maximization problem

$$\sup_{0 \leq \alpha \leq \rho} \left\{ \frac{J(\alpha)}{T} - \sqrt{\frac{V_\rho(\alpha)}{LT^2}}Q^{-1}(\epsilon) \right\}$$

(145)

we study the difference

$$J(\rho) - J(\alpha) - \sqrt{\frac{V(\rho)}{L}}Q^{-1}(\epsilon) + \sqrt{\frac{V_\rho(\alpha)}{L}}Q^{-1}(\epsilon) \geq g_\rho(\alpha) - \sqrt{\frac{V_\rho(\alpha)}{L}}Q^{-1}(\epsilon).$$

(146)

Clearly, every $\alpha$ for which the RHS of (146) is nonnegative can be discarded without loss of optimality. Recall that, by assumption, $0 < \epsilon < \frac{1}{2}$, so $Q^{-1}(\epsilon) > 0$. We thus continue by lower-bounding $\bar{V}_{\rho}(\alpha) \geq 0$ and by using that, for sufficiently large $\rho_0$, we have $\bar{V}(\rho) \leq \bar{V}_{UB}(T, \rho_0)$ for some constant $\bar{V}_{UB}(T, \rho_0)$ that is independent of $\rho$ (Lemma 5):

$$g_\rho(\alpha) \geq g_\rho(\alpha) - \sqrt{\frac{V_{UB}(T, \rho_0)}{L}}Q^{-1}(\epsilon) \equiv f_{L,\rho}(\alpha).$$

(147)

Again, the values of $\alpha$ for which $f_{L,\rho}(\alpha) \geq 0$ are suboptimal and can be discarded without loss of optimality.

Let us write $f_{L,\rho}(\alpha)$ as $f_{L,\rho}(\alpha) \approx g_\rho(\alpha) - \omega_L$, where $\omega_L = \sqrt{\frac{V_{UB}(T, \rho_0)}{L}}Q^{-1}(\epsilon)$. Further let $\delta_L = 1 - \frac{\omega_L}{\rho}$, where $\alpha_0$ is the unique real root of $\alpha \mapsto f_{L,\rho}(\alpha)$, i.e., $f_{L,\rho}(\alpha_0) = 0$. Indeed, we know that $\alpha \mapsto f_{L,\rho}(\alpha)$ has only one zero crossing because $g_\rho'(\alpha) = f_{L,\rho}'(\alpha)$, and by Parts 1) and 3) of Lemma 8, for sufficiently large $\rho$, $\alpha \mapsto f_{L,\rho}(\alpha)$ is either unimodal with maximum or strictly decreasing. Furthermore, since $\omega_L \geq 0$ and $\lim_{L \to \infty} \omega_L = 0$, we have that $f_{L,\rho}(\rho) = -\omega_L \leq 0$ and $f_{L,\rho}(0) > 0$ for $L$ and $\rho$ sufficiently large, hence the claim follows. By the same line of arguments, we also conclude that all $\alpha$’s between 0 and $\rho(1 - \delta_L)$ can be discarded, since for such $\alpha$’s the function $f_{L,\rho}(\alpha)$ is nonnegative.

To study the behavior of $\delta_L$, we next note that $\omega_L = -(f_{L,\rho}(\rho) - f_{L,\rho}(\alpha_0))$, so by the Mean Value Theorem there exists an $x_0 \in [\alpha_0, \rho]$ such that

$$\omega_L = -\int_{\alpha_0}^{\rho} f_{L,\rho}'(\alpha)d\alpha = -\rho \delta_L f_{L,\rho}'(x_0).$$

(148)

It follows that

$$\omega_L \geq \delta_L \inf_{x \in [\alpha_0, \rho]} |\rho f_{L,\rho}'(x)|$$

(149)

where we use that $x \mapsto f_{L,\rho}'(x)$ is negative for $\alpha_0 < x < \rho$. By Lemma 8, $\alpha \mapsto g_\rho(\alpha)$ has exactly one maximizer which we shall denote by $\alpha^*$. Since $\omega_L$ does not depend on $\alpha$, $\alpha^*$ also maximizes $\alpha \mapsto f_{L,\rho}(\alpha)$. Note that $\alpha^*$ does not depend on $\rho$, since by Part 2) of Lemma 8 the derivative of $\alpha \mapsto g_\rho(\alpha)$ does not depend on $\rho$. 

23
We next show that there exists an \( \tilde{\alpha} \) independent of \( L \) and \( \rho \), such that \( \alpha_0 \geq \tilde{\alpha} > \alpha^* \). Indeed, by Lemma 8 we have that for sufficiently large \( \rho \), the function \( \alpha \mapsto g_\rho(\alpha) \) is either unimodal with maximum or strictly decreasing and \( g_\rho(0) > 0 \). This implies that \( g_\rho(\alpha^*) > 0 \), which in turn implies that

\[
\inf_{L \geq L_0, \rho \geq \rho_0} f_{L, \rho}(\alpha^*) > 0
\]

for \( L_0 \) and \( \rho_0 \) sufficiently large, since \( \lim_{L \to \infty} \omega_L = 0 \). We next note that

\[
\inf_{L \geq L_0, \rho \geq \rho_0} f_{L, \rho}(\tilde{\alpha}) \geq \inf_{L \geq L_0, \rho \geq \rho_0} f_{L, \rho}(\alpha^*) - |f_{L, \rho}(\tilde{\alpha}) - f_{L, \rho}(\alpha^*)|
\]

since the difference

\[
f_{L, \rho}(\tilde{\alpha}) - f_{L, \rho}(\alpha^*) = g_\rho(\tilde{\alpha}) - g_\rho(\alpha^*)
\]

\[
= \log \left( \frac{1 + T\tilde{\alpha}}{1 + T\alpha^*} \right) + (T - 1)E \left[ \log \left( \frac{(1 + T\alpha^*)Z_1 + T - 1 + \beta(T, \rho_0)}{(1 + T\tilde{\alpha})Z_1 + T - 1 + \beta(T, \rho_0)} \right) \right]
\]

is independent of \( L \) and \( \rho \). By the continuity of \( \alpha \mapsto g_\rho(\alpha) \), it follows from (150)–(152) that there exists an \( \tilde{\alpha} \in (\alpha^*, \rho] \) that is independent of \( L \) and \( \rho \) and that satisfies

\[
\inf_{L \geq L_0, \rho \geq \rho_0} f_{L, \rho}(\tilde{\alpha}) > 0
\]

for \( L_0 \) and \( \rho_0 \) sufficiently large. Hence, \( \tilde{\alpha} \) lies between \( \alpha^* \) and \( \alpha_0 \). It follows that \( \omega_L \) in (149) can be lower-bounded by

\[
\omega_L \geq \delta_L \inf_{x \in [\alpha, \rho]} |\rho f'_{L, \rho}(x)|.
\]

We next show that \( \inf_{x \in [\tilde{\alpha}, \rho]} |\rho f'_{L, \rho}(x)| \) is bounded away from zero. As above, we use that \( f'_{L, \rho}(x) = g'_\rho(x) \) and

\[
g'_\rho(\alpha) = \frac{\rho T \rho}{1 + T\alpha} \left\{ -(T - 2) + (T - 1) \frac{T - 1 + \beta(T, \rho_0)}{1 + T\alpha} e^{\frac{T - 1 + \beta(T, \rho_0)}{1 + T\alpha}} E_1 \left( \frac{T - 1 + \beta(T, \rho_0)}{1 + T\alpha} \right) \right\}.
\]

Recall that the term inside the curly brackets is independent of \( L \) and \( \rho \) and strictly negative for every \( \alpha \in [\tilde{\alpha}, \rho] \). Moreover, the term outside the curly brackets is bounded away from zero for every \( \alpha \in [\tilde{\alpha}, \rho], \rho \geq \rho_0 \). We conclude that

\[
F \triangleq \inf_{\rho \geq \rho_0, x \in [\tilde{\alpha}, \rho]} |\rho f'_{L, \rho}(x)| > 0.
\]

It follows from (154) and the definition of \( \omega_L \) that, for sufficiently large \( L_0 \) and \( \rho_0 \),

\[
\delta_L \leq \sqrt{\frac{V_{UB}(T, \rho_0)Q^{-1}(\epsilon)}{F}} \frac{1}{\sqrt{L}}.
\]

By assuming that, without loss of optimality, \( \rho(1 - \delta_L) \leq \alpha \leq \rho \), we can derive a tighter lower bound on \( \hat{V}_\rho(\alpha) \) using the lower bound given in Appendix VIII instead of lower-bounding it by zero. Indeed, by (168) in Appendix VIII

\[
\sqrt{\frac{V_{\rho}(\alpha)}{L}} \geq \sqrt{\frac{V(\rho) - \Upsilon \delta_L}{L}} \geq \sqrt{\frac{V(\rho)}{L}} - \sqrt{\frac{\Upsilon \delta_L}{L}}, \quad \rho(1 - \delta_L) \leq \alpha \leq \rho.
\]

We then obtain

\[
\hat{J}(\rho) - \hat{J}(\alpha) \geq g_\rho(\alpha) - \sqrt{\frac{\Upsilon \delta_L}{L}} Q^{-1}(\epsilon)
\]

\[
\triangleq \tilde{f}_{L, \rho}(\alpha), \quad \rho(1 - \delta_L) \leq \alpha \leq \rho
\]

Again, the values of \( \alpha \) for which \( \tilde{f}_{L, \rho}(\alpha) \geq 0 \) are suboptimal and can be discarded without loss of optimality.
Let us write \( \tilde{f}_{L,\rho}(\alpha) = g_{\rho}(\alpha) - \tilde{\omega}_L \), where \( \tilde{\omega}_L = \sqrt{\frac{\delta_L}{L}} Q^{-1}(\epsilon) \). Further let \( \delta_L = 1 - \frac{\tilde{\omega}_L}{\rho} \), where \( \tilde{\omega}_0 \) is the unique real root of \( \alpha \mapsto \tilde{f}_{L,\rho}(\alpha) \). As above, it can be shown that all \( \alpha \)'s between 0 and \( \rho(1 - \delta_L) \) can be discarded, since for such \( \alpha \)'s the function \( \tilde{f}_{L,\rho}(\alpha) \) is nonnegative. By repeating the steps (148)–(157) but with \( \omega_L \) replaced by \( \tilde{\omega}_L \), we obtain that, for sufficiently large \( L_0 \) and \( \rho_0 \),

\[
\tilde{\delta}_L \leq \frac{\tilde{\omega}_L}{\rho} \leq \left( \frac{Q^{-1}(\epsilon)}{\rho} \right)^{3/2} \sqrt{\gamma} \sqrt{\frac{\tilde{\omega}_L(T, \rho_0)}{\gamma}} \frac{1}{L^{3/4}}
\]

(160)

where the last inequality follows from (157).

If we perform the above steps \( N \) times, then we can show that, without loss of optimality,

\[
\alpha \geq \rho \left( 1 - \delta^{(N)}_L \right)
\]

(161)

where \( \delta^{(N)}_L \) satisfies

\[
0 \leq \delta^{(N)}_L \leq \left( \frac{Q^{-1}(\epsilon) \sqrt{\gamma}}{\rho} \right)^{2 - 2^{-N+1}} \left( \frac{\tilde{\omega}_L(T, \rho_0)}{\gamma} \right)^{2^{-N}} \left( \frac{1}{L^{1 - 2^{-N}}} \right).
\]

(162)

By letting \( N \) tend to infinity, we conclude that we can assume, without loss of optimality, that

\[
\alpha \geq \rho \left( 1 - \delta^{(\infty)}_L \right)
\]

(163)

where \( \delta^{(\infty)} \) satisfies

\[
0 \leq \delta^{(\infty)} \leq K \frac{L}{\gamma}
\]

(164)

for some positive constant that is independent of \((L, \rho, \alpha)\). This concludes the proof of Part 2) of Lemma 8.

**APPENDIX VII**

**PROOF OF LEMMA 8**

The derivative of \( \alpha \mapsto g_{\rho}(\alpha) \) can be expressed as

\[
g'_{\rho}(\alpha) = \frac{T}{1 + T\alpha} - (T - 1)E \left[ \frac{T Z_1}{(1 + T\alpha)Z_1 + (T - 1) + \beta(T, \rho_0)} \right]
\]

\[
= T \left[ \frac{1}{1 + T\alpha} - \frac{T - 1}{1 + T\alpha} + \frac{T - 1}{1 + T\alpha} + \beta(T, \rho_0) e^{\frac{T - 1 + \beta(T, \rho_0)}{1 + T\alpha}} E_1 \left( \frac{T - 1 + \beta(T, \rho_0)}{1 + T\alpha} \right) \right]
\]

\[
= \frac{T}{1 + T\alpha} \left[ -(T - 2) + (T - 1) \frac{T - 1 + \beta(T, \rho_0)}{1 + T\alpha} e^{\frac{T - 1 + \beta(T, \rho_0)}{1 + T\alpha}} E_1 \left( \frac{T - 1 + \beta(T, \rho_0)}{1 + T\alpha} \right) \right]
\]

(165)

where the first step follows because, by [22] App. A.9, we can swap derivative and expected value; the second step follows by solving the expected value using [16 Sec. 3.353-5.7]. Note that the RHS of (165) does not depend on \( \rho \). Hence Part 2) of Lemma 8 follows immediately.

We next prove Part 1) of Lemma 8. Because \( T/(1 + T\alpha) \) in (165) is non-negative, the sign of \( \alpha \mapsto g'_{\rho}(\alpha) \) is determined by the terms inside the square brackets. Let \( \vartheta \triangleq \frac{1 + T\alpha}{T - 1 + \beta(T, \rho_0)} \). Note that \( \vartheta \mapsto \frac{1}{\vartheta} \exp \left( \frac{1}{\vartheta} \right) E_1 \left( \frac{1}{\vartheta} \right) \) is strictly decreasing since, by [16 Eq. 3.353-3],

\[
\frac{1}{\vartheta} e^{\frac{1}{\vartheta}} E_1 \left( \frac{1}{\vartheta} \right) = 1 - \int_{0}^{1} e^{-\frac{t}{1 - \vartheta t}} dt
\]

(166)

and \( \vartheta \mapsto e^{-\frac{1}{1 - \vartheta t}} \) is strictly positive and strictly increasing in \( \vartheta \). This implies that \( \alpha \mapsto g'_{\rho}(\alpha) \) changes its sign at most once, and if it does then it changes from positive to negative. Consequently, \( \alpha \mapsto g_{\rho}(\alpha) \) achieves its minimum either at 0 or at \( \rho \).
We finally prove Part 3) of Lemma 8 by showing that \( \lim_{\rho \to \infty} g_\rho(0) > 0 \) for \( T > 2 \). To this end, we express \( g_\rho(0) \) as

\[
g_\rho(0) = (T - 2)E \left[ \log \left( 1 + \frac{T \rho Z_1}{Z_1 + (T - 1) + \beta(T, \rho_0)} \right) \right] + E \left[ \log \left( Z_1 + \frac{T - 1 + \beta(T, \rho_0)}{1 + T \rho} \right) \right] - E \left[ \log (Z_1 + T - 1 + \beta(T, \rho_0)) \right].
\]  
(167)

The first expected value on the RHS of (167) tends to infinity as \( \rho \to \infty \), whereas the other expected values are bounded in \( \rho \). It thus follows that, for \( T > 2 \), the RHS of (167) tends to infinity as \( \rho \to \infty \), hence the claim follows. This concludes the proof of Lemma 8.

**APPENDIX VIII**

**LOWER BOUND ON \( \bar{V}_\rho(\alpha) \)**

We show that for \( \alpha \in [\rho(1 - \delta), \rho] \) and \( 0 \leq \delta \leq 1/2 \)

\[
\bar{V}_\rho(\alpha) \geq \bar{V}(\rho) - T \delta
\]  
(168)

where \( \Upsilon \) is a constant that is independent of \( (L, \rho, \alpha) \). Let \( \Omega(\alpha) \triangleq \bar{j}_\alpha(\alpha) - \bar{j}(\alpha) \), i.e.,

\[
\Omega(\alpha) = -T \rho - T \alpha \left( Z_1 - 1 \right) - \frac{T \rho}{1 + T \rho} \left( Z_2 - (T - 1) \right)
+ (T - 1) \log \left( 1 + T \alpha \right) Z_1 + Z_2 + \beta(T, \rho)) - (T - 1) E \left[ \log \left( 1 + T \alpha \right) Z_1 + Z_2 + \beta(T, \rho) \right].
\]  
(169)

Hence, \( \bar{V}_\rho(\alpha) = E[\Omega^2(\alpha)] \). We next analyze the difference

\[
\bar{V}(\rho) - \bar{V}_\rho(\alpha) = E \left[ \Omega^2(\rho) - \Omega^2(\alpha) \right] = E \left[ (\Omega(\rho) - \Omega(\alpha)) (\Omega(\rho) + \Omega(\alpha)) \right] \leq \sqrt{E[\Omega^2(\rho) - \Omega^2(\alpha)]} E[(\Omega(\rho) + \Omega(\alpha))^2]
\]  
(170)

where the last inequality follows from the Cauchy-Schwarz inequality. On the one hand, using that \((a_1 + a_2)^2 \leq 2a_1^2 + 2a_2^2\), we have for any arbitrary \( \bar{\rho} \),

\[
\sup_{\alpha \geq 0} \sup_{\rho \geq \bar{\rho}} E \left[ (\Omega(\rho) + \Omega(\alpha))^2 \right] \leq \sup_{\rho \geq \bar{\rho}} E \left[ \Omega^2(\rho) \right] + 2 \sup_{\rho \geq \bar{\rho}} E \left[ \Omega^2(\alpha) \right] = 2 \sup_{\rho \geq \bar{\rho}} \bar{V}(\rho) + 2 \sup_{\rho \geq \bar{\rho}} \bar{V}_\rho(\alpha)
\]  
(171)

which, by Lemma 8, is bounded. On the other hand, using that \((a_1 + a_2 + a_3)^2 \leq c(a_1^2 + a_2^2 + a_3^2)\) for some positive constant \( c \) and \( E[X^2] \geq E[X]^2 \) for any random variable \( X \), we obtain

\[
E \left[ (\Omega(\rho) - \Omega(\alpha))^2 \right] = E \left[ \left( \frac{T \rho - T \alpha}{1 + T \rho} \right) \left( Z_1 - 1 \right) - (T - 1) E \left[ \log \left( \left( 1 + T \rho \right) Z_1 + Z_2 + \beta(T, \rho) \right) \right] \right] \leq c \left( \frac{T \rho - T \alpha}{1 + T \rho} \right)^2 + 2c(T - 1)^2 E \left[ \log^2 \left( \left( 1 + T \rho \right) Z_1 + Z_2 + \beta(T, \rho) \right) \right].
\]  
(172)

For \( \alpha \in [\rho(1 - \delta), \rho] \), this can be further upper-bounded as

\[
E \left[ (\Omega(\rho) - \Omega(\alpha))^2 \right] \leq c \delta^2 + 2c(T - 1)^2 \log^2 \left( 1 + \frac{\delta}{1 - \delta} \right) \leq (c + 8c(T - 1)^2) \delta^2
\]  
(173)

where the last inequality follows because, by assumption, \( \delta \leq 1/2 \), so \( \frac{\delta^2}{1 - \delta} \leq 4\delta^2 \). Combining (170) with (171) and (173) proves (168).
REFERENCES


