# FUNDAMENTAL LIMITS OF

# SHORT-PACKET WIRELESS COMMUNICATIONS

by

Alejandro Lancho Serrano

in partial fulfillment of the requirements for the degree of Doctor in

Multimedia and Communications

Universidad Carlos III de Madrid

Tutor/ Advisor:

Tobias Koch

June 24, 2019

*To Noelia and my family*

*"Caminante, no hay camino,*
*se hace camino al andar."*
Antonio Machado (1875-1939)

# Acknowledgements

This thesis starts with a quote that perfectly describes what has brought me here. In english, it says, "*Traveler, there is no road, you make your own path as you walk*". I had never said, when I started my undergraduate studies in 2009, that I was going to doctorate, and even less in something theoretical. But here I am, presenting my information-theoretical thesis. Hence, this piece of an Antonio Machado's poem could not explain better what has been my professional life so far. I guess that the person responsible for the change my mind has suffered during these last years has been my advisor, Tobias Koch, also known as "Tobi".

During my master studies, I could work with the uc3m Communications Group in the field that I most liked at that moment and that I still enjoy today, wireless communications. During that year I could enjoy a very special environment with my lab mates, where I would like to mention Alex, Borja, Javier, Juanjo and Máximo. Thanks to this group of people and the great environment we had, I realized that my initial plan consisting of doing my master studies and going to industry had changed. At that moment, I knew I wanted to do a Ph.D. Before moving to the next stage of the story, I have also to say that this group of people helped me somehow to discover one of my current hobbies: endurance sports, and in particular, running.

When I was finishing my master studies, an opportunity to work with Tobi arised. I knew him from the master classes of information theory. To be frank, I was not sure whether working on something theoretical was a good idea or not, but I was completely sure that I had a lot of things to learn from him, so with effort, the result could be great. Today, I can say that it was the best choice I could make. Tobi has taught me innumerable things. I learnt from him a lot of technical stuff, how to write a paper, how to give a talk, and in general, his idea about how good research must be performed. I really appreciate all your effort, so thanks Tobi!

I want also to say thank you to all the lab mates that I had during this period of time. I can not list all the names here, but thank you all for having shared really good moments.

I would like to say thank you to Gonzalo Vazquez-Vilar for the many discussions we had and for the very nice collaborations that we could also perform. Furthermore, I would like to say thank you to Prof. Giuseppe Durisi and his team for receiving me in Gothenburg during the fall of 2017 and make me feel like at home from the very first day. Especially, I would like to thank Johan for the many productive discussions we had, and Sven, for trusting me and renting his apartment to me during the time I was in Gothenburg.

Since there is life beyond work, I would like to thank all my friends that have been with me during this time. Especially, I would like to mention first Borja and Sergio. The first, a friend from highschool who studied also the bachelor degree with me, and who is also about to get his Ph.D. in the same university department. The second, I can say that he is part of my family, I know him since I was born, and he also studied with me the bachelor degree. I shared really good moments with both of them. I also would like to mention César and Dani. With them, I share the passion of running, which is my main hobby and outlet. They have showed me they are really good friends.

And last but not least, I would like to thank my family, especially, my parents, my brother Alberto, and Noelia. I have to thank my parents and brother for giving me their unconditional support in every stage of my studies. I still remember how my brother asked for permission to speak at home when I was studying. Thanks to that respect I could study at home in the best environment one could imagine. Finally, I would like to thank Noelia for having supported me since we started our relationship more than eight years ago. Having someone who is willing to be with you and to move with you all around the world if necessary is worth its weight in gold. For all of this, I can say that, without you, none of this would have been possible.

# PUBLISHED AND SUBMITTED CONTENT

The technical contributions presented in this thesis have been published in the following journals and conference papers:

1. A. Lancho, J. Östman, G. Durisi, T. Koch and G. Vazquez-Vilar, "Saddlepoint approximations for Rayleigh block-fading channels", in Proc. IEEE Int. Symp. Inf. Theory (ISIT), Paris, France, Jul. 2019. **Accepted for publication**.
   This work is wholly included in Chapters 4 and 6.

2. A. Lancho, T. Koch and G. Durisi. "Normal Approximations for Fading Channels", in Proc. IEEE 52th Annual Conference on Information Sciences and Systems, Princeton, NJ, Mar. 2018. **Invited paper** `https://doi.org/10.1109/CISS.2018.8362297`.
   This work is partly included in Chapters 4 and 5.

3. A. Lancho, T. Koch and G. Durisi. "A High-SNR Normal Approximation for Single-Antenna Rayleigh Block-Fading Channels", in Proc. IEEE Int. Symp. Inf. Theory (ISIT), Aachen, Germany, Jun. 2017. **Finalist IEEE Jack Keil Wolf ISIT Student Paper Award** `https://doi.org/10.1109/ISIT.2017.8006834`.
   This work is wholly included in Chapters 4 and 5.

**Preprints**

1. A. Lancho, J. Östman, G. Durisi, T. Koch and G. Vazquez-Vilar, "Saddlepoint approximations for Rayleigh block-fading channels", Apr. 2019. (Available: `https://arxiv.org/abs/1904.10442`).
   This work is wholly included in Chapters 4 and 6.

2. A. Lancho, T. Koch and G. Durisi , "On single-antenna Rayleigh block-fading channels at finite blocklength", Jul. 2017. (Available: `https://arxiv.org/abs/1706.07778`). **Submitted to the IEEE Transactions on Information Theory. Status: "Second round of revision"**.
   This work is wholly included in Chapters 4 and 5.

# MERITS NOT INCLUDED IN THIS THESIS

1. G. Vazquez-Vilar, A. Guillén i Fàbregas, T. Koch and A. Lancho. "Saddlepoint approximation of the error probability of binary hypothesis testing", in Proc. IEEE Int. Symp. Inf. Theory (ISIT). Vail, USA, Jun 17-22, 2018. https://doi.org/10.1109/ISIT.2018.8437503.

2. J. Font-Segura, G. Vazquez-Vilar, A. Martinez, A. Guillén i Fàbregas and A. Lancho. "Saddlepoint Approximations of Lower and Upper Bounds to the Error Probability in Channel Coding", in Proc. 52th Annual Conference on Information Sciences and Systems, Princeton, NJ, March 21-23, 2018. **Invited paper** https://doi.org/10.1109/CISS.2018.8362304.

3. V. P. Gil Jiménez, A. Lancho Serrano, B. Genovés Guzmán and A. Garcia Armada, "Learning Mobile Communications Standards Through Flexible Software Defined Radio Base Stations“, IEEE Commun. Mag. vol. 55, no. 6, pp. 116-123, May 2017. https://doi.org/10.1109/MCOM.2017.1601219.

4. B. Genovés Guzmán, A. Lancho Serrano and V. P. Gil Jiménez , "Cooperative Optical Wireless Transmission for improving performance in indoor scenarios for Visible Light Communications“, IEEE Trans. on Consum. Electron. vol. 61, no. 4, pp. 393-401, Nov. 2015. **First and second authors equal contribution**. https://doi.org/10.1109/TCE.2015.7389772.

# Abstract

This thesis concerns the maximum coding rate at which data can be transmitted over a noncoherent, 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. This is an emerging problem motivated, *inter alia*, by the next generation of wireless communications, where the understanding of the fundamental limits in the transmission of short packets is crucial. For this setting, traditional information-theoretical metrics of performance that rely on the transmission of long packets, such as *capacity* or *outage capacity*, are not good benchmarks anymore, and the study of the maximum coding rate as a function of the blocklength is needed. For the noncoherent Rayleigh block-fading channel model, to study the maximum coding rate as a function of the blocklength, only nonasymptotic bounds that must be evaluated numerically were available in the literature. The principal drawback of the nonasymptotic bounds is their high computational cost, which increases linearly with the number of blocks (also called throughout this thesis *coherence intervals*) needed to transmit a given codeword. By means of different asymptotic expansions in the number of blocks, this thesis provides an alternative way of studying the maximum coding rate as a function of the blocklength for the noncoherent, single-antenna, Rayleigh block-fading channel.

The first approximation on the maximum coding rate derived in this thesis is a high-SNR *normal approximation*. This central-limit-theorem-based approximation becomes accurate as the signal-to-noise ratio (SNR) and the number of coherence intervals $L$ of size $\mathsf{T}$ tend to infinity. We show that the high-SNR normal approximation is roughly equal to the normal approximation one obtains by transmitting one pilot symbol per coherence block to estimate the fading coefficient, and by then transmitting $\mathsf{T}-1$ symbols per coherence block over a coherent fading channel. This suggests that, at high SNR, one pilot symbol per coherence block suffices to achieve both the capacity and the channel dispersion. While the approximation was derived under the assumption that the number of coherence intervals and the SNR tend to infinity, numerical analyses suggest that it becomes accurate already at SNR values of 15 dB, for 10 coherence intervals or more, and probabilities of error of $10^{-3}$ or more.

The derived normal approximation is not only useful because it complements the nonasymptotic bounds available in the literature, but also because it lays the foundation for analytical studies that analyze the behavior of the maximum coding rate as a function of system parameters such as SNR, number of coherence intervals, or blocklength. An example of such a study concerns the optimal design of a simple

slotted-ALOHA protocol, which is also given in this thesis.

Since a big amount of services and applications in the next generation of wireless communication systems will require to operate at low SNRs and small probabilities of error (for instance, SNR values of 0 dB and probabilities of error of $10^{-6}$), the second half of this thesis presents *saddlepoint approximations* of upper and lower nonasymptotic bounds on the maximum coding rate that are accurate in that regime. Similar to the normal approximation, these approximations become accurate as the number of coherence intervals $L$ increases, and they can be calculated efficiently. Indeed, compared to the nonasymptotic bounds, which require the evaluation of $L$-dimensional integrals, the saddlepoint approximations only require the evaluation of four one-dimensional integrals. Although developed under the assumption of large $L$, the saddlepoint approximations are shown to be accurate even for $L = 1$ and SNR values of 0 dB or more. The small computational cost of these approximations can be further avoided by performing high-SNR saddlepoint approximations that can be evaluated in closed form. These approximations are shown to be accurate for 10 dB or more.

In our analysis, the saddlepoint method is applied to the tail probabilities appearing in the nonasymptotic bounds. These probabilities often depend on a set of parameters, such as the SNR. Existing saddlepoint expansions do not consider such dependencies. Hence, they can only characterize the behavior of the expansion error in function of the number of coherence intervals $L$, but not in terms of the remaining parameters. In contrast, we derive a saddlepoint expansion for random variables whose distribution depends on an extra parameter, carefully analyze the error terms, and demonstrate that they are uniform in such an extra parameter. We then apply the expansion to the Rayleigh block-fading channel and obtain approximations in which the error terms depend only on the blocklength and are uniform in the remaining parameters.

The proposed approximations are shown to recover the normal approximation and the *reliability function* of the channel, thus providing a unifying tool for the two regimes, which are usually considered separately in the literature. Specifically, we show that the high-SNR normal approximation can be recovered from the normal approximation derived from the saddlepoint approximations. By means of an error-exponent analysis that recovers the reliability function of the channel, we also obtain easier-to-evaluate approximations of the saddlepoint approximations consisting of the error exponent of the channel multiplied by a subexponential factor. Numerical evidence suggests that these approximations are nearby as accurate as the saddlepoint approximations.

**Keywords:** Channel dispersion, fifth generation, finite blocklength, high SNR, information theory, machine-type communications, noncoherent setting, normal approximation, Rayleigh block-fading channel, saddlepoint approximation, ultra-reliable low-latency communications, wireless communications.

# Contents

# List of Figures

# List of Tables

# List of Acronyms

| | |
|---|---|
| 4G | fourth generation |
| 5G | fifth generation |
| APSK | amplitude phase-shift keying |
| ARJA | accumulate-repeat-jagged-accumulate |
| AWGN | additive white Gaussian noise |
| CDF | cumulative distribution function |
| CGF | cumulant generating function |
| DT | dependence testing |
| eMBB | enhanced mobile broadband |
| i.i.d. | independent and identically distributed |
| IoT | Internet of Things |
| LDPC | low-density parity-check |
| LHS | left-hand side |
| LTE | Long-Term Evolution |
| MC | meta converse |
| MGF | moment generating function |
| MIMO | multiple-input multiple-output |
| MISO | multiple-input single-output |
| mMTC | massive machine-type communications |
| OFDM | orthogonal frequency-division multiplexing |

| | |
|---|---|
| pdf | probability density function |
| RCU | random coding union |
| RHS | right-hand side |
| SNR | signal-to-noise ratio |
| URLLC | ultra-reliable low-latency communications |
| USTM | unitary space-time modulation |

# 1 Introduction

## 1.1 Motivation

Under the paradigm of the Internet of Things (IoT), next generation wireless communication systems are expected to interconnect a great variety of devices, ranging from vehicles or drones, which will operate in high-mobility scenarios, to autonomous machines or static sensors, operating in low-mobility scenarios [3, 4]. Traditional wireless communication technologies, such as the fourth generation (4G) Long-Term Evolution (LTE) or WiFi, focus on increasing the transmission data rates with no stringent latency constraints. Thus, a long-packet assumption is deemed feasible and, hence, *capacity* and *outage capacity* provide accurate benchmarks for the throughput achievable in such systems. Furthermore, when transmitting long packets, the length of *metadata*—extra information included in packets for the correct operation of the communication protocols—is negligible compared to the length of information payload contained in each packet. Thus, suboptimal encoding of metadata does not imply an impact in terms of efficiency. However, motivated by emerging services and applications that require low latency and high reliability, the fifth generation (5G) of wireless communication systems targets not only increased data rates, but also transmission of short-packets, where metadata can play an important role since its size may be comparable to the size of the information payload [3]. Specifically, 5G systems will support three main services: enhanced mobile broadband (eMBB), massive machine-type communications (mMTC), and ultra-reliable low-latency communications (URLLC) [4].

In eMBB, very high data rates as well as moderate rates for cell-edge users are to be supported maintaining a moderate reliability, i.e., probabilities of error of around $10^{-3}$ [4]. This service can be seen as a natural extension of 4G, where the devices are expected to be activated during long periods of time. As aforementioned, under these requirements, capacity and outage capacity provide good benchmarks.

In mMTC, a massive number of devices operating at low rates will be activated intermittently during very short periods of time with probabilities of error of around $10^{-1}$ [4]. Hence, this service will require the transmission of very short packets.

In URLLC, the devices will transmit short packets at low rates aiming for probabilities of error smaller than or equal to $10^{-5}$ [4]. In URLLC, the devices could also transmit intermittently with periodic control messages, but the main difference with respect to mMTC resides in the smaller number of devices that will be connected to the network.

For mMTC and URLLC, which require the transmission of short packets, traditional asymptotic information theoretical analyses, based on capacity and outage capacity, do not provide good benchmarks. Thus, for low-latency wireless communications, a more refined analysis of the maximum coding rate as a function of the blocklength, commonly named finite-blocklength analysis, is needed. Such an analysis is provided in this thesis.

## 1.2 State of the Art

Several techniques can be used to characterize the finite-blocklength performance. One possibility is to fix a reliability constraint and study the maximum coding rate as a function of the blocklength in the limit as the blocklength tends to infinity. Under this category falls the work on normal approximations for various communication channels. Specifically, among other people, Polyanskiy *et al.* [1] showed that, for various channels with positive capacity $C$, the maximum coding rate $R^*(n,\epsilon)$ at which data can be transmitted using an error-correcting code of fixed length $n$ with a block-error probability not larger than $\epsilon$ can be tightly approximated by

$$R^*(n,\epsilon) = C - \sqrt{\frac{V}{n}}Q^{-1}(\epsilon) + \mathcal{O}\left(\log n/n\right) \tag{1.1}$$

where $V$ denotes the channel dispersion, a quantity that measures the stochastic variability of the channel compared to a deterministic channel with identical capacity; $Q^{-1}(\cdot)$ denotes the inverse Gaussian $Q$-function; and $\mathcal{O}\left(n^{-1}\log n\right)$ comprises terms that decay no slower than $n^{-1}\log n$. The approximation that follows from (1.1) by ignoring the $\mathcal{O}(n^{-1}\log n)$ terms is commonly referred to as *normal approximation*. The work by Polyanskiy *et al.* [1] has been generalized to some wireless channels. For instance, the channel dispersion of coherent fading channels—where the receiver has perfect knowledge of the realizations of the fading coefficients—was studied by Polyanskiy and Verdú for the single-antenna case [5], and by Collins and Polyanskiy for the multiple-input single-output (MISO) Rayleigh block-fading [6] and the multiple-input multiple-output (MIMO) Rayleigh block-fading case [7, 8]. The channel dispersion of single-antenna quasistatic fading channels when both transmitter and receiver have perfect knowledge of the realization of the fading coefficients and the transmitter satisfies a long-term power constraint was obtained by Yang *et al.* [9]. In the noncoherent setting—where neither the transmitter nor the receiver have *a priori* knowledge of the realizations of the fading coefficients—the channel dispersion is only known in the quasistatic case, where it is zero [10, 11]. Upper and lower bounds on the second-order coding rate of quasistatic MIMO Rayleigh-fading channels have further been reported in [12] for the asymptotically-ergodic setup where the number of antennas grows linearly with the blocklength. For noncoherent Rayleigh block-fading channels, nonasymptotic bounds on the maximum coding rate were presented by Yang *et al.* for the single-antenna case [13] and by Östman *et al.* for the MIMO case [14, 15]. For further references see [3].

In a nutshell, in the noncoherent setting the channel dispersion is only known in the quasistatic case. For general block-fading channels, the maximum coding rate needs to be assessed by means of nonasymptotic bounds, whose evaluation is often computationally demanding. Obtaining the channel dispersion of noncoherent block-fading channels is difficult because the capacity-achieving input distribution is in general unknown. Thus, the standard approach to obtain expressions of the form (1.1), which entails an analysis of nonasymptotic upper and lower bounds on $R^*(n,\epsilon)$ based on the capacity-achieving input and output distributions in the limit as $n\to\infty$, cannot be followed. However, the behavior of capacity at high signal-to-noise ratio (SNR) is well understood for such channels. Indeed, it was demonstrated that an input distribution referred to as unitary space-time modulation (USTM) yields a lower bound on the capacity that is asymptotically tight [16, 17, 18]. Thus, a characterization of the channel dispersion at high SNR is feasible.

An alternative analysis of the finite-blocklength performance follows from fixing the coding rate and studying the exponential decay of the error probability as the blocklength grows large. The resulting error exponent is usually referred to as the *reliability function* [19, Ch. 5]. Error exponent results for the fading channel can be found in [20] and [21], where a lower bound on the reliability function is derived for multiple-antenna fading channels and for single-antenna Rician block-fading channels, respectively.

Both the exponential and sub-exponential behavior of the error probability can be characterized via the *saddlepoint method* [22, Ch. XVI]. This method has been applied in [23, 24, 25] to obtain approximations of the random coding union (RCU) bound [1, Th. 16], the RCU bound with parameter $s$ ($\text{RCU}_\text{s}$) [2, Th. 1], and the meta converse (MC) bound [1, Th. 31] for some memoryless channels.

## 1.3 Outline and Contributions

This thesis is organized as follows. Chapter 2 presents the system model used throughout Chapters 3–6. Chapter 3 presents a review of the nonasymptotic bounds on the maximum coding rate (or minimum probability of error) used in Chapters 5 and 6. Chapter 4 introduces the Rayleigh block-fading channel model and the preliminary definitions and results that will be useful throughout Chapters 5–6. Chapter 5 derives a high-SNR normal approximation for noncoherent Rayleigh block-fading channels. Chapter 6 presents saddlepoint approximations for noncoherent Rayleigh block-fading channels. This chapter further demonstrates that the derived approximations recover both the normal approximation and the reliability function of the channel. Chapter 7 concludes the thesis with a summary and discussion of the results.

### Chapter 5: A high-SNR Normal Approximation

In this chapter, we present an expression similar to (1.1) for the maximum coding rate $R^*(L, \epsilon, \rho)$ achievable over noncoherent, single-antenna, Rayleigh block-fading channels using error-correcting codes that span $L$ coherence intervals, have a block-error probability no larger than $\epsilon$, and satisfy the per-coherence-interval maximum 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 nonasymptotic bounds provided in [13, 14, 15]. On the other hand, it allows for a mathematical analysis of $R^*(L, \epsilon, \rho)$.

### Chapter 6: Saddlepoint Approximations

In this chapter, we apply the saddlepoint method to derive approximations of the MC upper bound and the $\text{RCU}_\text{s}$ lower bound on the maximum coding rate $R^*(L, \epsilon, \rho)$ (or *vice-versa* on the minimum probability of error $\epsilon^*(L, R, \rho)$) for noncoherent, single-antenna, Rayleigh block-fading channels using error-correcting codes that span $L$ coherence intervals, have a block-error probability no larger than $\epsilon$, and satisfy the per-coherence-interval equal power constraint $\rho$. While these approximations must be evaluated numerically, the computational complexity is independent of the number of diversity branches $L$. This is in stark contrast to the nonasymptotic MC and $\text{RCU}_\text{s}$ bounds, whose evaluation has a computational complexity that grows linearly in $L$. Numerical evidence suggests that the saddlepoint approximations, although developed under the assumption of large $L$, are accurate even for $L = 1$ if the SNR is greater than or equal to 0 dB. Furthermore, the proposed approximations are shown to recover the normal approximation and the reliability function of the channel, thus providing a unifying tool for the two regimes, which are usually considered separately in the literature.

## 1.4 Notation

We denote scalar random variables by upper case letters such as $X$, and their realizations by lower case letters such as $x$. Likewise, we use boldface upper case letters to denote random vectors, i.e., $\mathbf{X}$, and we use boldface lower case letters such as $\mathbf{x}$ to denote their realizations. We use upper case letters with the standard font to denote distributions, and lower case letters with the standard font to denote probability density functions (pdfs). We denote by $\mathsf{E}[\cdot]$ the expectation operator, and we use $\mathsf{P}[\cdot]$ for probabilities.

We use the letter $\imath$ to denote the imaginary unit, i.e., $\imath = \sqrt{-1}$. The superscripts $(\cdot)^\mathrm{T}$, $(\cdot)^*$ and $(\cdot)^\mathsf{H}$ denote transposition, complex conjugation and Hermitian transposition, respectively. The complement of a set $\mathscr{A}$ is denoted as $\mathscr{A}^\mathsf{c}$. We use "$\stackrel{\mathscr{L}}{=}$" to denote equality in distribution.

We further use $\mathbb{R}$ to denote the set of real numbers, $\mathbb{C}$ to denote the set of complex numbers, $\mathbb{Z}$ to denote the set of integers, $\mathbb{Z}^+$ for the set of positive integers, and $\mathbb{Z}_0^+$ for the set of nonnegative integers.

We denote by $\log(\cdot)$ the natural logarithm, by $\cos(\cdot)$ the cosine function, by $\sin(\cdot)$ the sine function, by $\mathbb{I}\{\cdot\}$ the indicator function, by $Q(\cdot)$ the Gaussian Q-function, by $\Gamma(\cdot)$ the Gamma function [26, Sec. 6.1.1], by $\tilde{\gamma}(\cdot,\cdot)$ the regularized lower incomplete gamma function [26, Sec. 6.5], by $\psi(\cdot)$ the digamma function [26, Sec. 6.3.2], by ${}_2F_1(\cdot,\cdot;\cdot;\cdot)$ the Gauss hypergeometric function [27, Sec. 9.1], by $\mathrm{E}_1(\cdot)$ the exponential integral function [26, Sec. 5.1.1] and by $\zeta(z,q)$ Riemann's zeta function [27, Sec. 9.511]. The gamma distribution with parameters $z$ and $q$ is denoted by $\text{Gamma}(z,q)$. We use $(x)^+$ to denote $\max\{0,x\}$, and $\lceil\cdot\rceil$ to denote the ceiling function. We denote by $\gamma$ Euler's constant.

We use the notation $o_\xi(1)$ to describe terms that vanish as $\xi \to \infty$ and are uniform in the rest of parameters involved. For example, we say that a function

$f(L,\rho)$ is $o_\rho(1)$ if it satisfies

$$\lim_{\rho\to\infty} \sup_{L\geq L_0} |f(L,\rho)| = 0 \tag{1.2}$$

for some $L_0 > 0$ independent of $\rho$. Similarly, we use the notation $\mathcal{O}_\xi(f(\xi))$ to describe terms that are of order $f(\xi)$ and are uniform in the rest of parameters. For example, we say that a function $g(L,\rho)$ is $\mathcal{O}_L\left(\frac{\log L}{L}\right)$ if it satisfies

$$\sup_{\rho\geq\rho_0} |g(L,\rho)| \leq \mathsf{K}\frac{\log L}{L}, \quad L \geq L_0 \tag{1.3}$$

for some $\mathsf{K}$, $L_0$, and $\rho_0$ independent of $L$ and $\rho$.

Double limits such as

$$\lim_{\substack{L\to\infty,\\ \rho\to\infty}} f(L,\rho) = \mathsf{K} \tag{1.4}$$

indicate that for every $\epsilon > 0$ there exists a pair $(L_0,\rho_0)$ independent of $(L,\rho)$ such that for every $L \geq L_0$ and $\rho \geq \rho_0$ we have $|f(L,\rho) - \mathsf{K}| \leq \epsilon$. We denote by $\varliminf$ the *limit inferior* and by $\varlimsup$ the *limit superior*. *Double limit inferiors* and *double limit superiors* are defined accordingly using the above definition of a double limit. For example,

$$\varliminf_{\substack{L\to\infty,\\ \rho\to\infty}} f(L,\rho) = \lim_{\substack{L_0\to\infty,\\ \rho_0\to\infty}} \inf_{L\geq L_0} \inf_{\rho\geq\rho_0} f(L,\rho). \tag{1.5}$$

# 2

# System Model

Consider the communication system depicted in Fig. 2.1, where a transmitter wishes to send a message $A$ to a receiver by encoding it in a length-$n$ sequence $\mathbf{X}^n = [X_1, \ldots, X_n]$, where $n$ is called the *blocklength*. This sequence is sent through a *channel*, which can be viewed as a mathematical representation of the noisy communication medium over which the message is transmitted. For the sake of simplicity, we shall assume that the channel is *memoryless* in the sense that the channel output at a given time instant only depends on the channel input at that given time instant, i.e.,

$$\mathrm{P}_{\mathbf{Y}^n|\mathbf{X}^n}(\mathbf{y}^n|\mathbf{x}^n) = \prod_{k=1}^{n} \mathrm{W}(y_k|x_k) \tag{2.1}$$

for some conditional distribution W independent of $k$. The channel outputs the sequence $\mathbf{Y}^n = [Y_1, \ldots, Y_n]$, based on which the receiver produces an estimate of $A$, denoted as $\hat{A}$. A successful communication occurs when $A = \hat{A}$, and an error occurs when $A \neq \hat{A}$.

We next introduce the notion of a channel code. An $(M, n, \epsilon)$-code consist of:

1. An encoder $f$: $\{1, \ldots, M\} \to \mathcal{X}^n$ where $\mathcal{X}$ denotes the set of possible channel inputs. Hence, the encoder maps the message $A$, which is uniformly distributed on $\{1, \ldots, M\}$, to a codeword $\mathbf{X}^n = [X_1, \ldots, X_n]$.

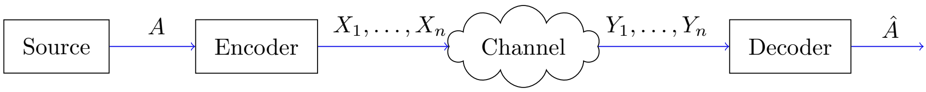


Figure 2.1: Schema of a communication system.

2. A decoder $g\colon \mathcal{Y}^n \to \{1, \ldots, M\}$ that maps the received channel output $\mathbf{Y}^n = [Y_1, \ldots, Y_n]$ to the estimated message $g(\mathbf{Y}^n) = \hat{A} \in \{1, \ldots, M\}$. Here, $\mathcal{Y}$ denotes the set of possible channel outputs. The decoder must satisfy one of the following error probability constraints:

   (a) The maximum error probability constraint
$$\max_{1 \le a \le M} \mathrm{P}\left[\hat{A} \neq A \middle| A = a\right] \le \epsilon. \tag{2.2a}$$

   (b) The average error probability constraint
$$\mathsf{P}\left[\hat{A} \neq A\right] \le \epsilon. \tag{2.2b}$$

The *maximum coding rate* and *minimum error probability* are respectively defined as

$$R^*(n, \epsilon) \triangleq \sup\left\{ \frac{\log(M)}{n} \,:\, \exists (M, n, \epsilon)\text{-code} \right\} \tag{2.3a}$$
$$\epsilon^*(n, R) \triangleq \inf\left\{ \epsilon \,:\, \exists (2^{nR}, n, \epsilon)\text{-code} \right\}. \tag{2.3b}$$

In words, $R^*(n, \epsilon)$ describes the largest data rate at which a message can be transmitted over a channel with a channel code of blocklength $n$ achieving an error probability not larger than $\epsilon$. Likewise, $\epsilon^*(n, R)$ describes the smallest probability of error with which a message can be transmitted over a channel with a channel code of blocklength $n$ achieving a rate not smaller than $R$.

It is common to impose a power constraint on the channel inputs. When $\mathcal{X}$ and $\mathcal{Y}$ are the set of real or complex numbers, perhaps the most common power constraints are:

1. The average-power constraint:
$$\mathsf{E}\left[\|\mathbf{X}^n\|^2\right] \le n\rho. \tag{2.4a}$$

2. The peak-power constraint:
$$|X_k|^2 \le \rho, \quad k = 1, \ldots, n. \tag{2.4b}$$

The peak-power constraint can be incorporated in the set of possible channel inputs by defining

$$\mathcal{X} = \{x \in \mathbb{C} : |x| \leq \rho\}\,. \tag{2.5}$$

On the contrary, the average-power constraint limits the entire codeword $\mathbf{X}^n$ and cannot be described by a Cartesian product $\mathcal{X}^n$.

In this thesis, we consider a single-antenna Rayleigh block-fading channel with coherence interval $\mathsf{T}$ (see Chapter 4). In a fading channel, there is both additive and multiplicative noise. In a block-fading channel, the multiplicative noise remains constant during the coherence interval $\mathsf{T}$ and then changes independently to a new value. Such a channel can be modelled as a block-memoryless channel. More precisely, we can set $\mathcal{X} = \mathcal{Y} = \mathbb{C}^{\mathsf{T}}$ and treat the codeword of length $n$ as a length-$L$ codeword of $\mathsf{T}$-dimensional symbols, i.e., $\mathbf{X}^n = \mathbf{X}^L = [\mathbf{X}_1, \ldots, \mathbf{X}_L]$. For simplicity, we shall restrict ourselves to codes whose blocklength $n$ satisfies $n = L\mathsf{T}$, where $L$ denotes the number of coherence intervals of length $\mathsf{T}$ needed to transmit the entire codeword. We shall consider the following power constraints:

1. The per-coherence-interval maximum power constraint:

$$\|\mathbf{X}_\ell\|^2 \leq \mathsf{T}\rho, \quad \ell = 1, \ldots, L. \tag{2.6a}$$

2. The per-coherence-interval equal power constraint:

$$\|\mathbf{X}_\ell\|^2 = \mathsf{T}\rho, \quad \ell = 1, \ldots, L. \tag{2.6b}$$

As already mentioned above, the power constraint can be incorporated in the set of possible channel inputs. For the power constraint (2.6a), this gives

$$\mathcal{X} = \{\mathbf{x} \in \mathbb{C}^{\mathsf{T}} \colon \|\mathbf{x}\|^2 \leq \mathsf{T}\rho\}. \tag{2.7a}$$

For the power constraint (2.6b), this gives

$$\mathcal{X} = \{\mathbf{x} \in \mathbb{C}^{\mathsf{T}} \colon \|\mathbf{x}\|^2 = \mathsf{T}\rho\}. \tag{2.7b}$$

We shall denote by $(M, L, \epsilon, \rho)$ an $(M, L, \epsilon)$-code that satisfies one of the power constraints (2.6a) or (2.6b). The *maximum coding rate* and *minimum error probability* for the Rayleigh block-fading channel are respectively defined as

$$R^*(L, \epsilon, \rho) \triangleq \sup\left\{\frac{\log(M)}{L\mathsf{T}} \ : \ \exists(M, L, \epsilon, \rho)\text{-code}\right\} \tag{2.8a}$$

$$\epsilon^*(L, R, \rho) \triangleq \inf\left\{\epsilon \, : \, \exists(2^{L\mathsf{T}R}, L, \epsilon, \rho)\text{-code}\right\}. \tag{2.8b}$$

Note that, upper (lower) bounds on $\epsilon^*(L, R, \rho)$ can be translated into lower (upper) bounds on $R^*(L, \epsilon, \rho)$ and *vice versa.* Thus, throughout this thesis we shall present our results in the form that is more convenient for the application at hand.

# 3

# Nonasymptotic Bounds

This chapter presents nonasymptotic bounds on the maximum coding rate as a function of the blocklength and probability of error. These bounds will be the starting point to derive asymptotic approximations of the maximum coding rate (or minimum probability of error) presented in Chapters 5 and 6. As mentioned in Chapter 1.1, traditional asymptotic information theoretical analyses, based on capacity or outage capacity, only capture the behaviour of the maximum coding rate in the limit as the blocklength tends to infinity. The nonasymptotic bounds presented in this chapter allow for more refined asymptotic approximations (see, for example, (1.1)). To facilitate their use in Chapters 5 and 6, we shall particularize the nonasymptotic bounds to the block-fading channel to be introduced in Chapter 4.

Throughout this chapter, we assume that $\mathbf{X}^L \in \mathcal{X}^L$, where $\mathcal{X}$ is given by (2.7a) or (2.7b) depending on the imposed power constraint. We further assume that $\mathrm{P}_{\mathbf{Y}^L|\mathbf{X}^L}$ is absolutely continuous with respect to the Lebesgue measure, so the pdf $\mathrm{p}_{\mathbf{Y}^L|\mathbf{X}^L}$ exists. This also implies that the output pdf $\mathrm{p}_{\mathbf{Y}^L}$ induced by $\mathrm{p}_{\mathbf{Y}^L|\mathbf{X}^L}$ and $\mathrm{P}_{\mathbf{X}^L}$ exists.

## 3.1 Achievability Bounds

This section reviews the achievability bounds that will be used in Chapters 5 and 6. Note that an achievability bound for average probability of error provides automatically a bound for maximum probability of error. Indeed, for maximum probability of error, every codeword in the codebook must satisfy the error constraint, while for average probability of error, the error constraint must be satisfied only in average over all codewords in the codebook. Nonetheless, we introduce specific bounds for maximum probability of error for those cases where tighter bounds can be obtained when the maximum probability of error is considered.

### 3.1.1 RCU Bound [1, Th. 16]

#### 3.1.1.1 Average Probability of Error

Fix an input distribution $\mathrm{P}_{\mathbf{X}^L}$. Assume that the transmitted codeword $\mathbf{X}^L$ is distributed according to $\mathrm{P}_{\mathbf{X}^L}$, and let $\bar{\mathbf{X}}^L$ be independent of $\mathbf{X}^L$ but also distributed according to $\mathrm{P}_{\mathbf{X}^L}$. Then, there exists a code with $M$ codewords, blocklength $L\mathsf{T}$, and average probability of error $\epsilon$ not exceeding

$$\epsilon \leq \mathsf{E}\big[\min\big\{1, (M-1)\mathsf{P}\big[i(\bar{\mathbf{X}}^L;\mathbf{Y}^L) \geq i(\mathbf{X}^L;\mathbf{Y}^L)|\mathbf{X}^L,\mathbf{Y}^L\big]\big\}\big] \tag{3.1}$$

where $i(\mathbf{X}^L;\mathbf{Y}^L)$ is defined as

$$i(\mathbf{X}^L;\mathbf{Y}^L) \triangleq \log\left(\frac{\mathrm{p}_{\mathbf{Y}^L|\mathbf{X}^L}\left(\mathbf{Y}^L \mid \mathbf{X}^L\right)}{\mathrm{p}_{\mathbf{Y}^L}\left(\mathbf{Y}^L\right)}\right) \tag{3.2}$$

and

$$\mathrm{P}_{\mathbf{X}^L,\mathbf{Y}^L,\bar{\mathbf{X}}^L}(\mathbf{x}^L,\mathbf{y}^L,\bar{\mathbf{x}}^L) = \mathrm{P}_{\mathbf{X}^L}(\mathbf{x}^L)\mathrm{p}_{\mathbf{Y}^L|\mathbf{X}^L}(\mathbf{y}^L|\mathbf{x}^L)\mathrm{P}_{\mathbf{X}^L}(\bar{\mathbf{x}}^L). \tag{3.3}$$

The bound given in (3.1) is in general hard to evaluate analytically. In the following sections, we provide two alternative achievability bounds. While these bounds are weaker than the RCU bound, they are in general analytically more tractable. For this reason, these are the bounds we shall use in Chapters 5 and 6.

### 3.1.2 RCU$_\mathrm{s}$ Bound [2, Th. 1]

#### 3.1.2.1 Average Probability of Error

Fix an input distribution $\mathrm{P}_{\mathbf{X}^L}$. For any $s > 0$, there exists a code with $M$ codewords, blocklength $L\mathsf{T}$, and average probability of error $\epsilon$ not exceeding

$$\epsilon \leq \mathsf{E}\big[\min\big\{1, \log M - i_s(\mathbf{X}^L;\mathbf{Y}^L)\big\}\big] \tag{3.4}$$

where $i_s(\mathbf{X}^L;\mathbf{Y}^L)$ is defined as

$$i_s(\mathbf{X}^L;\mathbf{Y}^L) \triangleq \log \frac{\mathrm{p}_{\mathbf{Y}^L|\mathbf{X}^L}(\mathbf{Y}^L|\mathbf{X}^L)^s}{\mathsf{E}\Big[\mathrm{p}_{\mathbf{Y}^L|\mathbf{X}^L}(\mathbf{Y}^L|\mathbf{X}^L)^s|\mathbf{Y}^L\Big]}. \tag{3.5}$$

Using that for any random variable $A$, $\mathsf{E}[\min\{1, A\}] = \mathsf{P}[A \geq U]$, where $U$ is uniformly distributed on the interval $[0, 1]$, (3.4) can be alternatively written as

$$\epsilon \leq \mathsf{P}\big[i_s(\mathbf{X}^L; \mathbf{Y}^L) \leq \log M - \log(U)\big] \tag{3.6}$$

which is a more tractable form to obtain closed form solutions or asymptotic approximations.

### 3.1.3 DT Bound

#### 3.1.3.1 Average Probability of Error [1, Th. 17]

Fix an input distribution $\mathrm{P}_{\mathbf{X}^L}$. Then, there exists a code with $M$ codewords, blocklength $L\mathsf{T}$, and average probability of error $\epsilon$ not exceeding

$$\epsilon \leq \mathsf{E}\left[\exp\left\{-\left[i(\mathbf{X}^L; \mathbf{Y}^L) - \log\frac{M-1}{2}\right]^+\right\}\right] \tag{3.7}$$

where $i(\mathbf{X}^L; \mathbf{Y}^L)$ is given in (3.2) particularized for $s = 1$. After a standard change of measure, (3.7) can be written as

$$\begin{aligned}\epsilon \leq \mathsf{P}&\left[i(\mathbf{X}^L; \mathbf{Y}^L) \leq \log\frac{M-1}{2}\right] \\ &+ (M-1)\mathsf{E}\left[e^{-i(\mathbf{X}^L; \mathbf{Y}^L)}\mathbb{I}\left\{i(\mathbf{X}^L; \mathbf{Y}^L) > \log\frac{M-1}{2}\right\}\right]\end{aligned} \tag{3.8}$$

which is more tractable analytically.

#### 3.1.3.2 Maximum Probability of Error [1, Th. 22]

Fix an input distribution $\mathrm{P}_{\mathbf{X}^L}$. Assume that the cumulative distribution function (CDF) $\mathsf{P}\big[i(\mathbf{x}^L; \mathbf{Y}^L) \leq \gamma\big]$ does not depend on $\mathbf{x}^L$. (Here, $\mathbf{Y}^L$ is distributed according to the output pdf $\mathrm{p}_{\mathbf{Y}^L}$ induced by the input distribution $\mathrm{P}_{\mathbf{X}^L}$ and the channel $\mathrm{p}_{\mathbf{Y}^L|\mathbf{X}^L}$.) Then, there exists a code with $M$ codewords, blocklength $L\mathsf{T}$, and maximum probability of error $\epsilon$ not exceeding

$$\epsilon \leq \mathsf{E}\left[\exp\left\{-\big[i(\mathbf{X}^L; \mathbf{Y}^L) - \log(M-1)\big]^+\right\}\right] \tag{3.9}$$

where $i(\mathbf{X}^L; \mathbf{Y}^L)$ is given in (3.2) particularized for $s = 1$. Again, after a standard change of measure, (3.9) can be written as

$$\begin{aligned}\epsilon \leq \mathsf{P}&\big[i(\mathbf{X}^L; \mathbf{Y}^L) \leq \log(M-1)\big] \\ &+ (M-1)\mathsf{E}\left[e^{-i(\mathbf{X}^L; \mathbf{Y}^L)}\mathbb{I}\{i(\mathbf{X}^L; \mathbf{Y}^L) > \log(M-1)\}\right]\end{aligned} \tag{3.10}$$

which again is a more tractable form to obtain closed form solutions or asymptotic approximations.

## 3.2 Converse Bounds

This section reviews the converse bounds that are used later in Chapters 5 and 6. As in Section 3.1, we distinguish the cases of average probability of error and maximum probability of error.

### 3.2.1 MC Bound

#### 3.2.1.1 Average Probability of Error [1, Th. 27]

Let $\mathrm{P}_{\mathbf{X}^L}$ be some input distribution. Further let $\mathrm{Q}_{\mathbf{Y}^L}$ be any output distribution (not necessarily the one induced by the input distribution and the channel). Then, every code with $M$ codewords, average probability of error $\epsilon$, and blocklength $L\mathsf{T}$, satisfies

$$M \leq \sup_{\mathrm{P}_{\mathbf{X}^L}} \inf_{\mathrm{Q}_{\mathbf{Y}^L}} \log\left(\frac{1}{\beta(\mathrm{P}_{\mathbf{X}^L,\mathbf{Y}^L}, \mathrm{P}_{\mathbf{X}^L}\mathrm{Q}_{\mathbf{Y}^L})}\right) \tag{3.11}$$

where $\beta(\mathrm{P}_{\mathbf{X}^L,\mathbf{Y}^L}, \mathrm{P}_{\mathbf{X}^L}\mathrm{Q}_{\mathbf{Y}^L})$ denotes the minimum probability of error under hypothesis $\mathrm{P}_{\mathbf{X}^L}\mathrm{Q}_{\mathbf{Y}^L}$ if the probability of error under hypothesis $\mathrm{P}_{\mathbf{X}^L,\mathbf{Y}^L}$ does not exceed $\epsilon$ [1, Eq. (100)]. The expression (3.11) may be intractable, since it requires the evaluation of the $\beta(\cdot,\cdot)$ function. To sidestep this problem, we can use [1, Eq. (106)] to relax (3.11) as follows:

$$M \leq \sup_{\mathrm{P}_{\mathbf{X}^L}} \inf_{\mathrm{q}_{\mathbf{Y}^L}} \sup_{\xi>0} \Big\{\log\xi - \log\Big(\mathsf{P}\big[j(\mathbf{X}^L;\mathbf{Y}^L) \leq \log\xi\big] - \epsilon\Big)\Big\} \tag{3.12}$$

where $j(\mathbf{X}^L;\mathbf{Y}^L)$ is defined as

$$j(\mathbf{X}^L;\mathbf{Y}^L) \triangleq \log\frac{\mathrm{p}_{\mathbf{Y}^L|\mathbf{X}^L}(\mathbf{Y}^L|\mathbf{X}^L)}{\mathrm{q}_{\mathbf{Y}^L}(\mathbf{Y}^L)}. \tag{3.13}$$

This relaxation of the MC bound coincides with the Verdú-Han bound [28, Th. 4] with the only difference that the true output pdf $\mathrm{p}_{\mathbf{Y}}$ is replaced by an arbitrary output pdf $\mathrm{q}_{\mathbf{Y}}$. This bound (3.12) for an arbitrary output pdf $\mathrm{q}_{\mathbf{Y}}$ coincides also with the Hayashi-Nagaoka lemma for classical quantum channels [29, Lemma 4]. Even though (3.12) is a relaxation of (3.11), throughout this thesis we shall refer to (3.12) simply as the MC bound.

Note that the bound (3.12) still requires the maximization over $\mathrm{P}_{\mathbf{X}^L}$, which makes its evaluation difficult. However, there are special cases, including the Rayleigh block-fading channel to be introduced in Chapter 4 (see also Chapter 6.2.2) with USTM channel inputs, where $\beta(\mathrm{P}_{\mathbf{X}^L,\mathbf{Y}^L}, \mathrm{P}_{\mathbf{X}^L}\mathrm{Q}_{\mathbf{Y}^L})$ does not depend on $\mathrm{P}_{\mathbf{X}^L}$. In those cases, we have that

$$\beta(\mathrm{P}_{\mathbf{X}^L,\mathbf{Y}^L}, \mathrm{P}_{\mathbf{X}^L}\mathrm{Q}_{\mathbf{Y}^L}) = \beta(\mathrm{P}_{\mathbf{Y}^L|\mathbf{X}^L=\mathbf{x}^L}, \mathrm{Q}_{\mathbf{Y}^L}) \tag{3.14}$$

where the right-hand side (RHS) is independent of the choice of $\mathbf{x}^L \in \mathcal{X}^L$. Thus, by fixing an auxiliary output distribution $\mathrm{Q}_{\mathbf{Y}^L}$, we obtain [1, Th. 28]

$$M \leq \log\left(\frac{1}{\beta(\mathrm{P}_{\mathbf{Y}^L|\mathbf{X}^L=\mathbf{x}^L}, \mathrm{Q}_{\mathbf{Y}^L})}\right) \tag{3.15}$$

as well as the relaxed version

$$M \leq \sup_{\xi>0}\left\{\log\xi - \log\left(\mathsf{P}\left[j(\mathbf{x}^L;\mathbf{Y}^L) \leq \log\xi\right] - \epsilon\right)\right\} \tag{3.16}$$

where $j(\mathbf{x}^L;\mathbf{Y}^L)$ is given in (3.13). This is the form of the bound that will be used later in Chapter 6.

#### 3.2.1.2 Maximum Probability of Error [1, Th. 31]

Choose an auxiliary output distribution $\mathrm{Q}_{\mathbf{Y}^L}$ and assume that the transmitted codeword $\mathbf{x}^L$ belongs to the set $\mathcal{X}^L$. Then, every code with $M$ codewords, maximum probability of error $\epsilon$, and blocklength $L\mathsf{T}$, satisfies

$$M \leq \sup_{\mathbf{x}^L\in\mathcal{X}^L} \log\left(\frac{1}{\beta(\mathbf{x}^L, \mathrm{Q}_{\mathbf{Y}^L})}\right) \tag{3.17}$$

where $\beta(\mathbf{x}^L, \mathrm{Q}_{\mathbf{Y}^L})$ denotes the minimum probability of error under hypothesis $\mathrm{Q}_{\mathbf{Y}^L}$ if the probability of error under hypothesis $\mathrm{P}_{\mathbf{X}^L}$ does not exceed $\epsilon$ [1, Eq. (100)]. Note that the maximization in (3.17) is over all possible transmitted codewords $\mathbf{x}^L \in \mathcal{X}^L$, rather than over all possible input distributions $\mathrm{P}_{\mathbf{X}^L}$. Hence, the main difficulty in evaluating (3.17) lies in the evaluation of the $\beta(\cdot,\cdot)$ function. As in the previous section, we can use [1, Eq. (106)] to obtain the following relaxation of (3.11), which avoids the evaluation of the $\beta(\cdot,\cdot)$ function:

$$M \leq \sup_{\mathbf{x}^L\in\mathcal{X}^L} \sup_{\xi>0}\left\{\log\xi - \log\left(\mathsf{P}\left[j(\mathbf{x}^L;\mathbf{Y}^L) \leq \log\xi\right] - \epsilon\right)\right\} \tag{3.18}$$

where $j(\mathbf{x}^L;\mathbf{Y}^L)$ is given in (3.13). This bound will be used in Chapter 5.

# 4

# The Rayleigh Block-Fading Channel

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

$$\mathbf{Y}_\ell = H_\ell \mathbf{X}_\ell + \mathbf{W}_\ell \tag{4.1}$$

where $\mathbf{X}_\ell$ and $\mathbf{Y}_\ell$ are $\mathsf{T}$-dimensional, complex-valued, random vectors containing the input and output signals, respectively; $\mathbf{W}_\ell$ is the additive noise with independent and identically distributed (i.i.d.), zero-mean, unit-variance, circularly-symmetric, complex Gaussian entries; and $H_\ell$ is a zero-mean, unit-variance, circularly-symmetric, complex Gaussian random variable. We assume that $H_\ell$ and $\mathbf{W}_\ell$ are mutually independent and take on independent realizations over successive coherence intervals. Moreover, the joint law of $(H_\ell, \mathbf{W}_\ell)$ does not depend on the channel inputs. We consider a noncoherent setting where the transmitter and the receiver are aware of the distribution of $H_\ell$ but not of its realization. As aforementioned in Chapter 2, we denote the input codeword as $\mathbf{X}^L = [\mathbf{X}_1, \ldots, \mathbf{X}_L]$ and the channel output induced by the transmitted codeword as $\mathbf{Y}^L = [\mathbf{Y}_1, \ldots, \mathbf{Y}_L]$.

According to (4.1), conditioned on $\mathbf{X}^L = \mathbf{x}^L$, the output vector $\mathbf{Y}^L$ is blockwise i.i.d. Gaussian. Thus, the conditional pdf of $\mathbf{Y}_\ell$ given $\mathbf{X}_\ell = \mathbf{x}$ is independent of $\ell$

and satisfies

$$\mathrm{p}_{\mathbf{Y}|\mathbf{X}}(\mathbf{y}|\mathbf{x}) = \frac{1}{\pi^{\mathsf{T}}(1+\|\mathbf{x}\|^2)} \exp\left\{-\|\mathbf{y}\|^2 + \frac{|\mathbf{y}^{\mathsf{H}}\mathbf{x}|^2}{1+\|\mathbf{x}\|^2}\right\}, \quad \mathbf{y} \in \mathbb{C}^{\mathsf{T}}. \tag{4.2}$$

Here and throughout the thesis, we omit the subscript $\ell$ when immaterial. We shall evaluate the achievability bounds for inputs of the form $\mathbf{X}^L = \sqrt{\mathsf{T}\rho}\mathbf{U}^L$, where the components of $\mathbf{U}^L = [\mathbf{U}_1, \dots, \mathbf{U}_L]$ are i.i.d. and uniformly distributed on the unit sphere in $\mathbb{C}^{\mathsf{T}}$. This distribution of $\mathbf{X}^L$ can be viewed as a single-antenna particularization of USTM [16]. In the following, we shall write $\bar{\mathrm{P}}_{\mathbf{X}}$ to denote the distribution of $\mathbf{X}_\ell = \sqrt{\mathsf{T}\rho}\mathbf{U}_\ell$. Note that, since the variance of $H_\ell$ and of the entries of $\mathbf{W}_\ell$ are normalized to one, $\rho$ can be interpreted as the average SNR at the receiver. The USTM distribution is relevant because it gives rise to a lower bound on capacity that is asymptotically tight at high SNR [17, 18]. In fact, it can be shown that this lower bound accurately approximates capacity already for intermediate SNR values. For example, [13, Fig. 1] illustrates that the lower bound is indistinguishable from the upper bound on capacity given in [13, Eq. (17)] for $\rho \geq 10\,\mathrm{dB}$.

The outputs $\mathbf{Y}^L$ induced by the USTM input distribution have the pdf

$$\mathrm{q}^{(\mathrm{U})}_{\mathbf{Y}^L}(\mathbf{y}^L) = \prod_{\ell=1}^{L} \mathrm{q}^{(\mathrm{U})}_{\mathbf{Y}}(\mathbf{y}_\ell) \tag{4.3}$$

where [13, Eq. (18)]

$$\mathrm{q}^{(\mathrm{U})}_{\mathbf{Y}}(\mathbf{y}) = \frac{e^{\frac{-\|\mathbf{y}\|^2}{1+\mathsf{T}\rho}} \|\mathbf{y}\|^{2(1-\mathsf{T})}\Gamma(\mathsf{T})}{\pi^{\mathsf{T}}(1+\mathsf{T}\rho)} \tilde{\gamma}\left(\mathsf{T}-1, \frac{\mathsf{T}\rho\|\mathbf{y}\|^2}{1+\mathsf{T}\rho}\right)\left(\frac{1+\mathsf{T}\rho}{\mathsf{T}\rho}\right)^{\mathsf{T}-1}, \mathbf{y} \in \mathbb{C}^{\mathsf{T}}. \tag{4.4}$$

Note that (4.4) contains the regularized lower incomplete gamma function which is difficult to analyze. The following lemma presents bounds on the logarithm of this function, which we shall use throughout this thesis.

**Lemma 4.1** *The logarithm of the regularized lower incomplete gamma function can be bounded as*

$$0 \leq \log \frac{1}{\tilde{\gamma}(\mathsf{T}-1, x)} \leq (\mathsf{T}-1)\log\left(1 + \frac{\Gamma(\mathsf{T})^{\frac{1}{\mathsf{T}-1}}}{x}\right), \quad x > 0. \tag{4.5}$$

*Proof:* See Appendix A.1. ■

In the remainder of this thesis, we shall denote by $\mathbf{Y}^L$ a blockwise i.i.d. Gaussian random vector whose conditional pdf, conditioned on $\mathbf{X}^L = \mathbf{x}^L$, is given by $\prod_{\ell=1}^{L} \mathrm{p}_{\mathbf{Y}|\mathbf{X}}(\mathbf{y}_\ell|\mathbf{x}_\ell)$ with $\mathrm{p}_{\mathbf{Y}|\mathbf{X}}(\mathbf{y}|\mathbf{x})$ as in (4.2). We shall denote by $\tilde{\mathbf{Y}}^L$ a blockwise i.i.d. Gaussian random vector that is independent of $\mathbf{X}^L$ and has pdf $\mathrm{q}^{(\mathrm{U})}_{\mathbf{Y}^L}$.

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

$$\begin{aligned} |\mathbf{Y}_\ell^{\mathsf{H}}\mathbf{X}_\ell|^2 &\overset{\mathscr{L}}{=} |H_\ell^* \mathsf{T}\alpha_\ell + \boldsymbol{W}_\ell^*(1)\sqrt{\mathsf{T}\alpha_\ell}|^2 \\ &\overset{\mathscr{L}}{=} \mathsf{T}\alpha_\ell(1+\mathsf{T}\alpha_\ell)Z_{1,\ell} \qquad (4.6) \\ \|\mathbf{Y}_\ell\|^2 &\overset{\mathscr{L}}{=} \|H_\ell\sqrt{\mathsf{T}\alpha}\mathbf{e}_1 + \boldsymbol{W}_\ell\|^2 \\ &\overset{\mathscr{L}}{=} (1+\mathsf{T}\alpha_\ell)Z_{1,\ell} + Z_{2,\ell} \qquad (4.7) \end{aligned}$$

where $\mathbf{W}_\ell(1)$ denotes the first componet of $\mathbf{W}_\ell$ and $\mathbf{e}_1$ is the unitary vector $[1,0,\ldots,0]^{\mathrm{T}}$ of dimension $\mathsf{T}\times 1$. Furthermore, $\{Z_{1,\ell}, \ell\in\mathbb{Z}\}$ 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$(\mathsf{T}-1,1)$-distributed random variables.

Conditioned on $\|\mathbf{X}_\ell\|^2 = \mathsf{T}\alpha_\ell$, the distributions of $|\tilde{\mathbf{Y}}_\ell^{\mathsf{H}}\mathbf{X}_\ell|^2$ and $\|\tilde{\mathbf{Y}}_\ell\|^2$ can be written as

$$|\tilde{\mathbf{Y}}_\ell^{\mathsf{H}}\mathbf{X}_\ell|^2 \overset{\mathscr{L}}{=} |(H_\ell^*\sqrt{\mathsf{T}\rho}\mathbf{U}_\ell(1) + \boldsymbol{W}_\ell^*(1))\sqrt{\mathsf{T}\alpha_\ell}|^2 \qquad (4.8)$$

$$\|\tilde{\mathbf{Y}}_\ell\|^2 \overset{\mathscr{L}}{=} \|H_\ell\sqrt{\mathsf{T}\rho}\mathbf{U}_\ell + \boldsymbol{W}_\ell\|^2. \qquad (4.9)$$

In (4.6)–(4.9), the parameter $\alpha_\ell$ can be thought of as the power allocated over the coherence interval $\ell$.

In the following sections we introduce some quantities that we shall need in the remainder of the thesis.

## 4.1 Information Densities

The *generalized information density* random variable for $\mathbf{X}^L$ and $\mathbf{Y}^L$ is defined for any $s>0$ as

$$i_s(\mathbf{X}^L;\mathbf{Y}^L) \triangleq \log\frac{\mathrm{p}_{\mathbf{Y}^L|\mathbf{X}^L}(\mathbf{Y}^L|\mathbf{X}^L)^s}{\mathrm{p}^s_{\mathbf{Y}^L}(\mathbf{Y}^L)}, \qquad s>0 \qquad (4.10a)$$

$$\mathrm{p}^s_{\mathbf{Y}^L}(\mathbf{y}^L) \triangleq \int \mathrm{p}_{\mathbf{Y}^L|\mathbf{X}^L}(\mathbf{Y}^L|\mathbf{X}^L)^s \mathrm{dP}_{\mathbf{X}}(\mathbf{x}^L), \quad s>0. \qquad (4.10b)$$

When the input distribution is USTM, the generalized information density $i_s(\mathbf{X}^L;\mathbf{Y}^L)$ can be expressed as

$$i_s(\mathbf{X}^L;\mathbf{Y}^L) = \sum_{\ell=1}^{L} i_{\ell,s}(\rho) \qquad (4.11)$$

where

$$\begin{aligned} i_{\ell,s}(\rho) \stackrel{\mathscr{L}}{=} &(\mathsf{T}-1)\log(s\mathsf{T}\rho) - \log\Gamma(\mathsf{T}) - s\frac{\mathsf{T}\rho Z_{2,\ell}}{1+\mathsf{T}\rho} \\ &+ (\mathsf{T}-1)\log\left(\frac{(1+\mathsf{T}\rho)Z_{1,\ell}+Z_{2,\ell}}{1+\mathsf{T}\rho}\right) \\ &- \log\tilde{\gamma}\left(\mathsf{T}-1, s\frac{\mathsf{T}\rho((1+\mathsf{T}\rho)Z_{1,\ell}+Z_{2,\ell})}{1+\mathsf{T}\rho}\right). \end{aligned} \tag{4.12}$$

For $s=1$, $i_s(\mathbf{X}^L;\mathbf{Y}^L)$ can be written as

$$i_1(\mathbf{X}^L;\mathbf{Y}^L) = i(\mathbf{X}^L;\mathbf{Y}^L) = \log\left(\frac{\mathrm{p}_{\mathbf{Y}^L|\mathbf{X}^L}\left(\mathbf{Y}^L \mid \mathbf{X}^L\right)}{\mathrm{p}_{\mathbf{Y}^L}\left(\mathbf{Y}^L\right)}\right) \tag{4.13}$$

where $\mathrm{p}_{\mathbf{Y}^L}\left(\mathbf{Y}^L\right)$ is the output pdf induced by the input distribution.[1] When the input distribution is USTM, $i(\mathbf{X}^L;\mathbf{Y}^L)$ can be expressed as

$$i(\mathbf{X}^L;\mathbf{Y}^L) = \sum_{\ell=1}^{L} i_\ell(\rho) \tag{4.14}$$

where $i_\ell(\rho) = i_{\ell,1}(\rho)$. Using the left-most inequality in Lemma 4.1, we can lower-bound $i_\ell(\rho)$ by

$$\underline{i}_\ell(\rho) \triangleq (\mathsf{T}-1)\log(\mathsf{T}\rho) - \log\Gamma(\mathsf{T}) - \frac{\mathsf{T}\rho Z_{2,\ell}}{1+\mathsf{T}\rho} + (\mathsf{T}-1)\log\left(\frac{(1+\mathsf{T}\rho)Z_{1,\ell}+Z_{2,\ell}}{1+\mathsf{T}\rho}\right). \tag{4.15}$$

## 4.2 Mismatched Information Densities

Next, we consider the *mismatched information density*,[2] which is defined as

$$j(\mathbf{X}^L;\mathbf{Y}^L) \triangleq \log\frac{\mathrm{P}_{\mathbf{Y}^L|\mathbf{X}^L}(\mathbf{Y}^L|\mathbf{X}^L)}{\mathrm{q}_{\mathbf{Y}^L}(\mathbf{Y}^L)} \tag{4.16}$$

where $\mathrm{q}_{\mathbf{Y}^L}$ is an arbitrary auxiliary output pdf. When $\mathrm{q}_{\mathbf{Y}^L}$ is the pdf induced by USTM channel inputs, i.e., $\mathrm{q}^{(\mathrm{U})}_{\mathbf{Y}^L}$ given in (4.3), $j(\mathbf{X}^L;\mathbf{Y}^L)$ can be expressed as

$$j(\mathbf{X}^L;\mathbf{Y}^L) = \sum_{\ell=1}^{L} j(\mathbf{X}_\ell;\mathbf{Y}_\ell) \tag{4.17}$$

[1] The existence of the conditional pdf $\mathrm{p}_{\mathbf{Y}^L|\mathbf{X}^L}$ implies that the output pdf $\mathrm{p}_{\mathbf{Y}^L}$ exists for every input distribution.

[2] We use the word "mismatched" to indicate that the output pdf $\mathrm{q}_{\mathbf{Y}^L}(\mathbf{Y}^L)$ in the denominator is not the one induced by the input distribution and the channel.

where

$$j(\mathbf{X}_\ell;\mathbf{Y}_\ell) = \log\left(\frac{1+\mathsf{T}\rho}{\Gamma(\mathsf{T})}\right) + \frac{|\mathbf{Y}_\ell^{\mathsf{H}}\mathbf{X}_\ell|^2}{1+\|\mathbf{X}_\ell\|^2} - \frac{\mathsf{T}\rho\|\mathbf{Y}_\ell\|^2}{1+\mathsf{T}\rho} + (\mathsf{T}-1)\log\left(\frac{\mathsf{T}\rho\|\mathbf{Y}_\ell\|^2}{1+\mathsf{T}\rho}\right)$$
$$- \log(1+\|\mathbf{X}_\ell\|^2) - \log\tilde{\gamma}\left(\mathsf{T}-1, \frac{\mathsf{T}\rho\|\mathbf{Y}_\ell\|^2}{1+\mathsf{T}\rho}\right). \tag{4.18}$$

By (4.6) and (4.7), $j(\mathbf{X}_\ell;\mathbf{Y}_\ell)$ depends on $\mathbf{X}_\ell$ only via $\|\mathbf{X}_\ell\|^2 = \mathsf{T}\alpha_\ell$. We can thus express $j(\mathbf{X}_\ell;\mathbf{Y}_\ell)$ conditioned on $\|\mathbf{X}_\ell\|^2 = \mathsf{T}\alpha_\ell$ as

$$j_\ell(\alpha_\ell) \triangleq (\mathsf{T}-1)\log(\mathsf{T}\rho) - \log\Gamma(\mathsf{T}) - \frac{(\mathsf{T}\rho-\mathsf{T}\alpha_\ell)Z_{1,\ell}}{1+\mathsf{T}\rho} - \frac{\mathsf{T}\rho Z_{2,\ell}}{1+\mathsf{T}\rho} + \log\left(\frac{1+\mathsf{T}\rho}{1+\mathsf{T}\alpha_\ell}\right)$$
$$+ (\mathsf{T}-1)\log\left(\frac{(1+\mathsf{T}\alpha_\ell)Z_{1,\ell}+Z_{2,\ell}}{1+\mathsf{T}\rho}\right)$$
$$- \log\tilde{\gamma}\left(\mathsf{T}-1, \frac{\mathsf{T}\rho((1+\mathsf{T}\alpha_\ell)Z_{1,\ell}+Z_{2,\ell})}{1+\mathsf{T}\rho}\right). \tag{4.19}$$

Define $\beta(\rho) \triangleq \Gamma(\mathsf{T})^{\frac{1}{\mathsf{T}-1}}\frac{1+\mathsf{T}\rho}{\mathsf{T}\rho}$, and let

$$\bar{j}_\ell(\alpha_\ell) \triangleq (\mathsf{T}-1)\log(\mathsf{T}\rho) - \log\Gamma(\mathsf{T}) - \frac{(\mathsf{T}\rho-\mathsf{T}\alpha_\ell)Z_{1,\ell}}{1+\mathsf{T}\rho} - \frac{\mathsf{T}\rho Z_{2,\ell}}{1+\mathsf{T}\rho} + \log\left(\frac{1+\mathsf{T}\rho}{1+\mathsf{T}\alpha_\ell}\right)$$
$$+ (\mathsf{T}-1)\log\left(\frac{(1+\mathsf{T}\alpha_\ell)Z_{1,\ell}+Z_{2,\ell}}{1+\mathsf{T}\rho}\right)$$
$$+ (\mathsf{T}-1)\log\left(1+\frac{\beta(\rho)}{(1+\mathsf{T}\alpha_\ell)Z_{1,\ell}+Z_{2,\ell}}\right). \tag{4.20}$$

By Lemma 4.1, we have that, with probability one,

$$j_\ell(\alpha_\ell) \le \bar{j}_\ell(\alpha_\ell), \quad \alpha_\ell \in [0,\rho]. \tag{4.21}$$

We next consider an auxiliary output pdf that will be useful for the derivation of the saddlepoint approximations in Chapter 6. Specifically, let

$$\mathrm{q}^s_{\mathbf{Y}^L}(\mathbf{y}_\ell) \triangleq \prod_{\ell=1}^{L}\mathrm{q}^s_{\mathbf{Y}_\ell}(\mathbf{y}_\ell) \tag{4.22}$$

where

$$\mathrm{q}^s_{\mathbf{Y}_\ell}(\mathbf{y}_\ell) \triangleq \frac{1}{\mu(s)}\left(\int \mathrm{p}_{\mathbf{Y}_\ell|\mathbf{X}_\ell}(\mathbf{y}_\ell|\mathbf{x}_\ell)^s \mathrm{d}\bar{\mathrm{P}}_{\mathbf{X}}(\mathbf{x}_\ell)\right)^{1/s} \tag{4.23}$$

and $\mu(s)$ is a normalizing factor. Based on (4.22) and (4.23), we define the *generalized mismatched information density* $j_s(\mathbf{X}^L;\mathbf{Y}^L)$ as

$$j_s(\mathbf{X}^L;\mathbf{Y}^L) \triangleq \log\frac{\mathrm{p}_{\mathbf{Y}^L|\mathbf{X}^L}(\mathbf{Y}^L|\mathbf{X}^L)}{\mathrm{q}^s_{\mathbf{Y}^L}(\mathbf{y}_\ell)}. \tag{4.24}$$

Using this definition together with (4.2) and (4.22), the mismatched information density $j_s(\mathbf{X}^L;\mathbf{Y}^L)$ can be written as

$$j_s(\mathbf{X}^L;\mathbf{Y}^L) = \sum_{\ell=1}^{L} j_s(\mathbf{X}_\ell;\mathbf{Y}_\ell). \tag{4.25}$$

It holds that

$$j_s(\mathbf{X}_\ell;\mathbf{Y}_\ell) = \log\mu(s) + \frac{1}{s} i_s(\mathbf{X}_\ell;\mathbf{Y}_\ell). \tag{4.26}$$

Note that for USTM inputs, we have $j_\ell(\rho) = i_\ell(\rho)$.

## 4.3 Information Rates and Dispersions

We define the expectation and variance of $i_{\ell,s}(\rho)$ by

$$I_s(\rho) \triangleq \mathsf{E}\big[i_{\ell,s}(\rho)\big] \tag{4.27}$$

$$V_s(\rho) \triangleq \mathsf{Var}\big[i_{\ell,s}(\rho)\big]. \tag{4.28}$$

Note that $I_s(\rho)$ evaluated at $s=1$ corresponds to the mutual information between $\mathbf{X}_\ell$ and $\mathbf{Y}_\ell$. We further define the expectation of $j_{\ell,s}(\rho)$ and $j_\ell(\alpha_\ell)$ as

$$J_s(\rho) \triangleq \mathsf{E}[j_{\ell,s}(\rho)] \tag{4.29}$$

$$J(\alpha_\ell) \triangleq \mathsf{E}[j_\ell(\alpha_\ell)], \quad 0 \le \alpha_\ell \le \rho. \tag{4.30}$$

Note that $J_1(\rho) = I_1(\rho)$, in which case we omit the subscript and simply write $I(\rho)$. We next compute the expected value of (4.15), denoted by $\underline{I}(\rho) \triangleq \mathsf{E}\big[\underline{i}_\ell(\rho)\big]$, as

$$\begin{aligned}\underline{I}(\rho) &= (\mathsf{T}-1)\log(\mathsf{T}\rho) - \log\Gamma(\mathsf{T}) - \frac{(\mathsf{T}-1)\mathsf{T}\rho}{1+\mathsf{T}\rho} \\ &\quad - (\mathsf{T}-1)\log(1+\mathsf{T}\rho) + (\mathsf{T}-1)\mathsf{E}\big[\log((1+\mathsf{T}\rho)Z_1 + Z_2)\big] &\text{(4.31a)}\\ &= (\mathsf{T}-1)\log(\mathsf{T}\rho) - \log\Gamma(\mathsf{T}) - (\mathsf{T}-1)\left[\log(1+\mathsf{T}\rho) + \frac{\mathsf{T}\rho}{1+\mathsf{T}\rho} - \psi(\mathsf{T}-1)\right] \\ &\quad + {}_2F_1\left(1, \mathsf{T}-1; \mathsf{T}; \frac{\mathsf{T}\rho}{1+\mathsf{T}\rho}\right) &\text{(4.31b)}\end{aligned}$$

where the expected value has been solved using [27, Sec. 4.337-1] to integrate with respect to $Z_1$ and [27, Sec. 4.352-1], [27, Sec. 3.381-4], and [27, Sec. 6.228-2] to integrate with respect to $Z_2$. Clearly,

$$I(\rho) \ge \underline{I}(\rho). \tag{4.32}$$

The conditional expected value of (4.20) given $\|\mathbf{X}_\ell\|^2 = \mathsf{T}\alpha_\ell$, denoted by $\bar{J}(\alpha_\ell) \triangleq \mathsf{E}[\bar{j}_\ell(\alpha_\ell)]$, can be evaluated as

$$\begin{aligned}\bar{J}(\alpha_\ell) = (\mathsf{T}-1)\log(\mathsf{T}\rho) - \log\Gamma(\mathsf{T}) - \frac{\mathsf{T}\rho - \mathsf{T}\alpha_\ell}{1+\mathsf{T}\rho} - \frac{(\mathsf{T}-1)\mathsf{T}\rho}{1+\mathsf{T}\rho}\\ + \log\left(\frac{1+\mathsf{T}\rho}{1+\mathsf{T}\alpha_\ell}\right) - (\mathsf{T}-1)\log(1+\mathsf{T}\rho)\\ + (\mathsf{T}-1)\mathsf{E}\big[\log((1+\mathsf{T}\alpha_\ell)Z_1 + Z_2 + \beta(\rho))\big].\end{aligned} \tag{4.33}$$

It can be shown that $J(\cdot)$ and $I(\cdot)$ bound the capacity [30]

$$C(\rho) = \sup_{\mathrm{P}_{\mathbf{X}^L}:\,\|\mathbf{X}_\ell\|^2\le\mathsf{T}\rho} \frac{\mathsf{E}[i(\mathbf{X}_\ell;\mathbf{Y}_\ell)]}{\mathsf{T}}. \tag{4.34}$$

Indeed, on the one hand we have

$$C(\rho) \le \sup_{0\le\alpha\le\rho} \frac{J(\alpha)}{\mathsf{T}} \le \sup_{0\le\alpha\le\rho} \frac{\bar{J}(\alpha)}{\mathsf{T}} \tag{4.35}$$

where the first inequality follows from [31, Th. 5.1], and the second inequality follows from (4.21). On the other hand,

$$C(\rho) \ge \frac{I(\rho)}{\mathsf{T}} \ge \frac{\underline{I}(\rho)}{\mathsf{T}} \tag{4.36}$$

where the first inequality follows because USTM is a valid input distribution, and the second inequality follows by (4.32). It can be further shown that

$$\lim_{\rho\to\infty}\left\{\sup_{0\le\alpha\le\rho} \bar{J}(\alpha) - \underline{I}(\rho)\right\} = 0. \tag{4.37}$$

Thus, USTM yields an asymptotically tight lower bound on capacity, as already mentioned before.

Let

$$V(\rho) \triangleq \mathsf{E}\Big[\big(i_\ell(\rho) - I(\rho)\big)^2\Big] \tag{4.38a}$$

$$\bar{V}_\rho(\alpha) \triangleq \mathsf{E}\Big[\big(\bar{j}_\ell(\alpha) - \bar{J}(\alpha)\big)^2\Big] \tag{4.38b}$$

where the subscript $\rho$ in $\bar{V}_\rho(\alpha)$ is introduced to highlight that $\bar{V}_\rho(\alpha)$ depends both on $\alpha$ and $\rho$, but it is omitted when $\alpha = \rho$. In Lemma A.2 (Appendix A.9) and Lemma A.3 (Appendix A.10), we show that $I(\rho)$, $V(\rho)$, $\bar{J}(\rho)$, and $\bar{V}_\rho(\rho)$ can be approximated as

$$I(\rho) = \underline{I}(\rho) + o_\rho(1) \tag{4.39a}$$

$$V(\rho) = \tilde{V} + o_\rho(1) \tag{4.39b}$$

$$\bar{J}(\rho) = \underline{I}(\rho) + o_\rho(1) \tag{4.39c}$$

$$\bar{V}(\rho) = \tilde{V} + o_\rho(1). \tag{4.39d}$$

A closed form expression for $\underline{I}(\rho)$ is given in (4.31b). Moreover, $\tilde{V}$ in (4.39b) and (4.39d) is defined as

$$\tilde{V} \triangleq (\mathsf{T}-1)^2\frac{\pi^2}{6} + (\mathsf{T}-1). \tag{4.40}$$

## 4.4 The Moment Generating Function

The moment generating function (MGF) of $I_s(\rho) - i_{\ell,s}(\rho)$ is given by

$$m_{\rho,s}(\tau) = \mathsf{E}\Big[e^{\tau(I_s(\rho)-i_{\ell,s}(\rho))}\Big] \tag{4.41}$$

and its cumulant generating function (CGF) is given by

$$\psi_{\rho,s}(\tau) = \log m_{\rho,s}(\tau). \tag{4.42}$$

The region of convergence (RoC) of $m_{\rho,s}(\tau)$ is defined as

$$\mathcal{S}_m(\rho,s) \triangleq \{\tau\in\mathbb{R}\colon m_{\rho,s}(\tau)<\infty\}. \tag{4.43}$$

Similarly, we shall say that a set $\mathcal{S}$ is in the RoC of the family of MGFs $m_{\rho,s}(\tau)$ (parametrized by $(\rho,s)$) if

$$\sup_{(\tau,\rho,s)\in\mathcal{S}_m} m_{\rho,s}^{(k)}(\tau) < \infty, \quad k\in\mathbb{Z}_0^+. \tag{4.44}$$

The following lemma presents two sets that are in the RoC of the family of MGFs $m_{\rho,s}(\tau)$ (parametrized by $(\rho,s)$).

**Lemma 4.2 (Region of Convergence)**
Part 1)*: For every $\rho_0>0$, $s_0>0$, and $0<a<1/(\mathsf{T}-1)$ independent of $(L,\rho,s,\tau)$, we have that*

$$\sup_{\substack{-a(\mathsf{T}-1)<\tau\le a,\\ s\in[s_0,1],\\ \rho\ge\rho_0}} m_{\rho,s}^{(k)}(\tau) < \infty, \quad k\in\mathbb{Z}_0^+. \tag{4.45}$$

Part 2)*: For every $0<s_0<s_{\max}<\infty$, $0<\rho_0<\rho_{\max}<\infty$, $0<a<1$, and $0<b<\min\Big\{\frac{\mathsf{T}}{\mathsf{T}-1},\frac{1+\mathsf{T}\rho_{\max}}{\mathsf{T}\rho_{\max}s_{\max}}\Big\}$ independent of $(L,\rho,s,\tau)$, we have that*

$$\sup_{\substack{-a<\tau\le b,\\ s\in[s_0,s_{\max}],\\ \rho_0\le\rho\le\rho_{\max}}} m_{\rho,s}^{(k)}(\tau) < \infty, \quad k\in\mathbb{Z}_0^+. \tag{4.46}$$

*Proof:* See Appendix B.5. ■

# 5

# A high-SNR Normal Approximation

In this chapter, we present an expression similar to (1.1) for the maximum coding rate $R^*(L, \epsilon, \rho)$ achievable over the noncoherent, single-antenna, Rayleigh block-fading channel introduced in (4.1) using error-correcting codes that span $L$ coherence intervals, have a block-error probability no larger than $\epsilon$, and satisfy the per-coherence-interval maximum power constraint (2.6a). 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 nonasymptotic bounds provided in [13, 14, 15]. On the other hand, it allows for a mathematical analysis of $R^*(L, \epsilon, \rho)$.

## 5.1 Main Results

The main result of this chapter is a high-SNR normal approximation on $R^*(L, \epsilon, \rho)$ presented in Section 5.1.1. In Section 5.1.2, we assess the accuracy of this approximation by means of numerical examples. Possible applications are discussed in Section 5.1.3.

### 5.1.1 A High-SNR Normal Approximation

**Theorem 5.1** *Assume that* $\mathsf{T} > 2$ *and that* $0 < \epsilon < \frac{1}{2}$. *Then, in the limit as* $L \to \infty$ *and* $\rho \to \infty$, *the maximum coding rate* $R^*(L, \epsilon, \rho)$ *can be approximated as*

$$R^*(L, \epsilon, \rho) = \frac{\underline{I}(\rho)}{\mathsf{T}} + o_\rho(1) - \sqrt{\frac{\tilde{V} + o_\rho(1)}{L\mathsf{T}^2}} Q^{-1}(\epsilon) + \mathcal{O}_L\left(\frac{\log L}{L}\right) \tag{5.1}$$

*where* $\underline{I}(\rho)$ *and* $\tilde{V}$ *are defined in* (4.31b) *and* (4.40), *respectively.*

*Proof:* See Section 5.2. ■

**Remark 5.1** *A common approach to deal with limits in two parameters is to couple them so as to reduce the double limit to a single limit. For example, one could set* $\rho = g(L)$ *for some increasing function* $g(\cdot)$ *and then study the maximum coding rate* $R^*(L, \epsilon, g(L))$ *in the limit as* $L \to \infty$. *While this approach sidesteps the difficulties in dealing with double limits, it gives rise to results that are hard to interpret, especially if the asymptotic behavior of* $R^*(L, \epsilon, g(L))$ *depends critically on* $g(\cdot)$. *Indeed,* $L$ *describes the blocklength of the error-correcting code, and* $\rho$ *specifies the SNR at which messages are sent over the channel. There is no physical reason why these two parameters should be coupled, hence it is unclear which coupling* $g(\cdot)$ *describes the communication system best. In contrast, the approximation presented in Theorem 5.1 is interpretable and more robust, since it holds for any sufficiently large* $L$ *and* $\rho$ *(irrespective of their relation). In fact, since the* $o_\rho(1)$ *terms are uniform in* $L$, *and the* $\mathcal{O}(\log L/L)$ *term is uniform in* $\rho$, *the approximation* (5.1) *applies also for* any *(strictly increasing) coupling between* $L$ *and* $\rho$.

**Remark 5.2** *The assumption that* $0 < \epsilon < 1/2$ *is required to ensure that* $Q^{-1}(\epsilon)$ *is nonnegative, which simplifies the manipulations of the channel dispersion. Treating the case* $1/2 < \epsilon < 1$ *would require a separate analysis. For the sake of compactness, we decided to omit such an analysis, since we believe that* $0 < \epsilon < 1/2$ *covers all cases of practical interest.*

Ignoring the $\mathcal{O}_L(\log L/L)$ and the $o_\rho(1)$ terms in (5.1), we obtain the following high-SNR normal approximation:

$$R^*(L, \epsilon, \rho) \approx \frac{\underline{I}(\rho)}{\mathsf{T}} - \sqrt{\frac{\tilde{V}}{L\mathsf{T}^2}} Q^{-1}(\epsilon). \tag{5.2}$$

The closed form expression for $\underline{I}(\rho)$ in (4.31b) contains a hypergeometric function, which is difficult to analyze mathematically. We therefore present also a simplified expression that is less accurate than (4.31b) but easier to analyze. Specifically, it follows from Lemma A.2 (Appendix A.9) that

$$\underline{I}(\rho) = (\mathsf{T}-1)\log(\mathsf{T}\rho) - \log\Gamma(\mathsf{T}) - (\mathsf{T}-1)(1+\gamma) + o_\rho(1). \tag{5.3}$$

The quantity $\underline{I}(\rho)/\mathsf{T}$ is a high-SNR approximation of the information rate achievable with i.i.d. USTM inputs; cf. [32, Eq. (12)] (see also [13, Eq. (5)]). It is shown in [16, Th. 4] that $\underline{I}(\rho)/\mathsf{T}$ is an asymptotically-tight lower bound on the capacity $C(\rho)$ in the sense that

$$\lim_{\rho\to\infty}\left\{C(\rho) - \frac{\underline{I}(\rho)}{\mathsf{T}}\right\} = 0. \tag{5.4}$$

For comparison, the capacity of the additive white Gaussian noise (AWGN) channel is given by [3, Eq. (7)]

$$C_{\mathrm{AWGN}}(\rho) = \log(1+\rho) = \log\rho + o_\rho(1). \tag{5.5}$$

The capacity of the coherent Rayleigh block-fading channel (when the channel state information is available at the receiver) is given by [33]

$$C_c(\rho) = \mathsf{E}\big[\log(1+\rho Z_1)\big] = \log\rho - \gamma + o_\rho(1) \tag{5.6}$$

and in the noncoherent case (cf. (5.3))

$$\frac{\underline{I}(\rho)}{\mathsf{T}} = \frac{\mathsf{T}-1}{\mathsf{T}}\Big[\log(\rho) - \gamma\Big] + \mathcal{O}_\rho(1). \tag{5.7}$$

It can be shown that the $o_\rho(1)$ and $\mathcal{O}_\rho(1)$ terms in (5.6) and (5.7) are uniform in $T$.

The channel dispersion of the AWGN channel is given by [3, Eq. (8)]

$$V_{\mathrm{AWGN}}(\rho) = \rho\frac{2+\rho}{(1+\rho)^2} = 1 + o_\rho(1). \tag{5.8}$$

For the coherent Rayleigh block-fading channel, the channel dispersion $V_c(\rho)$ is given by [5, Th. 2]

$$V_c(\rho) = \mathrm{Var}\big[\log(1+\rho Z_1)\big] + \frac{1}{\mathsf{T}} - \frac{1}{\mathsf{T}}\mathsf{E}\left[\frac{1}{1+\rho Z_1}\right]^2 = \frac{\pi^2}{6} + \frac{1}{\mathsf{T}} + o_\rho(1). \tag{5.9}$$

According to Theorem 5.1, the ratio $\tilde{V}/\mathsf{T}^2$ can be viewed as a high-SNR approximation of the channel dispersion.

By comparing (5.7) and (5.6), we see that $\underline{I}(\rho)/\mathsf{T}$ is, up to a $\mathcal{O}_\rho(1)$ term, equal to $(1-1/\mathsf{T})C_c(\rho)$. Further observe that $\tilde{V}/\mathsf{T}^2$ corresponds to the dispersion one obtains by transmitting one pilot symbol per coherence block to estimate the fading coefficient, and by then transmitting $\mathsf{T}-1$ symbols per coherence block over a coherent fading channel. This suggests the heuristic that, at high SNR, one pilot symbol per coherence block suffices to achieve both capacity and channel dispersion. However, this heuristic may be misleading since it is *prima facie* unclear whether one pilot symbol per coherence block suffices to obtain a fading estimate that is (almost) perfect. A more refined analysis of the maximum coding rate achievable with pilot assisted transmission has been recently performed by Östman *et al.* [21].

Further observe that, as $\mathsf{T}$ tends to infinity, $\underline{I}(\rho)/\mathsf{T}$ converges to $C_c(\rho)$ and $\tilde{V}/\mathsf{T}^2$ converges to $V_c(\rho)$. Thus, as the coherence interval grows to infinity, both capacity and channel dispersion of the noncoherent block-fading channel converge to the corresponding quantities for the coherent channel. This agrees with the intuition that the cost of estimating the channel vanishes as the coherence interval tends to infinity.

Finally, observe that $C_{\text{AWGN}}(\rho)$ is larger than $\underline{I}(\rho)/\mathsf{T}$ and $C_c(\rho)$, and $V_{\text{AWGN}}(\rho)$ is smaller than $\tilde{V}/\mathsf{T}^2$ and $V_c(\rho)$ (except for $\mathsf{T}=3$, where $\tilde{V}/\mathsf{T}^2 < V_{\text{AWGN}}(\rho)$). Thus, the presence of fading results in a less favorable channel.

### 5.1.2 Numerical Examples

We illustrate the accuracy of the high-SNR normal approximation (5.2) by means of numerical examples. In Figs. 5.1 and 5.2 we show the approximation (5.2) as a function of $n = L\mathsf{T}$ for a fixed coherence interval $\mathsf{T}$ and for different SNR values. In the normal approximation, we evaluate $\underline{I}(\rho)$ using both the exact expression (4.31b) as well as the approximation (5.3). For comparison, we also plot the coherent normal approximation

$$R^*(L,\epsilon,\rho) \approx C_c(\rho) - \sqrt{\frac{V_c(\rho)}{L}}Q^{-1}(\epsilon) \tag{5.10}$$

where $C_c(\rho)$ was defined in (5.6) and $V_c(\rho)$ was defined in (5.9). We further plot a nonasymptotic (in $\rho$ and $L$) lower bound on $R^*(L,\epsilon,\rho)$ that is based on the dependence testing (DT) lower bound (3.10) with USTM channel inputs (see (5.21) below) and computed by Monte Carlo simulations. We further plot a nonasymptotic (in $\rho$ and $L$) upper bound on $R^*(L,\epsilon,\rho)$ that is based on the MC upper bound (3.17) with auxiliary output pdf (4.3) (see (5.35) below). Specifically, we plot the weakened version

$$R^*(L,\epsilon,\rho) \leq \inf_{\xi>0}\left\{\frac{\log\xi}{L\mathsf{T}} - \inf_{\boldsymbol{\alpha}\in[o,\rho]^L}\frac{\log\Bigl(1-\epsilon-\mathsf{P}\bigl[\sum_{\ell=1}^{L} j_\ell(\alpha_\ell)\geq\log\xi\bigr]\Bigr)}{L\mathsf{T}}\right\} \tag{5.11}$$

which is obtained using (3.18) and was evaluated by Monte Carlo simulations. In (5.11), $\boldsymbol{\alpha} = (\alpha_1,\dots,\alpha_L)$ denotes the vector of power allocations. We finally plot $\underline{I}(\rho)/\mathsf{T}$ as given by (4.31b). Observe that the high-SNR normal approximation of $R^*(L,\epsilon,\rho)$ is fairly accurate already for $\rho = 15\,\text{dB}$ and $n \geq 200$ when we use the exact

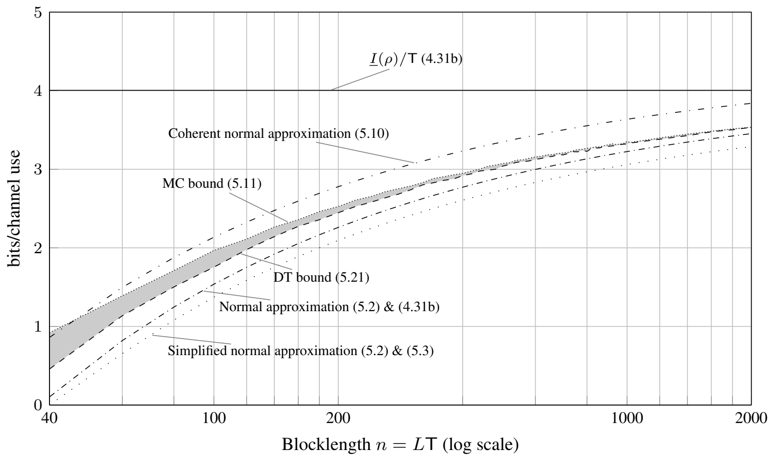


Figure 5.1: Bounds on $R^*(L, \epsilon, \rho)$ for $\rho = 15\,\mathrm{dB}$, $\mathsf{T} = 20$, $\epsilon = 10^{-3}$. The shaded area indicates the area in which $R^*(L, \epsilon, \rho)$ lies.

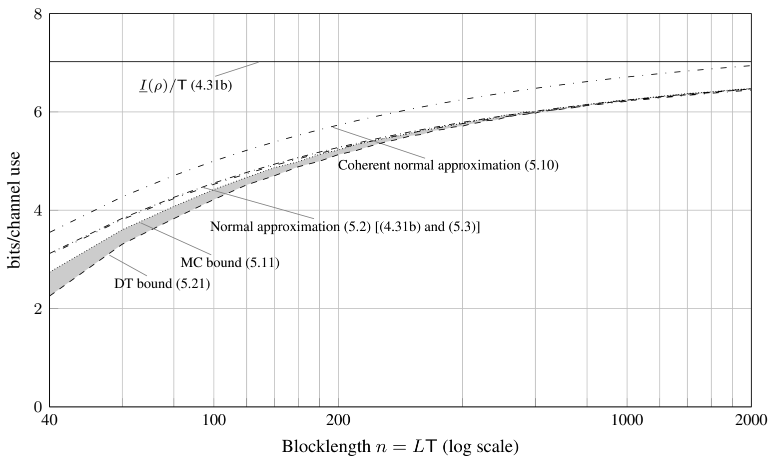


Figure 5.2: Bounds on $R^*(L, \epsilon, \rho)$ for $\rho = 25\,\mathrm{dB}$, $\mathsf{T} = 20$, $\epsilon = 10^{-3}$. The shaded area indicates the area in which $R^*(L, \epsilon, \rho)$ lies.

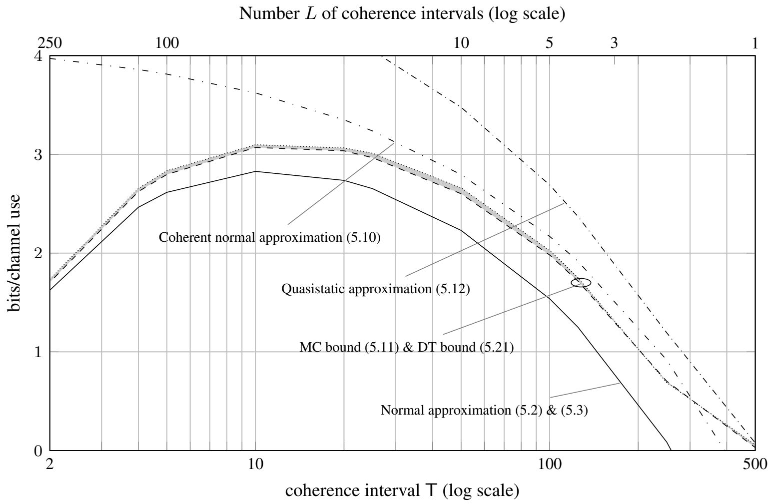


Figure 5.3: Bounds on $R^*(L, \epsilon, \rho)$ for $n = L\mathsf{T} = 500$, $\epsilon = 10^{-3}$, $\rho = 15\,\text{dB}$. The MC bound and the DT-USTM bound are almost indistinguishable. The shaded area indicates the area in which $R^*(L, \epsilon, \rho)$ lies.

expression (4.31b) for $\underline{I}(\rho)$. For $\rho = 25\,\text{dB}$ and $n \geq 200$, the normal approximation is accurate even when we approximate $\underline{I}(\rho)$ using the simplified expression (5.3). Further observe that the normal approximation is pessimistic for $\rho = 15\,\text{dB}$ and optimistic for $\rho = 25\,\text{dB}$. As expected, the coherent normal approximation is strictly larger than the noncoherent high-SNR normal approximation. The gap between the two normal approximations appears to be independent of $n$. This agrees with the intuition that the cost for estimating the channel mostly depends on the coherence interval $\mathsf{T}$. Finally observe that the DT lower bound on $R^*(L, \epsilon, \rho)$, computed for USTM channel inputs, is fairly close to the MC upper bound, which holds for any input distribution satisfying the power constraint (2.6a), for $n \geq 5$ and $\rho = 15\,\text{dB}$ or $n \geq 2$ and $\rho = 25\,\text{dB}$. Thus, while it was shown that USTM channel inputs achieve the capacity asymptotically as the SNR tends to infinity, they also give rise to lower bounds on $R^*(L, \epsilon, \rho)$ that are impressively tight for moderate SNR values and short blocklengths. A similar observation was also made in [13].

In Figs. 5.3 and 5.4, we show the high-SNR normal approximation (5.2) (with $\underline{I}(\rho)/\mathsf{T}$ evaluated using the approximation (5.3)) as a function of the coherence interval $\mathsf{T}$ for a fixed blocklength $n$ (hence $L$ is inversely proportional to $\mathsf{T}$). We further plot the coherent normal approximation (5.10). For comparison, we also show the DT bound (see (5.21) below), evaluated for an USTM input distribution, and the weakened version of the MC bound (5.11) evaluated by Monte Carlo simulations. Finally, we present the normal approximation that was proposed in [10] for quasistatic

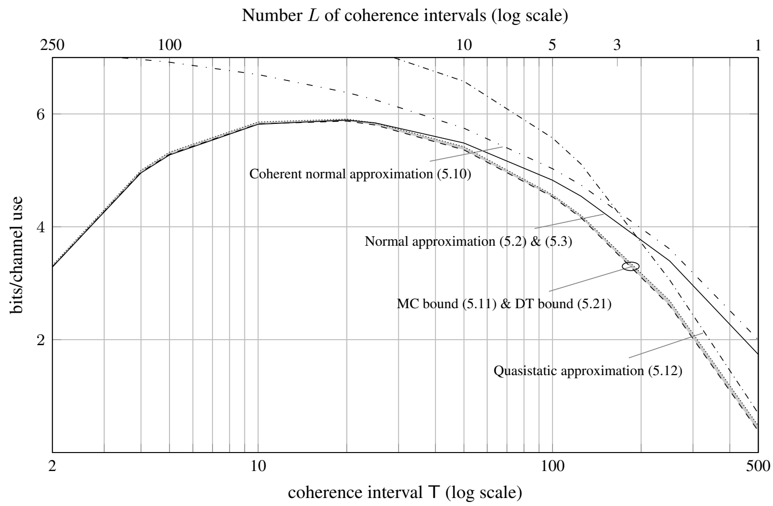


Figure 5.4: Bounds on $R^*(L, \epsilon, \rho)$ for $n = L\mathsf{T} = 500$, $\epsilon = 10^{-3}$, $\rho = 25\,\text{dB}$. The MC bound and the DT-USTM bound are almost indistinguishable. The shaded area indicates the area in which $R^*(L, \epsilon, \rho)$ lies.

MIMO block-fading channels. To adapt the quasistatic MIMO block-fading channel to our system model, we replace $\mathbb{H}$ in [10] by an $L \times L$ diagonal matrix with diagonal entries $H_1, \ldots, H_L$. Thus, specializing [10, Eq. (95)] to our case, we obtain

$$\epsilon \approx \mathbb{E}\left[Q\left(\frac{C(\mathbb{H}) - L\, R^*(L, \epsilon, \rho)}{\sqrt{V(\mathbb{H})/\mathsf{T}}}\right)\right] \tag{5.12}$$

where

$$C(\mathbb{H}) \triangleq \sum_{j=1}^{L} \log(1 + \rho |H_j|^2) \tag{5.13a}$$

$$V(\mathbb{H}) \triangleq L - \sum_{j=1}^{L} \frac{1}{\log(1 + \rho |H_j|^2)^2}. \tag{5.13b}$$

As already observed in Figs. 5.1 and 5.2, the high-SNR normal approximation is fairly accurate for $\rho = 15$ dB and $L \geq 10$, and it is indistinguishable from the DT and MC bounds for $\rho = 25$ dB and $L \geq 10$. The high-SNR normal approximation becomes less accurate as $L$ decreases. Observe that the coherent normal approximation (5.10) provides a good approximation when $\mathsf{T}$ is large but becomes inaccurate when $\mathsf{T} \leq 100$. Further observe that the normal approximation for the quasistatic case (5.12), which is tailored towards the case where $L$ is small, becomes accurate only for $L \leq 3$ in both figures. The figures show that there is an optimal tradeoff between $L$ and $\mathsf{T}$ for

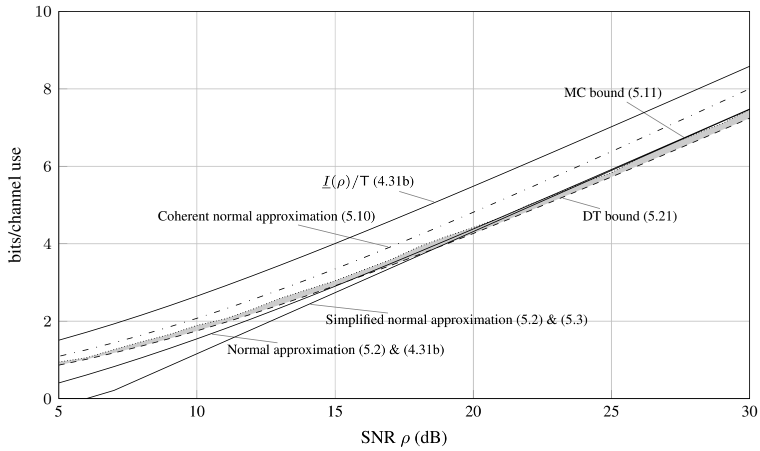


Figure 5.5: Bounds on $R^*(L,\epsilon,\rho)$ for $\mathsf{T}=20$, $L=25$ and $\epsilon=10^{-3}$. The shaded area indicates the area in which $R^*(L,\epsilon,\rho)$ lies.

a fixed blocklength $n$. This is, for example, of relevance for the design of orthogonal frequency-division multiplexing (OFDM) systems, where the duration of a codeword is smaller than the coherence time, hence only frequency diversity is available. The system designer can then determine the number of diversity branches $L$ available to each user by assigning OFDM symbols from different time and frequency slots. Figs. 5.1 and 5.2 indicate the optimal value of $L$ for $\epsilon=10^{-3}$ and $\rho=\{15,25\}$ dB. We refer to [34] for a more detailed discussion.

In Fig. 5.5, we plot the high-SNR normal approximation (5.2), evaluating $\underline{I}(\rho)$ using both (4.31b) and (5.3), as a function of the SNR $\rho$ for fixed $\mathsf{T}$ and $L$. Again, we also plot the coherent normal approximation (5.10). For comparison, we further plot the DT bound (see (5.21) below) evaluated for an USTM input distribution, the weakened version of the MC bound (5.11), and $\underline{I}(\rho)/\mathsf{T}$ using (4.31b). Observe that the normal approximation that uses (4.31b) becomes accurate already at SNR values of 15 dB, while the normal approximation that uses $\underline{I}(\rho)$ in (5.3) is accurate from SNR values of 20 dB. Further observe that the normal approximation is pessimistic for $\rho<20$ dB and optimistic for $\rho\geq 20$ dB. As expected, the coherent normal approximation is strictly larger than the noncoherent high-SNR normal approximation. Observe that the gap between the coherent normal approximation and the nonasymptotic bounds stays constant for $\rho\geq 15$ dB but decreases as $\rho$ becomes small. This is because, for small values of $\rho$, knowledge of the fading coefficients is less essential. Finally, we again observe that USTM channel input, which achieve the capacity asymptotically as the SNR tends to infinity, also give rise to lower bounds on $R^*(L,\epsilon,\rho)$ that are impressively tight for all SNR values

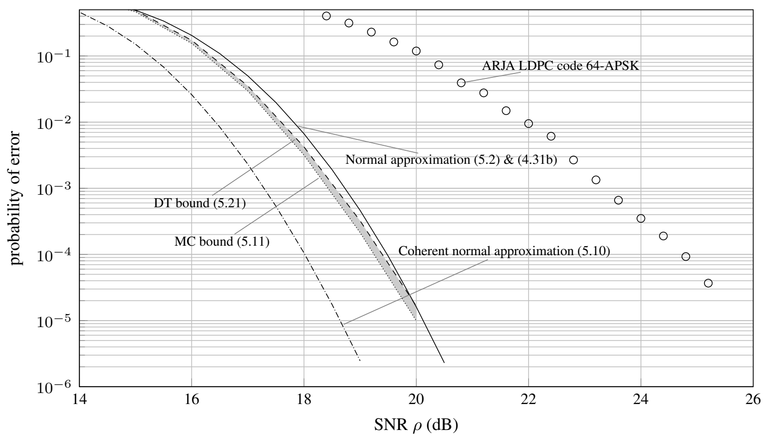


Figure 5.6: Bounds on the probability of error $\epsilon$ for $R = 4$, $\mathsf{T} = 20$ and $L = 25$. The shaded area indicates the area in which the true probability of error $\epsilon$ lies.

considered in the plot.

In Fig. 5.6, we plot the probability of error as a function of the SNR $\rho$ for $R = 4$, $\mathsf{T} = 20$, and $L = 25$. Specifically, we show the high-SNR normal approximation (5.2), with $\underline{I}(\rho)$ evaluated using (4.31b), the coherent normal approximation (5.10), the DT bound (see (5.21) below) evaluated for an USTM input distribution, and the weakened version of the MC bound (5.11). For comparison, we further show the performance of an accumulate-repeat-jagged-accumulate (ARJA) low-density parity-check (LDPC) (3000,2000)-code combined with a 64-amplitude phase-shift keying (APSK) modulation [35, Figure 3(b)], pilot-assisted transmission (2 pilot symbols per coherence block), and maximum likelihood channel estimation followed by mismatched nearest-neighbor decoding at the receiver. For details see [36, Sec. 4]. Observe that the high-SNR normal approximation is accurate for the whole range of SNRs evaluated. Further observe that the gap between the presented real code and the rest of curves is substantial. This suggests that more sophisticated joint channel-estimation decoding procedures together with shaping techniques need to be adopted to close the gap (see e.g., [37]).

### 5.1.3 Engineering Wisdom

As argued, e.g., in [3], the normal approximation can be used to analyze the performance of communication protocols. For example, let us consider the uplink scenario in [3, Sec. IV-C], where $\mathsf{d}$ devices intend to send $\mathsf{k}$ information bits to a base station within the time corresponding to $n$ channel uses. The $n$ channel uses are divided into $\mathsf{s}$ equally-sized slots of $n_{\mathsf{s}} \triangleq n/\mathsf{s}$ channels uses. The devices apply a simple slotted-ALOHA protocol: each device picks randomly one of the $\mathsf{s}$ slots in the frame and sends its packet. If two or more devices pick the same slot, then a collision occurs and none of their packets is received correctly. If only one device picks a particular slot (singleton slot), then the error probability is calculated using the normal approximation. Specifically, in [3, Sec. IV-C] the normal approximation for the AWGN channel was considered, i.e.,[1]

$$R^*(n,\epsilon) \approx C_{\text{AWGN}}(\rho) - \sqrt{\frac{V_{\text{AWGN}}(\rho)}{n}} Q^{-1}(\epsilon) + \frac{1}{2}\frac{\log n}{n} \tag{5.14}$$

where

$$C_{\text{AWGN}}(\rho) = \log(1+\rho) \tag{5.15a}$$

$$V_{\text{AWGN}}(\rho) = \rho\frac{2+\rho}{(1+\rho)^2}. \tag{5.15b}$$

By solving (5.14) for $\epsilon$, we obtain an approximation for the packet error probability as a function of the packet length $n$, the number of information bits $\mathsf{k} = nR$ to be conveyed in a packet, and the SNR $\rho$, i.e.,

$$\epsilon^*(\mathsf{k}, n, \rho) \approx Q\left(\frac{nC_{\text{AWGN}}(\rho) - \mathsf{k}\log 2 + (\log n)/2}{\sqrt{nV_{\text{AWGN}}(\rho)}}\right). \tag{5.16}$$

By replacing (5.14) by our high-SNR normal approximation (5.2), we obtain the following approximation for the packet error probability when packets are transmitted over a noncoherent single-antenna Rayleigh block-fading channel of coherence interval $\mathsf{T}$:

$$\epsilon^*(\mathsf{k}, n, \rho) \approx Q\left(\frac{n\underline{I}(\rho) - \mathsf{k}\mathsf{T}\log 2}{\sqrt{n\mathsf{T}\tilde{V}}}\right). \tag{5.17}$$

Likewise, replacing (5.14) by the normal approximation for the coherent Rayleigh block-fading channel [5, Eq. (34)], we obtain

$$\epsilon^*(\mathsf{k}, n, \rho) \approx Q\left(\frac{nC_c(\rho) - \mathsf{k}\log 2}{\sqrt{n\mathsf{T}V_c(\rho)}}\right) \tag{5.18}$$

where

$$C_c(\rho) \triangleq \mathsf{E}\big[\log(1+\rho Z_1)\big] \tag{5.19a}$$

$$V_c(\rho) \triangleq \mathsf{T}\mathrm{Var}\big[\log(1+\rho Z_1)\big] + 1 - \mathsf{E}\left[\frac{1}{1+\rho Z_1}\right]. \tag{5.19b}$$

[1]For the AWGN channel, the $\mathcal{O}(\log n/n)$ in (1.1) can be replaced by $(\log n)/(2n) + \mathcal{O}(1/n)$ [1, 38].

Table 5.1: Optimal slot size for different channel models and $n = L\mathsf{T} = 480$, $\mathsf{k} = 256$, $\mathsf{d} = 12$.

| SNR | coherence interval $\mathsf{T}$ | optimal number of slots $\mathsf{s}$ | | | |
|---|---|---|---|---|---|
| | | noncoherent Rayleigh block-fading | coherent Rayleigh block-fading | AWGN | classic slotted-ALOHA |
| $\rho = 15$ dB | $\mathsf{T} = 5$ | $\mathsf{s} = 4$ | $\mathsf{s} = 6$ | $\mathsf{s} = 8$ | $\mathsf{s} = 12$ |
| | $\mathsf{T} = 20$ | $\mathsf{s} = 6$ | $\mathsf{s} = 6$ | $\mathsf{s} = 8$ | $\mathsf{s} = 12$ |
| $\rho = 25$ dB | $\mathsf{T} = 5$ | $\mathsf{s} = 8$ | $\mathsf{s} = 12$ | $\mathsf{s} = 12$ | $\mathsf{s} = 12$ |
| | $\mathsf{T} = 20$ | $\mathsf{s} = 8$ | $\mathsf{s} = 8$ | $\mathsf{s} = 12$ | $\mathsf{s} = 12$ |

The probability of successful transmission is given by [3, Eq. (24)], namely,

$$P_{\text{success}} = \frac{\mathsf{d}}{\mathsf{s}}\left(1 - \frac{1}{\mathsf{s}}\right)^{\mathsf{d}-1}\left(1 - \epsilon^*(\mathsf{k}, n_{\mathsf{s}}, \rho)\right) \tag{5.20}$$

where $(\mathsf{d}/\mathsf{s})(1 - 1/\mathsf{s})^{\mathsf{d}-1}$ is the probability that only one device transmits in a given slot [39, Sec. 5.3.2]. Our goal is to choose $\mathsf{s}$ such that the probability of successful transmission is maximized given $\mathsf{d}$, $\mathsf{k}$, $n$ and $\rho$. This problem entails a tradeoff between the probability of collision and the number of channel uses available for each packet, which affects the achievable error probability in a singleton slot.

As a concrete example, we consider the case when $n = 480$, $\mathsf{d} = 12$, and $\mathsf{k} = 256$.[2] In Table 5.1, we show the optimal number of slots $\mathsf{s}$ for the noncoherent Rayleigh block-fading channel (with $\epsilon^*(\mathsf{k}, n_{\mathsf{s}}, \rho)$ approximated by (5.17)), the coherent Rayleigh block-fading channel (with $\epsilon^*(\mathsf{k}, n_{\mathsf{s}}, \rho)$ approximated by (5.18)), the AWGN channel (with $\epsilon^*(\mathsf{k}, n_{\mathsf{s}}, \rho)$ approximated by (5.16)), and the classic slotted-ALOHA protocol ($\epsilon^*(\mathsf{k}, n_{\mathsf{s}}, \rho) = 0$) for the SNR values $\rho = 15$ dB and $\rho = 25$ dB and coherence intervals $\mathsf{T} = 5$ and $\mathsf{T} = 20$. To be consistent with our system model, for the Rayleigh block-fading channel (both coherent and noncoherent) we only consider slot sizes $n_{\mathsf{s}}$ that are integer multiples of $\mathsf{T}$. Observe that the optimal number of slots $\mathsf{s}$ depends critically on the SNR, the coherence interval, and the considered channel model. For example, for the classic slotted-ALOHA protocol, the optimal number of slots is $\mathsf{s} = 12$, which coincides with the total number of devices $\mathsf{d} = 12$. In contrast, for the AWGN channel, the optimal number of slots is $\mathsf{s} = 8$ for $\rho = 15$ dB and coincides with the one of the classic slotted-ALOHA for $\rho = 25$ dB. In most cases, the optimal number of slots $\mathsf{s}$ for the Rayleigh block-fading channel (both coherent and noncoherent) is yet again smaller and depends both on the SNR and the coherence interval $\mathsf{T}$. When $\mathsf{T} = 20$, the optimal number of slots $\mathsf{s}$ for the noncoherent Rayleigh

[2] The fact that $n$ is fixed implies that the number of coherence intervals $L$ changes inversely proportional to $\mathsf{T}$ for the block-fading cases.

block-fading channel coincides with that for the coherent channel. This agrees with the intuition that, when $\mathsf{T}$ is sufficiently large, the fading coefficients can be learned with little training overhead. In general, the optimal number of slots $\mathsf{s}$ decreases as the channel becomes less favorable. Intuitively, larger codes are required to combat the impairments due to AWGN and fading. Hence, the packet length $n_{\mathsf{s}}$ must be increased or, equivalently, the number of slots $\mathsf{s} = n/n_{\mathsf{s}}$ must be reduced.

## 5.2 Proof of Theorem 5.1

The proof of Theorem 5.1 is based on a lower bound on $R^*(L,\epsilon,\rho)$, given in Section 5.2.1, and on an upper bound on $R^*(L,\epsilon,\rho)$, given in Section 5.2.2. Since these bounds coincide up to terms of order $\mathcal{O}_L(\log L/L)$ and $o_\rho(1)$ (compare (5.22) with (5.53) below, using (4.39a) and (4.39b)) they prove (5.1).

### 5.2.1 DT Lower Bound

To obtain a lower bound on $R^*(L,\epsilon,\rho)$, we evaluate the DT bound defined in Chapter 3.1.3.2 for the USTM input distribution defined in Chapter 4. Thus, assume that $\mathbf{X}^L \sim \bar{\mathrm{P}}_{\mathbf{X}^L}$, which implies $\mathbf{Y}^L \sim \mathrm{q}^{(\mathrm{U})}_{\mathbf{Y}^L}$. One can show (see [15, App. A]) that the CDF $\mathsf{P}[i(\mathbf{x}^L;\tilde{\mathbf{Y}}^L) \le \alpha]$ does not depend on $\mathbf{x}^L$. Furthermore, the USTM input distribution satisfies the power constraint (2.6a) with probability one. A lower bound on $R^*(L,\epsilon,\rho)$ follows therefore from the DT bound (maximum probability of error) (see Chapter 3.1.3.2), which, after a standard change of measure, can be stated as follows: there exists a code with $M$ codewords, blocklength $L\mathsf{T}$, and maximum probability of error $\epsilon$ not exceeding

$$\begin{aligned}\epsilon \le \mathsf{P}\big[i(\mathbf{X}^L;\mathbf{Y}^L) \le \log(M-1)\big]\\ + (M-1)\mathsf{E}\Big[e^{-i(\mathbf{X}^L;\mathbf{Y}^L)}\mathrm{I}\{i(\mathbf{X}^L;\mathbf{Y}^L) > \log(M-1)\}\Big].\end{aligned} \tag{5.21}$$

To show that (5.21) yields the lower bound

$$R^*(L,\epsilon,\rho) \ge \frac{I(\rho)}{\mathsf{T}} - \sqrt{\frac{V(\rho)}{L\mathsf{T}^2}}Q^{-1}(\epsilon) + \mathcal{O}_L\!\left(\frac{1}{L}\right) \tag{5.22}$$

we follow almost verbatim the steps in [1, Eqs. (258)–(267)] (with $\gamma$ in [1] replaced by $M-1$). The main difference is that, in our case, $V(\rho)$ defined in (4.38a) and $B(\rho)$ (cf. [1, Eq. (254)]) defined as

$$B(\rho) \triangleq \frac{6\mathsf{E}\Big[\big|i_\ell(\rho) - I(\rho)\big|^3\Big]}{V(\rho)^{3/2}} \tag{5.23}$$

depend on $\rho$. To ensure that the term $\mathcal{O}_L(1/L)$ in (5.22) is uniform in $\rho$, we will show that both $V(\rho)$ and $B(\rho)$ are bounded in $\rho$. We then apply the Berry-Esseen theorem [22, Ch. XVI.5] to obtain [1, Eq. (259)] with $B(\rho)$ replaced by an upper bound $B(\rho_0)$ that holds for all $\rho \geq \rho_0$ and a sufficiently large $\rho_0$, followed by [1, Eqs. (261)–(265)], which gives

$$R^*(L,\epsilon,\rho) \geq \frac{I(\rho)}{\mathsf{T}} - \sqrt{\frac{V(\rho)}{L\mathsf{T}^2}}Q^{-1}(\tau) \tag{5.24}$$

where

$$\tau = \epsilon - \left(\frac{2\log 2}{\sqrt{2\pi}} + 5B(\rho_0)\right)\frac{1}{\sqrt{L}}. \tag{5.25}$$

A Taylor-series expansion of $Q^{-1}(\tau)$ around $\epsilon$ yields then

$$Q^{-1}(\tau) = Q^{-1}(\epsilon) + \mathcal{O}_L\left(\frac{1}{\sqrt{L}}\right) \tag{5.26}$$

which in turn gives (5.22).

To show that $V(\rho)$ and $B(\rho)$ are bounded in $\rho$, we resort to the following lemmas:

**Lemma 5.2** *Let $\bar{V}_\rho(\alpha)$ be defined as in* (4.38b) *and let $0 \leq \delta \leq 1/2$. For every $\rho(1-\delta) \leq \alpha \leq \rho$, we have*

$$\bar{V}_\rho(\alpha) \geq \left(\frac{\mathsf{T}\rho}{1+\mathsf{T}\rho}\right)^2 (\mathsf{T}-1) - \Xi\delta + o_\rho(1) \tag{5.27}$$

*where $\Xi$ is a positive constant that only depends on $\mathsf{T}$.*

*Proof:* See Appendix A.2. ■

**Lemma 5.3** *For every $\rho_0 > 0$, we have*

$$\sup_{\substack{\alpha\geq 0,\\ \rho\geq\rho_0}} \bar{V}_\rho(\alpha) < \infty \tag{5.28a}$$

$$\sup_{\rho\geq\rho_0} V(\rho) < \infty. \tag{5.28b}$$

*Proof:* See Appendix A.3. ■

**Lemma 5.4** *For every $\rho_0 > 0$, we have*

$$\sup_{\substack{\alpha\geq 0,\\ \rho\geq\rho_0}} \mathsf{E}\left[\left|\bar{\jmath}_\ell(\alpha) - \bar{J}(\alpha)\right|^3\right] < \infty \tag{5.29a}$$

$$\sup_{\rho\geq\rho_0} \mathsf{E}\left[\left|i_\ell(\rho) - I(\rho)\right|^3\right] < \infty. \tag{5.29b}$$

*Proof:* See Appendix A.4. ■

For $\delta = 0$, Lemma 5.2 yields

$$\bar{V}(\rho) \geq \left(\frac{\mathsf{T}\rho}{1+\mathsf{T}\rho}\right)^2 (\mathsf{T}-1) + o_\rho(1). \tag{5.30}$$

Together with (4.39b) and (4.39d), this implies that

$$V(\rho) \geq \left(\frac{\mathsf{T}\rho_0}{1+\mathsf{T}\rho_0}\right)^2 \frac{\mathsf{T}-1}{2}, \quad \rho \geq \rho_0 \tag{5.31}$$

for a sufficiently large $\rho_0$. Furthermore, Lemma 5.3 implies that, for every $\rho_0 > 0$, there exists an $V_{\mathrm{UB}}(\rho_0)$ that is independent of $\rho$ and that satisfies

$$V(\rho) \leq V_{\mathrm{UB}}(\rho_0), \quad \rho \geq \rho_0. \tag{5.32}$$

Finally, Lemma 5.4 implies that for every $\rho_0 > 0$ there exists an $S(\rho_0)$ that is independent of $\rho$ and satisfies

$$\mathsf{E}\left[\left|i_\ell(\rho) - I(\rho)\right|^3\right] \leq S(\rho_0), \quad \rho \geq \rho_0. \tag{5.33}$$

Combining (5.31) and (5.33), it follows that for a sufficiently large $\rho_0 > 0$ there exists a $B(\rho_0)$ that is independent of $\rho$ and that satisfies

$$B(\rho) \leq \frac{6S(\rho_0)}{\left(\frac{\mathsf{T}\rho_0}{1+\mathsf{T}\rho_0}\right)^3 \left(\frac{\mathsf{T}-1}{2}\right)^{3/2}} \triangleq B(\rho_0), \quad \rho \geq \rho_0. \tag{5.34}$$

This concludes the proof of the lower bound (5.22).

### 5.2.2 MC Upper Bound

An upper bound on $R^*(L,\epsilon,\rho)$ follows from the MC bound defined in Chapter 3.2.1.2 computed for the auxiliary pdf $\mathrm{q}^{(\mathrm{U})}_{\mathbf{Y}^L}$ defined in (4.3), i.e.,

$$R^*(L,\epsilon,\rho) \leq \frac{1}{LT} \sup_{\boldsymbol{\alpha}\in[0,\rho]^L} \log\left(\frac{1}{\beta(\boldsymbol{\alpha}, \mathrm{q}^{(\mathrm{U})}_{\mathbf{Y}^L})}\right). \tag{5.35}$$

Here, $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_L)$ denotes the vector of power allocations, and $\beta(\boldsymbol{\alpha}, \mathrm{q}^{(\mathrm{U})}_{\mathbf{Y}^L})$ denotes the minimum probability of error under hypothesis $\mathrm{q}^{(\mathrm{U})}_{\mathbf{Y}^L}$ if the probability of error under hypothesis $\mathrm{p}_{\mathbf{Y}^L|\mathbf{X}^L=\boldsymbol{x}^L}$ does not exceed $\epsilon$ [1, Eq. (100)]. Note that, by (4.6)–(4.9), $\beta(\boldsymbol{\alpha}, \mathrm{q}^{(\mathrm{U})}_{\mathbf{Y}^L})$ depends on $\boldsymbol{x}^L$ only via $\boldsymbol{\alpha}$ (recall that $\|\mathbf{X}_\ell\|^2 = \mathsf{T}\alpha_\ell$).

For $0 < \delta < 1$, let $L_\delta(\boldsymbol{\alpha})$ denote the number of $\alpha_\ell$'s in $\boldsymbol{\alpha}$ that satisfy $\rho(1-\delta) \leq \alpha_\ell \leq \rho$. The following lemma demonstrates that we can assume without loss of optimality that $L_\delta(\boldsymbol{\alpha}) \geq L/2$, i.e., in at least half of the coherence intervals $\alpha_\ell$ is larger than $\rho(1-\delta)$.

**Lemma 5.5** *For every $0 < \delta < 1$, $\mathsf{T} > 2$, and $0 < \epsilon < 1/2$, there exists a pair $(L_0, \rho_0)$ such that, for $L \geq L_0$ and $\rho \geq \rho_0$, the supremum in* (5.35) *can be replaced without loss of optimality by a supremum over $\boldsymbol{\alpha} \in \mathcal{A}_{\rho,\delta}$, where*

$$\mathcal{A}_{\rho,\delta} \triangleq \{\boldsymbol{\alpha} \in [0,\rho]^L : L_\delta(\boldsymbol{\alpha}) \geq L/2\}. \tag{5.36}$$

*Proof:* See Appendix A.5. ∎

In the following, we implicitly assume that $L \geq L_0$ and $\rho \geq \rho_0$ for some sufficiently large $L_0$ and $\rho_0$ so that Lemma 5.5 holds. Applying Lemma 5.5 to (5.35), and upper-bounding the RHS of (5.35) using (3.18) and (4.21), we obtain

$$R^*(L,\epsilon,\rho) \leq \sup_{\boldsymbol{\alpha}\in\mathcal{A}_{\rho,\delta}} \left\{ \frac{\log \xi(\boldsymbol{\alpha})}{L\mathsf{T}} - \frac{\log\Big(\mathsf{P}\big[\sum_{\ell=1}^{L} \bar{\jmath}_\ell(\alpha_\ell) \leq \log\xi(\boldsymbol{\alpha})\big] - \epsilon\Big)}{L\mathsf{T}} \right\} \tag{5.37}$$

for every $\xi\colon [0,\rho]^L \to (0,\infty)$.

Let

$$\bar{B}(\boldsymbol{\alpha}) \triangleq \frac{6\sum\limits_{\ell=1}^{L} \mathsf{E}\Big[\big|\bar{\jmath}_\ell(\alpha_\ell) - \bar{J}(\alpha_\ell)\big|^3\Big]}{\left(\sum\limits_{\ell=1}^{L} \bar{V}_\rho(\alpha_\ell)\right)^{3/2}}. \tag{5.38}$$

By Lemma 5.4, the expectation $\mathsf{E}\big[|\bar{\jmath}_\ell(\alpha) - \bar{J}(\alpha)|^3\big]$ can be upper-bounded by a constant $\bar{S}(\rho_0)$ that is independent of $\alpha$ and $\rho$. Furthermore, by the nonnegativity of $\bar{V}_\rho(\alpha_\ell)$,

$$\sum_{\ell=1}^{L} \bar{V}_\rho(\alpha_\ell) \geq \sum_{\ell\in\mathscr{L}_\delta(\boldsymbol{\alpha})} \bar{V}_\rho(\alpha_\ell) \tag{5.39}$$

where $\mathscr{L}_\delta(\boldsymbol{\alpha}) \triangleq \{\ell = 1,\dots,L : \alpha_\ell \geq \rho(1-\delta)\}$. Lemma 5.2 demonstrates that, for $\alpha \geq \rho(1-\delta)$,

$$\bar{V}_\rho(\alpha) \geq \left(\frac{\mathsf{T}\rho}{1+\mathsf{T}\rho}\right)^2 (\mathsf{T}-1) - \Xi\delta + o_\rho(1). \tag{5.40}$$

Thus, for

$$\delta = \left(\frac{\mathsf{T}\rho_0}{1+\mathsf{T}\rho_0}\right)^2 \frac{\mathsf{T}-1}{3\Xi} \tag{5.41}$$

and $\rho_0$ sufficiently large, we have

$$\sum_{\ell=1}^{L} \bar{V}_\rho(\alpha_\ell) \geq L_\delta(\boldsymbol{\alpha}) \left(\frac{\mathsf{T}\rho_0}{1+\mathsf{T}\rho_0}\right)^2 \frac{\mathsf{T}-1}{2}, \quad \rho \geq \rho_0. \tag{5.42}$$

Hence, for every $\boldsymbol{\alpha} \in \mathcal{A}_{\rho,\delta}$ and $\delta$ as chosen in (5.41),

$$\bar{B}(\boldsymbol{\alpha}) \leq \frac{6L\bar{S}(\rho_0)}{\left(\frac{(\mathsf{T}-1)L}{4}\right)^{3/2} \left(\frac{\mathsf{T}\rho_0}{1+\mathsf{T}\rho_0}\right)^3} \triangleq \frac{\bar{B}(\rho_0)}{\sqrt{L}}. \tag{5.43}$$

Let

$$\lambda = Q^{-1}\left(\epsilon + \frac{2\bar{B}(\rho_0)}{\sqrt{L}}\right) \tag{5.44}$$

and

$$\log \xi(\boldsymbol{\alpha}) = \sum_{\ell=1}^{L} \bar{J}(\alpha_\ell) - \lambda\sqrt{\sum_{\ell=1}^{L} \bar{V}_\rho(\alpha_\ell)}. \tag{5.45}$$

With this choice, the Berry-Esseen theorem and (5.43) imply that, for every $\boldsymbol{\alpha} \in \mathcal{A}_{\rho,\delta}$,

$$\left|\mathsf{P}\left[\sum_{\ell=1}^{L} \bar{\jmath}_\ell(\alpha_\ell) \le \log \xi(\boldsymbol{\alpha})\right] - Q(\lambda)\right| \le \bar{B}(\boldsymbol{\alpha}) \le \frac{\bar{B}(\rho_0)}{\sqrt{L}}. \tag{5.46}$$

Thus, for such $\boldsymbol{\alpha}$,

$$\mathsf{P}\left[\sum_{\ell=1}^{L} \bar{\jmath}_\ell(\alpha_\ell) \le \log \xi(\boldsymbol{\alpha})\right] \ge \epsilon + \frac{\bar{B}(\rho_0)}{\sqrt{L}}. \tag{5.47}$$

Substituting (5.47) into the upper bound (5.37), we obtain

$$R^*(L,\epsilon,\rho) \le \sup_{\boldsymbol{\alpha}\in\mathcal{A}_{\rho,\delta}} \left\{ \frac{\sum_{\ell=1}^{L} \bar{J}(\alpha_\ell)}{L\mathsf{T}} - \sqrt{\frac{\sum_{\ell=1}^{L} \bar{V}_\rho(\alpha_\ell)}{L^2\mathsf{T}^2}}\, Q^{-1}\left(\epsilon + \frac{2\bar{B}(\rho_0)}{\sqrt{L}}\right) \right\} - \frac{\log \bar{B}(\rho_0)}{L\mathsf{T}} + \frac{1}{2}\frac{\log L}{L\mathsf{T}}. \tag{5.48}$$

By the assumption $0 < \epsilon < \frac{1}{2}$, the inverse $Q$-function on the RHS of (5.48) is positive for sufficiently large $L$. It follows by the concavity of $x \mapsto \sqrt{x}$ and Jensen's inequality that (5.48) can be further upper-bounded as

$$\begin{aligned} R^*(L,\epsilon,\rho) &\le \frac{1}{L}\sum_{\ell=1}^{L} \sup_{0\le\alpha_\ell\le\rho} \left\{ \frac{\bar{J}(\alpha_\ell)}{\mathsf{T}} - \sqrt{\frac{\bar{V}_\rho(\alpha_\ell)}{L\mathsf{T}^2}}\, Q^{-1}\left(\epsilon + \frac{2\bar{B}(\rho_0)}{\sqrt{L}}\right) \right\} \\ &\quad - \frac{\log \bar{B}(\rho_0)}{L\mathsf{T}} + \frac{1}{2}\frac{\log L}{L\mathsf{T}} \\ &= \sup_{0\le\alpha\le\rho} \left\{ \frac{\bar{J}(\alpha)}{\mathsf{T}} - \sqrt{\frac{\bar{V}_\rho(\alpha)}{L\mathsf{T}^2}}\, Q^{-1}\left(\epsilon + \frac{2\bar{B}(\rho_0)}{\sqrt{L}}\right) \right\} \\ &\quad - \frac{\log \bar{B}(\rho_0)}{L\mathsf{T}} + \frac{1}{2}\frac{\log L}{L\mathsf{T}} \end{aligned} \tag{5.49}$$

where the second step follows because the channel is blockwise i.i.d., so the terms inside the curly brackets do not depend on $\ell$.

Applying a Taylor-series expansion of $Q^{-1}(\epsilon + 2\bar{B}(\rho_0)/\sqrt{L})$ around $\epsilon$, we obtain

$$Q^{-1}\left(\epsilon + \frac{2\bar{B}(\rho_0)}{\sqrt{L}}\right) = Q^{-1}(\epsilon) + \mathcal{O}_L\left(\frac{1}{\sqrt{L}}\right). \tag{5.50}$$

Further using that, by Lemma 5.3, $\bar{V}_\rho(\alpha)$ is bounded in $\rho$ and $\alpha$, and collecting terms of order $\log L/L$, we can rewrite (5.49) as

$$R^*(L,\epsilon,\rho) \le \sup_{0\le\alpha\le\rho}\left\{\frac{\bar{J}(\alpha)}{\mathsf{T}} - \sqrt{\frac{\bar{V}_\rho(\alpha)}{L\mathsf{T}^2}}Q^{-1}(\epsilon)\right\} + \mathcal{O}_L\left(\frac{\log L}{L}\right). \tag{5.51}$$

We next show that

$$\sup_{0\le\alpha\le\rho}\left\{\frac{\bar{J}(\alpha)}{\mathsf{T}} - \sqrt{\frac{\bar{V}_\rho(\alpha)}{L\mathsf{T}^2}}Q^{-1}(\epsilon)\right\} = \frac{\bar{J}(\rho)}{\mathsf{T}} - \sqrt{\frac{\bar{V}(\rho)}{L\mathsf{T}^2}}Q^{-1}(\epsilon) + \mathcal{O}_L\left(\frac{1}{L}\right). \tag{5.52}$$

We then obtain the desired upper bound

$$R^*(L,\epsilon,\rho) \le \frac{\underline{I}(\rho)+o_\rho(1)}{\mathsf{T}} - \sqrt{\frac{\tilde{V}+o_\rho(1)}{L\mathsf{T}^2}}Q^{-1}(\epsilon) + \mathcal{O}_L\left(\frac{\log L}{L}\right) \tag{5.53}$$

from (4.39c) and (4.39d).

To prove (5.52), we first present the following auxiliary results.

**Lemma 5.6** *1. Assume that $\mathsf{T} > 2$. For sufficiently large $\rho$, we have*

$$\sup_{0\le\alpha\le\rho} \bar{J}(\alpha) = \bar{J}(\rho). \tag{5.54}$$

*2. Assume that $\mathsf{T} > 2$ and $0 < \epsilon < \frac{1}{2}$. Consider the supremum on the left-hand side (LHS) of (5.52). For sufficiently large $L$ and $\rho$, we can assume without loss of optimality that $\alpha \in [\rho(1-\frac{\mathsf{K}}{L}),\rho]$ for some nonnegative constant $\mathsf{K}$ that is independent of $(L,\rho,\alpha)$.*

*Proof:* See Appendix A.6. ■

We next set out to prove (5.52). By Part 2) of Lemma 5.6, we can assume without loss of optimality that

$$\alpha \ge \rho\left(1-\frac{\mathsf{K}}{L}\right). \tag{5.55}$$

Furthermore, we show in Appendix A.8 that

$$\bar{V}_\rho(\alpha) \ge \bar{V}(\rho) - \Upsilon\delta, \quad \rho(1-\delta) \le \alpha \le \rho \tag{5.56}$$

where $\Upsilon$ is a positive constant that only depends on $\mathsf{T}$. Particularizing this bound for $\delta = \mathsf{K}/L$, we obtain

$$\bar{V}_\rho(\alpha) \ge \bar{V}(\rho) - \Upsilon\frac{\mathsf{K}}{L}, \quad \rho\left(1-\frac{\mathsf{K}}{L}\right) \le \alpha \le \rho. \tag{5.57}$$

Combining (5.57) with Part 1) of Lemma 5.6, and using that by the assumption $0 < \epsilon < \frac{1}{2}$ we have $Q^{-1}(\epsilon) > 0$, we obtain

$$\sup_{0\leq\alpha\leq\rho}\left\{\frac{\bar{J}(\alpha)}{L\mathsf{T}} - \sqrt{\frac{\bar{V}_\rho(\alpha)}{L\mathsf{T}^2}}Q^{-1}(\epsilon)\right\} \leq \frac{\bar{J}(\rho)}{\mathsf{T}} - \sqrt{\frac{\bar{V}(\rho) - \frac{\Upsilon\mathsf{K}}{L}}{L\mathsf{T}^2}}Q^{-1}(\epsilon)$$
$$= \frac{\bar{J}(\rho)}{\mathsf{T}} - \sqrt{\frac{\bar{V}(\rho)}{L\mathsf{T}^2}}Q^{-1}(\epsilon) + \mathcal{O}_L\left(\frac{1}{L}\right). \qquad (5.58)$$

This proves (5.52) and concludes the proof of the upper bound.

## 5.3 Conclusion

We presented a high-SNR normal approximation for the maximum coding rate $R^*(L, \epsilon, \rho)$ achievable over noncoherent, single-antenna, Rayleigh block-fading channels using an error-correcting code that spans $L$ coherence intervals, has a block-error probability no larger than $\epsilon$, and satisfies the power constraint $\rho$. The high-SNR normal approximation is roughly equal to the normal approximation one obtains by transmitting one pilot symbol per coherence block to estimate the fading coefficient, and by then transmitting $\mathsf{T} - 1$ symbols per coherence block over a coherent fading channel. This suggests that, at high SNR, one pilot symbol per coherence block suffices to achieve both the capacity and the channel dispersion. 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 at SNR values of 15 dB and for 10 coherence intervals or more.

The obtained normal approximation is useful in two ways. First, it complements the nonasymptotic bounds provided in Chapter 3 ( see also [13, 14, 15]), whose evaluation is computationally demanding. Second, it lays the foundation for analytical studies that analyze the behavior of the maximum coding rates as a function of system parameters such as SNR, number of coherence intervals, or blocklength. An example of such a study was illustrated in Section 5.1.3 concerning the optimal design of a simple slotted-ALOHA protocol. Needless to say, the obtained normal approximation can also be used to study more sophisticated communication protocols.

# 6
# Saddlepoint Approximations

In this chapter, we apply the saddlepoint method to derive approximations of the MC upper bound and the $\mathrm{RCU_s}$ lower bound introduced in Chapter 3 on the maximum coding rate $R^*(L,\epsilon,\rho)$ (or *vice-versa* on the minimum probability of error $\epsilon^*(L,R,\rho)$) for the noncoherent, single-antenna, Rayleigh block-fading channel introduced in (4.1) using error-correcting codes that span $L$ coherence intervals, have a block-error probability no larger than $\epsilon$, and satisfy the per-coherence-interval equal power constraint (2.6b). While these approximations must be evaluated numerically, the computational complexity is independent of the number of diversity branches $L$. This is in stark contrast to the nonasymptotic MC and $\mathrm{RCU_s}$ bounds, whose evaluation has a computational complexity that grows linearly in $L$. Numerical evidence suggests that the saddlepoint approximations, although developed under the assumption of large $L$, are accurate even for $L = 1$ if the SNR is greater than or equal to 0 dB. Furthermore, the proposed approximations are shown to recover the normal approximation and the reliability function of the channel, thus providing a unifying tool for the two regimes, which are usually considered separately in the literature.

In our analysis, the saddlepoint method is applied to the tail probabilities appearing in the nonasymptotic bounds. These probabilities often depend on a set of parameters, such as the SNR. Existing saddlepoint expansions do not consider such dependencies. Hence, they can only characterize the behavior of the expansion error in function of the number of coherence intervals $L$, but not in terms of the remaining parameters. In contrast, we derive a saddlepoint expansion for random variables whose distribution depends on an extra parameter, carefully analyze the error terms, and demonstrate that they are uniform in such an extra parameter. We then apply the expansion to the Rayleigh block-fading channel and obtain approximations in which the error terms depend only on the blocklength and are uniform in the remaining parameters.

## 6.1 Saddlepoint Expansion

Let $\{X_k\}_{k=1}^n$ be a sequence of i.i.d., real-valued, zero-mean, random variables, whose distribution depends on $\theta \in \Theta$, where $\Theta$ denotes the set of possible values of $\theta$.

The MGF of $X_k$ is defined as

$$m_\theta(\zeta) \triangleq \mathsf{E}\left[e^{\zeta X_k}\right] \tag{6.1}$$

the CGF is defined as

$$\psi_\theta(\zeta) \triangleq \log m_\theta(\zeta) \tag{6.2}$$

and the characteristic function is defined as

$$\varphi_\theta(\zeta) \triangleq \mathsf{E}\left[e^{\imath\zeta X_k}\right]. \tag{6.3}$$

We denote by $m_\theta^{(k)}(\zeta)$ and $\psi_\theta^{(k)}(\zeta)$ the $k$-th derivative of $\zeta \mapsto m_\theta(\zeta)$ and $\zeta \mapsto \psi_\theta(\zeta)$, respectively. For the first, second, and third derivatives we sometimes use the notation $m'_\theta(\zeta)$, $m''_\theta(\zeta)$, $m'''_\theta(\zeta)$, $\psi'_\theta(\zeta)$, $\psi''_\theta(\zeta)$, and $\psi'''_\theta(\zeta)$.

A random variable $X_k$ is said to be *lattice* if it is supported on the points $b$, $b \pm h$, $b \pm 2h\ldots$ for some $b$ and $h$. A random variable that is not lattice is said to be *nonlattice*. It can be shown that a random variable is nonlattice if, and only if [22, Ch. XV.1, Lemma 4]

$$|\varphi_\theta(\zeta)| < 1, \quad \text{for every } \zeta \neq 0. \tag{6.4}$$

We shall say that a family of random variables $X_k$ (parametrized by $\theta$) is nonlattice if

$$\sup_{\theta\in\Theta} |\varphi_\theta(\zeta)| < 1, \quad \text{for every } \zeta \neq 0. \tag{6.5}$$

Similarly, we shall say that a family of distributions (parametrized by $\theta$) is nonlattice if the corresponding family of random variables is nonlattice.

**Proposition 6.1** *Let the family of i.i.d. random variables $\{X_k\}_{k=1}^n$ (parametrized by $\theta$) be nonlattice. Suppose that there exists a $\zeta_0 > 0$ such that*

$$\sup_{\substack{\theta\in\Theta,\\ |\zeta|<\zeta_0}} \left|m_\theta^{(k)}(\zeta)\right| < \infty, \quad k = 0, 1, 2, 3, 4 \tag{6.6}$$

*and*

$$\inf_{\substack{\theta\in\Theta,\\ |\zeta|<\zeta_0}} \left|\psi''_\theta(\zeta)\right| > 0. \tag{6.7}$$

*Then, we have the following results:*

Part 1)*: If for a given $\gamma \geq 0$ there exists a $\tau \in [0,\zeta_0)$ such that $n\psi'_\theta(\tau) = \gamma$, then*

$$\mathsf{P}\left[\sum_{k=1}^{n} X_k \geq \gamma\right] = e^{n[\psi_\theta(\tau)-\tau\psi'_\theta(\tau)]}\left[f_\theta(\tau,\tau) + \frac{K_\theta(\tau,n)}{\sqrt{n}} + o_n\left(\frac{1}{\sqrt{n}}\right)\right] \tag{6.8}$$

*where $o_n(1/\sqrt{n})$ comprises terms that vanish faster than $1/\sqrt{n}$ and are uniform in $\tau$ and $\theta$. Here,*

$$f_\theta(u,\tau) \triangleq e^{n\frac{u^2}{2}\psi''_\theta(\tau)} Q\left(u\sqrt{n\psi''_\theta(\tau)}\right) \tag{6.9a}$$

$$K_\theta(\tau,n) \triangleq \frac{\psi'''_\theta(\tau)}{6\psi''_\theta(\tau)^{3/2}}\left(-\frac{1}{\sqrt{2\pi}} + \frac{\tau^2\psi''_\theta(\tau)n}{\sqrt{2\pi}} - \tau^3\psi''_\theta(\tau)^{3/2}n^{3/2}f_\theta(\tau,\tau)\right). \tag{6.9b}$$

Part 2)*: Let $U$ be uniformly distributed on $[0,1]$. If for a given $\gamma \geq 0$ there exists a $\tau \in [0,\zeta_0)$ such that $n\psi'_\theta(\tau) = \gamma$, then*

$$\mathsf{P}\left[\sum_{k=1}^{n} X_k \geq \gamma + \log U\right]$$
$$= e^{n[\psi_\theta(\tau)-\tau\psi'_\theta(\tau)]}\left[f_\theta(\tau,\tau) + f_\theta(1-\tau,\tau) + \frac{\tilde{K}_\theta(\tau,n)}{\sqrt{n}} + o_n\left(\frac{1}{\sqrt{n}}\right)\right] \tag{6.10}$$

*where $\tilde{K}_\theta(\tau,n)$ is defined as*

$$\tilde{K}_\theta(\tau,n) \triangleq \frac{\psi'''_\theta(\tau)}{6\psi''_\theta(\tau)^{3/2}}\left[\frac{(2\tau-1)\,\psi''_\theta(\tau)n}{\sqrt{2\pi}} - \left(\psi''_\theta(\tau)n\right)^{3/2}\left(\tau^3 f_\theta(\tau,\tau) - (1-\tau)^3 f_\theta(1-\tau,\tau)\right)\right] \tag{6.11}$$

*and $o_n(1/\sqrt{n})$ is uniform in $\tau$ and $\theta$.*

**Corollary 6.2** *Assume that there exists a $\zeta_0 > 0$ satisfying* (6.6) *and* (6.7)*. If for a given $\gamma \geq 0$ there exists a $\tau \in [0, \min\{\zeta_0, 1-\delta\})$ (for some arbitrary $\delta > 0$ independent of $n$ and $\theta$) such that $n\psi'_\theta(\tau) = \gamma$, then the saddlepoint expansion* (6.10) *can be upper-bounded as*

$$\mathsf{P}\left[\sum_{k=1}^{n} X_k \geq \gamma + \log U\right]$$
$$\leq e^{n[\psi_\theta(\tau)-\tau\psi'_\theta(\tau)]}\left[f_\theta(\tau,\tau) + f_\theta(1-\tau,\tau) + \frac{\hat{K}_\theta(\tau)}{\sqrt{n}} + o_n\left(\frac{1}{\sqrt{n}}\right)\right] \tag{6.12}$$

*where $\hat{K}_\theta(\tau)$ is independent of $n$, and is defined as*

$$\hat{K}_\theta(\tau) \triangleq \frac{1}{\sqrt{2\pi}} \frac{\psi_\theta'''(\tau)}{6\psi_\theta''(\tau)^{3/2}} \tag{6.13}$$

*and $o_n(1/\sqrt{n})$ is uniform in $\tau$ and $\theta$.*

**Remark 6.1** *Since $X_k$ is zero-mean by assumption, we have that $m_\theta(\zeta) \geq 1$ by Jensen's inequality. Together with (6.6), this implies that*

$$\sup_{\substack{\theta\in\Theta,\\ |\zeta|<\zeta_0}} \left|\psi_\theta^{(k)}(\zeta)\right| < \infty, \quad k = 0, 1, 2, 3, 4. \tag{6.14}$$

**Remark 6.2** *When the nonnegative $\gamma$ grows sublinearly in $n$, for sufficiently large $n$, one can always find a $\tau \in (-\zeta_0, \zeta_0)$ such that $n\psi_\theta'(\tau) = \gamma$. Indeed, it follows by (6.6) and Remark 6.1 that $\tau \mapsto \psi_\theta(\tau)$ is an analytic function on $(-\zeta_0, \zeta_0)$ with power series*

$$\psi_\theta(\tau) = \frac{1}{2}\psi_\theta''(0)\tau^2 + \frac{1}{6}\psi_\theta'''(0)\tau^3 + \ldots \tag{6.15}$$

*Here, we have used that $\psi_\theta(0) = 0$ by definition and $\psi_\theta'(0) = 0$ because $X_k$ is zero-mean. By assumption (6.7), the function $\tau \mapsto \psi_\theta(\tau)$ is strictly convex. Together with $\psi_\theta'(0) = 0$, this implies that $\psi_\theta'(\tau)$ strictly increases for $\tau > 0$. Hence, the choice*

$$\psi_\theta'(\tau) = \frac{\gamma}{n} \tag{6.16}$$

*establishes a one-to-one mapping between $\tau$ and $\gamma$, and $\gamma/n \to 0$ implies that $\tau \to 0$. Thus, for sufficiently large $n$, $\tau$ is inside the region of convergence $(-\zeta_0, \zeta_0)$.*

*Proof of Proposition 6.1, Part 1):* The proof follows closely the steps by Feller [22, Ch. XVI]. Since we consider a slightly more involved setting, where the distribution of $X_k$ depends on a parameter $\theta$, we reproduce all the steps here. Let $F_\theta$ denote the distribution of $Y_k \triangleq X_k - \tilde{\gamma}$, where $\tilde{\gamma} \triangleq \gamma/n$. The CGF of $Y_k$ is given by

$$\tilde{\psi}_\theta(\zeta) \triangleq \psi_\theta(\zeta) - \zeta\tilde{\gamma}. \tag{6.17}$$

We consider a tilted random variable $V_k$ with distribution

$$\vartheta_{\theta,\tau}(x) = e^{-\tilde{\psi}_\theta(\tau)} \int_{-\infty}^{x} e^{\tau t} dF_\theta(t) = e^{-\psi_\theta(\tau)+\tau\tilde{\gamma}} \int_{-\infty}^{x} e^{\tau t} dF_\theta(t) \tag{6.18}$$

where the parameter $\tau$ lies in $(-\zeta_0, \zeta_0)$. Note that the exponential term $e^{-\psi_\theta(\tau)+\tau\tilde{\gamma}}$ on the RHS of (6.18) is a normalizing factor that guarantees that $\vartheta_{\theta,\tau}$ is a distribution.

Let $v_{\theta,\tau}(\zeta)$ denote the MGF of the tilted random variable $V_k$, which is given by

$$\begin{aligned}
v_{\theta,\tau}(\zeta) &= \int_{-\infty}^{\infty} e^{\zeta x} d\vartheta_{\theta,\tau}(x) \\
&= \int_{-\infty}^{\infty} e^{\zeta x} e^{-\psi_\theta(\tau)+\tau\tilde{\gamma}} e^{\tau x} dF_\theta(x) \\
&= e^{-\psi_\theta(\tau)+\tau\tilde{\gamma}} \int_{-\infty}^{\infty} e^{(\zeta+\tau)x} dF_\theta(x) \\
&= e^{-\psi_\theta(\tau)+\tau\tilde{\gamma}} \mathsf{E}\left[e^{(\zeta+\tau)(X_k-\tilde{\gamma})}\right] \\
&= e^{-\psi_\theta(\tau)} \mathsf{E}\left[e^{(\zeta+\tau)X_k}\right] e^{-\zeta\tilde{\gamma}} \\
&= \frac{m_\theta(\zeta+\tau)}{m_\theta(\tau)} e^{-\zeta\tilde{\gamma}}.
\end{aligned} \tag{6.19}$$

Together with $\mathsf{E}[V_k] = v'_{\theta,\tau}(0)$, this yields

$$\begin{aligned}
\mathsf{E}[V_k] &= \left.\frac{\partial v_{\theta,\tau}(\zeta)}{\partial \zeta}\right|_{\zeta=0} \\
&= e^{-\psi_\theta(\tau)} \left.\left(\mathsf{E}\left[X_k e^{(\zeta+\tau)X_k}\right] e^{-\zeta\tilde{\gamma}} - \tilde{\gamma} e^{-\zeta\tilde{\gamma}} \mathsf{E}\left[e^{(\zeta+\tau)X_k}\right]\right)\right|_{\zeta=0} \\
&= e^{-\psi_\theta(\tau)} \left(\mathsf{E}\left[X_k e^{\tau X_k}\right] - \tilde{\gamma} e^{\psi_\theta(\tau)}\right) \\
&= e^{-\psi_\theta(\tau)} \mathsf{E}\left[X_k e^{\tau X_k}\right] - \tilde{\gamma} \\
&= \psi'_\theta(\tau) - \tilde{\gamma}.
\end{aligned} \tag{6.20}$$

Note that, by (6.6), derivative and expected value can be swapped as long as $|\zeta+\tau| < \zeta_0$. This condition is, in turn, satisfied for sufficiently small $\zeta$ as long as $|\tau| < \zeta_0$. Following along similar lines, one can show that

$$\begin{aligned}
\mathsf{Var}[V_k] &= \mathsf{E}\left[V_k^2\right] - \mathsf{E}[V_k]^2 \\
&= v''_{\theta,\tau}(0) - v'_{\theta,\tau}(0)^2 \\
&= \psi''_\theta(\tau)
\end{aligned} \tag{6.21}$$

$$\begin{aligned}
\mathsf{E}\left[(V_k - \mathsf{E}[V_k])^3\right] &= \mathsf{E}\left[V_k^3\right] + 2\mathsf{E}[V_k]^3 - 3\mathsf{E}\left[V_k^2\right]\mathsf{E}[V_k] \\
&= v'''_{\theta,\tau}(0) + 2v'_{\theta,\tau}(0)^3 - 3v''_{\theta,\tau}(0)v'_{\theta,\tau}(0) \\
&= \psi'''_\theta(\tau)
\end{aligned} \tag{6.22}$$

and

$$\begin{aligned}
\mathsf{E}\left[(V_k - \mathsf{E}[V_k])^4\right] &= \mathsf{E}\left[V_k^4\right] - 3\mathsf{E}[V_k]^4 - 4\mathsf{E}\left[V_k^3\right]\mathsf{E}[V_k] + 6\mathsf{E}\left[V_k^2\right]\mathsf{E}[V_k]^2 \\
&= \psi^{(4)}_\theta(\tau) + 3\psi''_\theta(\tau)^2.
\end{aligned} \tag{6.23}$$

Let now $F_\theta^{\star n}$ denote the distribution of $\sum_{k=1}^n (X_k - \tilde{\gamma})$ and $\vartheta_{\theta,\tau}^{\star n}$ denote the distribution of $\sum_{k=1}^n V_k$. By (6.18) and (6.19), the distributions $F_\theta^{\star n}$ and $\vartheta_{\theta,\tau}^{\star n}$ again stand in the relationship (6.18) except that the term $e^{-\psi_\theta(\tau)}$ is replaced by $e^{-n\psi_\theta(\tau)}$ and $\tilde{\gamma}$ is replaced by $n\tilde{\gamma}$. Since $n\tilde{\gamma} = \gamma$, by inverting (6.18) we can establish the relationship

$$\mathsf{P}\left[\sum_{k=1}^n X_k \geq \gamma\right] = e^{n\psi_\theta(\tau)-\tau\gamma} \int_0^\infty e^{-\tau y} d\vartheta_{\theta,\tau}^{\star n}(y). \tag{6.24}$$

Furthermore, by choosing $\tau$ such that $n\psi_\theta'(\tau) = \gamma$, it follows from (6.20) that the distribution $\vartheta_{\theta,\tau}^{\star n}$ has zero mean. We next substitute in (6.24) the distribution $\vartheta_{\theta,\tau}^{\star n}$ by the zero-mean normal distribution with variance $n\psi_\theta''(\tau)$, denoted by $\mathfrak{N}_{n\psi_\theta''(\tau)}$, and analyze the error incurred by this substitution. To this end, we define

$$A_\tau \triangleq e^{n\psi_\theta(\tau)-\tau\gamma} \int_0^\infty e^{-\tau y} d\mathfrak{N}_{n\psi_\theta''(\tau)}(y). \tag{6.25}$$

By fixing $\tau$ according to (6.16), (6.25) becomes

$$\begin{aligned}
A_\tau &= \frac{e^{n\left[\psi_\theta(\tau)-\tau\psi_\theta'(\tau)\right]}}{\sqrt{2\pi n\psi_\theta''(\tau)}} \int_0^\infty e^{-\tau y} e^{-\frac{y^2}{2n\psi_\theta''(\tau)}} dy \\
&= \frac{e^{n\left[\psi_\theta(\tau)-\tau\psi_\theta'(\tau)\right]}}{\sqrt{2\pi}} \int_0^\infty e^{-\tau t\sqrt{n\psi_\theta''(\tau)}} e^{-\frac{t^2}{2}} dt \\
&= \frac{e^{n\left[\psi_\theta(\tau)-\tau\psi_\theta'(\tau)+\frac{\tau^2}{2}\psi_\theta''(\tau)\right]}}{\sqrt{2\pi}} \int_0^\infty e^{-\frac{1}{2}(t+\tau\sqrt{n\psi_\theta''(\tau)})^2} dt \\
&= \frac{e^{n\left[\psi_\theta(\tau)-\tau\psi_\theta'(\tau)+\frac{\tau^2}{2}\psi_\theta''(\tau)\right]}}{\sqrt{2\pi}} \int_{\tau\sqrt{n\psi_\theta''(\tau)}}^\infty e^{-\frac{x^2}{2}} dx \\
&= e^{n\left[\psi_\theta(\tau)-\tau\psi_\theta'(\tau)+\frac{\tau^2}{2}\psi_\theta''(\tau)\right]} Q\left(\tau\sqrt{n\psi_\theta''(\tau)}\right)
\end{aligned} \tag{6.26}$$

where the second equality follows by the change of variable $y = t\sqrt{n\psi_\theta''(\tau)}$, and the fourth equality follows by the change of variable $x = t + \tau\sqrt{n\psi_\theta''(\tau)}$.

We next show that the error incurred by substituting $\mathfrak{N}_{n\psi_\theta''(\tau)}$ for $\vartheta_{\theta,\tau}^{\star n}$ in (6.24) is small. To do so, we write

$$\begin{aligned}
\mathsf{P}\left[\sum_{k=1}^n X_k \geq n\psi_\theta'(\tau)\right] - A_\tau &= e^{n[\psi_\theta(\tau)-\tau\psi_\theta'(\tau)]} \int_0^\infty e^{-\tau y} \left(d\vartheta_{\theta,\tau}^{\star n}(y) - d\mathfrak{N}_{n\psi_\theta''(\tau)}(y)\right) \\
&= e^{n[\psi_\theta(\tau)-\tau\psi_\theta'(\tau)]} \Bigg[-\left(\vartheta_{\theta,\tau}^{\star n}(0) - \mathfrak{N}_{n\psi_\theta''(\tau)}(0)\right) \\
&\qquad + \tau \int_0^\infty \left(\vartheta_{\theta,\tau}^{\star n}(y) - \mathfrak{N}_{n\psi_\theta''(\tau)}(y)\right) e^{-\tau y} dy\Bigg]
\end{aligned} \tag{6.27}$$

where the last equality follows by integration by parts [22, Ch. V.6, Eq. (6.1)].

We next use [22, Sec. XVI.4, Th. 1] (stated as Lemma 6.3 below) to assess the error commited by replacing $\vartheta_{\theta,\tau}^{\star n}$ by $\mathfrak{N}_{n\psi_\theta''(\tau)}$. To state Lemma 6.3, we first introduce the following additional notation. Let $\{\tilde{X}_k\}_{k=1}^{n}$ be a sequence of i.i.d., real-valued, zero-mean, random variables with one-dimensional probability distribution $\tilde{F}_\theta$ that depends on an extra parameter $\theta \in \Theta$. We denote the $k$-th moment for any possible value of $\theta \in \Theta$ by

$$\mu_{k,\theta} = \int_{-\infty}^{\infty} x^k d\tilde{F}_\theta(x) \tag{6.28}$$

and we denote the second moment as $\mu_{2,\theta} = \sigma_\theta^2$.

For the distribution of the normalized $n$-fold convolution of a sequence of i.i.d., zero-mean, unit-variance random variables, we write

$$\tilde{F}_{n,\theta}(x) = \tilde{F}_\theta^{\star n}(x\sigma_\theta\sqrt{n}). \tag{6.29}$$

Note that $\tilde{F}_{n,\theta}$ has zero-mean and unit-variance. As above, we denote by $\mathfrak{N}$ the zero-mean, unit-variance, normal distribution, and we denote by $\mathfrak{n}$ the zero-mean, unit-variance, normal pdf.

**Lemma 6.3** *Assume that the family of distributions $\tilde{F}_{n,\theta}$ (parametrized by $\theta$) is nonlattice. Further assume that*

$$\sup_{\theta\in\Theta} \mu_{4,\theta} < \infty \tag{6.30}$$

*and*

$$\inf_{\theta\in\Theta} \sigma_\theta > 0. \tag{6.31}$$

*Then, for every $\theta \in \Theta$,*

$$\tilde{F}_{n,\theta}(x) - \mathfrak{N}(x) = \frac{\mu_{3,\theta}}{6\sigma_\theta^3\sqrt{n}}(1-x^2)\mathfrak{n}(x) + o_n\left(\frac{1}{\sqrt{n}}\right) \tag{6.32}$$

*where the $o_n(1/\sqrt{n})$ term is uniform in $x$ and $\theta$.*

*Proof:* See Appendix B.1. ∎

We next use (6.32) from Lemma 6.3 to expand (6.27). To this end, we first note that, as shown in Appendix B.2, if a family of distributions is nonlattice, then so is the corresponding family of tilted distributions. Consequently, the family of distributions $\vartheta_{\theta,\tau}^{\star n}$ (parametrized by $\theta$) is nonlattice since the family $F_\theta^{\star n}$ (parametrized by $\theta$) is nonlattice by assumption. We next note that the variable $y$ in (6.27) corresponds to $x\sigma_\theta\sqrt{n}$ in (6.29). Hence, $y = x\sqrt{\psi_\theta''(\tau)n}$, so applying (6.32) to (6.27) with

$\vartheta_{\theta,\tau}^{\star n}(y) = \tilde{F}_{n,\theta}(y/\sqrt{\psi_\theta''(\tau)n})$ and $\mathfrak{N}_{n\psi_\theta''(\tau)}(y) = \mathfrak{N}(y/\sqrt{\psi_\theta''(\tau)n})$, we obtain

$$
\begin{aligned}
&\mathsf{P}\left[\sum_{k=1}^{n} X_k \geq n\psi_\theta'(\tau)\right] - A_\tau \\
&= e^{n[\psi_\theta(\tau)-\tau\psi_\theta'(\tau)]}\Bigg[-\frac{1}{\sqrt{2\pi}}\frac{\psi_\theta'''(\tau)}{6\psi_\theta''(\tau)^{3/2}\sqrt{n}} + o_n\left(\frac{1}{\sqrt{n}}\right) \\
&\quad + \tau\int_0^\infty \left(\frac{\psi_\theta'''(\tau)}{6\psi_\theta''(\tau)^{3/2}\sqrt{n}}\left(1-\frac{y^2}{n\psi_\theta''(\tau)}\right)\mathfrak{n}\left(\frac{y}{\sqrt{\psi_\theta''(\tau)n}}\right) + o_n\left(\frac{1}{\sqrt{n}}\right)\right)e^{-\tau y}dy\Bigg] \\
&= e^{n[\psi_\theta(\tau)-\tau\psi_\theta'(\tau)]}\Bigg[o_n\left(\frac{1}{\sqrt{n}}\right) \\
&\quad + \frac{1}{\sqrt{2\pi}}\frac{\psi_\theta'''(\tau)}{6\psi_\theta''(\tau)^{3/2}\sqrt{n}}\left(-1 + \int_0^\infty \tau\sqrt{\psi_\theta''(\tau)n}\left(1-z^2\right)e^{-\tau\sqrt{\psi_\theta''(\tau)n}z-\frac{z^2}{2}}dz\right)\Bigg] \\
&= e^{n[\psi_\theta(\tau)-\tau\psi_\theta'(\tau)]}\Bigg[o_n\left(\frac{1}{\sqrt{n}}\right) \\
&\qquad + \frac{\psi_\theta'''(\tau)}{6\psi_\theta''(\tau)^{3/2}\sqrt{n}}\left(-\frac{1}{\sqrt{2\pi}} + \frac{\tau^2\psi_\theta''(\tau)n}{\sqrt{2\pi}} - \tau^3\psi_\theta''(\tau)^{3/2}n^{3/2}f_\theta(\tau,\tau)\right)\Bigg] \\
&= e^{n[\psi_\theta(\tau)-\tau\psi_\theta'(\tau)]}\left[\frac{K_\theta(\tau,n)}{\sqrt{n}} + o_n\left(\frac{1}{\sqrt{n}}\right)\right]
\end{aligned} \tag{6.33}
$$

with $f_\theta(\tau,\tau)$ defined in (6.9a), and $K_\theta(\tau,n)$ defined in (6.9b). Here we used that $\psi_\theta''(\tau)$ and $\psi_\theta'''(\tau)$ coincide with the second and third moments of the tilted random variable $V_k$, respectively; see (6.21) and (6.22). The second equality follows by the change of variable $y = z\sqrt{n\psi_\theta''(\tau)}$.

Finally, substituting $A_\tau$ in (6.26) into (6.33), and recalling that $n\psi_\theta'(\tau) = \gamma$, we obtain Part 1) of Proposition 6.1, namely

$$
\mathsf{P}\left[\sum_{k=1}^{n} X_k \geq n\psi_\theta'(\tau)\right] = e^{n[\psi_\theta(\tau)-\tau\psi_\theta'(\tau)]}\left[f_\theta(\tau,\tau) + \frac{K_\theta(\tau,n)}{\sqrt{n}} + o_n\left(\frac{1}{\sqrt{n}}\right)\right]. \tag{6.34}
$$

■

*Proof of Proposition 6.1, Part 2):* The proof of Part 2) follows along similar lines as the proof of Part 1). Hence, we will focus on describing what is different. Specifically, the LHS of (6.10) differs from the LHS of (6.8) by the additional term $\log U$. To account for this difference, we can follow the same steps as Scarlett *et al.* [23, Appendix E]. Since in our setting the distribution of $X_k$ depends on the parameter $\theta$, we repeat the main steps in the following:

$$
\mathsf{P}\left[\sum_{k=1}^{n} X_k \geq \gamma + \log U\right] = e^{n\psi_\theta(\tau)-\tau\gamma}\int_0^1\int_{\log u}^\infty e^{-\tau y}d\vartheta_{\theta,\tau}^{\star n}(y)du
$$

$$
\begin{aligned}
&= e^{n\psi_\theta(\tau)-\tau\gamma}\int_{-\infty}^{\infty}\int_0^{\min\{1,e^y\}} e^{-\tau y}du\, d\vartheta_{\theta,\tau}^{\star n}(y)\\
&= e^{n\psi_\theta(\tau)-\tau\gamma}\left(\int_0^{\infty} e^{-\tau y}d\vartheta_{\theta,\tau}^{\star n}(y)+\int_{-\infty}^{0} e^{(1-\tau)y}d\vartheta_{\theta,\tau}^{\star n}(y)\right)
\end{aligned}
\tag{6.35}
$$

where the second equality follows from Fubini's theorem [40, Ch. 2, Sec. 9.2]. We next proceed as in the proof of the previous part. The first term in (6.35) coincides with (6.24). We next focus on the second term, namely,

$$
e^{n\psi_\theta(\tau)-\tau\gamma}\int_{-\infty}^{0} e^{(1-\tau)y}d\vartheta_{\theta,\tau}^{\star n}(y). \tag{6.36}
$$

We substitute in (6.36) the distribution $\vartheta_{\theta,\tau}^{\star n}$ by the zero-mean normal distribution with variance $n\psi_\theta''(\tau)$, denoted by $\mathfrak{N}_{n\psi_\theta''(\tau)}$, which yields

$$
\tilde{A}_\tau \triangleq e^{n\psi_\theta(\tau)-\tau\gamma}\int_{0}^{\infty} e^{(1-\tau)y}d\mathfrak{N}_{n\psi_\theta''(\tau)}(y). \tag{6.37}
$$

By fixing $\tau$ according to (6.16), (6.37) can be computed as

$$
\begin{aligned}
\tilde{A}_\tau &= \frac{e^{n\left[\psi_\theta(\tau)-\tau\psi_\theta'(\tau)\right]}}{\sqrt{2\pi n\psi_\theta''(\tau)}}\int_{-\infty}^{0} e^{(1-\tau)y}e^{-\frac{y^2}{2n\psi_\theta''(\tau)}}dy\\
&= \frac{e^{n\left[\psi_\theta(\tau)-\tau\psi_\theta'(\tau)\right]}}{\sqrt{2\pi}}\int_{-\infty}^{0} e^{(1-\tau)t\sqrt{n\psi_\theta''(\tau)}}e^{-\frac{t^2}{2}}dt\\
&= \frac{e^{n\left[\psi_\theta(\tau)-\tau\psi_\theta'(\tau)+\frac{(1-\tau)^2}{2}\psi_\theta''(\tau)\right]}}{\sqrt{2\pi}}\int_{-\infty}^{0} e^{-\frac{1}{2}\left(t-(1-\tau)\sqrt{n\psi_\theta''(\tau)}\right)^2}dt\\
&= \frac{e^{n\left[\psi_\theta(\tau)-\tau\psi_\theta'(\tau)+\frac{(1-\tau)^2}{2}\psi_\theta''(\tau)\right]}}{\sqrt{2\pi}}\int_{-\infty}^{-(1-\tau)\sqrt{n\psi_\theta''(\tau)}} e^{-\frac{x^2}{2}}dx\\
&= \frac{e^{n\left[\psi_\theta(\tau)-\tau\psi_\theta'(\tau)+\frac{(1-\tau)^2}{2}\psi_\theta''(\tau)\right]}}{\sqrt{2\pi}}\int_{(1-\tau)\sqrt{n\psi_\theta''(\tau)}}^{\infty} e^{-\frac{x^2}{2}}dx\\
&= e^{n\left[\psi_\theta(\tau)-\tau\psi_\theta'(\tau)+\frac{(1-\tau)^2}{2}\psi_\theta''(\tau)\right]}Q\left((1-\tau)\sqrt{n\psi_\theta''(\tau)}\right)
\end{aligned}
\tag{6.38}
$$

where the second equality follows by the change of variable $y=t\sqrt{n\psi_\theta''(\tau)}$, and the fourth equality follows by the change of variable $x=t-(1-\tau)\sqrt{n\psi_\theta''(\tau)}$.

As we did in (6.27), we next evaluate the error incurred by substituting $\vartheta_{\theta,\tau}^{\star n}$ by $\mathfrak{N}_{n\psi_\theta''(\tau)}$ in (6.36). Indeed,

$$
e^{n\psi_\theta(\tau)-\tau\gamma}\int_{-\infty}^{0} e^{(1-\tau)y}d\vartheta_{\theta,\tau}^{\star n}(y)-\tilde{A}_\tau
$$

$$
\begin{aligned}
&= e^{n[\psi_\theta(\tau)-\tau\psi_\theta'(\tau)]} \int_{-\infty}^{0} e^{(1-\tau)y}\Big(d\vartheta_{\theta,\tau}^{\star n}(y) - d\mathfrak{N}_{n\psi_\theta''(\tau)}(y)\Big) \\
&= e^{n[\psi_\theta(\tau)-\tau\psi_\theta'(\tau)]} \bigg[\Big(\vartheta_{\theta,\tau}^{\star n}(0) - \mathfrak{N}_{n\psi_\theta''(\tau)}(0)\Big) \\
&\qquad\qquad - (1-\tau)\int_{-\infty}^{0}\Big(\vartheta_{\theta,\tau}^{\star n}(y) - \mathfrak{N}_{n\psi_\theta''(\tau)}(y)\Big)\, e^{(1-\tau)y}dy\bigg] \\
&= e^{n[\psi_\theta(\tau)-\tau\psi_\theta'(\tau)]} \bigg[o_n\Big(\frac{1}{\sqrt{n}}\Big) + \frac{1}{\sqrt{2\pi}}\frac{\psi_\theta'''(\tau)}{6\psi_\theta''(\tau)^{3/2}\sqrt{n}} \\
&\qquad \times \bigg(1 - \int_{-\infty}^{0}(1-\tau)\sqrt{\psi_\theta''(\tau)n}\big(1-z^2\big)\, e^{(1-\tau)\sqrt{\psi_\theta''(\tau)n}z-\frac{z^2}{2}}dz\bigg)\bigg] \\
&= e^{n[\psi_\theta(\tau)-\tau\psi_\theta'(\tau)]} \bigg[o_n\Big(\frac{1}{\sqrt{n}}\Big) + \frac{\psi_\theta'''(\tau)}{6\psi_\theta''(\tau)^{3/2}\sqrt{n}} \\
&\qquad \times \bigg(\frac{1}{\sqrt{2\pi}} - \frac{(1-\tau)^2\psi_\theta''(\tau)n}{\sqrt{2\pi}} + (1-\tau)^3\big(\psi_\theta''(\tau)n\big)^{3/2} f_\theta(1-\tau,\tau)\bigg)\bigg]
\end{aligned}
\tag{6.39}
$$

where the second step follows by integration by parts [22, Ch. V.6, Eq. (6.1)], and the second-to-last step by Lemma 6.3.

Combining (6.35) with (6.24), (6.34), (6.38), and (6.39), we obtain the desired result, namely,

$$
\begin{aligned}
&\mathsf{P}\bigg[\sum_{k=1}^{n} X_k \ge n\psi_\theta'(\tau) + \log U\bigg] \\
&\qquad = e^{n[\psi_\theta(\tau)-\tau\psi_\theta'(\tau)]}\bigg[f_\theta(\tau,\tau) + f_\theta(1-\tau,\tau) + \frac{\tilde{K}_\theta(\tau,n)}{\sqrt{n}} + o_n\Big(\frac{1}{\sqrt{n}}\Big)\bigg]
\end{aligned}
\tag{6.40}
$$

where $\tau$ is chosen according to (6.16) and $\tilde{K}_\theta(\tau,n)$ is defined in (6.11). ∎

*Proof of Corollary 6.2:* Using (6.35) with (6.25) and (6.37), and fixing $\tau$ according to (6.16), we can write

$$
\begin{aligned}
&\mathsf{P}\bigg[\sum_{k=1}^{n} X_k \ge \gamma + \log U\bigg] - A_\tau - \tilde{A}_\tau \\
&= e^{n[\psi_\theta(\tau)-\tau\psi_\theta'(\tau)]}\bigg[-\Big(\vartheta_{\theta,\tau}^{\star n}(0) - \mathfrak{N}_{n\psi_\theta''(\tau)}(0)\Big) + \Big(\vartheta_{\theta,\tau}^{\star n}(0) - \mathfrak{N}_{n\psi_\theta''(\tau)}(0)\Big) \\
&\qquad + \tau\int_{0}^{\infty}\Big(\vartheta_{\theta,\tau}^{\star n}(y) - \mathfrak{N}_{n\psi_\theta''(\tau)}(y)\Big)\, e^{-\tau y}dy \\
&\qquad - (1-\tau)\int_{-\infty}^{0}\Big(\vartheta_{\theta,\tau}^{\star n}(y) - \mathfrak{N}_{n\psi_\theta''(\tau)}(y)\Big)\, e^{(1-\tau)y}dy\bigg].
\end{aligned}
\tag{6.41}
$$

Using integration by parts as we did in (6.27) and (6.39), together with the change of variable $y = z\sqrt{\psi''_\theta(\tau)}$, the RHS of (6.41) can be written as

$$\begin{aligned}
&e^{n[\psi_\theta(\tau)-\tau\psi'_\theta(\tau)]}\Bigg[\frac{1}{\sqrt{2\pi n}}\frac{\psi'''_\theta(\tau)}{6\psi''_\theta(\tau)^{3/2}}\Bigg(\int_0^\infty \tau\sqrt{\psi''_\theta(\tau)n}\big(1-z^2\big)\,e^{-\tau\sqrt{\psi''_\theta(\tau)n}z-\frac{z^2}{2}}dz \\
&\qquad -\int_{-\infty}^0 (1-\tau)\sqrt{\psi''_\theta(\tau)n}\big(1-z^2\big)\,e^{(1-\tau)\sqrt{\psi''_\theta(\tau)n}z-\frac{z^2}{2}}dz\Bigg)+o_n\Big(\frac{1}{\sqrt{n}}\Big)\Bigg] \\
&= e^{n[\psi_\theta(\tau)-\tau\psi'_\theta(\tau)]}\Bigg[\frac{1}{\sqrt{2\pi n}}\frac{\psi'''_\theta(\tau)}{6\psi''_\theta(\tau)^{3/2}}\Bigg(\int_0^\infty \tau\sqrt{\psi''_\theta(\tau)n}\big(1-z^2\big)\,e^{-\tau\sqrt{\psi''_\theta(\tau)n}z-\frac{z^2}{2}}dz \\
&\qquad +\int_0^\infty (1-\tau)\sqrt{\psi''_\theta(\tau)n}\big(z^2-1\big)\,e^{-(1-\tau)\sqrt{\psi''_\theta(\tau)n}z-\frac{z^2}{2}}dz\Bigg)+o_n\Big(\frac{1}{\sqrt{n}}\Big)\Bigg]
\end{aligned} \tag{6.42}$$

where we replaced $z$ by $-z$ in the second integral. Keeping the positive part of each integral on the RHS of (6.42), it follows that

$$\begin{aligned}
&e^{n[\psi_\theta(\tau)-\tau\psi'_\theta(\tau)]}\Bigg[\frac{1}{\sqrt{2\pi n}}\frac{\psi'''_\theta(\tau)}{6\psi''_\theta(\tau)^{3/2}}\Bigg(\int_0^\infty \tau\sqrt{\psi''_\theta(\tau)n}\big(1-z^2\big)\,e^{-\tau\sqrt{\psi''_\theta(\tau)n}z-\frac{z^2}{2}}dz \\
&\qquad +\int_0^\infty (1-\tau)\sqrt{\psi''_\theta(\tau)n}\big(z^2-1\big)\,e^{-(1-\tau)\sqrt{\psi''_\theta(\tau)n}z-\frac{z^2}{2}}dz\Bigg)+o_n\Big(\frac{1}{\sqrt{n}}\Big)\Bigg] \\
&\le e^{n[\psi_\theta(\tau)-\tau\psi'_\theta(\tau)]}\Bigg[\frac{1}{\sqrt{2\pi}}\frac{\psi'''_\theta(\tau)}{6\psi''_\theta(\tau)^{3/2}\sqrt{n}}\Bigg(\int_0^1 \tau\sqrt{\psi''_\theta(\tau)n}\big(1-z^2\big)\,e^{-\tau\sqrt{\psi''_\theta(\tau)n}z-\frac{z^2}{2}}dz \\
&\qquad +\int_1^\infty (1-\tau)\sqrt{\psi''_\theta(\tau)n}\big(z^2-1\big)\,e^{-(1-\tau)\sqrt{\psi''_\theta(\tau)n}z-\frac{z^2}{2}}dz\Bigg)+o_n\Big(\frac{1}{\sqrt{n}}\Big)\Bigg].
\end{aligned} \tag{6.43}$$

We next bound each integral separately. The first integral on the RHS of (6.43) can be upper-bounded as

$$\begin{aligned}
\int_0^1 \tau\sqrt{\psi''_\theta(\tau)n}\big(1-z^2\big)\,e^{-\tau\sqrt{\psi''_\theta(\tau)n}z-\frac{z^2}{2}}dz &\le \int_0^1 \tau\sqrt{\psi''_\theta(\tau)n}e^{-\tau\sqrt{\psi''_\theta(\tau)n}z}dz \\
&= 1-e^{-\tau\sqrt{\psi''_\theta(\tau)n}} \\
&\le 1
\end{aligned} \tag{6.44}$$

where the first inequality follows by disregarding the quadratic exponent. The second integral on the RHS of (6.43) can be upper-bounded as

$$\begin{aligned}
&\int_1^\infty (1-\tau)\sqrt{\psi''_\theta(\tau)n}\big(z^2-1\big)\,e^{-(1-\tau)\sqrt{\psi''_\theta(\tau)n}z-\frac{z^2}{2}}dz \\
&\qquad\le \int_1^\infty (1-\tau)\sqrt{\psi''_\theta(\tau)n}z^2e^{-(1-\tau)\sqrt{\psi''_\theta(\tau)n}z}dz \\
&\qquad= \frac{\big[(1-\tau)\sqrt{\psi''_\theta(\tau)n}\big[(1-\tau)\sqrt{\psi''_\theta(\tau)n}+2\big]+2\big]\,e^{-(1-\tau)\sqrt{\psi''_\theta(\tau)n}}}{\Big((1-\tau)^2\,\psi''_\theta(\tau)n\Big)}.
\end{aligned} \tag{6.45}$$

If $\tau \in [0, \min\{\zeta_0, 1-\delta\})$ for some arbitrary $\delta$ independent of $n$ and $\theta$, then the RHS of (6.45) vanishes faster than $1/\sqrt{n}$ uniformly in $n$ and $\theta$. Substituting (6.44) and (6.45) into (6.43), we thus obtain the upper bound

$$\mathsf{P}\left[\sum_{k=1}^{n} X_k \geq n\psi_\theta'(\tau) + \log U\right] - A_\tau - \tilde{A}_\tau$$
$$\leq e^{n[\psi_\theta(\tau) - \tau\psi_\theta'(\tau)]}\left[\frac{1}{\sqrt{2\pi}}\frac{\psi_\theta'''(\tau)}{6\psi_\theta''(\tau)^{3/2}\sqrt{n}} + o_n\left(\frac{1}{\sqrt{n}}\right)\right]. \tag{6.46}$$

Consequently,

$$\mathsf{P}\left[\sum_{k=1}^{n} X_k \geq \gamma + \log U\right]$$
$$\leq e^{n[\psi_\theta(\tau) - \tau\psi_\theta'(\tau)]}\left[f_\theta(\tau,\tau) + f_\theta(1-\tau,\tau) + \frac{\hat{K}_\theta(\tau)}{\sqrt{n}} + o_n\left(\frac{1}{\sqrt{n}}\right)\right] \tag{6.47}$$

where $\hat{K}_\theta(\tau)$ was defined in (6.13). ■

## 6.2 Saddlepoint Expansions for RCU$_\mathrm{s}$ and MC

### 6.2.1 RCU$_\mathrm{s}$ Bound

As upper bound on $\epsilon^*(L,R,\rho)$, we use the RCU$_\mathrm{s}$ bound (3.6) which states that, for any $s > 0$,

$$\epsilon^*(L,R,\rho) \leq \mathsf{P}\left[\sum_{\ell=1}^{L}(I_s(\rho) - i_{\ell,s}(\rho)) \geq LI_s(\rho) + \log(U) - L\mathsf{T}R\right] \tag{6.48}$$

where $U$ is uniformly distributed on the interval $[0,1]$.

**Theorem 6.4 (Saddlepoint Expansion RCU$_\mathrm{s}$)** *Suppose that $\mathcal{S}$ is characterized either by* (4.45) *or by* (4.46)*. Then, the coding rate $R$ and minimum error probability $\epsilon^*$ can be parametrized by $(\tau,\rho,s) \in \mathcal{S}$ as*

$$R(\tau,s) = \frac{1}{\mathsf{T}}(I_s(\rho) - \psi_{\rho,s}'(\tau)) \tag{6.49a}$$
$$\epsilon^*(\tau,s) \leq e^{L\left[\psi_{\rho,s}(\tau) - \tau\psi_{\rho,s}'(\tau)\right]}\left[f_{\rho,s}(\tau,\tau) + f_{\rho,s}(1-\tau,\tau) + \frac{\hat{K}_{\rho,s}(\tau)}{\sqrt{L}} + o_L\left(\frac{1}{\sqrt{L}}\right)\right] \tag{6.49b}$$

*where $f(\cdot,\cdot)$ is defined in* (6.9a)*, $\hat{K}_{\rho,s}(\cdot)$ is defined in* (6.13)*, and $o_L(1/\sqrt{L})$ is uniform in $\tau$, $s$ and $\rho$.*

*Proof:* The desired result follows by applying Proposition 6.1 and Corollary 6.2 to (6.48). Indeed, by Lemma B.2 (Appendix B.3), the family of random variables $I_s(\rho) - i_{s,\ell}(\rho)$ (parametrized by $(\rho, s)$) is nonlattice. The first condition (6.6) required for Proposition 6.1 and Corollary 6.2 is satisfied by Lemma 4.2. It can be further observed that $V_s(\rho)$ is strictly increasing in $\rho$ (for a fixed $s$) and strictly increasing in $s$ (for a fixed $\rho$). Consequently, it is bounded away from zero for every $\rho \geq \rho_0$ and $s \geq s_0$ (for arbitrary $\rho_0 > 0$ and $s_0 > 0$). Since $\psi''_{\rho,s}(0) = V_s(\rho)$, it follows from Lemma B.3 that the second condition (6.7) required for Proposition 6.1 and Corollary 6.2 is satisfied, too. ■

**Remark 6.3** *The set $\mathcal{S}$ characterized by* (4.46) *with $s_{\max} = 1$ includes $0 \leq \tau < 1$. In this case, the identity* (6.49a) *with $s \in (0, 1]$ and $\tau \in [0, 1)$ characterizes all rates $R$ between the critical rate*

$$R_s^{cr}(\rho) \triangleq \frac{1}{\mathsf{T}}\big(I_s(\rho) - \psi'_{\rho,s}(1)\big) \tag{6.50}$$

*and $I_s(\rho)$. Solving* (6.49a) *for $\tau$, we obtain from Theorem 6.4 an upper bound on the minimum error probability $\epsilon^*(L, R, \rho)$ as a function of the rate $R \in (R_s^{cr}(\rho), I_s(\rho)]$, $s \in (0, 1]$.*

### 6.2.2 MC Bound

A lower bound on $\epsilon^*(L, R, \rho)$ follows by evaluating the MC bound (3.15) for the auxiliary distribution $\mathrm{q}^s_{\mathbf{Y}_\ell}(\mathbf{y}_\ell)$ given in (4.23) and using (3.16). This yields, for every $\xi > 0$ and $s > 0$,

$$\epsilon^*(L, R, \rho) \geq \mathsf{P}\left[\sum_{\ell=1}^{L}(I_s(\rho) - i_{\ell,s}(\rho)) \geq sLJ_s(\rho) - s\log\xi\right] - e^{(\log\xi - L\mathsf{T}R)} \tag{6.51}$$

where we have used (4.26) to express $j_{\ell,s}(\rho)$ in terms of $i_{\ell,s}(\rho)$.

**Theorem 6.5 (Saddlepoint Expansion MC)** *Suppose that $\mathcal{S}$ is characterized either by* (4.45) *or by* (4.46). *Then, for every rate $R$ and $(\tau, \rho, s) \in \mathcal{S}$*

$$\epsilon^*(L, R, \rho) \geq e^{L\left[\psi_{\rho,s}(\tau) - \tau\psi'_{\rho,s}(\tau)\right]}\left[f_{\rho,s}(\tau,\tau) + \frac{K_{\rho,s}(\tau, L)}{\sqrt{L}} + o_L\left(\frac{1}{\sqrt{L}}\right)\right] - e^{L\left[J_s(\rho) - \frac{\psi'_{\rho,s}(\tau)}{s} - \mathsf{T}R\right]} \tag{6.52}$$

*where $f(\cdot,\cdot)$ is defined in* (6.9a), *$K_{\rho,s}(\cdot,\cdot)$ is defined in* (6.11), *and the $o_L(1/\sqrt{L})$ term is uniform in $\tau$, $s$ and $\rho$.*

*Proof:* The inequality (6.52) follows by applying Proposition 6.1, Part 1) to (6.51) with $\xi = LJ_s(\rho) - L\psi'_{\rho,s}(\tau)/s$. Indeed, by Lemma B.2 (Appendix B.3), the family of random variables $I_s(\rho) - i_{s,\ell}(\rho)$ (parametrized by $(\rho, s)$) is nonlattice. The first condition (6.6) required for Proposition 6.1 is satisfied by Lemma 4.2. It can be further observed that $V_s(\rho)$ is strictly increasing in $\rho$ (for a fixed $s$) and strictly increasing in $s$ (for a fixed $\rho$). Consequently, it is bounded away from zero for every $\rho \geq \rho_0$ and $s \geq s_0$ (for arbitrary $\rho_0 > 0$ and $s_0 > 0$). Since $\psi''_{\rho,s}(0) = V_s(\rho)$, it follows from Lemma B.3 that the second condition (6.7) required for Proposition 6.1 is satisfied, too. ■

The expansions (6.49b) and (6.52) can be evaluated more efficiently than the nonasymptotic bounds (6.48) and (6.51). Indeed, (6.48) and (6.51) require the evaluation of the $L$-dimensional distribution of $\sum_{\ell=1}^{L} i_{\ell,s}(\rho)$, whereas (6.49b) and (6.52) depend only on the CGFs $\psi_{\rho,s}(\tau)$, $\psi'_{\rho,s}(\tau)$, $\psi''_{\rho,s}(\tau)$ and $\psi'''_{\rho,s}(\tau)$, which can be obtained by solving one-dimensional integrals.

## 6.3 Normal Approximation

The maximum coding rate can be expanded as

$$R^*(L, \epsilon, \rho) = \frac{I(\rho)}{\mathsf{T}} - \sqrt{\frac{V(\rho)}{L\mathsf{T}^2}} Q^{-1}(\epsilon) + \mathcal{O}_L\left(\frac{\log L}{L}\right) \tag{6.53}$$

where $I(\rho) = I_1(\rho)$ is defined in (4.27), and $V(\rho) = V_1(\rho)$ is defined in (4.28). This is usually referred to as *normal approximation.* As we shall show next, (6.53) can also be recovered from the expansions (6.49b) and (6.52).

### 6.3.1 Achievability Part

Let $\rho \geq \rho_0$ for some arbitrary $\rho_0 > 0$. To prove that the RHS of (6.53) is achievable, we consider (6.49a) and (6.49b) evaluated for $s = 1$, namely,

$$R(\tau, 1) = \frac{1}{\mathsf{T}}(I(\rho) - \psi'_{\rho,1}(\tau)) \tag{6.54a}$$

$$\epsilon^*(L, R, \rho) \leq e^{L[\psi_{\rho,1}(\tau) - \tau\psi'_{\rho,1}(\tau)]}\left[f_{\rho,1}(\tau, \tau) + f_{\rho,1}(1-\tau, \tau) + \frac{\hat{K}_{\rho,1}(\tau)}{\sqrt{L}} + o_L\left(\frac{1}{\sqrt{L}}\right)\right] \tag{6.54b}$$

and evaluate (6.54) for a sequence of $\tau$'s (as a function of $L$) defined as

$$\tau_L \triangleq \frac{Q^{-1}\left(\epsilon - \frac{k_{1,\rho}}{\sqrt{L}}\right)}{\sqrt{L\psi''_{\rho,1}(0)}}. \tag{6.55}$$

Here, $k_{1,\rho}$ is a positive constant independent of $L$ and uniform in $\rho$, which will be specified later. Note that, since $\tau_L$ decays to zero as $L \to \infty$, we have that $(\tau_L, \rho, 1) \in \mathcal{S}$ (with $\mathcal{S}$ characterized by (4.45)) for sufficiently large $L$. Consequently, Theorem 6.4 applies and the $o_L(1/\sqrt{L})$ term in (6.54b) is uniform in $\rho \geq \rho_0$ for some arbitrary $\rho_0 > 0$.

We next show that, for the choice of $\tau_L$ in (6.55), the RHS of (6.54a) equals the RHS of (6.53) and that the RHS of (6.54b) is less than $\epsilon$, which demonstrates that the rate $R(\tau, 1)$ is indeed a lower bound on $R^*(L, \epsilon, \rho)$.

To evaluate (6.49a), we first expand $\psi'_{\rho,1}(\tau)$ as the Taylor series

$$\psi'_{\rho,1}(\tau) = \tau\psi''_{\rho,1}(0) + \frac{\tau^2}{2}\psi'''_{\rho,1}(\tilde{\tau}) \tag{6.56}$$

for some $\tilde{\tau} \in (0, \tau)$. Applying then (6.55) and (6.56) to (6.54a), we obtain

$$R(\tau_L, 1) = \frac{I(\rho)}{\mathsf{T}} - \sqrt{\frac{\psi''_{\rho,1}(0)}{L\mathsf{T}^2}}Q^{-1}\left(\epsilon - \frac{k_{1,\rho}}{\sqrt{L}}\right) - \frac{Q^{-1}\left(\epsilon - \frac{k_{1,\rho}}{\sqrt{L}}\right)^2 \psi'''_{\rho,1}(\tilde{\tau})}{2L\psi''_{\rho,1}(0)}. \tag{6.57}$$

Using that $\psi''_{\rho,1}(0) = V(\rho)$, and expanding $Q^{-1}(\epsilon - k_{1,\rho}/\sqrt{L})$ around $\epsilon$, we can write (6.57) as

$$R(\tau_L, 1) = \frac{I(\rho)}{\mathsf{T}} - \sqrt{\frac{V(\rho)}{L\mathsf{T}^2}}Q^{-1}(\epsilon) + \mathcal{O}_L\left(\frac{1}{L}\right). \tag{6.58}$$

where the $\mathcal{O}_L(1/L)$ term is uniform in $\rho$ by Part 1) of Lemma 4.2, the assumption that $k_{1,\rho}$ is uniform in $\rho$, and the observation that $V_s(\rho)$ is bounded away from zero for every $\rho \geq \rho_0$ and $s \geq s_0$ (for arbitrary $\rho_0 > 0$ and $s_0 > 0$).

We next prove that, for the choice of $\tau_L$ in (6.55), and for $L$ sufficiently large, the RHS of (6.54b) is less than $\epsilon$. Consequently, $R(\tau_L, 1)$ is a lower bound on $R^*(L, \epsilon, \rho)$. To this end, we first show that

$$\begin{aligned}\epsilon^*(L, R, \rho) \leq{}& e^{L[\psi_{\rho,1}(\tau_L) - \tau_L\psi'_{\rho,1}(\tau_L)]} \\ &\times\left[f_{\rho,1}(\tau_L, \tau_L) + f_{\rho,1}(1 - \tau_L, \tau_L) + \frac{k_{2,\rho}}{\sqrt{L}} + o_L\left(\frac{1}{\sqrt{L}}\right)\right]\end{aligned} \tag{6.59}$$

where $k_{2,\rho} > 0$ is independent of $L$ and uniform in $\rho$. To obtain (6.59), we show that $\hat{K}_{\rho,1}(\tau_L)$ on the RHS of (6.54b) can be written as

$$\begin{aligned}\hat{K}_{\rho,1}(\tau_L) &= \frac{1}{\sqrt{2\pi}}\frac{\psi'''_{\rho,1}(\tau_L)}{6\psi''_{\rho,1}(\tau_L)^{3/2}} \\ &= \frac{1}{\sqrt{2\pi}}\frac{\psi'''_{\rho,1}(0)}{6\psi''_{\rho,1}(0)^{3/2}} + \mathcal{O}_L\left(\frac{1}{\sqrt{L}}\right) \\ &= k_{2,\rho} + \mathcal{O}_L\left(\frac{1}{\sqrt{L}}\right)\end{aligned} \tag{6.60}$$

where the $\mathcal{O}_L(1/\sqrt{L})$ term is uniform in $\rho$. Indeed, by using Taylor series expansions, it follows that

$$\psi''_{\rho,1}(\tau_L) = \psi''_{\rho,1}(0) + \tau_L \psi'''_{\rho,1}(0) + \frac{\tau_L^2}{2}\psi^{(4)}_{\rho,1}(\tau_2) \tag{6.61a}$$

$$\psi'''_{\rho,1}(\tau_L) = \psi'''_{\rho,1}(0) + \tau_L \psi^{(4)}_{\rho,1}(\tau_3) \tag{6.61b}$$

for some $\tau_2$, $\tau_3 \in (0, \tau_L)$. Using (6.61a) and (6.61b), it then follows that

$$\begin{aligned}\frac{\psi'''_{\rho,1}(\tau_L)}{\psi''_{\rho,1}(\tau_L)^{3/2}} &= \frac{\psi'''_{\rho,1}(0) + \tau_L \psi^{(4)}_{\rho,1}(\tau_3)}{\left(\psi''_{\rho,1}(0) + \tau_L \psi'''_{\rho,1}(0) + \frac{\tau_L^2}{2}\psi^{(4)}_{\rho,1}(\tau_2)\right)^{3/2}} \\ &= \frac{\psi'''_{\rho,1}(0)}{\left(\psi''_{\rho,1}(0) + \tau_L \psi'''_{\rho,1}(0) + \frac{\tau_L^2}{2}\psi^{(4)}_{\rho,1}(\tau_2)\right)^{3/2}} \\ &\quad + \frac{\tau_L \psi^{(4)}_{\rho,1}(\tau_3)}{\left(\psi''_{\rho,1}(0) + \tau_L \psi'''_{\rho,1}(0) + \frac{\tau_L^2}{2}\psi^{(4)}_{\rho,1}(\tau_2)\right)^{3/2}}.\end{aligned} \tag{6.62}$$

The second term on the RHS of (6.62) is $\mathcal{O}_L(1/\sqrt{L})$ uniformly in $\rho$ by Part 1) of Lemma 4.2 and because $V_s(\rho)$ is bounded away from zero for every $\rho \geq \rho_0$ and $s \geq s_0$ (for arbitrary $\rho_0 > 0$ and $s_0 > 0$). A Taylor series expansion over the first term on the RHS of (6.62) yields

$$\begin{aligned}&\frac{\psi'''_{\rho,1}(0)}{\left(\psi''_{\rho,1}(0) + \tau_L \psi'''_{\rho,1}(0) + \frac{\tau_L^2}{2}\psi^{(4)}_{\rho,1}(\tau_2)\right)^{3/2}} \\ &\qquad = \frac{\psi'''_{\rho,1}(0)}{\psi''_{\rho,1}(0)^{3/2}} - \frac{3}{2}\frac{\psi'''_{\rho,1}(0)\left(\tau_L \psi'''_{\rho,1}(0) + \frac{\tau_L^2}{2}\psi^{(4)}_{\rho,1}(\tau_2)\right)}{\left(\psi''_{\rho,1}(0) + \tilde{\tau}\psi'''_{\rho,1}(0) + \frac{\tilde{\tau}^2}{2}\psi^{(4)}_{\rho,1}(\tau_2)\right)^{5/2}}\end{aligned} \tag{6.63}$$

for some $\tilde{\tau} \in (0, \tau_L)$. The second term on the RHS of (6.63) is $\mathcal{O}_L(1/\sqrt{L})$ uniformly in $\rho$, again by Part 1) of Lemma 4.2 and because $V_s(\rho)$ is bounded away from zero for every $\rho \geq \rho_0$ and $s \geq s_0$ (for arbitrary $\rho_0 > 0$ and $s_0 > 0$). By the same arguments, the first term on the RHS of (6.63) is uniform in $\rho$. Combining (6.62) and (6.63), we thus obtain (6.60), from which (6.59) follows.

By using the definition of $f_\theta(u, \tau)$ in (6.9a), we next show that (6.59) can be written as

$$\begin{aligned}\epsilon^*(L, R, \rho) &\leq e^{L[\psi_{\rho,1}(\tau_L) - \tau_L \psi'_{\rho,1}(\tau_L)]} \\ &\quad \times \left[e^{L\frac{\tau_L^2}{2}\psi''_{\rho,1}(\tau_L)} Q\left(\tau_L\sqrt{L\psi''_{\rho,1}(\tau_L)}\right) + \frac{k_{2,\rho}}{\sqrt{L}} + \frac{k_{3,\rho}}{\sqrt{L}} + o_L\left(\frac{1}{\sqrt{L}}\right)\right]\end{aligned} \tag{6.64}$$

where $k_{3,\rho} > 0$ is independent of $L$ and uniform in $\rho$. To show this, we consider the upper bound

$$Q\Big((1-\tau_L)\sqrt{L\psi''_{\rho,1}(\tau_L)}\Big) \le \frac{1}{\sqrt{2\pi}(1-\tau_L)\sqrt{L\psi''_{\rho,1}(\tau_L)}} e^{-\frac{(1-\tau_L)^2}{2}\psi''_{\rho,1}(\tau_L)} \tag{6.65}$$

from which we obtain that, for sufficiently large $L$,

$$0 < e^{L\frac{(1-\tau_L)^2}{2}\psi''_{\rho,1}(\tau)} Q\Big((1-\tau_L)\sqrt{L\psi''_{\rho,1}(\tau_L)}\Big) \le \frac{1}{\sqrt{2\pi}(1-\tau_L)\sqrt{L\psi''_{\rho,1}(\tau_L)}} \le \frac{k_{3,\rho}}{\sqrt{L}} \tag{6.66}$$

where the right-most inequality follows because

$$\frac{1}{1-\tau_L} = \frac{1}{1-\frac{Q^{-1}\left(\epsilon-\frac{k_{1,\rho}}{\sqrt{L}}\right)}{\sqrt{L\psi''_{\rho,1}(0)}}} \le \frac{1}{1-\frac{Q^{-1}(\epsilon-k_{1,\rho})}{\sqrt{\psi''_{\rho,1}(0)}}} \tag{6.67}$$

and because $V_s(\rho)$ is bounded away from zero for every $\rho \ge \rho_0$ and $s \ge s_0$ (for arbitrary $\rho_0 > 0$ and $s_0 > 0$), which together with (6.61a) implies that also $\psi''_{\rho,1}(\tau_L)$ is bounded away from zero for sufficiently large $L$. Hence, we can find a positive constant $k_{3,\rho}$ that is independent of $L$ and uniform in $\rho$ and that satisfies (6.66) for sufficiently large $L$. It then follows that we can upper-bound $f_{\rho,1}(1-\tau_L,\tau_L)$ by $k_{3,\rho}/\sqrt{L}$.

We finally show that we can write (6.64) as

$$\begin{aligned}
\epsilon^*(L,R,\rho) \le{}& \left[1+\frac{k_{4,\rho}}{\sqrt{L}}+\mathcal{O}_L\Big(\frac{1}{L}\Big)\right]\left[\epsilon-\frac{k_{1,\rho}}{\sqrt{L}}-\frac{k_{5,\rho}}{\sqrt{L}}+\mathcal{O}_L\Big(\frac{1}{L}\Big)\right] \\
&+e^{-\frac{Q^{-1}(\epsilon)^2}{2}}\left[1+\mathcal{O}_L\Big(\frac{1}{\sqrt{L}}\Big)\right]\left[\frac{k_{2,\rho}}{\sqrt{L}}+\frac{k_{3,\rho}}{\sqrt{L}}+o_L\Big(\frac{1}{\sqrt{L}}\Big)\right] \\
={}& \epsilon-\frac{k_{1,\rho}}{\sqrt{L}}-\frac{k_{5,\rho}}{\sqrt{L}}+\frac{\epsilon k_{4,\rho}}{\sqrt{L}}+e^{-\frac{Q^{-1}(\epsilon)^2}{2}}\left[\frac{k_{2,\rho}}{\sqrt{L}}+\frac{k_{3,\rho}}{\sqrt{L}}\right]+o_L\Big(\frac{1}{\sqrt{L}}\Big)
\end{aligned} \tag{6.68}$$

where $k_{4,\rho} > 0$ and $k_{5,\rho} > 0$ are specified below and are independent of $L$ and uniform in $\rho$. Thus, if we choose $k_{1,\rho}$ larger than

$$k_{1,\rho} > \epsilon k_{4,\rho} + e^{-\frac{Q^{-1}(\epsilon)^2}{2}}[k_{2,\rho}+k_{3,\rho}] - k_{5,\rho} \tag{6.69}$$

then, for sufficiently large $L$, the RHS of (6.68) is upper-bounded by $\epsilon$ uniformly in $\rho$.

To show (6.68), we first consider the Taylor series expansions

$$\psi_{\rho,1}(\tau_L) = \frac{\tau_L^2}{2}\psi''_{\rho,1}(0) + \frac{\tau_L^3}{6}\psi'''_{\rho,1}(0) + \frac{\tau_L^4}{24}\psi^{(4)}_{\rho,1}(\tau_0) \tag{6.70a}$$

$$\psi'_{\rho,1}(\tau_L) = \tau_L\psi''_{\rho,1}(0) + \frac{\tau_L^2}{2}\psi'''_{\rho,1}(0) + \frac{\tau_L^3}{6}\psi^{(4)}_{\rho,1}(\tau_1) \tag{6.70b}$$

for $\tau_0, \tau_1 \in (0, \tau_L)$. By using (6.70) together with (6.61), and by following similar steps as above, it can be shown that

$$e^{L[\psi_{\rho,1}(\tau_L)-\tau_L\psi'_{\rho,1}(\tau_L)]} = e^{-\frac{Q^{-1}(\epsilon)^2}{2}}\left(1+\mathcal{O}_L\left(\frac{1}{\sqrt{L}}\right)\right) \tag{6.71a}$$

$$e^{L[\psi_{\rho,1}(\tau_L)-\tau_L\psi'_{\rho,1}(\tau_L)+\frac{\tau_L^2}{2}\psi''_{\rho,1}(\tau_L)]} = 1+\frac{k_{4,\rho}}{\sqrt{L}}+\mathcal{O}_L\left(\frac{1}{L}\right) \tag{6.71b}$$

$$Q\left(\tau_L\sqrt{L\psi''_{\rho,1}(\tau_L)}\right) = \epsilon-\frac{k_{1,\rho}}{\sqrt{L}}-\frac{k_{5,\rho}}{\sqrt{L}}+\mathcal{O}_L\left(\frac{1}{L}\right) \tag{6.71c}$$

where $k_{4,\rho} > 0$ and $k_{5,\rho} > 0$ are independent of $L$ and uniform in $\rho$, and both $\mathcal{O}_L(1/\sqrt{L})$ and $\mathcal{O}_L(1/L)$ are uniform in $\rho$. Substituting (6.71) into (6.64), the upper bound in (6.68) follows.

### 6.3.2 Converse Part

To show that the RHS of (6.53) is also a converse bound, we evaluate (6.52) for $s = 1$, namely,

$$\epsilon^*(L,R,\rho) \geq e^{L\left[\psi_{\rho,1}(\tau)-\tau\psi'_{\rho,1}(\tau)\right]}\left[f_{\rho,1}(\tau,\tau)+\frac{K_{\rho,1}(\tau,L)}{\sqrt{L}}+o_L\left(\frac{1}{\sqrt{L}}\right)\right] - e^{L\left[I(\rho)-\frac{\psi'_{\rho,1}(\tau)}{s}-\mathsf{T}R\right]} \tag{6.72}$$

and evaluate (6.72) for a sequence of $\tau$'s defined as

$$\tau_L \triangleq \frac{Q^{-1}(\tilde{\epsilon}_L)}{\sqrt{L\psi''_{\rho,1}(0)}}-\frac{\psi'''_{\rho,1}(0)\big(1+Q^{-1}(\tilde{\epsilon}_L)^2\big)}{3L\psi''_{\rho,1}(0)^2} \tag{6.73}$$

where

$$\tilde{\epsilon}_L \triangleq \epsilon+\frac{k_\rho}{\sqrt{L}} \tag{6.74}$$

and $k_\rho$ is a constant independent of $L$ and uniform in $\rho$, which will be specified later. Again, since $\tau_L$ vanish as $L \to \infty$, we have that $(\tau_L, \rho, 1) \in \mathcal{S}$ (with $\mathcal{S}$ characterized by (4.45)) for sufficiently large $L$. Consequently, Theorem 6.5 applies and the $o_L(1/\sqrt{L})$ term in (6.72) is uniform in $\rho$ for any $\rho \geq \rho_0$ with $\rho_0 > 0$. We next show that (6.72) can be written as

$$\begin{aligned}\epsilon^*(L,R,\rho) \geq &-e^{L\left[I(\rho)-\frac{\psi'_{\rho,1}(\tau_L)}{s}-\mathsf{T}R\right]}+e^{L\left[\psi_{\rho,1}(\tau_L)-\tau_L\psi'_{\rho,1}(\tau_L)\right]}\\ &\times\left[f_{\rho,1}(\tau_L,\tau_L)+\left(\frac{Q^{-1}(\tilde{\epsilon}_L)^2-1}{\sqrt{2\pi}}-\tilde{\epsilon}_LQ^{-1}(\tilde{\epsilon}_L)^3e^{\frac{Q^{-1}(\tilde{\epsilon}_L)}{2}}\right)\frac{\psi'''_{\rho,1}(0)}{6\sqrt{L}\psi''_{\rho,1}(0)^{3/2}}\right.\\ &\left.+o_L\left(\frac{1}{\sqrt{L}}\right)\right].\end{aligned} \tag{6.75}$$

Indeed, $K_{\rho,1}(\tau_L, L)$ is given by

$$K_{\rho,1}(\tau_L, L) \triangleq \frac{\psi'''_{\rho,1}(\tau_L)}{6\psi''_{\rho,1}(\tau_L)^{3/2}} \left( \frac{\tau_L^2 \psi''_{\rho,1}(\tau_L) L - 1}{\sqrt{2\pi}} - \tau_L^3 \psi''_{\rho,1}(\tau_L)^{3/2} L^{3/2} f_{\rho,1}(\tau_L, \tau_L) \right). \tag{6.76}$$

Using (6.62) and (6.63), and performing a Taylor series expansion of the terms inside the square brackets around zero, for our choice of $\tau_L$ in (6.73), it follows that

$$K_{\rho,1}(\tau_L, L) = \left( \frac{(Q^{-1}(\tilde{\epsilon}_L)^2 - 1)}{\sqrt{2\pi}} - \tilde{\epsilon}_L Q^{-1}(\tilde{\epsilon}_L)^3 e^{\frac{Q^{-1}(\tilde{\epsilon}_L)}{2}} \right) \frac{\psi'''_{\rho,1}(0)}{6\psi''_{\rho,1}(0)^{3/2}} + \mathcal{O}_L\left(\frac{1}{\sqrt{L}}\right) \tag{6.77}$$

where it can be shown that the $\mathcal{O}_L(1/\sqrt{L})$ term is uniform in $\rho$ by following similar steps as the ones used to analyze the error term in (6.60). Hence, substituting (6.77) into (6.72), we obtain the lower bound in (6.75).

We next show that

$$\begin{aligned}
\epsilon^*(L, R, \rho) \geq{} & \left[ 1 + \frac{Q^{-1}(\tilde{\epsilon}_L)^3 \psi'''_{\rho,1}(0)}{6\psi''_{\rho,1}(0)^{3/2}\sqrt{L}} + \mathcal{O}_L\left(\frac{1}{L}\right) \right] \\
& \times \left[ \tilde{\epsilon}_L + \frac{1}{\sqrt{2\pi}} \frac{\left(2 - Q^{-1}(\tilde{\epsilon}_L)^2\right) \psi'''_{\rho,1}(0)}{6\psi''_{\rho,1}(0)^{3/2}\sqrt{L}} e^{-\frac{Q^{-1}(\tilde{\epsilon}_L)^2}{2}} + \mathcal{O}_L\left(\frac{1}{L}\right) \right] \\
& + \left[ e^{-\frac{Q^{-1}(\tilde{\epsilon}_L)^2}{2}} \left( 1 + \mathcal{O}_L\left(\frac{1}{\sqrt{L}}\right) \right) \right] \\
& \times \left[ \left( \frac{1}{\sqrt{2\pi}} \left(Q^{-1}(\tilde{\epsilon}_L)^2 - 1\right) - \tilde{\epsilon}_L Q^{-1}(\tilde{\epsilon}_L)^3 e^{\frac{Q^{-1}(\tilde{\epsilon}_L)}{2}} \right) \frac{\psi'''_{\rho,1}(0)}{6\psi''_{\rho,1}(0)^{3/2}} \right. \\
& \left. + o_L\left(\frac{1}{\sqrt{L}}\right) \right] - e^{L\left(I(\rho) - \psi'_{\rho,1}(\tau_L) - \mathsf{T}R\right)} \\
={} & \tilde{\epsilon}_L + \frac{\tilde{\epsilon}_L Q^{-1}(\tilde{\epsilon}_L)^3 \psi'''_{\rho,1}(0)}{6\psi''_{\rho,1}(0)^{3/2}\sqrt{L}} + \frac{1}{\sqrt{2\pi}} \frac{\left(2 - Q^{-1}(\tilde{\epsilon}_L)^2\right) \psi'''_{\rho,1}(0)}{6\psi''_{\rho,1}(0)^{3/2}\sqrt{L}} e^{-\frac{Q^{-1}(\tilde{\epsilon}_L)^2}{2}} \\
& + e^{-\frac{Q^{-1}(\tilde{\epsilon}_L)^2}{2}} \left[ \frac{1}{\sqrt{2\pi}} \left(Q^{-1}(\tilde{\epsilon}_L)^2 - 1\right) - \tilde{\epsilon}_L Q^{-1}(\tilde{\epsilon}_L)^3 e^{\frac{Q^{-1}(\tilde{\epsilon}_L)}{2}} \right] \\
& \times \frac{\psi'''_{\rho,1}(0)}{6\psi''_{\rho,1}(0)^{3/2}\sqrt{L}} - e^{L\left(I(\rho) - \psi'_{\rho,1}(\tau_L) - \mathsf{T}R\right)} + o_L\left(\frac{1}{\sqrt{L}}\right) \\
={} & \tilde{\epsilon}_L + \frac{e^{-\frac{Q^{-1}(\tilde{\epsilon}_L)^2}{2}}}{\sqrt{2\pi}} \frac{\psi'''_{\rho,1}(0)}{6\psi''_{\rho,1}(0)^{3/2}\sqrt{L}} - e^{L\left(I(\rho) - \psi'_{\rho,1}(\tau_L) - \mathsf{T}R\right)} + o_L\left(\frac{1}{\sqrt{L}}\right)
\end{aligned} \tag{6.78}$$

where the first equality follows by keeping the terms up to order $1/\sqrt{L}$ and collecting the terms that vanish faster than $1/\sqrt{L}$ in the $o_L(1/\sqrt{L})$ term, and the second equality follows by simple algebra. To prove the inequality in (6.78), we use that, analogously to (6.71), for our choice of $\tau_L$ in (6.73) we have

$$e^{L[\psi_{\rho,1}(\tau_L)-\tau_L\psi'_{\rho,1}(\tau_L)]} = e^{-\frac{Q^{-1}(\tilde{\epsilon}_L)^2}{2}}\left(1+\mathcal{O}_L\left(\frac{1}{\sqrt{L}}\right)\right) \tag{6.79a}$$

$$e^{L[\psi_{\rho,1}(\tau_L)-\tau_L\psi'_{\rho,1}(\tau_L)+\frac{\tau_L^2}{2}\psi''_{\rho,1}(\tau_L)]} = 1+\frac{Q^{-1}(\tilde{\epsilon}_L)^3\psi'''_{\rho,1}(0)}{6\psi''_{\rho,1}(0)^{3/2}\sqrt{L}}+\mathcal{O}_L\left(\frac{1}{L}\right) \tag{6.79b}$$

$$\begin{aligned} Q\left(\tau_L\sqrt{L\psi''_{\rho,1}(\tau_L)}\right) &= \tilde{\epsilon}_L+\frac{1}{\sqrt{2\pi}}\frac{\left(2-Q^{-1}(\tilde{\epsilon}_L)^2\right)\psi'''_{\rho,1}(0)}{6\psi''_{\rho,1}(0)^{3/2}\sqrt{L}}e^{-\frac{Q^{-1}(\tilde{\epsilon}_L)^2}{2}} \\ &\quad +\mathcal{O}_L\left(\frac{1}{L}\right) \end{aligned} \tag{6.79c}$$

where it can be shown that the $\mathcal{O}_L(1/\sqrt{L})$ and $\mathcal{O}_L(1/L)$ terms are uniform in $\rho$ by following similar steps as the ones used to analyze the error term in (6.60). Thus, substituting the identities in (6.79) into (6.75), we obtain the inequality in (6.78).

We next show that

$$\begin{aligned} R^*(L,\rho,\epsilon) \leq{}& \frac{I(\rho)}{\mathsf{T}}-\frac{\psi'_{\rho,1}(\tau_L)}{\mathsf{T}} \\ &-\frac{1}{L\mathsf{T}}\log\left(\frac{k_\rho}{\sqrt{L}}+e^{-\frac{Q^{-1}(\epsilon)^2}{2}}\frac{1}{\sqrt{2\pi}}\frac{\psi'''_{\rho,1}(0)}{6\psi''_{\rho,1}(0)^{3/2}\sqrt{L}}+o_L\left(\frac{1}{\sqrt{L}}\right)\right). \end{aligned} \tag{6.80}$$

Indeed, a Taylor series expansion of $Q^{-1}(\tilde{\epsilon}_L)$ around $\epsilon$ yields

$$Q^{-1}\left(\epsilon+\frac{k_\rho}{\sqrt{L}}\right) = Q^{-1}(\epsilon)+\frac{k_\rho\sqrt{2\pi}}{\sqrt{L}}e^{\frac{Q^{-1}(\epsilon)^2}{2}}+\mathcal{O}_L\left(\frac{1}{L}\right) \tag{6.81}$$

where the $\mathcal{O}_L(1/L)$ term is uniform in $\rho$ because $k_\rho$ is uniform in $\rho$. Using (6.81), we can expand $e^{-\frac{Q^{-1}(\tilde{\epsilon}_L)^2}{2}}$ as

$$e^{-\frac{Q^{-1}(\tilde{\epsilon}_L)^2}{2}} = e^{-\frac{Q^{-1}(\epsilon)^2}{2}}-2Q^{-1}(\epsilon)\frac{k_\rho\sqrt{2\pi}}{\sqrt{L}}+\mathcal{O}_L\left(\frac{1}{L}\right) \tag{6.82}$$

where the $\mathcal{O}_L(1/L)$ term is uniform in $\rho$ again because $k_\rho$ is uniform in $\rho$. Hence, using (6.81) and (6.82) in (6.78), collecting the terms that vanish faster than $1/\sqrt{L}$ in the $o_L(1/\sqrt{L})$ term, and solving (6.78) for $R$, we obtain (6.80).

We finally show that

$$\begin{aligned} R^*(L,\epsilon,\rho) \leq{}& \frac{I(\rho)}{\mathsf{T}}-\sqrt{\frac{\psi''_{\rho,1}(0)}{L\mathsf{T}^2}}Q^{-1}(\epsilon)+\frac{1}{2\mathsf{T}}\frac{\log L}{L} \\ &-\frac{\sqrt{2\pi\psi''_{\rho,1}(0)}}{L\mathsf{T}}e^{\frac{Q^{-1}(\epsilon)^2}{2}}-\frac{\left(1-Q^{-1}(\epsilon)^2\right)\psi'''_{\rho,1}(0)}{6\psi''_{\rho,1}(0)L\mathsf{T}}+\mathcal{O}_L\left(\frac{1}{L^{3/2}}\right). \end{aligned} \tag{6.83}$$

Since $\psi''_{\rho,1}(0)=V(\rho)$, and since $V_s(\rho)$ is bounded away from zero for every $\rho\geq\rho_0$ and $s\geq s_0$ (for arbitrary $\rho_0>0$ and $s_0>0$), it can then be shown that (6.83) coincides with (6.53) upon collecting terms of order $\log L/L$.

To show (6.83), we expand the last term in (6.80) as

$$\log\left(\frac{k_\rho}{\sqrt{L}}+e^{-\frac{Q^{-1}(\epsilon)^2}{2}}\frac{1}{\sqrt{2\pi}}\frac{\psi'''_{\rho,1}(0)}{6\psi''_{\rho,1}(0)^{3/2}\sqrt{L}}+o_L\left(\frac{1}{\sqrt{L}}\right)\right)$$
$$=\log\left(\frac{k_\rho}{\sqrt{L}}+e^{-\frac{Q^{-1}(\epsilon)^2}{2}}\frac{1}{\sqrt{2\pi}}\frac{\psi'''_{\rho,1}(0)}{6\psi''_{\rho,1}(0)^{3/2}\sqrt{L}}\right)+o_L\left(\frac{1}{\sqrt{L}}\right). \quad (6.84)$$

(This expansion holds because the first two terms inside the logarithm on the LHS are positive.) We then choose

$$k_\rho = 1-e^{-\frac{Q^{-1}(\epsilon)^2}{2}}\frac{1}{\sqrt{2\pi}}\frac{\psi'''_{\rho,1}(0)}{6\psi''_{\rho,1}(0)^{3/2}} \quad (6.85)$$

which is independent of $L$ and uniform in $\rho$ because $V_s(\rho)$ is bounded away from zero for every $\rho \geq \rho_0$ and $s \geq s_0$ (for arbitrary $\rho_0 > 0$ and $s_0 > 0$) and $\psi'''_{\rho,1}(0)$ is bounded in $\rho$ by Part 1) of Lemma 4.2 and because $m_{\rho,s}(\tau) \geq 1$ (see also Remark 6.1). It follows that (6.80) can be written as

$$R^*(L,\rho,\epsilon) \leq \frac{I(\rho)}{\mathsf{T}}-\frac{\psi'_{\rho,1}(\tau_L)}{\mathsf{T}}+\frac{1}{2\mathsf{T}}\frac{\log L}{L}+\mathcal{O}_L\left(\frac{1}{L^{3/2}}\right). \quad (6.86)$$

By using (6.70b) with $\tau_L$ defined in (6.73), it follows that

$$\frac{\psi'_{\rho,1}(\tau_L)}{\mathsf{T}} = \sqrt{\frac{\psi''_{\rho,1}(0)}{L\mathsf{T}^2}}Q^{-1}(\tilde{\epsilon}_L)+\frac{\left(1+Q^{-1}(\tilde{\epsilon}_L)^2\right)\psi'''_{\rho,1}(0)}{3\psi''_{\rho,1}(0)L\mathsf{T}}$$
$$-\frac{Q^{-1}(\tilde{\epsilon}_L)^2\psi'''_{\rho,1}(0)}{2\psi''_{\rho,1}(0)L\mathsf{T}}+\mathcal{O}_L\left(\frac{1}{L^{3/2}}\right) \quad (6.87)$$

where it can be shown that the $\mathcal{O}_L(1/L^{3/2})$ term is uniform in $\rho$ by following similar steps as the ones used to analyze the error term in (6.60). Using (6.81) and (6.85), we can write (6.87) as

$$\begin{aligned}\frac{\psi'_{\rho,1}(\tau_L)}{\mathsf{T}} &= \sqrt{\frac{\psi''_{\rho,1}(0)}{L\mathsf{T}^2}}Q^{-1}(\epsilon)+\frac{k_\rho\sqrt{2\pi\psi''_{\rho,1}(0)}}{L\mathsf{T}}e^{\frac{Q^{-1}(\epsilon)^2}{2}}\\ &\quad+\frac{\left(2-Q^{-1}(\epsilon)^2\right)\psi'''_{\rho,1}(0)}{6\psi''_{\rho,1}(0)L\mathsf{T}}+\mathcal{O}_L\left(\frac{1}{L^{3/2}}\right)\\ &= \sqrt{\frac{\psi''_{\rho,1}(0)}{L\mathsf{T}^2}}Q^{-1}(\epsilon)+\frac{\sqrt{2\pi\psi''_{\rho,1}(0)}}{L\mathsf{T}}e^{\frac{Q^{-1}(\epsilon)^2}{2}}\\ &\quad-\frac{\psi'''_{\rho,1}(0)}{6L\mathsf{T}\psi''_{\rho,1}(0)}+\frac{\left(2-Q^{-1}(\epsilon)^2\right)\psi'''_{\rho,1}(0)}{6\psi''_{\rho,1}(0)L\mathsf{T}}+\mathcal{O}_L\left(\frac{1}{L^{3/2}}\right)\\ &= \sqrt{\frac{\psi''_{\rho,1}(0)}{L\mathsf{T}^2}}Q^{-1}(\epsilon)+\frac{\sqrt{2\pi\psi''_{\rho,1}(0)}}{L\mathsf{T}}e^{\frac{Q^{-1}(\epsilon)^2}{2}}\end{aligned}$$

$$+\frac{\left(1-Q^{-1}(\epsilon)^2\right)\psi'''_{\rho,1}(0)}{6\psi''_{\rho,1}(0)L\mathsf{T}}+\mathcal{O}_L\left(\frac{1}{L^{3/2}}\right). \tag{6.88}$$

Using (6.88) in (6.86), the upper bound (6.83) follows.

### 6.3.3 High-SNR Normal Approximation

Since the $\mathcal{O}_L(\log L/L)$ term in (6.53) is uniform in $\rho$, it is also possible to recover from (6.53) the high-SNR normal approximation presented in Theorem 5.1 (Chapter 5). To do so, we use (4.39a) and (4.39b) to write

$$\begin{aligned} I(\rho) &= (\mathsf{T}-1)\log(\mathsf{T}\rho)-\log\Gamma(\mathsf{T}) \\ &\quad -(\mathsf{T}-1)\left[\log(1+\mathsf{T}\rho)+\frac{\mathsf{T}\rho}{1+\mathsf{T}\rho}-\psi(\mathsf{T}-1)\right] \\ &\quad + {}_2F_1\left(1,\mathsf{T}-1;\mathsf{T};\frac{\mathsf{T}\rho}{1+\mathsf{T}\rho}\right)+o_\rho(1) && (6.89) \\ V(\rho) &= (\mathsf{T}-1)^2\frac{\pi^2}{6}+(\mathsf{T}-1)+o_\rho(1) && (6.90) \end{aligned}$$

where $o_\rho(1)$ comprises terms that are uniform in $L$ and vanish as $\rho\to\infty$. The desired result follows directly by substituting (6.89) and (6.90) into (6.53).

## 6.4 Error Exponent Analysis

The expansions (6.49b) and (6.52) can be written as an exponential term times a subexponential factor. As we show next, the exponential terms of both expansions coincide for rates $R^{\text{cr}}_{1/2}(\rho)<R<I(\rho)$, so they characterize the reliability function

$$E_r(R,\rho)\triangleq\lim_{L\to\infty}-\frac{1}{L}\log\epsilon^*(L,R,\rho). \tag{6.91}$$

**Theorem 6.6** *Let* $\rho_0\le\rho\le\rho_{\max}$ *and* $\underline{\tau}<\tau<\bar{\tau}$ *for some arbitrary* $0<\rho_0<\rho_{\max}<\infty$ *and* $0<\underline{\tau}<\bar{\tau}<1$. *Set* $s_\tau\triangleq 1/(1+\tau)$. *Then, the coding rate* $R$ *and the minimum error probability* $\epsilon^*$ *can be parametrized by* $\tau\in(\underline{\tau},\bar{\tau})$ *as*

$$R(\tau)=\frac{1}{\mathsf{T}}\left(I_{s_\tau}(\rho)-\psi'_{\rho,s_\tau}(\tau)\right) \tag{6.92a}$$

$$\underline{\mathsf{A}}_\rho(\tau)\le\epsilon^*(L,R,\rho)e^{-L\left[\psi_{\rho,s_\tau}(\tau)-\tau\psi'_{\rho,s_\tau}(\tau)\right]}\le\overline{\mathsf{A}}_\rho(\tau) \tag{6.92b}$$

*where*

$$\overline{\mathsf{A}}_\rho(\tau)\triangleq\frac{1}{\sqrt{2\pi L\tau^2\psi''_{\rho,s_\tau}(\tau)}}+\frac{|\hat{K}_{\rho,s_\tau}(\tau)|}{\sqrt{L}}+\frac{1}{\sqrt{2\pi L(1-\tau)^2\psi''_{\rho,s_\tau}(\tau)}}+o_L\left(\frac{1}{\sqrt{L}}\right) \tag{6.93a}$$

$$\underline{\mathsf{A}}_\rho(\tau)\triangleq\frac{1}{\tau(1+\tau)^{(1+\tau)}\left(2\pi L\psi''_{\rho,s_\tau}(\tau)\right)^{\frac{1}{2s_\tau}}}+o_L\left(\frac{1}{L^{\frac{1}{2s_\tau}}}\right). \tag{6.93b}$$

*The little-o term in* (6.93a) *is uniform in $\rho$ and $\tau$. The little-o term in* (6.93b) *is uniform in $\rho$ (for every given $\tau$).*[1]

*Proof:* The proof is divided into direct and converse parts. The direct part is given in Section 6.4.1, the converse part in Section 6.4.2. ∎

**Remark 6.4** *For $s_\tau = 1/(1+\tau)$ with $\tau \in (0,1)$, the identity* (6.92a) *characterizes all rates $R$ between the critical rate given in* (6.50) *particularized for $s = 1/2$, namely,*

$$R^{cr}_{1/2}(\rho) = \frac{1}{\mathsf{T}}\Big(I_{1/2}(\rho) - \psi'_{\rho,1/2}(1)\Big) \tag{6.94}$$

*and $I(\rho)$. Solving* (6.92a) *for $\tau$, we obtain approximations of upper and lower bounds on the minimum probability of error $\epsilon^*(L,R,\rho)$ as a function of the rate $R \in (R^{cr}_{1/2}(\rho)\,, I(\rho))$.*

The first three terms of $\overline{\mathsf{A}}_\rho(\tau)$ are positive and dominate the $o_L(1/\sqrt{L})$ term. Similarly, the first term of $\underline{\mathsf{A}}_\rho(\tau)$ is positive and of order $L^{-\frac{1+\tau}{2}}$. It thus follows from Theorem 6.6 that the reliability function $E_r(R,\rho)$ can be parametrized by $\tau \in (0,1)$ as

$$\begin{aligned} E_r(R,\rho) &= \tau\psi'_{\rho,\frac{1}{1+\tau}}(\tau) - \psi_{\rho,\frac{1}{1+\tau}}(\tau) && (6.95a)\\ R &= \frac{1}{\mathsf{T}}\left(I_{\frac{1}{1+\tau}}(\rho) - \psi'_{\rho,\frac{1}{1+\tau}}(\tau)\right). && (6.95b)\end{aligned}$$

### 6.4.1 Direct Part

We first note that $(\tau,\rho,s_\tau)$, as specified in Theorem 6.6, are in the set $\mathcal{S}$ characterized by (4.46). It thus follows from Theorem 6.4 that

$$\begin{aligned} R(\tau,s_\tau) &= \frac{1}{\mathsf{T}}(I_{s_\tau}(\rho) - \psi'_{\rho,s_\tau}(\tau)) && (6.96a)\\ \epsilon^*(L,R,\rho) &\le e^{L[\psi_{\rho,s_\tau}(\tau) - \tau\psi'_{\rho,s_\tau}(\tau)]} \\ &\quad\times\left[f_{\rho,s_\tau}(\tau,\tau) + f_{\rho,s_\tau}(1-\tau,\tau) + \frac{\hat{K}_{\rho,s_\tau}(\tau)}{\sqrt{L}} + o_L\!\left(\frac{1}{\sqrt{L}}\right)\right]. && (6.96b)\end{aligned}$$

Recall that $f_{\rho,s}(u,\tau)$ can be upper-bounded as (cf. (6.65) and (6.66))

$$f_{\rho,s}(u,\tau) \le \frac{1}{\sqrt{2\pi}u\sqrt{L\psi''_{\rho,s}(\tau)}}. \tag{6.97}$$

[1]Since $\tau$ may depend on $\rho$, the little-$o$ term in (6.93b) may depend on $\rho$ via $\tau$.

Hence, using (6.97), and upper-bounding $\hat{K}_{\rho,s_\tau}(\tau)/\sqrt{L}$ by its absolute value, we can upper-bound the RHS of (6.96b) as

$$\epsilon^*(L,R,\rho) \le e^{L\left[\psi_{\rho,s_\tau}(\tau)-\tau\psi'_{\rho,s_\tau}(\tau)\right]}\left[\frac{1}{\sqrt{2\pi\tau^2 L\psi''_{\rho,s_\tau}(\tau)}}+\frac{1}{\sqrt{2\pi(1-\tau)^2 L\psi''_{\rho,s_\tau}(\tau)}}\right.$$
$$\left.+\left|\frac{\hat{K}_{\rho,s_\tau}(\tau)}{\sqrt{L}}\right|+o_L\left(\frac{1}{\sqrt{L}}\right)\right]. \quad (6.98)$$

We thus obtain the right-most inequality in (6.92b) upon choosing $\tau$ to satisfy (6.92a).

### 6.4.2 Converse Part

Again, $(\tau,\rho,s_\tau)$ as specified in Theorem 6.6 are in the set $\mathcal{S}$ characterized by (4.46). It thus follows from Theorem 6.5 that

$$\epsilon^*(L,R,\rho) \ge e^{L\left[\psi_{\rho,s_\tau}(\tau)-\tau\psi'_{\rho,s_\tau}(\tau)\right]}\left[f_{\rho,s_\tau}(\tau,\tau)+\frac{K_{\rho,s_\tau}(\tau,L)}{\sqrt{L}}+o_L\left(\frac{1}{\sqrt{L}}\right)\right]$$
$$-e^{L\left[J_{s_\tau}(\rho)-\frac{\psi'_{\rho,s_\tau}(\tau)}{s_\tau}-\mathsf{T}R\right]}. \quad (6.99)$$

In Appendix B.7, we show that the constant $K_{\rho,s}(\tau,L)$ defined in (6.52) is of order $\mathcal{O}_L(1/L)$ uniformly in $(\tau,s,\rho)$. Consequently, (6.99) can be written as

$$\epsilon^*(L,R,\rho) \ge e^{L\left[\psi_{\rho,s_\tau}(\tau)-\tau\psi'_{\rho,s_\tau}(\tau)\right]}\left[f_{\rho,s_\tau}(\tau,\tau)+o_L\left(\frac{1}{\sqrt{L}}\right)\right]$$
$$-e^{L\left[J_{s_\tau}(\rho)-\frac{1}{s_\tau}\psi'_{\rho,s_\tau}(\tau)-\mathsf{T}R\right]}. \quad (6.100)$$

In principle, we would like to choose $\tau$ such that the two exponents in (6.100) are equal. This can be achieved by the $\tau$ satisfying (6.92a). Indeed, recall that, by (4.26),

$$J_{s_\tau}(\rho) = \log\mu(s_\tau)+(1+\tau)I_{s_\tau}(\rho). \quad (6.101)$$

It follows that

$$\log\mu(s_\tau) = \log\int \mathsf{E}\Big[\mathrm{p}_{\mathbf{Y}_\ell|\mathbf{X}_\ell}(\mathbf{Y}_\ell|\mathbf{X}_\ell)^{s_\tau}\Big]^{\frac{1}{s_\tau}}dy$$
$$= \psi_{\rho,s_\tau}(\tau)-\tau I_{s_\tau}(\rho) \quad (6.102)$$

which together with (6.92a) yields that

$$J_{s_\tau}(\rho)-(1+\tau)\psi'_{\rho,s_\tau}(\tau)-\mathsf{T}R = \psi_{\rho,s_\tau}(\tau)-\tau\psi'_{\rho,s_\tau}(\tau). \quad (6.103)$$

While this choice of $\tau$ yields the correct exponent, alas, it yields a negative subexponential factor. Indeed, it can be checked that, for such $\tau$, $f_{\rho,s_\tau}(\tau,\tau)-1$ becomes negative for sufficiently large $L$. Fortunately, we can sidestep this problem by choosing $\tau$ as a function of $L$.

Before we describe our choice of $\tau$, we first need to introduce some notation. The Gallager $E_0$-function [19, Eq. (5.6.14)] is defined as

$$E_{0,\rho}(\tau, s) \triangleq -\log \mathsf{E}\left[e^{-\tau i_{\ell,s}(\rho)}\right]. \tag{6.104}$$

Some simple algebra shows that

$$\psi_{\rho,s}(\tau) = \tau I_s(\rho) - E_{0,\rho}(\tau, s) \tag{6.105a}$$

$$\psi'_{\rho,s}(\tau) = I_s(\rho) - E'_{0,\rho}(\tau, s) \tag{6.105b}$$

where $E'_{0,\rho}(\tau, s)$ denotes the first derivative of $E_{0,\rho}(\tau, s)$ with respect to $\tau$. We next define

$$\Psi_{\rho,s}(\tau) \triangleq \tau^2 \psi''_{\rho,s}(\tau). \tag{6.106}$$

Note that, by Part 2) of Lemma 4.2, we have for every $0 < \rho_0 < \rho_{\max}$, $0 < b < 1$, and $0 < s_0 < s_{\max}$

$$\sup_{\rho_0 \le \rho \le \rho_{\max}} \sup_{\substack{|\tau|<b,\\ s\in(s_0,s_{\max}]}} \left|\Psi^{(k)}_{\rho,s}(\tau)\right| < \infty, \quad k \in \mathbb{Z}_0^+. \tag{6.107}$$

We further have

$$\inf_{\rho \ge \rho_0} \inf_{\substack{\underline{\tau}<\tau<\bar{\tau},\\ s\in(s_0,s_{\max}]}} \left|\Psi_{\rho,s}(\tau)\right| > 0. \tag{6.108}$$

Indeed, using Lemma B.3 together with the observation that $\psi''_{\rho,s}(0) = V_s(\rho)$ is bounded away from zero for every $\rho \ge \rho_0$ and $s \ge s_0$ (for arbitrary $\rho_0 > 0$ and $s_0 > 0$), it follows that

$$\inf_{\rho \ge \rho_0} \inf_{\substack{\underline{\tau}<\tau<\bar{\tau},\\ s\in(s_0,s_{\max}]}} \left|\psi''_{\rho,s}(\tau)\right| > 0. \tag{6.109}$$

Furthermore, for $\tau > \underline{\tau} > 0$, (6.108) and (6.109) are equivalent, so the claim follows.

To shorten notation, in the following we shall write $\Psi_\rho(\tau)$ for $\Psi_{\rho,s_\tau}(\tau)$ and $E_{0,\rho}(\tau)$ for $E_{0,\rho}(\tau, s_\tau)$. We denote the first, second and third derivatives of $\tau \mapsto E_{0,\rho}(\tau)$ with respect to $\tau$ by $E'_{0,\rho}(\tau)$, $E''_{0,\rho}(\tau)$, $E'''_{0,\rho}(\tau)$, respectively. In general, we denote

$$E^{(k)}_{0,\rho}(\tau) = \frac{\partial^k E_{0,\rho}(\tau, \frac{1}{1+\tau})}{\partial \tau^k}, \quad k = 1, 2, \ldots \tag{6.110}$$

While for $k = 1$ it can be shown that $E'_{0,\rho}(\tau) = E'_{0,\rho}(\tau, s)\,|_{s=\frac{1}{1+\tau}}$ (where $\tau \mapsto E'_{0,\rho}(\tau, s)$ denotes the derivative of $\tau \mapsto E_{0,\rho}(\tau, s)$ with respect to $\tau$ when $s$ is held fixed), for higher-order derivatives this is no longer true, i.e.,

$$E^{(k)}_{0,\rho}(\tau) \neq \left.\frac{\partial^k E_{0,\rho}(\tau, s)}{\partial \tau^k}\right|_{s=\frac{1}{1+\tau}}, \quad k = 2, 3, \ldots \tag{6.111}$$

Let the sequence $\{\tau_L\}$ be given by

$$\tau_L = \tau + \frac{\log\Big(A\sqrt{2\pi L \Psi_\rho(\tau)}\Big)}{-LE''_{0,\rho}(\tau)} \tag{6.112}$$

where $A > 0$ is a free parameter that will be optimized later and $\tau$ satisfies (6.92a). Observe that $\tau_L \to \tau$ as $L \to \infty$.

Setting $s_L = 1/(1+\tau_L)$ in (6.100), and analyzing the resulting expression as $L \to \infty$, we will obtain not only the correct exponential behavior, but we will also obtain a positive subexponential term. To this end, we first evaluate (6.100) with $\tau$ replaced by $\tau_L$ to obtain

$$\begin{aligned}\epsilon^*(L,R,\rho) &\geq e^{-L\left[E_{0,\rho}(\tau_L)-\tau_L E'_{0,\rho}(\tau_L)\right]}\\ &\quad\times\left[f_{\rho,\frac{1}{1+\tau_L}}(\tau_L,\tau_L) + o_L\left(\frac{1}{\sqrt{L}}\right) - e^{-L\left[\mathsf{T}R - E'_{0,\rho}(\tau_L)\right]}\right]\end{aligned} \tag{6.113}$$

where we have used (6.101), (6.102), and (6.105b) together with the observation that $E'_{0,\rho}(\tau) = E'_{0,\rho}(\tau,s)\,|_{s=\frac{1}{1+\tau}}$. We next show that

$$\begin{aligned}\epsilon^*(L,R,\rho) &\geq e^{-L\left[E_{0,\rho}(\tau_L)-\tau_L E'_{0,\rho}(\tau_L)\right]}\\ &\quad\times\left[\frac{1}{\sqrt{2\pi L\Psi_\rho(\tau)}} + o_L\left(\frac{1}{\sqrt{L}}\right) - e^{-L\left[\mathsf{T}R - E'_{0,\rho}(\tau_L)\right]}\right].\end{aligned} \tag{6.114}$$

Indeed, by using the bound $Q(x) \geq (\sqrt{2\pi}x)^{-1}(1-x^{-2})$, $x > 0$, we obtain that

$$\begin{aligned}&f_{\rho,\frac{1}{1+\tau_L}}(\tau_L,\tau_L)\\ &\quad\geq \frac{1}{\sqrt{2\pi L\Psi_\rho(\tau_L)}}\left(1-\frac{1}{L\Psi_\rho(\tau_L)}\right)\\ &\quad= \frac{1}{\sqrt{2\pi L\big(\Psi_\rho(\tau)+(\tau_L-\tau)\Psi'_\rho(\hat{\tau}_0)\big)}} - \frac{1}{\sqrt{2\pi}\Big(L\big(\Psi_\rho(\tau)+(\tau_L-\tau)\Psi'_\rho(\hat{\tau}_0)\big)\Big)^{3/2}}\\ &\quad= \frac{1}{\sqrt{2\pi L\Psi_\rho(\tau)}} - \frac{1}{2\sqrt{2\pi L}}\frac{(\tau_L-\tau)\Psi'_\rho(\hat{\tau}_0)}{\big(\Psi_\rho(\tau)+(\tilde{\tau}-\tau)\Psi'_\rho(\hat{\tau}_0)\big)^{3/2}}\\ &\qquad - \frac{1}{\sqrt{2\pi}\Big(L\big(\Psi_\rho(\tau)+(\tau_L-\tau)\Psi'_\rho(\hat{\tau}_0)\big)\Big)^{3/2}}\end{aligned} \tag{6.115}$$

where the second step follows by performing a Taylor series expansion of $\tau_L \mapsto \Psi_\rho(\tau_L)$ around $\tau$, and the third step follows by performing a Taylor series expansion over the first term. We next show that

$$f_{\rho,\frac{1}{1+\tau_L}}(\tau_L,\tau_L) \geq \frac{1}{\sqrt{2\pi L\Psi_\rho(\tau)}} + \mathcal{O}_L\left(\frac{\log L}{L^{3/2}}\right) \tag{6.116}$$

for some $\hat{\tau}_0 \in (\tau, \tau_L)$, where the $\mathcal{O}_L(\log L/L^{3/2})$ term is uniform in $\rho$. Indeed, by (6.107) and Lemma B.4 (Appendix B.8), the difference $\tau_L - \tau$ is of order $\log L/L$

(uniformly in $\rho$). Furthermore, by (6.107) and (6.108), we have that $\tau \mapsto \Psi_\rho(\tau)$ is bounded away from zero and bounded (in $\rho$), and $\tau \mapsto \Psi'_\rho(\tau)$ is bounded in $\rho$. It follows that the second term on the RHS of (6.115) is of order $\log L/L^{3/2}$, and the third term is of order $1/L^{3/2}$. We thus obtain (6.114) by using (6.116) in (6.113) and combining the $\mathcal{O}_L(\log L/L^{3/2})$ term with the $o_L(1/\sqrt{L})$ term.

We finally show that (6.114) can be written as

$$\begin{aligned}
&\epsilon^*(L,R,\rho)\\
&\quad\geq e^{-L\left[E_{0,\rho}(\tau)-\tau E'_{0,\rho}(\tau)\right]}\left[\frac{1}{A^\tau(2\pi L\Psi_\rho(\tau))^{\tau/2}}+\mathcal{O}_L\left(\frac{\log^2 L}{L^{(1+\tau/2)}}\right)\right]\\
&\qquad\qquad\times\left[\frac{1}{\sqrt{2\pi L\Psi_\rho(\tau)}}-\frac{1}{A\sqrt{2\pi L\Psi_\rho(\tau)}}+o_L\left(\frac{1}{\sqrt{L}}\right)\right]\\
&\quad= e^{-L\left[E_{0,\rho}(\tau)-\tau E'_{0,\rho}(\tau)\right]}\left[\frac{1}{(2\pi L\Psi_\rho(\tau))^{\frac{1+\tau}{2}}}\left(\frac{1}{A^\tau}-\frac{1}{A^{1+\tau}}\right)+o_L\left(\frac{1}{L^{\frac{1+\tau}{2}}}\right)\right] \qquad (6.117)
\end{aligned}$$

where $A$ was introduced in (6.112). By following along similar lines as the ones used to show (6.116), it can be shown that the big-$\mathcal{O}$ term and the small-$o$ terms are uniform in $\rho$. The value of $A$ that yields the tightest lower bound on $\epsilon^*(L,R,\rho)$ corresponds to the maximizing argument of the function

$$f_\tau(A)=\frac{1}{A^\tau}-\frac{1}{A^{1+\tau}} \qquad (6.118)$$

which is given by $A_*=(1+\tau)/\tau$. Using this value in (6.117), and applying (6.105), we obtain the left-most inequality in (6.92b).

To show (6.117), we start by performing a Taylor series expansion of $\tau_L \mapsto E_{0,\rho}(\tau_L)$ and $\tau_L \mapsto E'_{0,\rho}(\tau_L)$ around $\tau$ to obtain

$$E_{0,\rho}(\tau_L)=E_{0,\rho}(\tau)+(\tau_L-\tau)E'_{0,\rho}(\tau)+\frac{(\tau_L-\tau)^2}{2}E''_{0,\rho}(\hat{\tau}_1) \qquad (6.119a)$$

$$E'_{0,\rho}(\tau_L)=E'_{0,\rho}(\tau)+(\tau_L-\tau)E''_{0,\rho}(\tau)+\frac{(\tau_L-\tau)^2}{2}E'''_{0,\rho}(\hat{\tau}_2) \qquad (6.119b)$$

for some $\hat{\tau}_1,\hat{\tau}_2\in(\tau,\tau_L)$.

By (6.92a) and (6.105b), we have that $\mathsf{T}R=E'_{0,\rho}(\tau)$. Consequently,

$$\begin{aligned}
e^{-L\left[\mathsf{T}R-E'_{0,\rho}(\tau_L)\right]}&=e^{L\left[(\tau_L-\tau)E''_{0,\rho}(\tau)+\frac{(\tau_L-\tau)^2}{2}E'''_{0,\rho}(\hat{\tau}_1)\right]}\\
&=e^{L\left[\frac{\log\left(A\sqrt{2\pi L\Psi_\rho(\tau)}\right)}{-LE''_{0,\rho}(\tau)}E''_{0,\rho}(\tau)+\mathcal{O}_L\left(\frac{(\log L)^2}{L^2}\right)\right]}\\
&=\frac{1}{A\sqrt{2\pi L\Psi_\rho(\tau)}}e^{\mathcal{O}_L\left(\frac{(\log L)^2}{L}\right)}\\
&=\frac{1}{A\sqrt{2\pi L\Psi_\rho(\tau)}}+\mathcal{O}_L\left(\frac{\log^2 L}{L^{3/2}}\right) \qquad (6.120)
\end{aligned}$$

where the first step follows by (6.119b); the second step follows by (6.112), (6.107), and by Lemma B.4 (Appendix B.8); and the fourth step follows by using that

$$e^{\mathcal{O}_L\left(\frac{\log^2 L}{L}\right)} = 1 + \mathcal{O}_L\left(\frac{\log^2 L}{L}\right). \tag{6.121}$$

By following along similar lines as the ones used to show (6.116), it can be shown that the $\mathcal{O}(\cdot)$ terms in (6.120) are uniform in $\rho$.

We next use (6.119) to write

$$\begin{aligned} E_{0,\rho}(\tau_L) - \tau_L E'_{0,\rho}(\tau_L) &= E_{0,\rho}(\tau) - \tau E'_{0,\rho}(\tau) - \tau_L(\tau_L - \tau)E''_{0,\rho}(\tau) \\ &\quad + \frac{(\tau_L - \tau)^2}{2}\left(E''_{0,\rho}(\hat{\tau}_0) - \tau_L E'''_{0,\rho}(\hat{\tau}_1)\right) \\ &= E_{0,\rho}(\tau) - \tau E'_{0,\rho}(\tau) - \tau(\tau_L - \tau)E''_{0,\rho}(\tau) \\ &\quad + \frac{(\tau_L - \tau)^2}{2}\left(E''_{0,\rho}(\hat{\tau}_0) - \tau_L E'''_{0,\rho}(\hat{\tau}_1) - 2E''_{0,\rho}(\tau)\right). \end{aligned} \tag{6.122}$$

By substituting (6.112) in (6.122), we obtain

$$E_{0,\rho}(\tau_L) - \tau_L E'_{0,\rho}(\tau_L) = E_{0,\rho}(\tau) - \tau E'_{0,\rho}(\tau) + \frac{\tau}{L}\log\left(A\sqrt{2\pi L \Psi_\rho(\tau)}\right) + \mathcal{O}_L\left(\frac{\log^2 L}{L^2}\right) \tag{6.123}$$

where the $\mathcal{O}_L\left(\log^2 L/L^2\right)$ term is uniform in $\rho$ by (6.107) and Lemma B.4 (Appendix B.8).

Using (6.121) and (6.123), we conclude that

$$\begin{aligned} e^{-L\left[E_{0,\rho}(\tau_L) - \tau_L E'_{0,\rho}(\tau_L)\right]} &= e^{-L\left[E_{0,\rho}(\tau) - \tau E'_{0,\rho}(\tau)\right]} \\ &\quad \times \left(\frac{1}{A^\tau \left(2\pi L \Psi_\rho(\tau)\right)^{\tau/2}} + \mathcal{O}_L\left(\frac{\log^2 L}{L^{(1+\tau/2)}}\right)\right) \end{aligned} \tag{6.124}$$

where the big-$\mathcal{O}$ term is uniform in $\rho$. Finally, substituting (6.120) and (6.124) into (6.114), we obtain (6.117), which was the last step required to prove the left-most inequality in (6.92b).

## 6.5 High-SNR Approximations

The approximations presented in Theorems 6.4 and 6.5 are functions of the CGF $\psi_{\rho,s}(\tau)$ and its derivatives, which typically need to be evaluated numerically. However, if $(\tau, \rho, s_\tau)$ are in the set $\mathcal{S}$ characterized by (4.45) (Lemma 4.2) then, at high SNR, these functions can be approximated accurately. Let

$$\begin{aligned}\bar{\psi}_{\rho,s}(\tau) &\triangleq \tau\left(-\frac{s\mathsf{T}\rho}{1+\mathsf{T}\rho}\mathsf{E}[Z_2] + (\mathsf{T}-1)\mathsf{E}\left[\log\left(Z_1 + \frac{Z_2}{1+\mathsf{T}\rho}\right)\right]\right)\\ &\qquad + \log\left(\mathsf{E}\left[\exp\left\{-\tau\left(-\frac{s\mathsf{T}\rho Z_2}{1+\mathsf{T}\rho} + (\mathsf{T}-1)\log\left(Z_1 + \frac{Z_2}{1+\mathsf{T}\rho}\right)\right)\right\}\right]\right)\\ &= \tau\left(-(\mathsf{T}-1)\left[\log(1+\mathsf{T}\rho) + \frac{s\mathsf{T}\rho}{1+\mathsf{T}\rho} - \psi(\mathsf{T}-1)\right]\right.\\ &\qquad\qquad \left. + {}_2F_1\left(1, \mathsf{T}-1; \mathsf{T}; \frac{\mathsf{T}\rho}{1+\mathsf{T}\rho}\right)\right)\\ &\qquad + \log\left(\frac{\eta_\rho^{\nu(\tau)}\Gamma(\mathsf{T}-1+\nu(\tau))}{\Gamma(\mathsf{T})(\eta_\rho+\lambda_\rho(\tau))^{\mathsf{T}-1+\nu(\tau)}}\,{}_2F_1\left(1, \mathsf{T}-1+\nu(\tau); \mathsf{T}; \frac{\lambda_\rho(\tau)}{\eta_\rho+\lambda_\rho(\tau)}\right)\right)\end{aligned} \tag{6.125}$$

where $\eta_\rho = \frac{1}{1+\mathsf{T}\rho}$, $\nu(\tau) = -\tau(\mathsf{T}-1)+1$, and $\lambda_\rho(\tau) = \frac{\mathsf{T}\rho}{1+\mathsf{T}\rho}(1-\tau s)$. The second expected value has been solved using [27, Sec. 4.337-1] to integrate with respect to $Z_1$, and [27, Sec. 4.352-1], [27, Sec. 3.381-4], and [27, Sec. 6.228-2] to integrate with respect to $Z_2$. The third expected value has been solved using [27, Sec. 3.381-3.[8]] to integrate with respect to $Z_1$, and [27, Sec. 6.455-1] to integrate with respect to $Z_2$. Note that the third term on the RHS of (6.125) is unbounded in $\rho$ if $\tau \geq 1/(T-1)$.

**Lemma 6.7** *Assume that $(\tau, \rho, s)$ are in the set $\mathcal{S}$ characterized by (4.45) (Lemma 4.2), i.e., $-1 \leq \tau \leq a$, $\rho \geq \rho_0$, and $s_0 \leq s \leq 1$, for some arbitrary $\rho_0 > 0$, $s_0 > 0$, and $0 < a < 1/(\mathsf{T}-1)$ independent of $(L, \rho, s, \tau)$. Then, the CGF $\psi_{\rho,s}(\tau)$ given in (4.42), and its respective first, second and third derivatives, can be approximated as*

$$\psi_{\rho,s}(\tau) = \bar{\psi}_{\rho,s}(\tau) + o_\rho(1) \tag{6.126a}$$

$$\psi'_{\rho,s}(\tau) = \bar{\psi}'_{\rho,s}(\tau) + o_\rho(1) \tag{6.126b}$$

$$\psi''_{\rho,s}(\tau) = \bar{\psi}''_{\rho,s}(\tau) + o_\rho(1) \tag{6.126c}$$

$$\psi'''_{\rho,s}(\tau) = \bar{\psi}'''_{\rho,s}(\tau) + o_\rho(1) \tag{6.126d}$$

*where $\bar{\psi}'_{\rho,s}$, $\bar{\psi}''_{\rho,s}$ and $\bar{\psi}'''_{\rho,s}$ denote the first, second and third derivatives of $\tau \mapsto \bar{\psi}_{\rho,s}(\tau)$, respectively, and $o_\rho(1)$ collects terms that vanish as $\rho \to \infty$ and are uniform in $L$, $\tau$ and $s$.*

*Proof:* See Appendix B.6. ∎

By inserting $\bar{\psi}_{\rho,s}(\tau)$ and its corresponding derivatives into (6.49b) and (6.52), we obtain high-SNR saddlepoint approximations that can be evaluated in closed form.

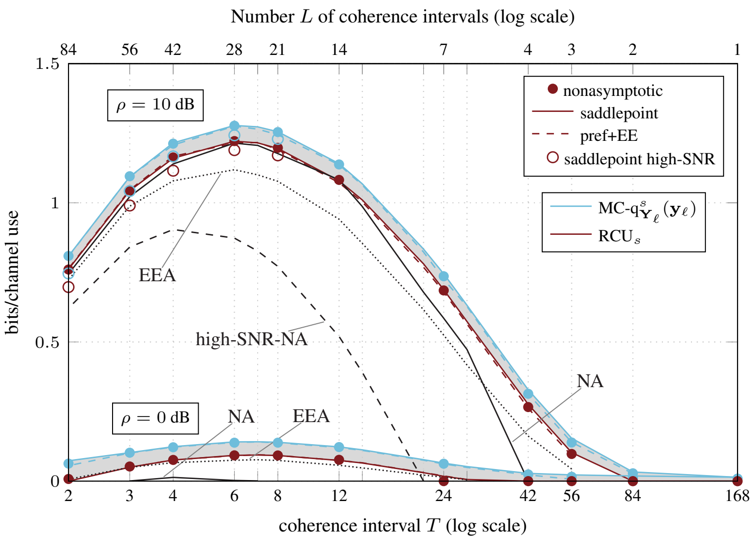


Figure 6.1: Bounds on $R^*(L, \epsilon, \rho)$ for $n = L\mathsf{T} = 168$, $\epsilon = 10^{-5}$, and SNR values $\rho = 0$ dB and $\rho = 10$ dB.

## 6.6 Numerical Results and Discussion

In Fig. 6.1, we study $R^*(L, \epsilon, \rho)$ as a function of $L$ for $n = L\mathsf{T} = 168$ (hence $\mathsf{T}$ is inversely proportional to $L$), $\epsilon = 10^{-5}$, and the SNR values 0 dB and 10 dB. We plot approximations of the RCU$_\mathrm{s}$ bound in red and approximations of the MC bound in blue by disregarding the $o_L(1/\sqrt{L})$ terms. Straight lines ("saddlepoint") depict the saddlepoint approximations (6.49b) and (6.52), dashed lines ("pref+EE") depict (6.92b). We further plot the nonasymptotic bounds (6.48) and (6.51) with dots. For 10 dB we also plot with circles the high-SNR versions of (6.49b) and (6.52) that are obtained by replacing $\psi_{\rho,s}(\cdot)$ and its derivatives by their high-SNR approximations presented in Section 6.5. Note that these approximations require that $0 \leq \tau < 1/(\mathsf{T} - 1)$, cf. Lemma 6.7. It turns out that, for such $\tau$ values, accurate high-SNR saddlepoint approximations can only be computed up to $\mathsf{T} = 8$, since for larger values of $\mathsf{T}$ the range of $\tau$ becomes too restricted. Finally, we plot the normal approximation (6.53) ("NA"), the high-SNR normal approximation given in Theorem 5.1.1 ("high-SNR-NA"), as well as the error exponent approximation that follows by solving $\epsilon^*(L, R, \rho) \approx \exp\{-LE_r(R, \rho)\}$ for $R$ ("EEA"). Observe that the approximations (6.49b), (6.52), and (6.92b) are almost indistinguishable from the nonasymptotic bounds. Further observe that the normal approximation "NA" is accurate for 10 dB and $L > 10$, but is loose for 0 dB. In contrast, the error exponent

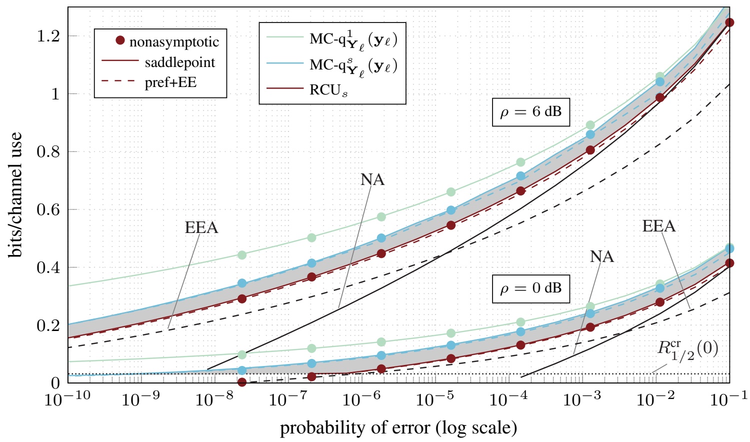


Figure 6.2: Bounds on $R^*(L, \epsilon, \rho)$ for $n = 168$, $\mathsf{T} = 12$, and SNR values $\rho = 0$ dB and $\rho = 6$ dB.

approximation "EEA" is loose for 10 dB, but is accurate for 0 dB. The high-SNR saddlepoint approximations are pessimistic, but are remarkably accurate for an SNR value as small as 10 dB.

In Fig. 6.2, we study $R^*(L, \epsilon, \rho)$ as a function of $\epsilon$ for $n = 168$, $\mathsf{T} = 12$, and the SNR values 6 dB and 0 dB. We plot approximations of the RCU$_\mathrm{s}$ bound in red and approximations of the MC bound in green (for $s = 1$) or in blue (when $s$ is numerically optimized). Straight lines ("saddlepoint") depict the saddlepoint approximations (6.49b) and (6.52), dashed lines ("pref+EE") show the approximations (6.92b). We further plot the nonasymptotic bounds (6.48) and (6.51) with dots. For $\rho = 0$ dB, we also show the critical rate $R^{\mathrm{cr}}_{1/2}(0)$. We finally plot the normal approximation (6.53) ("NA") and the error exponent approximation that follows by solving $\epsilon^*(L, R, \rho) \approx \exp\{-LE_r(R, \rho)\}$ for $R$ ("EEA"). Observe that the approximations (6.49b), (6.52), and (6.92b) are almost indistinguishable from the nonasymptotic bounds. Further observe how the normal approximation "NA" becomes accurate for large error probabilities, whereas the error exponent approximation "EEA" becomes accurate for small error probabilities. Finally note that the saddlepoint approximations can be evaluated for error probabilities less than $10^{-8}$, where the nonasymptotic bounds cannot be evaluated due to their computational complexity.

In Fig. 6.3, we study $R^*(L, \epsilon, \rho)$ as a function of the SNR $\rho$ for $n = 168$ ($\mathsf{T} = 12$, and $L = 14$) and $\epsilon = 10^{-5}$. We plot approximations of the RCU$_\mathrm{s}$ bound in red and approximations of the MC bound in blue (with $s$ numerically optimized). Straight lines ("saddlepoint") depict the saddlepoint approximations (6.49b) and

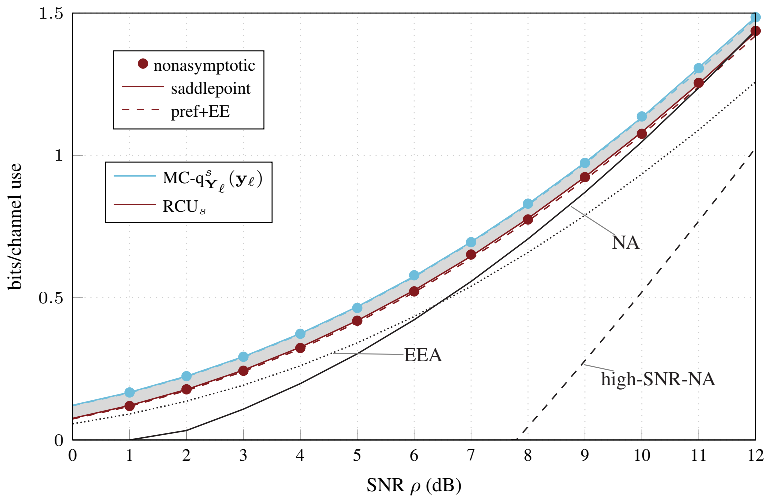


Figure 6.3: Bounds on $R^*(L, \epsilon, \rho)$ for $n = 168$, $\mathsf{T} = 12$, $L = 14$ and a probability of error of $10^{-5}$.

(6.52), dashed lines ("pref+EE") show the approximations (6.92b). We further plot the nonasymptotic bounds (6.48) and (6.51) with dots. We finally plot the normal approximation (6.53) ("NA") and the error exponent approximation that follows by solving $\epsilon^*(L, R, \rho) \approx \exp\{-LE_r(R, \rho)\}$ for $R$ ("EEA"). Observe that the approximations (6.49b), (6.52), and (6.92b) are almost indistinguishable from the nonasymptotic bounds. Observe also how the normal approximation "NA" becomes accurate as we increase the SNR. Note that the error exponent approximation "EEA" is accurate only for small SNR values.

In Figs. 6.4 and 6.5, we study $\epsilon^*(L, R, \rho)$ as a function of the SNR $\rho$. Specifically, in Fig. 6.4 we show $\epsilon^*(L, R, \rho)$ for $n = 168$ ($\mathsf{T} = 24$, and $L = 7$) and $R = 0.48$, and in Fig. 6.5 we show $\epsilon^*(L, R, \rho)$ for $n = 500$ ($\mathsf{T} = 20$, and $L = 25$) and $R = 4$. We plot approximations of the RCU$_\mathrm{s}$ bound in red and approximations of the MC bound in blue (with $s$ numerically optimized). Straight lines ("saddlepoint") depict the saddlepoint approximations (6.49b) and (6.52), dashed lines ("pref+EE") show the approximations (6.92b). We further plot the nonasymptotic bounds (6.48) and (6.51) with dots. We plot the normal approximation (6.53) ("NA") and the error exponent approximation that follows by solving $\epsilon^*(L, R, \rho) \approx \exp\{-LE_r(R, \rho)\}$ for $R$ ("EEA"). In Fig. 6.5, we further plot with circles the high-SNR versions of (6.49b) and (6.52) that are obtained by replacing $\psi_{\rho,s}(\cdot)$ and its derivatives by their high-SNR approximations presented in Section 6.5. In addition, we show the

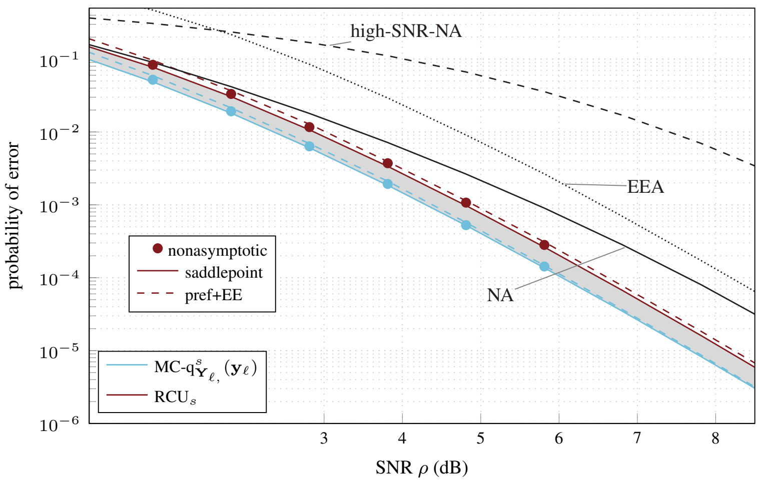


Figure 6.4: Bounds on $R^*(L, \epsilon, \rho)$ for $n = 168$, $\mathsf{T} = 24$, $L = 7$ and $R = 0.48$.

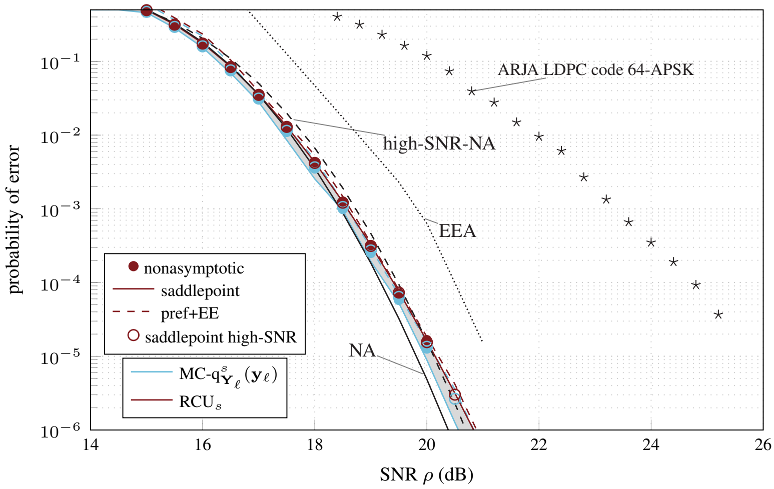


Figure 6.5: Bounds on $R^*(L, \epsilon, \rho)$ for $n = 500$, $\mathsf{T} = 20$, $L = 25$ and $R = 4$.

performance of an accumulate-repeat-jagged-accumulate (ARJA) low density parity check (LDPC) (3000,2000)-code combined with 64-APSK modulation, pilot-assisted transmission (2 pilot symbols per coherence block), and maximum likelihood channel estimation followed by mismatched nearest-neighbor decoding at the receiver ("ARJA LDPC code 64-APSK") [35, Figure 3(b)] (see [37]). Observe that the approximations

(6.49b), (6.52), (6.92b), and the high-SNR approximations of (6.49b) and (6.52) are almost indistinguishable from the nonasymptotic bounds. Observe also how, for this setting, the normal approximation "NA" is much more accurate as the error exponent approximation "EEA". Finally, the gap between the presented real code and the rest of the curves is substantial. This suggests that more sophisticated joint channel-estimation decoding procedures together with shaping techniques need to be adopted to close the gap; see e.g., [37].

# 7

# Summary and Conclusion

In this thesis, we studied the maximum coding rate at which data can be transmitted over noncoherent, single-antenna, Rayleigh block-fading channels using error-correcting codes that span $L$ coherence intervals and have a block-error probability no larger than $\epsilon$. For this specific channel model, only nonasymptotic bounds were available in the literature. We presented different asymptotic expansions of the nonasymptotic bounds leading to expressions that are analytically tractable, and whose computational cost is considerably reduced with respect to the nonasymptotic bounds.

Specifically, in **Chapter 5**, we presented a high-SNR normal approximation of the maximum coding rate $R^*(L, \epsilon, \rho)$ which can be evaluated in closed form. While we demonstrated that the approximation error vanishes as the number of coherence intervals and the SNR tend to infinity, by means of numerical examples we showed that it is accurate already at SNR values of 15 dB, for 10 coherence intervals or more, and probabilities of error no smaller than $10^{-3}$. The obtained normal approximation complements the nonasymptotic bounds provided in Chapter 3, and it allows for a mathematical analysis of $R^*(L, \epsilon, \rho)$. For example, we showed that the high-SNR normal approximation is roughly equal to the normal approximation one obtains by transmitting one pilot symbol per coherence block to estimate the fading coefficient, and by then transmitting $\mathsf{T} - 1$ symbols per coherence block over a coherent fading channel. This suggests that, at high SNR, one pilot symbol per coherence block suffices to achieve both the capacity and the channel dispersion. We finally showed an example where the normal approximation can be used to analyze a simple slotted-ALOHA protocol.

In **Chapter 6**, we applied the saddlepoint method to derive approximations of the nonasymptotic MC and the $\mathrm{RCU_s}$ bounds on the maximum coding rate $R^*(L, \epsilon, \rho)$ (or minimum probability of error $\epsilon^*(L, R, \rho)$) provided in Chapter 3. While these approximations must be evaluated numerically, the computational complexity is independent of the number of diversity branches $L$. This is in contrast to the nonasymptotic MC and $\mathrm{RCU_s}$ bounds, whose evaluation has a computational complexity that grows linearly in $L$. Numerical evidence suggests that the saddlepoint approximations are accurate for probabilities of error as small as $10^{-10}$, and although developed under the assumption of large $L$, are accurate even for $L = 1$ if the SNR is greater than or equal to 0 dB. Furthermore, we showed that the proposed approximations recover the normal approximation and the reliability function of the channel, thus providing a unifying tool for the two regimes, which traditionally have been considered separately in the literature.

Observe that the range of the parameters $(L, \epsilon, \rho)$ for which the saddlepoint approximations derived in Chapter 6 are accurate is bigger than the range of parameters for which the high-SNR normal approximation derived in Chapter 5 is accurate. Specifically, while the high-SNR normal approximation is accurate for SNR values larger than or equal to 15 dB, 10 coherence intervals or more, and probabilities of error larger than or equal to $10^{-3}$, the saddlepoint approximations are accurate for probabilities of error as small as $10^{-10}$ and for $L \geq 1$ if the SNR is larger than or equal to 0 dB. However, the high-SNR normal approximation was derived under the assumption of a more general power contraint than the saddlepoint approximations (compare (2.6a) and 2.6b). Furthermore, the high-SNR normal approximation can be evaluated in closed form and does not require any numerical evaluation.

# A

# Appendix to Chapter 5

## A.1 Proof of Lemma 4.1

The left-most inequality in (4.5) follows because the regularized lower incomplete gamma function is no larger than 1. For the right-most inequality in (4.5), consider the following bound by Alzer [41, Th. 1] (see also [42, Eq. (5.4)])

$$\tilde{\gamma}(a,x) > \left(1-e^{-s_a x}\right)^a, \quad (x \geq 0,\, a>0,\, a\neq 1) \tag{A.1}$$

where

$$s_a = \begin{cases} 1, & \text{if } 0<a<1 \\ \Gamma(a+1)^{-\frac{1}{a}}, & \text{if } a>1. \end{cases} \tag{A.2}$$

In order to obtain the right-most inequality in (4.5), we first lower-bound $\tilde{\gamma}(\cdot,\cdot)$ using (A.1)

$$\begin{aligned} \log\frac{1}{\tilde{\gamma}(\mathsf{T}-1,x)} &\leq (\mathsf{T}-1)\log\left(\frac{1}{1-e^{-x\Gamma(\mathsf{T})^{-\frac{1}{\mathsf{T}-1}}}}\right) \\ &= (\mathsf{T}-1)\log\left(1+\frac{1}{e^{x\Gamma(\mathsf{T})^{-\frac{1}{\mathsf{T}-1}}}-1}\right) \end{aligned} \tag{A.3}$$

where the second step follows by simple algebraic manipulations. Since $e^z \geq 1 + z$, this can be further upper-bounded as

$$\log \frac{1}{\tilde{\gamma}(\mathsf{T}-1, x)} \leq (\mathsf{T}-1)\log\left(1 + \frac{\Gamma(\mathsf{T})^{\frac{1}{\mathsf{T}-1}}}{x}\right). \tag{A.4}$$

This proves Lemma 4.1.

## A.2 Proof of Lemma 5.2

For every $\rho(1-\delta) \leq \alpha \leq \rho$,

$$\begin{aligned}
\bar{V}_\rho(\alpha) &= \mathsf{E}\Bigg[\bigg(-\frac{\mathsf{T}\rho - \mathsf{T}\alpha}{1+\mathsf{T}\rho}(Z_1 - 1) - \frac{\mathsf{T}\rho}{1+\mathsf{T}\rho}(Z_2 - (\mathsf{T}-1)) \\
&\quad + (\mathsf{T}-1)\log\big((1+\mathsf{T}\alpha)Z_1 + Z_2 + \beta(\rho)\big) \\
&\quad - (\mathsf{T}-1)\mathsf{E}\big[\log\big((1+\mathsf{T}\alpha)Z_1 + Z_2 + \beta(\rho)\big)\big]\bigg)^2\Bigg] \\
&\geq \mathsf{E}\left[\left(\frac{\mathsf{T}\rho - \mathsf{T}\alpha}{1+\mathsf{T}\rho}(Z_1 - 1) + \frac{\mathsf{T}\rho}{1+\mathsf{T}\rho}(Z_2 - (\mathsf{T}-1))\right)^2\right] \\
&\quad - 2\mathsf{E}\Bigg[\left(\frac{\mathsf{T}\rho - \mathsf{T}\alpha}{1+\mathsf{T}\rho}(Z_1 - 1) + \frac{\mathsf{T}\rho}{1+\mathsf{T}\rho}(Z_2 - (\mathsf{T}-1))\right) \\
&\qquad\qquad \times \bigg((\mathsf{T}-1)\log\big((1+\mathsf{T}\alpha)Z_1 + Z_2\big) \\
&\qquad\qquad\quad - (\mathsf{T}-1)\mathsf{E}\big[\log\big((1+\mathsf{T}\alpha)Z_1 + Z_2\big)\big] \\
&\qquad\qquad\quad + (\mathsf{T}-1)\log\left(1 + \frac{\beta(\rho)}{(1+\mathsf{T}\alpha)Z_1 + Z_2}\right) \\
&\qquad\qquad\quad - (\mathsf{T}-1)\mathsf{E}\left[\log\left(1 + \frac{\beta(\rho)}{(1+\mathsf{T}\alpha)Z_1 + Z_2}\right)\right]\bigg)\Bigg] \\
&\geq \left(\frac{\mathsf{T}\rho - \mathsf{T}\alpha}{1+\mathsf{T}\rho}\right)^2 + \left(\frac{\mathsf{T}\rho}{1+\mathsf{T}\rho}\right)^2(\mathsf{T}-1) \\
&\quad - 2(\mathsf{T}-1)\frac{\mathsf{T}\rho - \mathsf{T}\alpha}{1+\mathsf{T}\rho}\mathsf{E}\left[(Z_1 - 1)\log\left(Z_1 + \frac{Z_2}{1+\mathsf{T}\alpha}\right)\right] \\
&\quad - 2(\mathsf{T}-1)\frac{\mathsf{T}\rho}{1+\mathsf{T}\rho}\Bigg\{\mathsf{E}\left[(Z_2 - (\mathsf{T}-1))\log\left(Z_1 + \frac{Z_2}{1+\mathsf{T}\rho}\right)\right] \\
&\qquad\qquad - \mathsf{E}\left[\big(Z_2 - (\mathsf{T}-1)\big)\log\left(\frac{(1+\mathsf{T}\rho)Z_1 + Z_2}{(1+\mathsf{T}\alpha)Z_1 + Z_2}\right)\right]\Bigg\}
\end{aligned} \tag{A.5}$$

where the second inequality follows because $Z_1$ has mean and variance 1, $Z_2$ has mean and variance $\mathsf{T}-1$, and

$$\mathsf{E}\left[(Z_1-1)\log\left(1+\frac{\beta(\rho)}{(1+\mathsf{T}\alpha)Z_1+Z_2}\right)\right]\le 0 \tag{A.6a}$$

$$\mathsf{E}\left[(Z_2-(\mathsf{T}-1))\log\left(1+\frac{\beta(\rho)}{(1+\mathsf{T}\alpha)Z_1+Z_2}\right)\right]\le 0. \tag{A.6b}$$

The inequalities (A.6a) and (A.6b) follow because

$$(Z_1-1)\log\left(1+\frac{\beta(\rho)}{(1+\mathsf{T}\alpha)Z_1+Z_2}\right)\le (Z_1-1)\log\left(1+\frac{\beta(\rho)}{(1+\mathsf{T}\alpha)+Z_2}\right) \tag{A.7a}$$

$$\begin{aligned}(Z_2-(\mathsf{T}-1))&\log\left(1+\frac{\beta(\rho)}{(1+\mathsf{T}\alpha)Z_1+Z_2}\right)\\ &\le (Z_2-(\mathsf{T}-1))\log\left(1+\frac{\beta(\rho)}{(1+\mathsf{T}\alpha)Z_1+(\mathsf{T}-1)}\right)\end{aligned} \tag{A.7b}$$

and

$$\begin{aligned}\mathsf{E}&\left[(Z_1-1)\log\left(1+\frac{\beta(\rho)}{(1+\mathsf{T}\alpha)+Z_2}\right)\right]\\ &=\mathsf{E}\left[\big(Z_2-(\mathsf{T}-1)\big)\log\left(1+\frac{\beta(\rho)}{(1+\mathsf{T}\alpha)Z_1+(\mathsf{T}-1)}\right)\right]=0.\end{aligned} \tag{A.7c}$$

The first term on the RHS of (A.5) is nonnegative, so discarding it yields a lower bound. Furthermore, the third term in (A.5) can be lower-bounded by upper-bounding for $\rho(1-\delta)\le\alpha\le\rho$

$$\begin{aligned}2(\mathsf{T}-1)&\frac{\mathsf{T}\rho-\mathsf{T}\alpha}{1+\mathsf{T}\rho}\mathsf{E}\left[(Z_1-1)\log\left(Z_1+\frac{Z_2}{1+\mathsf{T}\alpha}\right)\right]\\ &\le 2(\mathsf{T}-1)\frac{\mathsf{T}\rho-\mathsf{T}\alpha}{1+\mathsf{T}\rho}\sqrt{\mathsf{E}\big[(Z_1-1)^2\big]\mathsf{E}\left[\log^2\left(Z_1+\frac{Z_2}{1+\mathsf{T}\alpha}\right)\right]}\\ &\le 2(\mathsf{T}-1)\delta\sqrt{\left(\frac{\pi^2}{6}+\gamma^2+\psi^2(\mathsf{T})+\zeta(2,\mathsf{T})\right)}.\end{aligned} \tag{A.8}$$

Here, the first inequality follows from the Cauchy-Schwarz inequality, and the last inequality follows because $\mathsf{E}\big[(Z_1-1)^2\big]=1$ and

$$\begin{aligned}\mathsf{E}\left[\log^2\left(Z_1+\frac{Z_2}{1+\mathsf{T}\alpha}\right)\right]&\le\mathsf{E}\big[\log^2(Z_1+Z_2)+\log^2(Z_1)\big]\\ &=\frac{\pi^2}{6}+\gamma^2+\zeta(2,\mathsf{T})+\psi^2(\mathsf{T})\end{aligned} \tag{A.9}$$

where we have evaluated the expected values using [27, Sec. 4.335-1] and [27, Sec. 4.358-2], respectively. The first inequality in (A.9) follows by treating the cases $Z_1 + Z_2/(1+T\rho) \le 1$ and $Z_1 + Z_2/(1+T\rho) > 1$ separately, and by lower-bounding in the former case $Z_1 + Z_2/(1+T\rho)$ by $Z_1$ and upper-bounding in the latter case $Z_1 + Z_2/(1+T\rho)$ by $Z_1 + Z_2$. Hence

$$\begin{aligned}
\log^2\left(Z_1 + \frac{Z_2}{1+T\rho}\right) &\le \log^2(Z_1) \\
&\le \log^2(Z_1) + \log^2(Z_1+Z_2), \quad \text{if } Z_1 + \frac{Z_2}{1+T\rho} \le 1 \qquad \text{(A.10a)} \\
\log^2\left(Z_1 + \frac{Z_2}{1+T\rho}\right) &\le \log^2(Z_1+Z_2) \\
&\le \log^2(Z_1) + \log^2(Z_1+Z_2), \quad \text{if } Z_1 + \frac{Z_2}{1+T\rho} > 1 \qquad \text{(A.10b)}
\end{aligned}$$

which yields the desired bound.

Finally, the fifth term on the RHS (A.5) can be lower-bounded by upper-bounding for $\rho(1-\delta) \le \alpha \le \rho$

$$\begin{aligned}
&\left|\mathsf{E}\left[(Z_2 - (\mathsf{T}-1))\log\left(\frac{(1+\mathsf{T}\rho)Z_1 + Z_2}{(1+\mathsf{T}\alpha)Z_1+Z_2}\right)\right]\right| \\
&\qquad \le \mathsf{E}\left[\big|Z_2 - (\mathsf{T}-1)\big|\log\left(\frac{(1+\mathsf{T}\rho)Z_1 + Z_2}{(1+\mathsf{T}\alpha)Z_1+Z_2}\right)\right] \\
&\qquad \le \mathsf{E}\big[|Z_2 - (\mathsf{T}-1)|\big]\log\left(\frac{\rho}{\alpha}\right) \\
&\qquad \le \mathsf{E}\big[|Z_2 - (\mathsf{T}-1)|\big]\log\left(\frac{1}{1-\delta}\right). \qquad \text{(A.11)}
\end{aligned}$$

Combining (A.8)–(A.11) with (A.5), we obtain the lower bound

$$\begin{aligned}
\bar{V}_\rho(\alpha) \ge{}& \left(\frac{\mathsf{T}\rho}{1+\mathsf{T}\rho}\right)^2(\mathsf{T}-1) - 2(\mathsf{T}-1)\delta\sqrt{\left(\frac{\pi^2}{6} + \gamma^2 + \psi^2(\mathsf{T}) + \zeta(2,\mathsf{T})\right)} \\
&- 2(\mathsf{T}-1)\frac{\mathsf{T}\rho}{1+\mathsf{T}\rho}\Bigg\{\mathsf{E}\left[\big(Z_2 - (\mathsf{T}-1)\big)\log\left(Z_1 + \frac{Z_2}{1+\mathsf{T}\rho}\right)\right] \\
&+ \mathsf{E}\big[|Z_2 - (\mathsf{T}-1)|\big]\log\left(\frac{1}{1-\delta}\right)\Bigg\}. \qquad \text{(A.12)}
\end{aligned}$$

Only the second and fourth term on the RHS of (A.12) depend on $\delta$. The former term is linear in $\delta$, the latter term can be upper-bounded by a linear term by using that, for $0 \le \delta \le 1/2$,

$$\log\left(\frac{1}{1-\delta}\right) \le \frac{\delta}{1-\delta} \le 2\delta. \qquad \text{(A.13)}$$

Hence, there exists a positive constant $\Xi$ that only depends on $\mathsf{T}$ such that

$$\bar{V}_\rho(\alpha) \ge \left(\frac{\mathsf{T}\rho}{1+\mathsf{T}\rho}\right)^2(\mathsf{T}-1) - \Xi\delta$$

$$- 2(\mathsf{T}-1)\frac{\mathsf{T}\rho}{1+\mathsf{T}\rho}\mathsf{E}\left[\left(Z_2-(\mathsf{T}-1)\right)\log\left(Z_1+\frac{Z_2}{1+\mathsf{T}\rho}\right)\right]. \tag{A.14}$$

We conclude the proof of Lemma 5.2 by demonstrating that

$$\mathsf{E}\left[(Z_2-(\mathsf{T}-1))\log\left(Z_1+\frac{Z_2}{1+\mathsf{T}\rho}\right)\right]=o_\rho(1). \tag{A.15}$$

This is a direct consequence of the dominated convergence theorem [43, Section 1.26], which can be applied because

$$\begin{aligned}\left|(Z_2-(\mathsf{T}-1))\log\left(Z_1+\frac{Z_2}{1+\mathsf{T}\rho}\right)\right|&\\ &\le \left|Z_2-(\mathsf{T}-1)\right|\left|\log\left(Z_1+\frac{Z_2}{1+\mathsf{T}\rho}\right)\right|\\ &\le \left|(Z_2-(\mathsf{T}-1))\right|\sqrt{\log^2(Z_1)+\log^2(Z_1+Z_2)} \end{aligned} \tag{A.16}$$

where the second inequality follows from the same steps as the first inequality in (A.9). Using the Cauchy-Schwarz inequality, the expected value of the RHS of (A.16) can be upper-bounded as

$$\begin{aligned}\mathsf{E}\left[\left|(Z_2-(\mathsf{T}-1))\right|\sqrt{\log^2(Z_1)+\log^2(Z_1+Z_2)}\right]&\\ &\le\sqrt{\mathsf{E}\left[(Z_2-(\mathsf{T}-1))^2\right]\mathsf{E}\left[\log^2(Z_1)+\log^2(Z_1+Z_2)\right]}\end{aligned} \tag{A.17}$$

which is finite by (A.11).

## A.3 Proof of Lemma 5.3

We shall first prove (5.28a). Using the definitions of $\bar{\jmath}_\ell(\alpha)$ and $\bar{J}(\alpha)$ in (4.20) and (4.33), respectively, we upper-bound $\bar{V}_\rho(\alpha)\triangleq\mathsf{E}\big[\big(\bar{\jmath}_\ell(\alpha)-\bar{J}(\alpha)\big)^2\big]$ as

$$\begin{aligned}\bar{V}_\rho(\alpha)=\mathsf{E}\Bigg[\Bigg(&\frac{\mathsf{T}\rho-\mathsf{T}\alpha}{1+\mathsf{T}\rho}(1-Z_1)+\frac{\mathsf{T}\rho}{1+\mathsf{T}\rho}(\mathsf{T}-1-Z_2)\\ &+(\mathsf{T}-1)\log\left(Z_1+\frac{Z_2}{1+\mathsf{T}\alpha}\right)-(\mathsf{T}-1)\mathsf{E}\left[\log\left(Z_1+\frac{Z_2}{1+\mathsf{T}\alpha}\right)\right]\\ &+(\mathsf{T}-1)\log\left(1+\frac{\beta(\rho)}{(1+\mathsf{T}\alpha)Z_1+Z_2}\right)\\ &-(\mathsf{T}-1)\mathsf{E}\left[\log\left(1+\frac{\beta(\rho)}{(1+\mathsf{T}\alpha)Z_1+Z_2}\right)\right]\Bigg)^2\Bigg]\\ \le c_{4,2}\Bigg(\left(\frac{\mathsf{T}\rho-\mathsf{T}\alpha}{1+\mathsf{T}\rho}\right)^2&\mathsf{E}\big[(Z_1-1)^2\big]+\left(\frac{\mathsf{T}\rho}{1+\mathsf{T}\rho}\right)^2\mathsf{E}\big[(Z_2-\mathsf{T}+1)^2\big]\end{aligned}$$

$$
\begin{aligned}
&+ 2(\mathsf{T}-1)^2\mathsf{E}\left[\log^2\left(Z_1 + \frac{Z_2}{1+\mathsf{T}\alpha}\right)\right] \\
&+ 2(\mathsf{T}-1)^2\mathsf{E}\left[\log^2\left(1 + \frac{\beta(\rho)}{(1+\mathsf{T}\alpha)Z_1 + Z_2}\right)\right]\Bigg)
\end{aligned}
$$

where we have used that

$$
|a_1 + \cdots + a_\eta|^\nu \le c_{\eta,\nu}(|a_1|^\nu + \cdots + |a_\eta|^\nu), \quad \eta,\nu \in \mathbb{Z}^+ \tag{A.18}
$$

for some positive constant $c_{\eta,\nu}$ that only depends on $\eta$ and $\nu$, and that $\mathsf{E}\big[(X - \mathsf{E}[X])^2\big] \le \mathsf{E}\big[X^2\big]$ for every real-valued random variable $X$.

We next show that each term on the RHS of (A.18) is bounded in $(\rho,\alpha)$. Indeed, we have $\mathsf{E}\big[(Z_1-1)^2\big] = 1$ and $\mathsf{E}\big[(Z_2 - (\mathsf{T}-1))^2\big] = (\mathsf{T}-1)$. Furthermore, since $0 \le (\mathsf{T}\rho - \mathsf{T}\alpha)/(1+\mathsf{T}\rho) \le 1$ and $0 \le \mathsf{T}\rho/(1+\mathsf{T}\rho) \le 1$, the first two terms on the RHS of (A.18) are bounded in $\rho$ and $\alpha$. The third term on the RHS of (A.18) can be upper-bounded by (see (A.9))

$$
\begin{aligned}
&(\mathsf{T}-1)^2\mathsf{E}\left[\log^2\left(Z_1 + \frac{Z_2}{1+\mathsf{T}\alpha}\right)\right] \\
&\qquad \le (\mathsf{T}-1)^2\mathsf{E}\big[\log^2(Z_1+Z_2)\big] + (\mathsf{T}-1)^2\mathsf{E}\big[\log^2(Z_1)\big] < \infty.
\end{aligned} \tag{A.19}
$$

Finally, for every $\rho_0 > 0$ and $\rho \ge \rho_0$, the fourth term on the RHS of (A.18) can be upper-bounded by

$$
\begin{aligned}
\mathsf{E}\left[(\mathsf{T}-1)^2\log^2\left(1 + \frac{\beta(\rho)}{(1+\mathsf{T}\alpha)Z_1 + Z_2}\right)\right] &\le (\mathsf{T}-1)^2\mathsf{E}\left[\log^2\left(1 + \frac{\beta(\rho)}{Z_1+Z_2}\right)\right] \\
&\le (\mathsf{T}-1)^2\mathsf{E}\left[\log^2\left(1 + \frac{\beta(\rho_0)}{Z_1+Z_2}\right)\right] \\
&< \infty
\end{aligned} \tag{A.20}
$$

where the last step follows because $\rho \mapsto \beta(\rho)$ is monotonically decreasing in $\rho$. This proves (5.28a).

The proof of (5.28b) follows along similar lines. Indeed, using the definitions of $i_\ell(\rho)$ and $I(\rho)$ for $s=1$ in (4.12) and (4.27), respectively, we can upper-bound $V(\rho) \triangleq \mathsf{E}\big[\big(i_\ell(\rho) - I(\rho)\big)^2\big]$ as

$$
\begin{aligned}
V(\rho) = \mathsf{E}\Bigg[\Bigg(&\frac{\mathsf{T}\rho}{1+\mathsf{T}\rho}(\mathsf{T}-1-Z_2) + (\mathsf{T}-1)\log\left(Z_1 + \frac{Z_2}{1+\mathsf{T}\rho}\right) \\
&- (\mathsf{T}-1)\mathsf{E}\left[\log\left(Z_1 + \frac{Z_2}{1+\mathsf{T}\rho}\right)\right] \\
&- \log\tilde{\gamma}\left(\mathsf{T}-1, \frac{\mathsf{T}\rho((1+\mathsf{T}\rho)Z_1 + Z_2)}{1+\mathsf{T}\rho}\right)
\end{aligned}
$$

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

$$
\begin{aligned}
&\leq c_{5,2}\Bigg(\left(\frac{\mathsf{T}\rho}{1+\mathsf{T}\rho}\right)^2 \mathsf{E}\big[(Z_2-\mathsf{T}+1)^2\big] + 2(\mathsf{T}-1)^2\mathsf{E}\left[\log^2\left(Z_1+\frac{Z_2}{1+\mathsf{T}\rho}\right)\right] \\
&\qquad + 2\mathsf{E}\left[\log^2 \tilde{\gamma}\left(\mathsf{T}-1, \frac{\mathsf{T}\rho((1+\mathsf{T}\rho)Z_1+Z_2)}{1+\mathsf{T}\rho}\right)\right]\Bigg). \qquad \text{(A.21)}
\end{aligned}
$$

We next show that each summand is bounded in $\rho$. Indeed, as shown before, the first and the second term on the RHS of (A.21) are bounded in $\rho$. To bound the third term on the RHS of (A.21), we use Lemma 4.1 and obtain

$$
\begin{aligned}
&\mathsf{E}\left[\log^2 \tilde{\gamma}\left(\mathsf{T}-1, \frac{\mathsf{T}\rho((1+\mathsf{T}\rho)Z_1+Z_2)}{1+\mathsf{T}\rho}\right)\right] \\
&\qquad\qquad \leq (\mathsf{T}-1)^2\mathsf{E}\left[\log^2\left(1+\frac{\beta(\rho)}{(1+\mathsf{T}\rho)Z_1+Z_2}\right)\right]. \qquad \text{(A.22)}
\end{aligned}
$$

By the monotonicity of $\rho \mapsto \beta(\rho)$, it follows that for every $\rho_0 > 0$ and $\rho \geq \rho_0$, the third term on the RHS of (A.21) is upper-bounded by

$$
\mathsf{E}\left[\log^2 \tilde{\gamma}\left(\mathsf{T}-1, \frac{\mathsf{T}\rho((1+\mathsf{T}\rho)Z_1+Z_2)}{1+\mathsf{T}\rho}\right)\right] \leq (\mathsf{T}-1)^2\mathsf{E}\left[\log^2\left(1+\frac{\beta(\rho_0)}{Z_1+Z_2}\right)\right] < \infty. \qquad \text{(A.23)}
$$

Combining the above steps with (A.21) we establish (5.28b).

## A.4 Proof of Lemma 5.4

We shall first prove (5.29a). Using the definitions of $\bar{j}_\ell(\alpha)$ and $\bar{J}(\alpha)$ in (4.20) and (4.33), respectively, we can upper-bound $\mathsf{E}\left[\left|\bar{j}_\ell(\alpha)-\bar{J}(\alpha)\right|^3\right]$ as

$$
\begin{aligned}
\mathsf{E}\left[\left|\bar{j}_\ell(\alpha)-\bar{J}(\alpha)\right|^3\right] &= \mathsf{E}\Bigg[\Bigg|\frac{\mathsf{T}\rho-\mathsf{T}\alpha}{1+\mathsf{T}\rho}(1-Z_1) + \frac{\mathsf{T}\rho}{1+\mathsf{T}\rho}(\mathsf{T}-1-Z_2) \\
&\qquad + (\mathsf{T}-1)\log\left(Z_1+\frac{Z_2}{1+\mathsf{T}\alpha}\right) \\
&\qquad - (\mathsf{T}-1)\mathsf{E}\left[\log\left(Z_1+\frac{Z_2}{1+\mathsf{T}\alpha}\right)\right] \\
&\qquad - \log\left(1+\frac{\beta(\rho)}{(1+\mathsf{T}\alpha)Z_1+Z_2}\right) \\
&\qquad + \mathsf{E}\left[\log\left(1+\frac{\beta(\rho)}{(1+\mathsf{T}\alpha)Z_1+Z_2}\right)\right]\Bigg|^3\Bigg] \\
&\leq c_{6,3}\Bigg(\left|\frac{\mathsf{T}\rho-\mathsf{T}\alpha}{1+\mathsf{T}\rho}\right|^3 \mathsf{E}\big[|Z_1-1|^3\big] + \left|\frac{\mathsf{T}\rho}{1+\mathsf{T}\rho}\right|^3 \mathsf{E}\big[|Z_2-\mathsf{T}+1|^3\big]
\end{aligned}
$$

$$+ 2(\mathsf{T}-1)^3\mathsf{E}\left[\left|\log\left(Z_1+\frac{Z_2}{1+\mathsf{T}\alpha}\right)\right|^3\right]$$
$$+ 2(\mathsf{T}-1)^3\mathsf{E}\left[\left|\log\left(1+\frac{\beta(\rho)}{(1+\mathsf{T}\alpha)Z_1+Z_2}\right)\right|^3\right]\Bigg) \tag{A.24}$$

where we have used (A.18) and that $\mathsf{E}[|X|^3] \geq |\mathsf{E}[X]|^3$ for every random variable $X$.

We next show that each term on the RHS of (A.24) is bounded in $\rho$ and $\alpha$. Indeed, the first two terms on the RHS of (A.24) are bounded because the third central moments of the Gamma-distributed random variables $Z_1$ and $Z_2$ are bounded, and because $0 \leq (\mathsf{T}\rho - \mathsf{T}\alpha)/(1+\mathsf{T}\rho) \leq 1$ and $0 \leq \mathsf{T}\rho/(1+\mathsf{T}\rho) \leq 1$. The third term on the RHS of (A.24) can be upper-bounded by using that

$$\left|\log\left(Z_1+\frac{Z_2}{1+\mathsf{T}\alpha}\right)\right| \leq |\log Z_1| + |\log(Z_1+Z_2)| \tag{A.25}$$

which follows from similar steps as the first inequality in (A.9). Hence,

$$\mathsf{E}\left[\left|\log\left(Z_1+\frac{Z_2}{1+\mathsf{T}\alpha}\right)\right|^3\right] \leq c_{2,3}\big(\mathsf{E}\big[|\log Z_1|^3\big] + \mathsf{E}\big[|\log(Z_1+Z_2)|^3\big]\big) < \infty \tag{A.26}$$

where the first inequality follows by (A.18). Finally, the fourth term on the RHS of (A.24) can be upper-bounded as

$$(\mathsf{T}-1)^3\mathsf{E}\left[\left|\log\left(1+\frac{\beta(\rho)}{(1+\mathsf{T}\alpha)Z_1+Z_2}\right)\right|^3\right] \leq (\mathsf{T}-1)^3\mathsf{E}\left[\left|\log\left(1+\frac{\beta(\rho)}{Z_1+Z_2}\right)\right|^3\right]. \tag{A.27}$$

By the monotonicity of $\rho \mapsto \beta(\rho)$, we thus have that for every $\rho_0 > 0$ and $\rho \geq \rho_0$,

$$(\mathsf{T}-1)^3\mathsf{E}\left[\left|\log\left(1+\frac{\beta(\rho)}{(1+\mathsf{T}\alpha)Z_1+Z_2}\right)\right|^3\right] \leq (\mathsf{T}-1)^3\mathsf{E}\left[\left|\log\left(1+\frac{\beta(\rho_0)}{Z_1+Z_2}\right)\right|^3\right] < \infty. \tag{A.28}$$

Combining the above steps with (A.24) we prove (5.29a).

We establish (5.29b) along similar lines. Using the definitions of $i_\ell(\rho)$ and $I(\rho)$ for $s=1$ in (4.12) and (4.27), respectively, we can upper-bound $\mathsf{E}\big[|i_\ell(\rho)-I(\rho)|^3\big]$ as

$$\begin{aligned}
&\mathsf{E}\left[\left|i_\ell(\rho)-I(\rho)\right|^3\right] \\
&\quad = \mathsf{E}\Bigg[\Bigg|\frac{\mathsf{T}\rho}{1+\mathsf{T}\rho}(\mathsf{T}-1-Z_2) + (\mathsf{T}-1)\log\left(Z_1+\frac{Z_2}{1+\mathsf{T}\rho}\right) \\
&\qquad - (\mathsf{T}-1)\mathsf{E}\left[\log\left(Z_1+\frac{Z_2}{1+\mathsf{T}\rho}\right)\right] - \log\tilde{\gamma}\left(\mathsf{T}-1, \frac{\mathsf{T}\rho((1+\mathsf{T}\rho)Z_1+Z_2)}{1+\mathsf{T}\rho}\right) \\
&\qquad + \mathsf{E}\left[\log\tilde{\gamma}\left(\mathsf{T}-1, \frac{\mathsf{T}\rho((1+\mathsf{T}\rho)Z_1+Z_2)}{1+\mathsf{T}\rho}\right)\right]\Bigg|^3\Bigg]
\end{aligned}$$

$$\leq c_{5,3}\left(\left|\frac{\mathsf{T}\rho}{1+\mathsf{T}\rho}\right|^3 \mathsf{E}\left[|Z_2-\mathsf{T}+1|^3\right] + 2(\mathsf{T}-1)^3\mathsf{E}\left[\left|\log\left(Z_1+\frac{Z_2}{1+\mathsf{T}\rho}\right)\right|^3\right]\right.$$
$$\left. + 2(\mathsf{T}-1)^3\mathsf{E}\left[\left|\log\tilde{\gamma}\left(\mathsf{T}-1,\frac{\mathsf{T}\rho((1+\mathsf{T}\rho)Z_1+Z_2)}{1+\mathsf{T}\rho}\right)\right|^3\right]\right) \tag{A.29}$$

where the inequality follows by (A.18) and because $\mathsf{E}[|X|^3] \geq |\mathsf{E}[X]|^3$ for every random variable $X$.

As shown before, the first two terms on the RHS of (A.29) are bounded in $\rho$. With respect to the third term, we first use Lemma 4.1 to obtain

$$\mathsf{E}\left[\left|\log\tilde{\gamma}\left(\mathsf{T}-1,\frac{\mathsf{T}\rho((1+\mathsf{T}\rho)Z_1+Z_2)}{1+\mathsf{T}\rho}\right)\right|^3\right]$$
$$\leq (\mathsf{T}-1)^3\mathsf{E}\left[\log^3\left(1+\frac{\beta(\rho)}{(1+\mathsf{T}\rho)Z_1+Z_2}\right)\right]. \tag{A.30}$$

By the monotonicity of $\rho \mapsto \beta(\rho)$, it follows that for every $\rho_0 > 0$ and $\rho \geq \rho_0$, the third term on the RHS of (A.29) is upper-bounded by

$$\mathsf{E}\left[\left|\log\tilde{\gamma}\left(\mathsf{T}-1,\frac{\mathsf{T}\rho((1+\mathsf{T}\rho)Z_1+Z_2)}{1+\mathsf{T}\rho}\right)\right|^3\right]$$
$$\leq (\mathsf{T}-1)^3\mathsf{E}\left[\log^3\left(1+\frac{\beta(\rho_0)}{(1+\mathsf{T}\rho)Z_1+Z_2}\right)\right] < \infty. \tag{A.31}$$

Combining the above steps with (A.29) we establish (5.29b).

## A.5 Proof of Lemma 5.5

Consider the upper bound (5.35), namely,

$$R^*(L,\epsilon,\rho) \leq \sup_{\boldsymbol{\alpha}\in[0,\rho]^L} \log\left(\frac{1}{\beta(\boldsymbol{\alpha}, \mathrm{q}^{(\mathrm{U})}_{\mathbf{Y}^L})}\right). \tag{A.32}$$

In the following, we show that, for sufficiently large $L$ and $\rho$, we can assume without loss of optimality that $\boldsymbol{\alpha} \in \mathcal{A}_{\rho,\delta}$. To this end, we demonstrate that for all $\boldsymbol{\alpha} \notin \mathcal{A}_{\rho,\delta}$ and sufficiently large $L$ and $\rho$, we can find a lower bound on $R^*(L,\epsilon,\rho)$ that exceeds an upper bound on (A.32). Hence, such $\boldsymbol{\alpha}$ cannot be optimal.

A lower bound on $R^*(L,\epsilon,\rho)$ follows from (5.22), and by bounding $I(\rho) \geq \underline{I}(\rho)$ and $V(\rho) \leq V_{\mathrm{UB}}(\rho_0)$, $\rho \geq \rho_0$, using (4.32) and (5.32), i.e.,

$$R^*(L,\epsilon,\rho) \geq \frac{\underline{I}(\rho)}{\mathsf{T}} - \sqrt{\frac{V_{\mathrm{UB}}(\rho_0)}{L\mathsf{T}^2}}Q^{-1}(\tau) \triangleq \frac{R_{\mathrm{LB}}(\rho)}{\mathsf{T}}, \quad \rho \geq \rho_0 \tag{A.33}$$

with $\tau$ defined in (5.25). Recall that, by the assumption $0 < \epsilon < \frac{1}{2}$, we have $Q^{-1}(\tau) > 0$ for $L$ sufficiently large.

It follows from (3.18) and (4.21) that the RHS of (A.32) can be upper-bounded as

$$\sup_{\boldsymbol{\alpha}\in[0,\rho]^L} \log\left(\frac{1}{\beta(\boldsymbol{\alpha}, \mathrm{q}^{(\mathrm{U})}_{\mathbf{Y}^L})}\right) \leq \sup_{\boldsymbol{\alpha}\in[0,\rho]^L} \left\{ \frac{\log\xi(\boldsymbol{\alpha})}{L\mathsf{T}} - \frac{\log\Big(1-\epsilon-\mathsf{P}\big[\sum_{\ell=1}^L \bar{\jmath}_\ell(\alpha_\ell) \geq \log\xi(\boldsymbol{\alpha})\big]\Big)}{L\mathsf{T}} \right\} \tag{A.34}$$

for every $\xi\colon [0,\rho]^L \to (0,\infty)$. By Lemma 5.3, for every $\rho_0 > 0$ there exists a $\bar{V}_{\mathrm{UB}}(\rho_0)$ that is independent of $\alpha$ and $\rho$ and that satisfies

$$\bar{V}_\rho(\alpha) \leq \bar{V}_{\mathrm{UB}}(\rho_0), \quad \alpha \geq 0,\ \rho \geq \rho_0. \tag{A.35}$$

Let

$$\log\xi(\boldsymbol{\alpha}) = \sum_{\ell=1}^L \bar{J}(\alpha_\ell) + \sqrt{\frac{L\bar{V}_{\mathrm{UB}}(\rho_0)}{(1-\epsilon)-\frac{1}{\sqrt{L}}}}. \tag{A.36}$$

By Chebyshev's inequality [22, Ch. V.7] and (A.35), we obtain

$$\mathsf{P}\left[\sum_{\ell=1}^L \bar{\jmath}_\ell(\alpha_\ell) \geq \log\xi(\boldsymbol{\alpha})\right] \leq \frac{\sum_{\ell=1}^L \bar{V}_\rho(\alpha_\ell)}{L\bar{V}_{\mathrm{UB}}(\rho_0)}\left(1-\epsilon-\frac{1}{\sqrt{L}}\right) \leq 1-\epsilon-\frac{1}{\sqrt{L}}, \quad \rho \geq \rho_0. \tag{A.37}$$

Combining (A.37) with (A.34), we obtain

$$\begin{aligned} R^*(L,\epsilon,\rho) &\leq \sup_{\boldsymbol{\alpha}\in[0,\rho]^L} \frac{\sum_{\ell=1}^L \bar{J}(\alpha_\ell)}{L\mathsf{T}} + \sqrt{\frac{\bar{V}_{\mathrm{UB}}(\rho_0)}{L\mathsf{T}^2(1-\epsilon)-\mathsf{T}^2\sqrt{L}}} + \frac{\log L}{2L\mathsf{T}} \\ &\triangleq \sup_{\boldsymbol{\alpha}\in[0,\rho]^L} \frac{1}{L}\sum_{\ell=1}^L \frac{R_{\mathrm{UB}}(\alpha_\ell)}{\mathsf{T}}, \quad \rho \geq \rho_0. \end{aligned} \tag{A.38}$$

The $\boldsymbol{\alpha}$'s for which $\frac{1}{L}\sum_{\ell=1}^L R_{\mathrm{UB}}(\alpha_\ell)/\mathsf{T}$ is smaller than (A.33) 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 $\boldsymbol{\alpha}$ that satisfy $\alpha_\ell \geq \rho(1-\delta)$ tends to 1 as $L$ and $\rho$ tend to infinity. Specifically, we consider the difference

$$\frac{1}{L}\sum_{\ell=1}^L \big[R_{\mathrm{LB}}(\rho) - R_{\mathrm{UB}}(\alpha_\ell)\big]$$
$$= \frac{1}{L}\sum_{\ell=1}^L \left[\frac{\mathsf{T}\rho - \mathsf{T}\alpha_\ell}{1+\mathsf{T}\rho} + \log\frac{1+\mathsf{T}\alpha_\ell}{1+\mathsf{T}\rho}\right.$$

$$+ (\mathsf{T}-1)\mathsf{E}\left[\log\frac{(1+\mathsf{T}\rho)Z_{1,\ell}+Z_{2,\ell}+\beta(\rho)}{(1+\mathsf{T}\alpha_\ell)Z_{1,\ell}+Z_{2,\ell}+\beta(\rho)}\right]$$
$$- (\mathsf{T}-1)\mathsf{E}\left[\log\left(1+\frac{\beta(\rho)}{(1+\mathsf{T}\rho)Z_{1,\ell}+Z_{2,\ell}}\right)\right]$$
$$\left.- \sqrt{\frac{V_{\mathrm{UB}}(\rho_0)}{L}}Q^{-1}(\tau) - \sqrt{\frac{\bar{V}_{\mathrm{UB}}(\rho_0)}{L(1-\epsilon)-\sqrt{L}}} - \frac{\log L}{2L}\right] \qquad (A.39)$$

where we have evaluated $R_{\mathrm{LB}}(\rho)$ and $R_{\mathrm{UB}}(\alpha_\ell)$ using (4.31a) and (4.33). We next fix a sufficiently large $\rho_0$ and assume $\rho \geq \rho_0$. Since $\rho \mapsto \beta(\rho)$ is decreasing in $\rho$, we can lower-bound the third-term on the RHS of (A.39) by replacing $\beta(\rho)$ by $\beta(\rho_0)$. We can further lower-bound (A.39) by omitting the first term on the RHS of (A.39), which is nonnegative since $\alpha_\ell \leq \rho$. This yields

$$\frac{1}{L}\sum_{\ell=1}^{L}\left[R_{\mathrm{LB}}(\rho) - R_{\mathrm{UB}}(\alpha_\ell)\right] \geq \frac{1}{L}\sum_{\ell=1}^{L}\left[\log\frac{1+\mathsf{T}\alpha_\ell}{1+\mathsf{T}\rho}\right.$$
$$+ (\mathsf{T}-1)\mathsf{E}\left[\log\frac{(1+\mathsf{T}\rho)Z_{1,\ell}+\mathsf{T}-1+\beta(\rho_0)}{(1+\mathsf{T}\alpha_\ell)Z_{1,\ell}+\mathsf{T}-1+\beta(\rho_0)}\right]$$
$$- (\mathsf{T}-1)\mathsf{E}\left[\log\left(1+\frac{\beta(\rho)}{(1+\mathsf{T}\rho)Z_{1,\ell}+Z_{2,\ell}}\right)\right]$$
$$\left.- \sqrt{\frac{V_{\mathrm{UB}}(\rho_0)}{L}}Q^{-1}(\tau) - \sqrt{\frac{\bar{V}_{\mathrm{UB}}(\rho_0)}{L(1-\epsilon)-\sqrt{L}}} - \frac{\log L}{2L}\right]$$
$$\triangleq \frac{1}{L}\sum_{\ell=1}^{L}\Delta_{L,\rho}(\alpha_\ell), \quad \rho \geq \rho_0.$$

In the following, we analyze the behaviour of the function $\alpha_\ell \mapsto \Delta_{L,\rho}(\alpha_\ell)$. Let

$$g_\rho(\alpha_\ell) \triangleq \log\frac{1+\mathsf{T}\alpha_\ell}{1+\mathsf{T}\rho} + (\mathsf{T}-1)\mathsf{E}\left[\log\frac{(1+\mathsf{T}\rho)Z_{1,\ell}+\mathsf{T}-1+\beta(\rho_0)}{(1+\mathsf{T}\alpha_\ell)Z_{1,\ell}+\mathsf{T}-1+\beta(\rho_0)}\right] \qquad (A.40)$$

and

$$\omega_{L,\rho} \triangleq (\mathsf{T}-1)\mathsf{E}\left[\log\left(1+\frac{\beta(\rho)}{(1+\mathsf{T}\rho)Z_1+Z_2}\right)\right]$$
$$+ \sqrt{\frac{V_{\mathrm{UB}}(\rho_0)}{L}}Q^{-1}(\tau) + \sqrt{\frac{\bar{V}_{\mathrm{UB}}(\rho_0)}{L(1-\epsilon)-\sqrt{L}}} + \frac{\log L}{2L}. \qquad (A.41)$$

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

$$\lim_{\substack{L\to\infty,\\ \rho\to\infty}} \omega_{L,\rho} = \lim_{\rho\to\infty}(\mathsf{T}-1)\mathsf{E}\left[\log\left(1+\frac{\beta(\rho)}{(1+\mathsf{T}\rho)Z_1+Z_2}\right)\right]$$
$$+ \lim_{L\to\infty}\left(\sqrt{\frac{V_{\mathrm{UB}}(\rho_0)}{L}}Q^{-1}(\tau) + \sqrt{\frac{\bar{V}_{\mathrm{UB}}(\rho_0)}{L(1-\epsilon)-\sqrt{L}}} + \frac{\log L}{2L}\right) = 0 \qquad (A.42)$$

where the first term in (A.42) vanishes by the dominated convergence theorem. The following lemma discusses the behavior of the function $\alpha_\ell \mapsto g_\rho(\alpha_\ell)$.

**Lemma A.1** *The function $\alpha \mapsto g_\rho(\alpha)$ has the following properties:*

1. *The derivative of $\alpha \mapsto g_\rho(\alpha)$ is either strictly positive, strictly negative, or changes its sign once from positive to negative. This implies that $g_\rho(\alpha)$, $0 \leq \alpha \leq \rho$ is minimized at the boundary of $[0, \rho]$, and it has a unique maximizer.*

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 $\mathsf{T} > 2$.*

4. *Let $\alpha^*$ denote the unique maximizer of $\alpha \mapsto g_\rho(\alpha)$. For $\mathsf{T} > 2$ and every $\alpha' > \alpha^*$, we have*
$$\sup_{\rho\geq\alpha'} \sup_{\alpha'\leq\alpha\leq\rho} \rho g'_\rho(\alpha) < 0. \tag{A.43}$$

*Proof:* See Appendix A.7. ■

We next study those $\boldsymbol{\alpha}$'s for which $\sum_{\ell=1}^{L} \Delta_{L,\rho}(\alpha_\ell) \geq 0$, since they can be discarded without loss of optimality. Let

$$\mathscr{L}_\delta(\boldsymbol{\alpha}) \triangleq \{\ell = 1, \ldots, L : \alpha_\ell \geq \rho(1-\delta)\} \tag{A.44}$$

and let $L_\delta(\boldsymbol{\alpha})$ denote the number of $\alpha_\ell$'s in $\boldsymbol{\alpha}$ that satisfy $\rho(1-\delta) \leq \alpha_\ell \leq \rho$, i.e., $L_\delta(\boldsymbol{\alpha})$ is the cardinality of $\mathscr{L}_\delta(\boldsymbol{\alpha})$. Further let

$$\Delta^*_{L,\rho}(\delta) \triangleq \inf_{0\leq\alpha\leq\rho(1-\delta)} \Delta_{L,\rho}(\alpha). \tag{A.45}$$

We can express (A.40) as

$$\sum_{\ell=1}^{L} \Delta_{L,\rho}(\alpha_\ell) = \sum_{\mathscr{L}_\delta(\boldsymbol{\alpha})} \Delta_{L,\rho}(\alpha_\ell) + \sum_{\mathscr{L}^c_\delta(\boldsymbol{\alpha})} \Delta_{L,\rho}(\alpha_\ell). \tag{A.46}$$

By Parts 1) and 3) of Lemma A.1,

$$\Delta_{L,\rho}(\alpha_\ell) \geq -\omega_{L,\rho}, \quad 0 \leq \alpha_\ell \leq \rho \tag{A.47}$$

for $\rho$ sufficiently large. Thus, we can lower-bound the first sum on the RHS of (A.46) by $-L_\delta(\boldsymbol{\alpha})\omega_{L,\rho}$ and the second sum on the RHS of (A.46) by $(L - L_\delta(\boldsymbol{\alpha}))\Delta^*_{L,\rho}(\delta)$, which yields

$$\sum_{\ell=1}^{L} \Delta_{L,\rho}(\alpha_\ell) \geq (L - L_\delta(\boldsymbol{\alpha}))\Delta^*_{L,\rho}(\delta) - L_\delta(\boldsymbol{\alpha})\omega_{L,\rho}. \tag{A.48}$$

This implies that we can discard without loss of optimality every $\boldsymbol{\alpha}$ for which

$$L\Delta^*_{L,\rho}(\delta) \geq L_\delta(\boldsymbol{\alpha})(\omega_{L,\rho} + \Delta^*_{L,\rho}(\delta)) \tag{A.49}$$

since for such $\boldsymbol{\alpha}$'s we also have that the RHS of (A.48) is nonnegative. Hence, an $\boldsymbol{\alpha}$ maximizing (A.32) must satisfy

$$\frac{L_\delta(\boldsymbol{\alpha})}{L} > 1 - \frac{\omega_{L,\rho}}{\omega_{L,\rho} + \Delta^*_{L,\rho}(\delta)}. \tag{A.50}$$

As we shall show below, for every $0 < \delta < 1$ we have

$$\omega_{L,\rho} + \Delta^*_{L,\rho}(\delta) \geq -\delta \sup_{\rho \geq \alpha'} \sup_{\alpha' \leq \alpha \leq \rho} \rho g'_\rho(\alpha) \tag{A.51}$$

for some $0 < \alpha' < \rho(1-\delta)$ that is independent of $\rho$. We further show that the RHS of (A.51) is independent of $L$ and $\rho$ and strictly positive. It follows that

$$\frac{L_\delta(\boldsymbol{\alpha})}{L} > 1 - \frac{\omega_{L,\rho}}{-\delta \sup_{\rho \geq \alpha'} \sup_{\alpha' \leq \alpha \leq \rho} \rho g'_\rho(\alpha)} \tag{A.52}$$

which, by (A.42), tends to one as $\rho$ and $L$ tend to infinity. Thus, for every $0 < \delta < 1$, there exist sufficiently large $L_0$ and $\rho_0$ such that

$$L_\delta(\boldsymbol{\alpha}) \geq L/2, \quad L \geq L_0,\ \rho \geq \rho_0. \tag{A.53}$$

This proves Lemma 5.5.

It remains to show (A.51). Let $\alpha_{\min} = \rho(1-\delta)$. By Part 1) of Lemma A.1, $\alpha \mapsto g_\rho(\alpha)$ 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 $0 \leq \alpha \leq \alpha_{\min}$ is either achieved at $\alpha = 0$ or at $\alpha_{\min}$.

By Part 3) of Lemma A.1 and by (A.42), we have

$$\varliminf_{\substack{L\to\infty,\\ \rho\to\infty}} \Delta_{L,\rho}(0) = \infty. \tag{A.54}$$

We next show that

$$\Delta_{L,\rho}(\alpha_{\min}) + \omega_{L,\rho} \geq -\delta \sup_{\rho \geq \alpha'} \sup_{\alpha' \leq \alpha \leq \rho} \rho g'_\rho(\alpha). \tag{A.55}$$

If $\alpha_{\min} \leq \alpha^*$, then this is clearly satisfied, since in this case $\Delta_{L,\rho}(\alpha_{\min}) \geq \Delta_{L,\rho}(0)$ and $\Delta_{L,\rho}(0)$ tends to infinity as $L \to \infty$ and $\rho \to \infty$. However, in general this case does not occur for large $\rho$ and $L$, since $\alpha_{\min}$ tends to infinity as $\rho \to \infty$ and, by Part 2) of Lemma A.1, $\alpha^*$ is not a function of $\rho$, which implies that $\alpha_{\min} > \alpha^*$ for $\rho$ sufficiently large. We thus focus on the case where $\alpha_{\min} > \alpha^*$. Note that

$$\Delta_{L,\rho}(\rho) - \Delta_{L,\rho}(\alpha_{\min}) = -\omega_{L,\rho} - \Delta_{L,\rho}(\alpha_{\min}) \tag{A.56}$$

since $g_\rho(\rho) = 0$. Thus, by the mean value theorem [44, Th. 5.10], there exists an $x_0 \in [\alpha_{\min}, \rho]$ such that

$$-\omega_{L,\rho} - \Delta_{L,\rho}(\alpha_{\min}) = \int_{\alpha_{\min}}^{\rho} \Delta'_{L,\rho}(\alpha) d\alpha = \rho\delta\Delta'_{L,\rho}(x_0) \tag{A.57}$$

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

$$\begin{aligned}\Delta_{L,\rho}(\alpha_{\min}) + \omega_{L,\rho} &\geq -\delta \sup_{\alpha_{\min}\leq\alpha\leq\rho} \rho\Delta'_{L,\rho}(\alpha)\\ &\geq -\delta \sup_{\rho\geq\alpha'} \sup_{\alpha'\leq\alpha\leq\rho} \rho g'_{\rho}(\alpha)\end{aligned} \tag{A.58}$$

for every $\alpha' \in (\alpha^*, \alpha_{\min})$, where the second inequality follows by noting that $\Delta'_{L,\rho}(x) = g'_{\rho}(x)$ and by further optimizing over $\rho$.[1] It remains to show that the RHS of (A.58) is independent of $L$ and $\rho$ and strictly positive. To this end, we first note that $\alpha \mapsto g_{\rho}(\alpha)$ is independent of $L$. Furthermore, by optimizing over $\rho \geq \alpha'$, the RHS of (A.58) becomes also independent of $\rho$ and, by Part 4) of Lemma A.1,

$$\sup_{\rho\geq\alpha'} \sup_{\alpha'\leq\alpha\leq\rho} \rho g'_{\rho}(\alpha) < 0, \quad \mathsf{T} > 2,\ \rho \geq \alpha' \tag{A.59}$$

for every $\alpha' \in (\alpha^*, \alpha_{\min})$. Thus, the claim (A.51) follows, which concludes the proof of Lemma 5.5.

## A.6 Proof of Lemma 5.6

#### A.6.0.1 Part 1)

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

$$\bar{J}(\rho) - \bar{J}(\alpha) \geq g_{\rho}(\alpha). \tag{A.60}$$

where the function $\alpha \mapsto g_{\rho}(\alpha)$ was defined in (A.40). By Parts 1) and 3) of Lemma A.1 (Appendix A.5), $g_{\rho}(\cdot)$ is nonnegative for sufficiently large $\rho$. It follows that, for such $\rho$,

$$\sup_{0\leq\alpha\leq\rho} \bar{J}(\alpha) = \bar{J}(\rho). \tag{A.61}$$

This proves Part 1) of Lemma 5.6.

#### A.6.0.2 Part 2)

To study

$$\sup_{0\leq\alpha\leq\rho} \left\{ \frac{\bar{J}(\alpha)}{\mathsf{T}} - \sqrt{\frac{\bar{V}_{\rho}(\alpha)}{L\mathsf{T}^2}} Q^{-1}(\epsilon) \right\} \tag{A.62}$$

[1] Since $\alpha^*$ is independent of $\rho$ and $\alpha_{\min} \to \infty$ as $\rho \to \infty$, it follows that we can find an $\alpha' \in (\alpha^*, \alpha_{\min})$ that is independent of $\rho$ and that satisfies (A.58).

we consider the difference

$$\bar{J}(\rho) - \sqrt{\frac{\bar{V}(\rho)}{L}}Q^{-1}(\epsilon) - \bar{J}(\alpha) + \sqrt{\frac{\bar{V}_\rho(\alpha)}{L}}Q^{-1}(\epsilon)$$
$$\geq g_\rho(\alpha) - \sqrt{\frac{\bar{V}(\rho)}{L}}Q^{-1}(\epsilon) + \sqrt{\frac{\bar{V}_\rho(\alpha)}{L}}Q^{-1}(\epsilon). \quad (A.63)$$

Clearly, every $\alpha$ for which the RHS of (A.63) is nonnegative is suboptimal and can be discarded without loss of optimality. We continue by lower-bounding $\bar{V}_\rho(\alpha) \geq 0$ and by using that $\bar{V}(\rho) \leq \bar{V}_{\mathrm{UB}}(\rho_0)$, $\rho \geq \rho_0$ for sufficiently large $\rho_0$ and for some constant $\bar{V}_{\mathrm{UB}}(\rho_0)$ that is independent of $\rho$ (Lemma 5.3, Appendix A.3). Since by the assumption $0 < \epsilon < \frac{1}{2}$ we have $Q^{-1}(\epsilon) > 0$, this yields

$$\begin{aligned} g_\rho(\alpha) - \sqrt{\frac{\bar{V}(\rho)}{L}}Q^{-1}(\epsilon) + \sqrt{\frac{\bar{V}_\rho(\alpha)}{L}}Q^{-1}(\epsilon) \\ \geq g_\rho(\alpha) - \sqrt{\frac{\bar{V}_{\mathrm{UB}}(\rho_0)}{L}}Q^{-1}(\epsilon) \\ \triangleq f_{L,\rho}(\alpha). \end{aligned} \quad (A.64)$$

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 $f_{L,\rho}(\alpha) \triangleq g_\rho(\alpha) - \omega_L$, where

$$\omega_L \triangleq \sqrt{\frac{\bar{V}_{\mathrm{UB}}(\rho_0)}{L}}Q^{-1}(\epsilon). \quad (A.65)$$

Note that $\Delta_{L,\rho}(\alpha)$ defined in (A.40) and $f_{L,\rho}(\alpha)$ only differ in terms that do not depend on $\alpha$ (namely, $\omega_{L,\rho}$ and $\omega_L$), so they have the same behavior with respect to $\alpha$ as summarized in Lemma A.1. Let $\delta_L \triangleq 1 - \alpha_0/\rho$, where $\alpha_0$ is the unique real root of $\alpha \mapsto f_{L,\rho}(\alpha)$. Indeed, we know that $\alpha \mapsto f_{L,\rho}(\alpha)$ has only one root because $\omega_L \geq 0$ and $\omega_L \to 0$ as $L \to \infty$, so $f_{L,\rho}(\rho) = -\omega_L \leq 0$ and $f_{L,\rho}(0) > 0$ for $L$ and $\rho$ sufficiently large. Furthermore, we have $f'_{L,\rho}(\alpha) = g'_\rho(\alpha)$ and $g'_\rho(\alpha)$ is either strictly positive, strictly negative, or changes its sign once from positive to negative (Part 1) of Lemma A.1). Consequently, $f_{L,\rho}(\alpha)$, $0 \leq \alpha \leq \rho$ is minimized at an endpoint of $[0, \rho]$ and it has a unique maximizer, so 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 without loss of optimality, since for such $\alpha$'s the function $f_{L,\rho}(\alpha)$ is nonnegative.

To study the behavior of $\delta_L$, we next note that $\omega_L = -(f_{L,\rho}(\rho) - f_{L,\rho}(\alpha_0))$. It follows then by similar steps as in (A.57)–(A.58) that

$$\omega_L \geq -\delta_L \sup_{\alpha_0 \leq \alpha \leq \rho} \rho f'_{L,\rho}(\alpha). \quad (A.66)$$

Let $\alpha^*$ denote the unique maximizer of $\alpha \mapsto f_{L,\rho}(\alpha)$. Recall that $\alpha^*$ does not depend on $\rho$, since by Part 2) of Lemma A.1, the derivative of $\alpha \mapsto g_\rho(\alpha)$ does not depend on

$\rho$. We next show that we can find an $\tilde{\alpha}$ independent of $L$ and $\rho$ such that $\alpha^* < \tilde{\alpha} < \alpha_0$. Indeed, by Lemma A.1, we have that $g_\rho(\alpha^*) > 0$ for sufficiently large $\rho$. This in turn implies that

$$\varliminf_{\substack{L\to\infty,\\ \rho\to\infty}} f_{L,\rho}(\alpha^*) > 0 \tag{A.67}$$

since $\lim_{L\to\infty}\omega_L = 0$. We next note that

$$\varliminf_{\substack{L\to\infty,\\ \rho\to\infty}} f_{L,\rho}(\tilde{\alpha}) \geq \varliminf_{\substack{L\to\infty,\\ \rho\to\infty}} f_{L,\rho}(\alpha^*) - \left|f_{L,\rho}(\tilde{\alpha}) - f_{L,\rho}(\alpha^*)\right| \tag{A.68}$$

where the difference

$$\begin{aligned} f_{L,\rho}(\tilde{\alpha}) - f_{L,\rho}(\alpha^*) &= g_\rho(\tilde{\alpha}) - g_\rho(\alpha^*) \\ &= \log\frac{1+\mathsf{T}\tilde{\alpha}}{1+\mathsf{T}\alpha^*} + (\mathsf{T}-1)\mathsf{E}\left[\log\frac{(1+\mathsf{T}\alpha^*)Z_1+\mathsf{T}-1+\beta(\rho_0)}{(1+\mathsf{T}\tilde{\alpha})Z_1+\mathsf{T}-1+\beta(\rho_0)}\right] \end{aligned} \tag{A.69}$$

is independent of $L$ and $\rho$. By the continuity of $\alpha \mapsto g_\rho(\alpha)$, it follows from (A.67)–(A.69) that there exists an $\tilde{\alpha} \in (\alpha^*, \rho]$ that is independent of $L$ and $\rho$ such that

$$\varliminf_{\substack{L\to\infty,\\ \rho\to\infty}} f_{L,\rho}(\tilde{\alpha}) > 0. \tag{A.70}$$

In other words, if $L$ and $\rho$ are sufficiently large, then we can find an $\tilde{\alpha} \in (\alpha^*, \alpha_0)$ that is independent of $L$ and $\rho$. Thus, in this case the RHS of (A.66) can be further lower-bounded by

$$\begin{aligned} \omega_L &\geq -\delta_L \sup_{\tilde{\alpha}\leq\alpha\leq\rho} \rho f'_{L,\rho}(\alpha) \\ &\geq -\delta_L \sup_{\rho\geq\tilde{\alpha}} \sup_{\tilde{\alpha}\leq\alpha\leq\rho} \rho f'_{L,\rho}(\alpha). \end{aligned} \tag{A.71}$$

We next argue that the constant

$$\mathsf{F} \triangleq -\sup_{\rho\geq\tilde{\alpha}} \sup_{\tilde{\alpha}\leq\alpha\leq\rho} \rho f'_{L,\rho}(\alpha) \tag{A.72}$$

is independent of $L$ and $\rho$ and strictly positive. Indeed, we have that $f'_{L,\rho}(x) = g'_\rho(x)$, which is independent of $L$. Furthermore, by optimizing over $\rho \geq \tilde{\alpha}$, the RHS of (A.72) becomes independent of $\rho$. Finally, setting $\alpha' = \tilde{\alpha}$ in (A.43) (Part 4) of Lemma A.1) yields

$$\sup_{\rho\geq\tilde{\alpha}} \sup_{\tilde{\alpha}\leq\alpha\leq\rho} \rho g'_\rho(\alpha) < 0, \quad \rho \geq \tilde{\alpha}. \tag{A.73}$$

Hence, the claim follows. Consequently, we obtain from (A.71) and the definition of $\omega_L$ and $\mathsf{F}$ that, for sufficiently large $L_0$ and $\rho_0$,

$$\delta_L \leq \frac{\sqrt{\bar{V}_{\mathrm{UB}}(\rho_0)}Q^{-1}(\epsilon)}{\mathsf{F}}\frac{1}{\sqrt{L}}, \quad \rho \geq \rho_0,\, L \geq L_0. \tag{A.74}$$

We next tighten this bound on $\delta_L$. Indeed, using that without loss of optimality we can assume $\rho(1-\delta_L) \leq \alpha \leq \rho$, we can derive a tighter lower bound on (A.63) by lower-bounding $\bar{V}_\rho(\alpha)$ using the lower bound given in Appendix A.8 instead of lower-bounding it by zero. Specifically, by (A.90) in Appendix A.8,

$$\sqrt{\frac{\bar{V}_\rho(\alpha)}{L}} \geq \sqrt{\frac{\bar{V}(\rho)-\Upsilon\delta_L}{L}} \geq \sqrt{\frac{\bar{V}(\rho)}{L}} - \sqrt{\frac{\Upsilon\delta_L}{L}}, \quad \rho(1-\delta_L) \leq \alpha \leq \rho. \tag{A.75}$$

We can thus lower-bound (A.63) as

$$\begin{aligned}\bar{J}(\rho) - \sqrt{\frac{\bar{V}(\rho)}{L}}Q^{-1}(\epsilon) - \bar{J}(\alpha) + \sqrt{\frac{\bar{V}_\rho(\alpha)}{L}}Q^{-1}(\epsilon)& \\ \geq g_\rho(\alpha) - \sqrt{\frac{\Upsilon\delta_L}{L}}Q^{-1}(\epsilon)& \\ \triangleq \tilde{f}_{L,\rho}(\alpha), \quad \rho(1-\delta_L) \leq \alpha \leq \rho.& \end{aligned} \tag{A.76}$$

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

$$\tilde{\omega}_L \triangleq \sqrt{\frac{\Upsilon\delta_L}{L}}Q^{-1}(\epsilon). \tag{A.77}$$

Further let $\tilde{\delta}_L \triangleq 1-\tilde{\alpha}_0/\rho$, where $\tilde{\alpha}_0$ 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 $\rho(1-\tilde{\delta}_L)$ can be discarded without loss of optimality, since for such $\alpha$'s the function $\tilde{f}_{L,\rho}(\alpha)$ is nonnegative. By repeating the steps (A.66)–(A.74) with $\omega_L$ replaced by $\tilde{\omega}_L$, we obtain for sufficiently large $L_0$ and $\rho_0$

$$\begin{aligned}\tilde{\delta}_L &\leq \frac{1}{\mathsf{F}}\sqrt{\frac{\Upsilon\delta_L}{L}}Q^{-1}(\epsilon) \\ &\leq \left(\frac{Q^{-1}(\epsilon)}{\mathsf{F}}\right)^{3/2}\sqrt{\Upsilon\sqrt{\bar{V}_{\mathrm{UB}}(\rho_0)}}\frac{1}{L^{3/4}}\end{aligned} \tag{A.78}$$

where the last inequality follows by upper-bounding $\delta_L$ using (A.74).

If we perform the above steps $N$ times, then we obtain that, without loss of optimality,

$$\alpha \geq \rho\left(1-\delta_L^{(N)}\right) \tag{A.79}$$

where $\delta_L^{(N)}$ satisfies

$$0 \leq \delta_L^{(N)} \leq \left(\frac{Q^{-1}(\epsilon)\sqrt{\Upsilon}}{\mathsf{F}}\right)^{2-2^{-N+1}}\left(\frac{\bar{V}_{\mathrm{UB}}(\rho_0)}{\Upsilon}\right)^{2^{-N}}\frac{1}{L^{1-2^{-N}}}. \tag{A.80}$$

Thus, by letting $N$ tend to infinity, we conclude that we can assume without loss of optimality that

$$\alpha \geq \rho\Big(1 - \delta_L^{(\infty)}\Big) \tag{A.81}$$

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

$$0 \leq \delta^{(\infty)} \leq \frac{\left(\frac{Q^{-1}(\epsilon)\sqrt{\Upsilon}}{\mathsf{F}}\right)^2}{L}. \tag{A.82}$$

This concludes the proof of Part 2) of Lemma 5.6.

## A.7 Proof of Lemma A.1

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

$$\begin{aligned}
g_\rho'(\alpha) &= \frac{\mathsf{T}}{1+\mathsf{T}\alpha} - (\mathsf{T}-1)\mathsf{E}\left[\frac{\mathsf{T}Z_1}{(1+\mathsf{T}\alpha)Z_1 + (\mathsf{T}-1) + \beta(\rho_0)}\right] \\
&= \mathsf{T}\Bigg[\frac{1}{1+\mathsf{T}\alpha} - \frac{\mathsf{T}-1}{1+\mathsf{T}\alpha} \\
&\qquad + \frac{\mathsf{T}-1}{1+\mathsf{T}\alpha}\frac{\mathsf{T}-1+\beta(\rho_0)}{1+\mathsf{T}\alpha} e^{\frac{\mathsf{T}-1+\beta(\rho_0)}{1+\mathsf{T}\alpha}} \mathrm{E}_1\left(\frac{\mathsf{T}-1+\beta(\rho_0)}{1+\mathsf{T}\alpha}\right)\Bigg] \\
&= \frac{\mathsf{T}}{1+\mathsf{T}\alpha}\Bigg[-(\mathsf{T}-2) \\
&\qquad + (\mathsf{T}-1)\frac{\mathsf{T}-1+\beta(\rho_0)}{1+\mathsf{T}\alpha} e^{\frac{\mathsf{T}-1+\beta(\rho_0)}{1+\mathsf{T}\alpha}} \mathrm{E}_1\left(\frac{\mathsf{T}-1+\beta(\rho_0)}{1+\mathsf{T}\alpha}\right)\Bigg].
\end{aligned} \tag{A.83}$$

The first equality follows because, by [45, App. A.9], we can swap derivative and expected value; the second equality follows by solving the expected value using [27, Sec. 3.353-5.[7]]. Note that the RHS of (A.83) does not depend on $\rho$. Hence Part 2) of Lemma A.1 follows immediately.

We next prove Part 1) of Lemma A.1. Because $\mathsf{T}/(1+\mathsf{T}\alpha)$ in (A.83) is nonnegative, the sign of $\alpha \mapsto g_\rho'(\alpha)$ is determined by the terms inside the square brackets. Let $\vartheta \triangleq \frac{1+\mathsf{T}\alpha}{\mathsf{T}-1+\beta(\rho_0)}$. Note that $\vartheta \mapsto \frac{1}{\vartheta}\exp\big(\frac{1}{\vartheta}\big)\mathrm{E}_1\big(\frac{1}{\vartheta}\big)$ is strictly decreasing since, by [27, Sec. 3.353-3],

$$\frac{1}{\vartheta}e^{\frac{1}{\vartheta}}\mathrm{E}_1\left(\frac{1}{\vartheta}\right) = 1 - \int_0^1 e^{-\frac{t}{(1-t)\vartheta}}\,\mathrm{d}t \tag{A.84}$$

and $\vartheta \mapsto e^{-\frac{t}{(1-t)\vartheta}}$ is strictly positive and strictly increasing in $\vartheta$. Hence, the function inside the squared brackets, defined as

$$\Xi(\alpha) \triangleq -(\mathsf{T}-2) + (\mathsf{T}-1)\frac{\mathsf{T}-1+\beta(\rho_0)}{1+\mathsf{T}\alpha} e^{\frac{\mathsf{T}-1+\beta(\rho_0)}{1+\mathsf{T}\alpha}} \mathrm{E}_1\left(\frac{\mathsf{T}-1+\beta(\rho_0)}{1+\mathsf{T}\alpha}\right) \tag{A.85}$$

is strictly decreasing. This implies that $\alpha \mapsto g'_\rho(\alpha)$ is either strictly positive, strictly negative, or changes its sign once from positive to negative.

We next prove Part 3) of Lemma A.1 by showing that $\lim_{\rho\to\infty} g_\rho(0) = \infty$ for $\mathsf{T} > 2$. To this end, we express $g_\rho(0)$ as

$$\begin{aligned} g_\rho(0) = (\mathsf{T}-2)\mathsf{E}\left[\log\left(1+\frac{\mathsf{T}\rho Z_1}{Z_1+(\mathsf{T}-1)+\beta(\rho_0)}\right)\right] \\ + \mathsf{E}\left[\log\left(Z_1+\frac{\mathsf{T}-1+\beta(\rho_0)}{1+\mathsf{T}\rho}\right)\right] - \mathsf{E}\big[\log\big(Z_1+\mathsf{T}-1+\beta(\rho_0)\big)\big]. \end{aligned} \tag{A.86}$$

The first expected value on the RHS of (A.86) tends to infinity as $\rho \to \infty$, whereas the other expected values are bounded in $\rho$. For $\mathsf{T} > 2$, it follows that the RHS of (A.86) tends to infinity as $\rho \to \infty$. Hence the claim follows.

We finally prove Part 4) of Lemma A.1 by analyzing $\rho g'_\rho(\alpha)$. It follows from (A.83) that

$$\rho g'_\rho(\alpha) = \frac{\mathsf{T}\rho}{1+\mathsf{T}\alpha}\left\{2-\mathsf{T}+(\mathsf{T}-1)\frac{\mathsf{T}-1+\beta(\rho_0)}{1+\mathsf{T}\alpha}e^{\frac{\mathsf{T}-1+\beta(\rho_0)}{1+\mathsf{T}\alpha}}\mathrm{E}_1\left(\frac{\mathsf{T}-1+\beta(\rho_0)}{1+\mathsf{T}\alpha}\right)\right\}. \tag{A.87}$$

As argued above, the function $\alpha \mapsto \Xi(\alpha)$ inside the curly brackets (cf. (A.85)) is independent of $L$ and $\rho$ and is strictly decreasing in $\alpha$. Hence, its supremum over $\alpha' \le \alpha \le \rho$ is achieved for $\alpha = \alpha'$. Further note that $\Xi(\alpha')$ is strictly negative for $\mathsf{T} > 2$ and $\alpha' > \alpha^*$. As for the term outside the curly brackets, we have for every $\alpha' > \alpha^*$

$$\inf_{\rho\ge\alpha'}\inf_{\alpha'\le\alpha\le\rho}\frac{\mathsf{T}\rho}{1+\mathsf{T}\alpha} = \frac{\mathsf{T}\alpha'}{1+\mathsf{T}\alpha'} > 0. \tag{A.88}$$

Combining these two results, we conclude that

$$\sup_{\rho\ge\alpha'}\sup_{\alpha'\le\alpha\le\rho}\rho g'_\rho(\alpha) < 0, \quad \mathsf{T} > 2,\ \alpha' > \alpha^*. \tag{A.89}$$

This proves Part 4) of Lemma A.1 and concludes the proof of Lemma A.1.

## A.8 Lower Bound on $\bar{V}_\rho(\alpha)$

We show that for all $\rho(1-\delta) \le \alpha \le \rho$, $0 \le \delta \le 1/2$, and $\rho \ge \rho_0$, we have

$$\bar{V}_\rho(\alpha) \ge \bar{V}(\rho) - \Upsilon\delta \tag{A.90}$$

where $\Upsilon$ is a positive constant that only depends on $\mathsf{T}$. Let $\Omega(\alpha) \triangleq \bar{\jmath}_\ell(\alpha) - \bar{J}(\alpha)$, i.e.,

$$\begin{aligned} \Omega(\alpha) = &-\frac{\mathsf{T}\rho-\mathsf{T}\alpha}{1+\mathsf{T}\rho}(Z_1-1) - \frac{\mathsf{T}\rho}{1+\mathsf{T}\rho}(Z_2-(\mathsf{T}-1)) \\ &+ (\mathsf{T}-1)\log\big((1+\mathsf{T}\alpha)Z_1+Z_2+\beta(\rho)\big) \\ &- (\mathsf{T}-1)\mathsf{E}\big[\log\big((1+\mathsf{T}\alpha)Z_1+Z_2+\beta(\rho)\big)\big]. \end{aligned} \tag{A.91}$$

It follows that $\bar{V}_\rho(\alpha) = \mathsf{E}[\Omega^2(\alpha)]$. We next analyze the difference

$$\begin{aligned}\bar{V}(\rho) - \bar{V}_\rho(\alpha) &= \mathsf{E}\big[\big(\Omega(\rho) - \Omega(\alpha)\big)\big(\Omega(\rho) + \Omega(\alpha)\big)\big] \\ &\leq \sqrt{\mathsf{E}\big[(\Omega(\rho) - \Omega(\alpha))^2\big]\mathsf{E}\big[(\Omega(\rho) + \Omega(\alpha))^2\big]} \end{aligned} \tag{A.92}$$

where the inequality follows from the Cauchy-Schwarz inequality. On the one hand, using (A.18), we have for every $\rho_0 > 0$,

$$\begin{aligned}\sup_{\substack{\alpha>0,\\ \rho\geq\rho_0}} \mathsf{E}\Big[\big(\Omega(\rho) + \Omega(\alpha)\big)^2\Big] &\leq c_{2,2} \sup_{\rho\geq\rho_0} \mathsf{E}\big[\Omega^2(\rho)\big] + c_{2,2} \sup_{\substack{\alpha\geq0,\\ \rho\geq\rho_0}} \mathsf{E}\big[\Omega^2(\alpha)\big] \\ &= c_{2,2} \sup_{\rho\geq\rho_0} \bar{V}(\rho) + c_{2,2} \sup_{\substack{\alpha\geq0,\\ \rho\geq\rho_0}} \bar{V}_\rho(\alpha) \end{aligned} \tag{A.93}$$

which, by Lemma 5.3, is bounded. On the other hand, using (A.18) and that $\mathsf{E}[X^2] \geq \mathsf{E}[X]^2$ for every random variable $X$, we obtain

$$\begin{aligned}\mathsf{E}\Big[\big(\Omega(\rho) - \Omega(\alpha)\big)^2\Big] &= \mathsf{E}\Bigg[\bigg(\frac{\mathsf{T}\rho - \mathsf{T}\alpha}{1+\mathsf{T}\rho}(Z_1 - 1) \\ &\qquad + (\mathsf{T}-1)\log\bigg(\frac{(1+\mathsf{T}\rho)Z_1 + Z_2 + \beta(\rho)}{(1+\mathsf{T}\alpha)Z_1 + Z_2 + \beta(\rho)}\bigg) \\ &\qquad - (\mathsf{T}-1)\mathsf{E}\bigg[\log\bigg(\frac{(1+\mathsf{T}\rho)Z_1 + Z_2 + \beta(\rho)}{(1+\mathsf{T}\alpha)Z_1 + Z_2 + \beta(\rho)}\bigg)\bigg]\bigg)^2\Bigg] \\ &\leq c_{3,2}\bigg(\frac{\mathsf{T}\rho - \mathsf{T}\alpha}{1+\mathsf{T}\rho}\bigg)^2 \\ &\qquad + 2c_{2,3}(\mathsf{T}-1)^2\mathsf{E}\bigg[\log^2\bigg(\frac{(1+\mathsf{T}\rho)Z_1 + Z_2 + \beta(\rho)}{(1+\mathsf{T}\alpha)Z_1 + Z_2 + \beta(\rho)}\bigg)\bigg]. \end{aligned} \tag{A.94}$$

When $\rho(1-\delta) \leq \alpha \leq \rho$, this can be further upper-bounded as

$$\begin{aligned}\mathsf{E}\Big[\big(\Omega(\rho) - \Omega(\alpha)\big)^2\Big] &\leq c_{3,2}\delta^2 + 2c_{3,2}(\mathsf{T}-1)^2\log^2\bigg(1 + \frac{\delta}{1-\delta}\bigg) \\ &\leq \big(c_{3,2} + 8c_{3,2}(\mathsf{T}-1)^2\big)\delta^2 \end{aligned} \tag{A.95}$$

where the last inequality follows because, by assumption, $\delta \leq 1/2$, hence $\frac{\delta^2}{(1-\delta)^2} \leq 4\delta^2$. Combining (A.93) and (A.95) with (A.92) we establish (A.90).

## A.9 High-SNR Approximations of Information Rates

**Lemma A.2** *The quantities $\bar{J}(\rho)$, $I(\rho)$ and $\underline{I}(\rho)$ can be approximated as*

$$\bar{J}(\rho) = (\mathsf{T}-1)\log(\mathsf{T}\rho) - \log\Gamma(\mathsf{T}) - (\mathsf{T}-1)(1+\gamma) + o_\rho(1) \tag{A.96a}$$

$$I(\rho) = (\mathsf{T}-1)\log(\mathsf{T}\rho) - \log\Gamma(\mathsf{T}) - (\mathsf{T}-1)(1+\gamma) + o_\rho(1) \tag{A.96b}$$

$$\underline{I}(\rho) = (\mathsf{T}-1)\log(\mathsf{T}\rho) - \log\Gamma(\mathsf{T}) - (\mathsf{T}-1)(1+\gamma) + o_\rho(1). \tag{A.96c}$$

*Proof:* We can express $\bar{J}(\rho)$, $I(\rho)$ and $\underline{I}(\rho)$ as (see (4.33), (4.12) and (4.31a))

$$\begin{aligned}\bar{J}(\rho) = (\mathsf{T}-1)\log(\mathsf{T}\rho) - \log\Gamma(\mathsf{T}) - (\mathsf{T}-1)\frac{\mathsf{T}\rho}{1+\mathsf{T}\rho} \\ + (\mathsf{T}-1)\mathsf{E}\left[\log\left(Z_1 + \frac{Z_2}{(1+\mathsf{T}\rho)}\right)\right] \\ + (\mathsf{T}-1)\mathsf{E}\left[\log\left(1 + \frac{\beta(\rho)}{(1+\mathsf{T}\rho)Z_1 + Z_2}\right)\right] \qquad \text{(A.97a)}\end{aligned}$$

$$\begin{aligned}I(\rho) = (\mathsf{T}-1)\log(\mathsf{T}\rho) - \log\Gamma(\mathsf{T}) - (\mathsf{T}-1)\frac{\mathsf{T}\rho}{1+\mathsf{T}\rho} \\ + (\mathsf{T}-1)\mathsf{E}\left[\log\left(Z_1 + \frac{Z_2}{(1+\mathsf{T}\rho)}\right)\right] \\ - \mathsf{E}\left[\log\tilde{\gamma}\left(\mathsf{T}-1, \frac{\mathsf{T}\rho((1+\mathsf{T}\rho)Z_1 + Z_2)}{1+\mathsf{T}\rho}\right)\right] \qquad \text{(A.97b)}\end{aligned}$$

$$\begin{aligned}\underline{I}(\rho) = (\mathsf{T}-1)\log(\mathsf{T}\rho) - \log\Gamma(\mathsf{T}) - (\mathsf{T}-1)\frac{\mathsf{T}\rho}{1+\mathsf{T}\rho} \\ + (\mathsf{T}-1)\mathsf{E}\left[\log\left(Z_1 + \frac{Z_2}{(1+\mathsf{T}\rho)}\right)\right]. \qquad \text{(A.97c)}\end{aligned}$$

Note that these expressions differ only in terms that vanish as $\rho \to \infty$. Indeed, we have

$$(\mathsf{T}-1)\mathsf{E}\left[\log\left(Z_1 + \frac{Z_2}{(1+\mathsf{T}\rho)}\right)\right] = -(\mathsf{T}-1)\gamma + o_\rho(1) \qquad \text{(A.98)}$$

$$(\mathsf{T}-1)\mathsf{E}\left[\log\left(1 + \frac{\beta(\rho)}{(1+\mathsf{T}\rho)Z_1 + Z_2}\right)\right] = o_\rho(1) \qquad \text{(A.99)}$$

$$\mathsf{E}\left[\log\tilde{\gamma}\left(\mathsf{T}-1, \frac{\mathsf{T}\rho((1+\mathsf{T}\rho)Z_1 + Z_2)}{1+\mathsf{T}\rho}\right)\right] = o_\rho(1). \qquad \text{(A.100)}$$

Here, (A.98) follows because, by the dominated convergence theorem,

$$\lim_{\rho\to\infty}\mathsf{E}\left[\log\left(Z_1 + \frac{Z_2}{(1+\mathsf{T}\rho)}\right)\right] = \mathsf{E}\left[\lim_{\rho\to\infty}\log\left(Z_1 + \frac{Z_2}{(1+\mathsf{T}\rho)}\right)\right] \qquad \text{(A.101)}$$

and because $\mathsf{E}[\log Z_1] = -\gamma$. The dominated convergence theorem can be applied since (see (A.25))

$$\left|\log\left(Z_1 + \frac{Z_2}{1+\mathsf{T}\rho}\right)\right| \le \left|\log(Z_1+Z_2)\right| + \left|\log(Z_1)\right| \qquad \text{(A.102)}$$

and $\mathsf{E}\big[\left|\log(Z_1+Z_2)\right| + \left|\log(Z_1)\right|\big] < \infty$.

Similarly, (A.99) and (A.100) follow by the dominated convergence theorem and by noting that the terms inside the expected values on the LHS of (A.99) and (A.100) vanish as $\rho \to \infty$. The dominated convergence theorem can be applied because for every $\rho_0 > 0$ and $\rho \geq \rho_0$

$$\left|\log \tilde{\gamma}\left(\mathsf{T}-1, \frac{\mathsf{T}\rho((1+\mathsf{T}\rho)Z_1+Z_2)}{1+\mathsf{T}\rho}\right)\right| \leq (\mathsf{T}-1)\log\left(1+\frac{\beta(\rho)}{(1+\mathsf{T}\rho)Z_1+Z_2}\right)$$
$$\leq (\mathsf{T}-1)\log\left(1+\frac{\beta(\rho_0)}{Z_1+Z_2}\right) \tag{A.103}$$

and because the expected value of the RHS of (A.103) is finite. Here, the first step follows from Lemma 4.1, and the last follows because $\rho \mapsto \beta(\rho)$ is monotonically decreasing in $\rho$.

Finally, $(\mathsf{T}-1)\frac{\mathsf{T}\rho}{1+\mathsf{T}\rho}$ in (A.97a)–(A.97c) can be expressed as

$$(\mathsf{T}-1)\frac{\mathsf{T}\rho}{1+\mathsf{T}\rho} = (\mathsf{T}-1) + o_\rho(1). \tag{A.104}$$

This establishes (A.96a)–(A.96c). ■

## A.10 High-SNR Approximations of Dispersions

**Lemma A.3** *The quantities $\bar{V}(\rho)$ and $V(\rho)$ defined in* (4.39d) *and* (4.39b)*, respectively, can be approximated as*

$$\bar{V}(\rho) = (\mathsf{T}-1)^2\frac{\pi^2}{6} + (\mathsf{T}-1) + o_\rho(1) \tag{A.105a}$$
$$V(\rho) = (\mathsf{T}-1)^2\frac{\pi^2}{6} + (\mathsf{T}-1) + o_\rho(1). \tag{A.105b}$$

*Proof:* We prove (A.105a) by analyzing $\bar{V}(\rho) \triangleq \mathsf{E}\left[\left(\bar{\jmath}_\ell(\rho) - \bar{J}(\rho)\right)^2\right]$ in the limit as $\rho \to \infty$. To this end, we first note that

$$\begin{aligned}\bar{\jmath}_\ell(\rho) - \bar{J}(\rho) = {} & \frac{\mathsf{T}\rho}{1+\mathsf{T}\rho}(\mathsf{T}-1-Z_2) + (\mathsf{T}-1)\log\left(Z_1 + \frac{Z_2}{1+\mathsf{T}\rho}\right) \\ & - (\mathsf{T}-1)\mathsf{E}\left[\log\left(Z_1+\frac{Z_2}{1+\mathsf{T}\rho}\right)\right] \\ & + (\mathsf{T}-1)\log\left(1+\frac{\beta(\rho)}{(1+\mathsf{T}\rho)Z_1+Z_2}\right) \\ & - (\mathsf{T}-1)\mathsf{E}\left[\log\left(1+\frac{\beta(\rho)}{(1+\mathsf{T}\rho)Z_1+Z_2}\right)\right]\end{aligned}$$

tends to

$$\mathsf{T}-1-Z_2 + (\mathsf{T}-1)\log(Z_1) - (\mathsf{T}-1)\mathsf{E}[\log Z_1] \tag{A.106}$$

as $\rho \to \infty$. (To obtain $\mathsf{E}[\log Z_1]$, we interchange limit and expectation, which can be justified by the dominated convergence theorem.) Since $Z_1$ and $Z_2$ are independent, we have that

$$\begin{aligned}\mathsf{E}\Big[\big(\mathsf{T}-1-Z_2+(\mathsf{T}-1)\log(Z_1)-(\mathsf{T}-1)\mathsf{E}[\log Z_1]\big)^2\Big]\\ &=\mathsf{E}\big[(\mathsf{T}-1-Z_2)^2\big]+(\mathsf{T}-1)^2\Big(\mathsf{E}\big[\log^2(Z_1)\big]-\mathsf{E}\big[\log Z_1\big]^2\Big)\\ &=(\mathsf{T}-1)+(\mathsf{T}-1)^2\frac{\pi^2}{6}. \end{aligned}\tag{A.107}$$

It remains to show that we can swap limit (as $\rho\to\infty$) and expectation. To this end, we next argue that the dominated convergence theorem applies. Indeed, proceeding similarly as in Appendix A.3, we conclude that for every $\rho_0>0$ and $\rho\geq\rho_0$

$$\begin{aligned}\big(\bar{\jmath}_\ell(\rho)-\bar{J}(\rho)\big)^2 &\leq c_{5,2}\Bigg(\left(\frac{\mathsf{T}\rho}{1+\mathsf{T}\rho}\right)^2(Z_2-\mathsf{T}+1)^2\\ &\quad+(\mathsf{T}-1)^2\log^2\left(Z_1+\frac{Z_2}{1+\mathsf{T}\rho}\right)\\ &\quad+(\mathsf{T}-1)^2\mathsf{E}\left[\log\left(Z_1+\frac{Z_2}{1+\mathsf{T}\rho}\right)\right]^2\\ &\quad+(\mathsf{T}-1)^2\log^2\left(1+\frac{\beta(\rho)}{(1+\mathsf{T}\rho)Z_1+Z_2}\right)\\ &\quad+(\mathsf{T}-1)^2\mathsf{E}\left[\log\left(1+\frac{\beta(\rho)}{(1+\mathsf{T}\rho)Z_1+Z_2}\right)\right]^2\Bigg)\\ &\leq c_{5,2}\Bigg((Z_2-\mathsf{T}+1)^2+(\mathsf{T}-1)^2\big(\log^2(Z_1+Z_2)+\log^2(Z_1)\big)\\ &\quad+(\mathsf{T}-1)^2\mathsf{E}\big[\big|\log(Z_1+Z_2)\big|+\big|\log(Z_1)\big|\big]^2\\ &\quad+(\mathsf{T}-1)^2\log^2\left(1+\frac{\beta(\rho_0)}{Z_1+Z_2}\right)\\ &\quad+(\mathsf{T}-1)^2\mathsf{E}\left[\log\left(1+\frac{\beta(\rho_0)}{Z_1+Z_2}\right)\right]^2\Bigg).\end{aligned}$$

To obtain the second inequality, we upper-bound the second term using that (see (A.9))

$$\log^2\left(Z_1+\frac{Z_2}{1+\mathsf{T}\rho}\right)\leq\log^2(Z_1+Z_2)+\log^2(Z_1),\tag{A.108}$$

the third term using that (see (A.25))

$$\left|\log\left(Z_1+\frac{Z_2}{1+\mathsf{T}\rho}\right)\right|\leq\big|\log(Z_1+Z_2)\big|+\big|\log(Z_1)\big|,\tag{A.109}$$

the fourth term using that, for every $\rho_0>0$ and $\rho\geq\rho_0$

$$\begin{aligned}\log^2\left(1+\frac{\beta(\rho)}{(1+\mathsf{T}\rho)Z_1+Z_2}\right)&\leq\log^2\left(1+\frac{\beta(\rho)}{Z_1+Z_2}\right)\\ &\leq\log^2\left(1+\frac{\beta(\rho_0)}{Z_1+Z_2}\right),\end{aligned}\tag{A.110}$$

and the fifth term using that, for every $\rho_0 > 0$ and $\rho \geq \rho_0$,

$$\log\left(1 + \frac{\beta(\rho)}{(1+\mathsf{T}\rho)Z_1 + Z_2}\right) \leq \log\left(1 + \frac{\beta(\rho_0)}{Z_1 + Z_2}\right). \tag{A.111}$$

Since the expected value of the RHS of (A.108) is finite, the dominated convergence theorem applies and (A.105a) follows.

To prove (A.105b) we proceed similarly. Indeed, by Lemma 4.1,

$$\begin{aligned} i_\ell(\rho) - I(\rho) = {} & \frac{\mathsf{T}\rho}{1+\mathsf{T}\rho}(\mathsf{T} - 1 - Z_2) + (\mathsf{T}-1)\log\left(Z_1 + \frac{Z_2}{1+\mathsf{T}\rho}\right) \\ & - (\mathsf{T}-1)\mathsf{E}\left[\log\left(Z_1 + \frac{Z_2}{1+\mathsf{T}\rho}\right)\right] \\ & - \log\tilde{\gamma}\left(\mathsf{T}-1, \frac{\mathsf{T}\rho((1+\mathsf{T}\rho)Z_1 + Z_2)}{1+\mathsf{T}\rho}\right) \\ & + \mathsf{E}\left[\log\tilde{\gamma}\left(\mathsf{T}-1, \frac{\mathsf{T}\rho((1+\mathsf{T}\rho)Z_1 + Z_2)}{1+\mathsf{T}\rho}\right)\right] \end{aligned}$$

tends to (A.106) as $\rho$ tends to infinity. It remains to show that limit (as $\rho \to \infty$) and expectation can be swapped. We next argue that this follows from dominated convergence theorem. Indeed, using (A.18), we obtain for every $\rho_0 > 0$ and $\rho \geq \rho_0$ that

$$\begin{aligned} \big(i_\ell(\rho) - I(\rho)\big)^2 \leq {} & c_{5,2}\Bigg(\left(\frac{\mathsf{T}\rho}{1+\mathsf{T}\rho}\right)^2 (Z_2 - \mathsf{T} + 1)^2 + (\mathsf{T}-1)^2\log^2\left(Z_1 + \frac{Z_2}{1+\mathsf{T}\rho}\right) \\ & + (\mathsf{T}-1)^2\mathsf{E}\left[\log\left(Z_1 + \frac{Z_2}{1+\mathsf{T}\rho}\right)\right]^2 \\ & + \log^2\tilde{\gamma}\left(\mathsf{T}-1, \frac{\mathsf{T}\rho((1+\mathsf{T}\rho)Z_1 + Z_2)}{1+\mathsf{T}\rho}\right) \\ & + \mathsf{E}\left[\log\tilde{\gamma}\left(\mathsf{T}-1, \frac{\mathsf{T}\rho((1+\mathsf{T}\rho)Z_1 + Z_2)}{1+\mathsf{T}\rho}\right)\right]^2\Bigg) \\ \leq {} & c_{5,2}\Bigg((Z_2 - \mathsf{T} + 1)^2 + (\mathsf{T}-1)^2\big(\log^2(Z_1 + Z_2) + \log^2(Z_1)\big) \\ & + (\mathsf{T}-1)^2\mathsf{E}\big[\big|\log(Z_1+Z_2)\big| + \big|\log(Z_1)\big|\big]^2 \\ & + (\mathsf{T}-1)^2\log^2\left(1 + \frac{\beta(\rho_0)}{Z_1 + Z_2}\right) \\ & + (\mathsf{T}-1)^2\mathsf{E}\left[\log\left(1 + \frac{\beta(\rho_0)}{Z_1 + Z_2}\right)\right]^2\Bigg). \end{aligned} \tag{A.112}$$

Here, we upper-bound the first three terms as in (A.108), and the fourth term using Lemma 4.1 and the monotonicity of $\rho \mapsto \beta(\rho)$ which yield

$$\log^2\tilde{\gamma}\left(\mathsf{T}-1, \frac{\mathsf{T}\rho((1+\mathsf{T}\rho)Z_1 + Z_2)}{1+\mathsf{T}\rho}\right) \leq (\mathsf{T}-1)^2\log^2\left(1 + \frac{\beta(\rho_0)}{Z_1 + Z_2}\right) \tag{A.113}$$

for every $\rho_0 > 0$ and $\rho \geq \rho_0$. Furthermore, the last term is upper-bounded using Lemma 4.1 and the monotonicity of $\rho \mapsto \beta(\rho)$:

$$\left| \log \tilde{\gamma}\left( \mathsf{T} - 1, \frac{\mathsf{T}\rho((1+\mathsf{T}\rho)Z_1 + Z_2)}{1+\mathsf{T}\rho} \right) \right| \leq (\mathsf{T}-1) \log\left( 1 + \frac{\beta(\rho_0)}{Z_1 + Z_2} \right). \tag{A.114}$$

Since the expected value of the RHS of (A.112) is finite, the dominated convergence theorem applies and (A.105b) follows. ■

# B

# Appendix to Chapter 6

## B.1 Proof of Lemma 6.3

The proof follows along similar lines as the proof of [22, Ch. XVI.4, Theorem 1]. The main particularity of our result is that it holds uniformly in the parameter $\theta$ of the distribution $\tilde{F}_\theta$, which makes the conditions of our lemma slightly more restrictive in the sense that we require the first four moments of $\tilde{F}_\theta$ to exist, whereas in the original theorem this is required only up to the third moment. In any case, the steps are almost analogous, and we will focus on explaining in detail those steps which require special treatment.

Let us denote the characteristic function of the distribution $\tilde{F}_\theta$ by

$$\tilde{\varphi}_\theta(\zeta) \triangleq \mathsf{E}\Big[e^{\imath\zeta\tilde{X}_k}\Big], \quad \zeta \in \mathbb{R} \tag{B.1}$$

where $\tilde{X}_k \sim \tilde{F}_\theta$, and define

$$G_\theta(x) \triangleq \mathfrak{N}(x) - \frac{\mu_{3,\theta}}{6\sigma_\theta^3\sqrt{n}}(x^2-1)\mathfrak{n}(x), \quad x \in \mathbb{R}. \tag{B.2}$$

Note that (6.30) implies that

$$\sup_{\theta\in\Theta}|\mu_{3,\theta}| < \infty \tag{B.3}$$

since, by Jensen's inequality, $|\mu_{3,\theta}| \leq \mu_{4,\theta}^{3/4}$. Using (B.3) and (6.31), one can show that the derivative of $G_\theta(x)$ is bounded in $\theta \in \Theta$, namely,

$$\begin{aligned}
\sup_{\substack{\theta\in\Theta,\\ x\in\mathbb{R}}} |G'_\theta(x)| &= \sup_{\substack{\theta\in\Theta,\\ x\in\mathbb{R}}} \left| \mathfrak{n}(x) - \frac{\mu_{3,\theta}}{6\sigma_\theta^3\sqrt{n}}\big(2x\mathfrak{n}(x) - (x^2-1)\mathfrak{n}'(x)\big) \right| \\
&\leq \sup_{x\in\mathbb{R}} \mathfrak{n}(x) + \frac{\sup_{\theta\in\Theta}|\mu_{3,\theta}|}{6\inf_{\theta\in\Theta}\sigma_\theta^3\sqrt{n}} \sup_{x\in\mathbb{R}}\big|2x\mathfrak{n}(x) - (x^2-1)\mathfrak{n}'(x)\big| < \infty. \qquad \text{(B.4)}
\end{aligned}$$

The characteristic function of $G_\theta$ is given by

$$\gamma_\theta(\zeta) = e^{-\frac{1}{2}\zeta^2}\left[1 + \frac{\mu_{3,\theta}}{6\sigma_\theta^3\sqrt{n}}(\imath\zeta)^3\right]. \qquad \text{(B.5)}$$

From (B.4) and (B.5), it follows that $G_\theta$ satisfies the conditions of [22, Ch. XVI.3, Lemma 2], namely, that for some positive constant $m$,

$$\sup_{\substack{\theta\in\Theta,\\ x\in\mathbb{R}}} \big|G'_\theta(x)\big| \leq m < \infty \qquad \text{(B.6)}$$

and that $G_\theta$ has a continuously-differentiable characteristic function $\gamma_\theta(\zeta)$ satisfying $\gamma_\theta(0) = 1$ and $\gamma'_\theta(0) = 0$. Then, the inequality [22, Ch. XVI.3, Eq. (3.13)]

$$|\tilde{F}_{n,\theta}(x) - G(x)| \leq \frac{1}{\pi}\int_{-T}^{T} \left| \frac{\tilde{\varphi}_\theta^n\left(\frac{\zeta}{\sigma_\theta\sqrt{n}}\right) - \gamma_\theta(\zeta)}{\zeta} \right| d\zeta + \frac{24m}{\pi T} \qquad \text{(B.7)}$$

and holds for all $x$ and $T > 0$.

Using (B.7) with $T = a\sqrt{n}$, where the constant $a$ is chosen sufficiently large such that $\frac{24m}{\pi} < \epsilon a$ for some $\epsilon$ independent of $x$ and $\theta$, we can write

$$|\tilde{F}_{n,\theta}(x) - G(x)| \leq \frac{1}{\pi}\int_{-a\sqrt{n}}^{a\sqrt{n}} \left| \frac{\tilde{\varphi}_\theta^n\left(\frac{\zeta}{\sigma_\theta\sqrt{n}}\right) - \gamma_\theta(\zeta)}{\zeta} \right| d\zeta + \frac{\epsilon}{\sqrt{n}}. \qquad \text{(B.8)}$$

Choose some $\delta > 0$ independent of $x$ and $\theta$. By assumption, the family of distributions $\tilde{F}_\theta$ (parametrized by $\theta$) is nonlattice, so $\sup_{\theta\in\Theta}|\tilde{\varphi}_\theta(\zeta)|$ is strictly smaller than 1 for every $|\zeta| \geq \delta$. Furthermore, (6.30) implies that the function $\zeta \mapsto \sup_{\theta\in\Theta}\tilde{\varphi}_\theta(\zeta)$ is continuous. Consequently, there exists a number $q_{\delta,\bar{\zeta}} < 1$ (independent of $\theta$) such that

$$\sup_{\theta\in\Theta}|\tilde{\varphi}_\theta(\zeta)| \leq q_{\delta,\bar{\zeta}}, \quad \delta \leq |\zeta| \leq \bar{\zeta} \qquad \text{(B.9)}$$

for some arbitrary $\bar{\zeta} \geq a/\inf_{\theta\in\Theta}\sigma_\theta$.

To prove that $\zeta \mapsto \sup_{\theta\in\Theta}\tilde{\varphi}_\theta(\zeta)$ is continuous, note that, by [22, Ch. XV.4, Lemma 2],

$$\sup_{\theta\in\Theta}|\tilde{\varphi}'_\theta(\zeta)| \leq \sup_{\theta\in\Theta}\mathsf{E}\Big[|\tilde{X}_k|\Big], \quad \zeta\in\mathbb{R} \qquad \text{(B.10)}$$

which by (6.30) is finite. Moreover, for every $\zeta_1, \zeta_2 \in \mathbb{R}$,

$$\begin{aligned}\left|\sup_{\theta\in\Theta} \tilde{\varphi}_\theta(\zeta_1) - \sup_{\theta\in\Theta} \tilde{\varphi}_\theta(\zeta_2)\right| &\leq \sup_{\theta\in\Theta} |\tilde{\varphi}_\theta(\zeta_1) - \tilde{\varphi}_\theta(\zeta_2)| \\ &\leq \sup_{\theta\in\Theta} \mathsf{E}\left[|\tilde{X}_k|\right] |\zeta_1 - \zeta_2| \end{aligned} \tag{B.11}$$

where the second step follows by expanding $\zeta_1 \mapsto \tilde{\varphi}_\theta(\zeta_1)$ as

$$\tilde{\varphi}_\theta(\zeta_1) = \tilde{\varphi}_\theta(\zeta_2) + \tilde{\varphi}'_\theta\big(\tilde{\zeta}\big)(\zeta_1 - \zeta_2) \tag{B.12}$$

for some $\tilde{\zeta} \in (\zeta_1, \zeta_2)$ and by (B.10). Since $\sup_{\theta\in\Theta} \mathsf{E}\left[|\tilde{X}_k|\right]$ is finite by (B.10), it follows that for every $\epsilon > 0$ there exists a $\delta > 0$ such that

$$|\zeta_1 - \zeta_2| \leq \delta \quad \Longrightarrow \quad \left|\sup_{\theta\in\Theta} \tilde{\varphi}_\theta(\zeta_1) - \sup_{\theta\in\Theta} \tilde{\varphi}_\theta(\zeta_2)\right| \leq \epsilon. \tag{B.13}$$

Thus, $\zeta \mapsto \sup_{\theta\in\Theta} \tilde{\varphi}_\theta(\zeta)$ is continuous.

Using (B.9), the contribution of the intervals $|\zeta| > \delta\sigma_\theta\sqrt{n}$ to the integral in (B.8) can be bounded as

$$\begin{aligned} &\frac{2}{\pi}(a\sqrt{n} - \delta\sigma_\theta\sqrt{n}) q^n_{\delta,\bar{\zeta}} + \frac{1}{\pi}\int_{\delta\sigma_\theta\sqrt{n}<|\zeta|<a\sqrt{n}} \left|\frac{\gamma_\theta(\zeta)}{\zeta}\right| d\zeta \\ &\quad \leq \frac{2}{\pi} a\sqrt{n} q^n_{\delta,\bar{\zeta}} + \frac{1}{\pi}\int_{|\zeta|>\delta\sigma_\theta\sqrt{n}} \left|\frac{\gamma_\theta(\zeta)}{\zeta}\right| d\zeta \\ &\quad \leq \frac{2}{\pi} a\sqrt{n} q^n_{\delta,\bar{\zeta}} + \frac{1}{\pi}\int_{|\zeta|>\delta\sqrt{n}\inf_{\theta\in\Theta}\sigma_\theta} \frac{e^{-\frac{1}{2}\zeta^2}\left[1 + \frac{\sup_{\theta\in\Theta}\mu_{3,\theta}}{6\inf_{\theta\in\Theta}\sigma_\theta^3\sqrt{n}}\zeta^3\right]}{\zeta} d\zeta \end{aligned} \tag{B.14}$$

where the last step follows from the definition of $\gamma_\theta$ in (B.5) and by lower-bounding $\sigma_\theta$ by $\inf_{\theta\in\Theta}\sigma_\theta$. The RHS of (B.14) tends to zero faster than any power of $1/n$ uniformly in $\theta$.

We next define

$$\kappa_\theta(\zeta) \triangleq \log\tilde{\varphi}_\theta(\zeta) + \frac{1}{2}\sigma_\theta^2\zeta^2. \tag{B.15}$$

Using (B.14) and (B.15), we can write the RHS of (B.8) as

$$\frac{1}{\pi}\int_{|\zeta|<\delta\sigma_\theta\sqrt{n}} e^{-\frac{1}{2}\zeta^2}\left|\frac{\exp\Big(n\kappa_\theta\Big(\frac{\zeta}{\sigma_\theta\sqrt{n}}\Big)\Big) - 1 - \frac{n\mu_{3,\theta}}{6}\Big(\frac{\imath\zeta}{\sigma_\theta\sqrt{n}}\Big)^3}{\zeta}\right| d\zeta + o_n\bigg(\frac{1}{n}\bigg). \tag{B.16}$$

To estimate the integral in (B.16), we will use that [22, Ch. XVI.2, Eq. (2.8)]

$$\left|e^\alpha - 1 - \beta\right| = \left|\big(e^\alpha - e^\beta\big) + \big(e^\beta - 1 - \beta\big)\right| \leq \left(|\alpha - \beta| + \frac{1}{2}\beta^2\right)e^\gamma \tag{B.17}$$

for any $\gamma \geq \max(|\alpha|, |\beta|)$.

Recall that, by assumption (6.30), the fourth moment $\mu_{4,\theta}$ of the distribution $\tilde{F}_\theta$ satisfies

$$\sup_{\theta\in\Theta} \mu_{4,\theta} < \infty. \tag{B.18}$$

This implies that

$$\sup_{\theta\in\Theta} \int_{-\infty}^{\infty} |x|^\ell d\tilde{F}_\theta(x) < \infty, \quad \ell = 1, 2, 3 \tag{B.19}$$

since, by Jensen's inequality, $|\mu_{\ell,\theta}| \le \mu_{4,\theta}^{\ell/4}$, $\ell = 1, 2, 3$. Then, given an $\epsilon > 0$ independent of $\theta$ and $\zeta$, it is possible to choose $\tilde{\delta}$ (again, independent of $\theta$ and $\zeta$) such that, for $|\zeta| < \tilde{\delta}$,

$$\left|\kappa_\theta(\zeta) - \frac{1}{6}\mu_{3,\theta}(\imath\zeta)^3\right| < \epsilon|\zeta|^3 \tag{B.20}$$

and

$$|\kappa_\theta(\zeta)| < \frac{1}{4}\sigma_\theta^2\zeta^2, \qquad \left|\frac{1}{6}\mu_{3,\theta}(\imath\zeta)^3\right| \le \frac{1}{4}\sigma_\theta^2\zeta^2. \tag{B.21}$$

Indeed, after a Taylor series expansion of $\zeta \mapsto \kappa_\theta(\zeta)$ around $\zeta = 0$, and noting that $\kappa_\theta(0) = \kappa'_\theta(0) = \kappa''_\theta(0)$, the LHS of (B.20) becomes

$$\left|\kappa_\theta(\zeta) - \frac{1}{6}\mu_{3,\theta}(\imath\zeta)^3\right| = \left|\frac{1}{6}\kappa'''_\theta(\tilde{\zeta})\zeta^3 - \frac{1}{6}\mu_{3,\theta}(\imath\zeta)^3\right| \tag{B.22}$$

for some $\tilde{\zeta} \in (0, \zeta)$. Equation (B.19) implies that $\tilde{\varphi}'''_\theta(0)$ exists and [22, Ch. XV.4, Lemma 2]

$$\tilde{\varphi}'''_\theta(0) = \kappa'''_\theta(0) = \imath^3\mu_{3,\theta}. \tag{B.23}$$

Furthermore, (B.18) implies that, for every $\epsilon > 0$, there exists a $\xi > 0$ such that

$$\sup_{\substack{\theta\in\Theta,\\ |\zeta|<\xi}} \left|\tilde{\varphi}_\theta^{(k)}(\zeta) - \tilde{\varphi}_\theta^{(k)}(0)\right| \le \epsilon, \quad k = 0, 1, 2, 3. \tag{B.24}$$

For $k = 0$, this follows from (B.11) and (B.19). In general, following the steps (B.10)–(B.11), it can be shown that

$$\sup_{\substack{\theta\in\Theta,\\ |\zeta|<\xi}} \left|\tilde{\varphi}_\theta^{(k)}(\zeta) - \tilde{\varphi}_\theta^{(k)}(0)\right| \le \sup_{\theta\in\Theta} \mathsf{E}\left[|\tilde{X}_k|^k\right]\xi, \quad k = 0, 1, 2, 3 \tag{B.25}$$

from which (B.24) follows because $\sup_{\theta\in\Theta} \mathsf{E}\left[|\tilde{X}_k|^k\right]$ is, by (B.18), finite. By the definition of $\kappa_\theta(\zeta)$ in (B.15), the $k$-th derivative $\kappa_\theta^{(k)}(\zeta)$ is given by the ratio between a linear combination of derivatives of $\tilde{\varphi}_\theta(\zeta)$ up to order $k$ in the numerator, and $\tilde{\varphi}_\theta(\zeta)^k$ in the denominator. Since $\tilde{\varphi}_\theta(0) = 1$, it follows that (B.24) implies that, for every $\epsilon$, there exists a $\tilde{\delta} > 0$ satisfying

$$\sup_{\substack{\theta\in\Theta,\\ |\zeta|<\tilde{\delta}}} \left|\kappa'''_\theta(\zeta) - \kappa'''_\theta(0)\right| \le 6\epsilon. \tag{B.26}$$

Combining (B.26) with (B.23), (B.22) can be bounded as

$$\left|\frac{1}{6}\kappa_\theta'''(\tilde{\zeta})\zeta^3 - \frac{1}{6}\mu_{3,\theta}(\imath\zeta)^3\right| = \frac{1}{6}|\zeta|^3\left|\kappa_\theta'''(\tilde{\zeta}) - \imath^3\mu_{3,\theta}\right| \le \epsilon|\zeta|^3, \quad |\zeta| < \tilde{\delta}. \tag{B.27}$$

This proves (B.20). The inequalities in (B.21) follow along similar lines.

Finally, using (B.17) together with (B.20) and (B.21), and replacing $\zeta$ by $\frac{\zeta}{n\sigma_\theta}$, we obtain that the integrand in (B.16) is upper-bounded by

$$\begin{aligned}\frac{e^{-\frac{1}{4}\zeta^2}}{|\zeta|}\left(\frac{\epsilon}{\sigma_\theta^3\sqrt{n}}|\zeta|^3 + \frac{\mu_{3,\theta}^2}{72n}\zeta^6\right) &= e^{-\frac{1}{4}\zeta^2}\left(\frac{\epsilon}{\sigma_\theta^3\sqrt{n}}\zeta^2 + \frac{\mu_{3,\theta}^2}{72n}|\zeta|^5\right)\\ &\le e^{-\frac{1}{4}\zeta^2}\left(\frac{\epsilon}{\inf\limits_{\theta\in\Theta}\sigma_\theta^3\sqrt{n}}\zeta^2 + \frac{\sup\limits_{\theta\in\Theta}\mu_{3,\theta}^2}{72n}|\zeta|^5\right), |\zeta| < \delta\sigma_\theta\sqrt{n}.\end{aligned} \tag{B.28}$$

Integrating over $\zeta$, this yields that (B.16) decays faster than $1/\sqrt{n}$ uniformly in $x$ and $\theta$. This concludes the proof of Lemma 6.3.

## B.2 Lattice Distributions and Exponential Tilting

**Lemma B.1** *Let $\varphi_\theta$ denote the characteristic function of some distribution $F_\theta$, and let $\tilde{\varphi}_\theta$ denote the characteristic function of the tilted distribution $\vartheta_{\theta,\tau}$ (cf.* (6.18)*). Then, for every $\zeta \neq 0$,*

$$\sup_{\theta\in\Theta}|\varphi_\theta(\zeta)| < 1 \tag{B.29}$$

*implies that*

$$\sup_{\theta\in\Theta}|\tilde{\varphi}_\theta(\zeta)| < 1. \tag{B.30}$$

*Thus, if a family of distributions is nonlattice, then so is the family of tilted distributions.*

*Proof:* The characteristic function of the tilted random variable $V_\theta$ can be written as

$$\begin{aligned}\tilde{\varphi}_\theta(\zeta) &\triangleq \int_{-\infty}^{\infty} e^{\imath\zeta x} d\vartheta_{\theta,\tau}(x)\\ &= \int_{-\infty}^{\infty} e^{\imath\zeta x} e^{-\psi_\theta(\tau)+\tau\tilde{\gamma}} e^{\tau x} dF_\theta(x)\\ &= e^{-\psi_\theta(\tau)+\tau\tilde{\gamma}} \int_{-\infty}^{\infty} e^{(\imath\zeta+\tau)x} dF_\theta(x)\\ &= \mathsf{E}\left[e^{(\imath\zeta+\tau)(X_\theta-\tilde{\gamma})}\right] e^{-\psi_\theta(\tau)+\tau\tilde{\gamma}}\\ &= \mathsf{E}\left[e^{(\imath\zeta+\tau)X_\theta}\right] e^{-\imath\zeta\tilde{\gamma}} e^{-\psi_\theta(\tau)}\\ &= \mathsf{E}\left[e^{(\imath\zeta+\tau)X_\theta}\right] e^{-\imath\zeta\tilde{\gamma}} \frac{1}{m_\theta(\tau)}\end{aligned} \tag{B.31}$$

where $m_\theta(\tau)$ denotes the MGF of $X_\theta$. It then follows that

$$|\tilde{\varphi}_\theta(\zeta)| = \left|\mathsf{E}\left[e^{(\imath\zeta+\tau)X_\theta}\right]\right| \frac{1}{m_\theta(\tau)}. \tag{B.32}$$

Let $\alpha \in \mathbb{C}$ be such that $|\alpha| = 1$ and

$$|\varphi_\theta(\zeta)| = \alpha \mathsf{E}\left[e^{\imath\zeta X_\theta}\right]. \tag{B.33}$$

Hence, if we write $\alpha = e^{\imath\phi}$ for some phase $\phi$,

$$\begin{aligned}
\left|\mathsf{E}\left[e^{\imath\zeta X_\theta}\right]\right| &= \mathsf{E}\left[\alpha e^{\imath\zeta X_\theta}\right] \\
&= \mathsf{E}[\cos(\zeta X_\theta + \phi)] + \imath \mathsf{E}[\sin(\zeta X_\theta + \phi)] \\
&= \mathsf{E}[\cos(\zeta X_\theta + \phi)]
\end{aligned} \tag{B.34}$$

where $\mathsf{E}[\sin(\zeta X_\theta + \phi)] = 0$ because the absolute value of $\mathsf{E}\left[e^{\imath\zeta X_\theta}\right]$ is real-valued. Likewise, for the tilted random variable $V_\theta$, there exists an $\tilde{\alpha} \in \mathbb{C}$ satisfying $|\tilde{\alpha}| = 1$ and

$$|\tilde{\varphi}_\theta(\zeta)| = \tilde{\alpha} \frac{\mathsf{E}\left[e^{(\imath\zeta+\tau)X_\theta}\right]}{m_\theta(\tau)}. \tag{B.35}$$

Writing $\tilde{\alpha} = e^{\imath\tilde{\phi}}$ for some phase $\tilde{\phi}$, we thus obtain

$$\begin{aligned}
\left|\frac{\mathsf{E}\left[e^{(\imath\zeta+\tau)X_\theta}\right]}{m_\theta(\tau)}\right| &= \frac{\mathsf{E}\left[\tilde{\alpha} e^{(\imath\zeta+\tau)X_\theta}\right]}{m_\theta(\tau)} \\
&= \frac{\mathsf{E}\left[e^{\tau X_\theta}\left(\cos\left(\zeta X_\theta + \tilde{\phi}\right) + \imath \sin\left(\zeta X_\theta + \tilde{\phi}\right)\right)\right]}{m_\theta(\tau)} \\
&= \frac{\mathsf{E}\left[e^{\tau X_\theta} \cos\left(\zeta X_\theta + \tilde{\phi}\right)\right]}{m_\theta(\tau)}
\end{aligned} \tag{B.36}$$

where again $\mathsf{E}\left[e^{\tau X_\theta} \sin(\zeta X_\theta + \phi)\right] = 0$ because the absolute value of $\mathsf{E}\left[e^{(\imath\zeta+\tau)X_\theta}\right]$ is real-valued. It further follows that

$$\begin{aligned}
\mathsf{E}\left[\cos\left(\zeta X_\theta + \tilde{\phi}\right)\right]^2 &\le \mathsf{E}\left[\cos\left(\zeta X_\theta + \tilde{\phi}\right)\right]^2 + \mathsf{E}\left[\sin\left(\zeta X_\theta + \tilde{\phi}\right)\right]^2 \\
&= \left|\mathsf{E}\left[e^{\imath\left(\zeta X_\theta + \tilde{\phi}\right)}\right]\right|^2 \\
&= \left|e^{\imath\tilde{\phi}} \mathsf{E}\left[e^{(i\zeta X_\theta)}\right]\right|^2 \\
&= \left|\mathsf{E}\left[e^{(\imath\zeta X_\theta)}\right]\right|^2 \\
&= \mathsf{E}[\cos(\zeta X_\theta + \phi)]^2
\end{aligned} \tag{B.37}$$

where the last step is due to (B.34). Clearly, $\left|\mathsf{E}\left[e^{\imath\zeta X_\theta}\right]\right| = \mathsf{E}[\cos(\zeta X_\theta + \phi)] \ge 0$. Thus, we have that

$$\mathsf{E}\left[\cos\left(\zeta X_\theta + \tilde{\phi}\right)\right] \le \mathsf{E}[\cos(\zeta X_\theta + \phi)]. \tag{B.38}$$

Let now

$$f(X_\theta) \triangleq 1 - \cos(\zeta X_\theta + \phi) \tag{B.39}$$

and

$$\tilde{f}(X_\theta) \triangleq 1 - \cos\Big(\zeta X_\theta + \tilde{\phi}\Big). \tag{B.40}$$

Note that (B.29) is equivalent to

$$\inf_{\theta \in \Theta} \mathsf{E}[f(X_\theta)] > 0. \tag{B.41}$$

Similarly, (B.30) is implied by

$$\inf_{\theta \in \Theta} \mathsf{E}\Big[e^{\tau X_\theta} \tilde{f}(X_\theta)\Big] > 0 \tag{B.42}$$

because

$$\begin{aligned} 1 - \sup_{\theta \in \Theta} \frac{\mathsf{E}\Big[e^{\tau X_\theta} \cos\Big(\zeta X_\theta + \tilde{\phi}\Big)\Big]}{m_\theta(\tau)} &= \inf_{\theta \in \Theta} \left\{ \frac{\mathsf{E}\big[e^{\tau X_\theta}\big] - \mathsf{E}\Big[e^{\tau X_\theta} \cos\Big(\zeta X_\theta + \tilde{\phi}\Big)\Big]}{m_\theta(\tau)} \right\} \\ &\geq \frac{\inf_{\theta \in \Theta} \Big\{\mathsf{E}\Big[e^{\tau X_\theta} \tilde{f}(X_\theta)\Big]\Big\}}{\sup_{\theta \in \Theta} m_\theta(\tau)} \end{aligned} \tag{B.43}$$

and $\sup_{\theta \in \Theta} m_\theta(\tau) < \infty$ by assumption (6.6).

We next show that

$$\inf_{\theta \in \Theta} \mathsf{E}\Big[e^{\tau X_\theta} \tilde{f}(X_\theta)\Big] = 0 \quad \Longrightarrow \quad \inf_{\theta \in \Theta} \mathsf{E}\Big[\tilde{f}(X_\theta)\Big] = 0. \tag{B.44}$$

Further note that, by (B.38),

$$\mathsf{E}\Big[\tilde{f}(X_\theta)\Big] = \mathsf{E}\Big[1 - \cos\Big(\zeta X_\theta + \tilde{\phi}\Big)\Big] \geq \mathsf{E}[1 - \cos(\lambda X_\theta + \phi)] = \mathsf{E}[f(X_\theta)]. \tag{B.45}$$

Since $f(\cdot)$ is nonnegative, $\inf_{\theta \in \Theta} \mathsf{E}\Big[\tilde{f}(X_\theta)\Big] = 0$ implies that $\inf_{\theta \in \Theta} \mathsf{E}[f(X_\theta)] = 0$. Hence, by reverse logic,

$$\inf_{\theta \in \Theta} \mathsf{E}[f(X_\theta)] > 0 \quad \Longrightarrow \quad \inf_{\theta \in \Theta} \mathsf{E}\Big[e^{\tau X_\theta} \tilde{f}(X_\theta)\Big] > 0 \tag{B.46}$$

which concludes the proof of (B.30).

To prove (B.44), we first note that, for every arbitrary $\delta > 0$,

$$\begin{aligned} \mathsf{E}\Big[e^{\tau X_\theta} \tilde{f}(X_\theta)\Big] &= \mathsf{E}\Big[e^{\tau X_\theta} \tilde{f}(X_\theta) \mathbb{I}\{|X_\theta| \leq \delta\}\Big] + \mathsf{E}\Big[e^{\tau X_\theta} \tilde{f}(X_\theta) \mathbb{I}\{|X_\theta| > \delta\}\Big] \\ &\geq \mathsf{E}\Big[e^{\tau X_\theta} \tilde{f}(X_\theta) \mathbb{I}\{|X_\theta| \leq \delta\}\Big] \\ &\geq \mathsf{E}\Big[\tilde{f}(X_\theta) \mathbb{I}\{|X_\theta| \leq \delta\}\Big] e^{-\tau\delta}. \end{aligned} \tag{B.47}$$

Next note that

$$\begin{aligned}\mathsf{E}\Big[\tilde{f}(X_\theta)\Big] &= \mathsf{E}\Big[\tilde{f}(X_\theta)\mathbb{I}\{|X_\theta| \le \delta\}\Big] + \mathsf{E}\Big[\tilde{f}(X_\theta)\mathbb{I}\{|X_\theta| > \delta\}\Big] \\ &\le \mathsf{E}\Big[\tilde{f}(X_\theta)\mathbb{I}\{|X_\theta| \le \delta\}\Big] + 2\frac{\sup_{\theta\in\Theta}\mathsf{E}\big[|X_\theta|^2\big]}{\delta^2} \end{aligned} \tag{B.48}$$

where the inequality follows because $\tilde{f}(X_\theta)$ is bounded by 2 and by Chebyshev's inequality. By assumption (6.6), we have that $\sup_{\theta\in\Theta}\mathsf{E}\big[X_\theta^2\big] < \infty$. Using (B.47) and (B.48), it follows that

$$\begin{aligned}\inf_{\theta\in\Theta}\mathsf{E}\Big[\tilde{f}(X_\theta)\Big] &\le \inf_{\theta\in\Theta}\mathsf{E}\Big[\tilde{f}(X_\theta)\mathbb{I}\{|X_\theta| \le \delta\}\Big] + 2\frac{\sup_{\theta\in\Theta}\mathsf{E}\big[|X_\theta|^2\big]}{\delta^2} \\ &\le e^{\tau\delta}\inf_{\theta\in\Theta}\mathsf{E}\Big[e^{\tau X_\theta}\tilde{f}(X_\theta)\Big] + 2\frac{\sup_{\theta\in\Theta}\mathsf{E}\big[|X_\theta|^2\big]}{\delta^2}. \end{aligned} \tag{B.49}$$

If $\inf_{\theta\in\Theta}\mathsf{E}\Big[e^{\tau X_\theta}\tilde{f}(X_\theta)\Big] = 0$ then, for every arbitrary $\delta > 0$,

$$\inf_{\theta\in\Theta}\mathsf{E}\Big[\tilde{f}(X_\theta)\Big] \le 2\frac{\sup_{\theta\in\Theta}\mathsf{E}\big[|X_\theta|^2\big]}{\delta^2}. \tag{B.50}$$

Thus, by letting $\delta \to \infty$, we obtain that $\inf_{\theta\in\Theta}\mathsf{E}\Big[\tilde{f}(X_\theta)\Big] \le 0$. Since $\tilde{f}(\cdot)$ is nonnegative, we conclude that $\inf_{\theta\in\Theta}\mathsf{E}\Big[\tilde{f}(X_\theta)\Big] = 0$, hence (B.44) follows. ■

## B.3 $I_s(\rho) - i_{s,\ell}(\rho)$ Is Nonlattice

Consider $I_s(\rho)$ defined in (4.27) and $i_{s,\ell}(\rho)$ defined in (4.12), and let

$$\varphi_{\rho,s}(\tau) \triangleq \mathsf{E}\Big[e^{\imath\tau(I_s(\rho)-i_{s,\ell}(\rho))}\Big]. \tag{B.51}$$

We have the following result.

**Lemma B.2** *For every $\rho_0 > 0$, $0 < s_0 < s_{\max}$, and $\delta > 0$, we have*

$$\sup_{\substack{\rho\ge\rho_0,\\ s\in[s_0,s_{\max}]}} |\varphi_{\rho,s}(\tau)| < 1, \quad |\tau| \ge \delta. \tag{B.52}$$

*Proof:* We prove (B.52) in two steps. We first show that, for every $\rho_{\max}$, we have

$$\sup_{\rho_0\le\rho\le\rho_{\max}} \sup_{s\in[s_0,s_{\max}]} |\varphi_{\rho,s}(\tau)| < 1, \quad |\tau| \ge \delta. \tag{B.53}$$

We then show that

$$\lim_{\rho\to\infty} \sup_{s\in[s_0,s_{\max}]} |\varphi_{\rho,s}(\tau)| < 1, \quad |\tau| \ge \delta. \tag{B.54}$$

To prove (B.53), we note that $I_s(\rho) - i_{s,\ell}(\rho)$ is a continuous function of the gamma-distributed random variables $Z_1$ and $Z_2$. Consequently, $I_s(\rho) - i_{s,\ell}(\rho)$ is nonlattice, so $|\varphi_{\rho,s}(\tau)| < 1$, $|\tau| \geq \delta$ for every $\delta > 0$. Since $\tau \mapsto \varphi_{\rho,s}(\tau)$ is continuous and the suprema in (B.53) are over the bounded intervals $[\rho_0, \rho_{\max}]$ and $[s_0, s_{\max}]$, the claim (B.53) follows.

We next prove (B.54). Define

$$B_\rho \triangleq (\mathsf{T}-1)\log(s\mathsf{T}\rho) - \log\Gamma(\mathsf{T}) \tag{B.55}$$

and note that

$$|\varphi_{\rho,s}(\tau)| = \left|e^{\imath\tau(I_s(\rho)-B_\rho)}\mathsf{E}\left[e^{-\imath\tau(i_{s,\ell}(\rho)-B_\rho)}\right]\right| = \left|\mathsf{E}\left[e^{-\imath\tau(i_{s,\ell}(\rho)-B_\rho)}\right]\right|. \tag{B.56}$$

Let

$$\Lambda_{s,\tau}(Z_1,Z_2) \triangleq -\tau(-sZ_2 + (\mathsf{T}-1)\log(Z_1)) \tag{B.57a}$$

$$\begin{aligned}\Pi_{\rho,s,\tau}(Z_1,Z_2) \triangleq -\tau\Bigg(&sZ_2\left(1-\frac{\mathsf{T}\rho}{1+\mathsf{T}\rho}\right) - (\mathsf{T}-1)\log(Z_1) + \log\left(Z_1 + \frac{Z_2}{1+\mathsf{T}\rho}\right)\\ &- \log\tilde{\gamma}\left(\mathsf{T}-1, s\frac{\mathsf{T}\rho((1+\mathsf{T}\rho)Z_1+Z_2)}{1+\mathsf{T}\rho}\right)\Bigg).\end{aligned} \tag{B.57b}$$

Using (B.57a) and (B.57b), we can write the RHS of (B.56) as

$$\left|\mathsf{E}\left[e^{-\imath\tau(i_{s,\ell}(\rho)-B_\rho)}\right]\right| = \left|\mathsf{E}\left[e^{\imath\Lambda_{s,\tau}(Z_1,Z_2)}e^{\imath\Pi_{\rho,s,\tau}(Z_1,Z_2)}\right]\right|. \tag{B.58}$$

We next show that

$$\lim_{\rho\to\infty}\sup_{s\in[s_0,s_{\max}]}\left|\left|\mathsf{E}\left[e^{\imath\Lambda_{s,\tau}(Z_1,Z_2)}e^{\imath\Pi_{\rho,s,\tau}(Z_1,Z_2)}\right]\right| - \left|\mathsf{E}\left[e^{\imath\Lambda_{s,\tau}(Z_1,Z_2)}\right]\right|\right| = 0. \tag{B.59}$$

It then follows that

$$\begin{aligned}\lim_{\rho\to\infty}\sup_{s\in[s_0,s_{\max}]}\left|\mathsf{E}\left[e^{-\imath\tau(i_{s,\ell}(\rho)-B_\rho)}\right]\right| &= \sup_{s\in[s_0,s_{\max}]}\left|\mathsf{E}\left[e^{\imath\Lambda_{s,\tau}(Z_1,Z_2)}\right]\right|\\ &= \sup_{s\in[s_0,s_{\max}]}\frac{|\Gamma(1-\imath\tau(\mathsf{T}-1))|}{|(1-\imath\tau s)^{\mathsf{T}-1}|}\\ &< 1, \quad |\tau| \geq \delta\end{aligned} \tag{B.60}$$

where the second equality follows because

$$\mathsf{E}\left[e^{\imath\Lambda_{s,\tau}(Z_1,Z_2)}\right] = \mathsf{E}\left[e^{\imath\tau sZ_2}e^{-\imath\tau(\mathsf{T}-1)\log Z_1}\right] = \frac{\Gamma(1-\imath\tau(\mathsf{T}-1))}{(1-i\tau s)^{\mathsf{T}-1}} \tag{B.61}$$

and the inequality follows because

$$|\Gamma(1-\imath\tau(\mathsf{T}-1))| \leq \Gamma(1) = 1 \tag{B.62a}$$

and

$$\inf_{s\in[s_0,s_{\max}]}\left|(1-\imath\tau s)^{\mathsf{T}-1}\right| = \left(1+\tau^2 s_0^2\right)^{\frac{\mathsf{T}-1}{2}} > 1, \quad |\tau|\geq\delta. \tag{B.62b}$$

This concludes the proof of (B.54).

It remains to prove (B.59). In the following, we shorten the notation of $\Lambda_{s,\tau}(Z_1,Z_2)$ and $\Pi_{\rho,s,\tau}(Z_1,Z_2)$ by omitting the arguments $(Z_1,Z_2)$ and the subindexes $(\rho,s,\tau)$. The LHS of (B.59) can be upper-bounded as

$$\begin{aligned}
&\lim_{\rho\to\infty}\sup_{s\in[s_0,s_{\max}]}\left|\left|\mathsf{E}\left[e^{\imath\Lambda}e^{\imath\Pi}\right]\right|-\left|\mathsf{E}\left[e^{\imath\Lambda}\right]\right|\right| \\
&\quad\leq \lim_{\rho\to\infty}\sup_{s\in[s_0,s_{\max}]}\left|\mathsf{E}\left[e^{\imath\Lambda}e^{\imath\Pi}\right]-\mathsf{E}\left[e^{\imath\Lambda}\right]\right| \\
&\quad= \lim_{\rho\to\infty}\sup_{s\in[s_0,s_{\max}]}\left|\mathsf{E}[\cos(\Lambda+\Pi)-\cos(\Lambda)]+\imath\mathsf{E}[\sin(\Lambda+\Pi)-\sin(\Lambda)]\right| 
\end{aligned}\tag{B.63}$$

where the inequality follows by the triangle inequality. Evaluating the absolute value, the RHS of (B.63) can be upper-bounded as

$$\begin{aligned}
&\lim_{\rho\to\infty}\sup_{s\in[s_0,s_{\max}]}\left|\mathsf{E}[\cos(\Lambda+\Pi)-\cos(\Lambda)]+\imath\mathsf{E}[\sin(\Lambda+\Pi)-\sin(\Lambda)]\right| \\
&\quad= \lim_{\rho\to\infty}\sup_{s\in[s_0,s_{\max}]}\sqrt{\left|\mathsf{E}[\cos(\Lambda+\Pi)-\cos(\Lambda)]\right|^2+\left|\mathsf{E}[\sin(\Lambda+\Pi)-\sin(\Lambda)]\right|^2} \\
&\quad\leq \lim_{\rho\to\infty}\sup_{s\in[s_0,s_{\max}]}\sqrt{\mathsf{E}^2[|\cos(\Lambda+\Pi)-\cos(\Lambda)|]+\mathsf{E}^2[|\sin(\Lambda+\Pi)-\sin(\Lambda)|]}
\end{aligned}\tag{B.64}$$

where the last step follows by the triangle inequality. We next perform Taylor series expansions to express $\cos(\Lambda+\Pi)$ and $\sin(\Lambda+\Pi)$ as

$$\cos(\Lambda+\Pi) = \cos(\Lambda)-\Pi\sin(\theta_1) \tag{B.65a}$$

$$\sin(\Lambda+\Pi) = \sin(\Lambda)+\Pi\cos(\theta_2) \tag{B.65b}$$

for some $\theta_1,\theta_2\in(0,\Pi)$. Substituting (B.65) into the RHS of (B.64) we obtain that

$$\begin{aligned}
&\lim_{\rho\to\infty}\sup_{s\in[s_0,s_{\max}]}\sqrt{\mathsf{E}^2[|\cos(\Lambda+\Pi)-\cos(\Lambda)|]+\mathsf{E}^2[|\sin(\Lambda+\Pi)-\sin(\Lambda)|]} \\
&\quad= \lim_{\rho\to\infty}\sup_{s\in[s_0,s_{\max}]}\sqrt{\mathsf{E}^2[|\Pi\sin(\theta_1)|]+\mathsf{E}^2[|\Pi\cos(\theta_2)|]} \\
&\quad\leq \lim_{\rho\to\infty}\sup_{s\in[s_0,s_{\max}]}\sqrt{2}\mathsf{E}[|\Pi|]
\end{aligned}\tag{B.66}$$

where the last step follows because $|\sin(\cdot)|\leq 1$ and $|\cos(\cdot)|\leq 1$. We next show that

$$\lim_{\rho\to\infty}\sup_{s\in[s_0,s_{\max}]}\mathsf{E}[|\Pi_{\rho,s,\tau}(Z_1,Z_2)|]=0 \tag{B.67}$$

which then together with (B.63)–(B.66) yields (B.59). To show (B.67), we first note that

$$\mathsf{E}[|\Pi_{\rho,s,\tau}(Z_1,Z_2)|] = \tau\bigg\{s(\mathsf{T}-1)\left(1-\frac{\mathsf{T}\rho}{1+\mathsf{T}\rho}\right) + \mathsf{E}\left[\log\left(Z_1+\frac{Z_2}{1+\mathsf{T}\rho}\right)\right] \\ - \mathsf{E}[\log(Z_1)] + \mathsf{E}\left[-\log\tilde{\gamma}\left(\mathsf{T}-1, s\frac{\mathsf{T}\rho((1+\mathsf{T}\rho)Z_1+Z_2)}{1+\mathsf{T}\rho}\right)\right]\bigg\} \tag{B.68}$$

since

$$\log\left(Z_1+\frac{Z_2}{(1+\mathsf{T}\rho)}\right) - \log(Z_1) \geq 0 \tag{B.69a}$$

$$-\log\tilde{\gamma}\left(\mathsf{T}-1, s\frac{\mathsf{T}\rho((1+\mathsf{T}\rho)Z_1+Z_2)}{1+\mathsf{T}\rho}\right) \geq 0. \tag{B.69b}$$

Thus, by the monotonicity of the regularized lower incomplete gamma function,

$$\sup_{s\in[s_0,s_{\max}]} \mathsf{E}[|\Pi_{\rho,s,\tau}(Z_1,Z_2)|] \\ \leq \tau\bigg\{s_{\max}(\mathsf{T}-1)\left(1-\frac{\mathsf{T}\rho}{1+\mathsf{T}\rho}\right) + \mathsf{E}\left[\log\left(Z_1+\frac{Z_2}{1+\mathsf{T}\rho}\right)\right] - \mathsf{E}[\log(Z_1)] \\ + \mathsf{E}\left[-\log\tilde{\gamma}\left(\mathsf{T}-1, s_0\frac{\mathsf{T}\rho((1+\mathsf{T}\rho)Z_1+Z_2)}{1+\mathsf{T}\rho}\right)\right]\bigg\}. \tag{B.70}$$

We next use that

$$\lim_{\rho\to\infty} s_{\max}(\mathsf{T}-1)\left(1-\frac{\mathsf{T}\rho}{1+\mathsf{T}\rho}\right) = 0. \tag{B.71}$$

Furthermore, by the dominated convergence theorem,

$$\lim_{\rho\to\infty} \mathsf{E}\left[\log\left(Z_1+\frac{Z_2}{1+\mathsf{T}\rho}\right)\right] = \mathsf{E}\left[\lim_{\rho\to\infty}\log\left(Z_1+\frac{Z_2}{1+\mathsf{T}\rho}\right)\right] = \mathsf{E}[\log(Z_1)] \tag{B.72}$$

and

$$\lim_{\rho\to\infty} \mathsf{E}\left[-\log\tilde{\gamma}\left(\mathsf{T}-1, s_0\frac{\mathsf{T}\rho((1+\mathsf{T}\rho)Z_1+Z_2)}{1+\mathsf{T}\rho}\right)\right] \\ = \mathsf{E}\left[\lim_{\rho\to\infty} -\log\tilde{\gamma}\left(\mathsf{T}-1, s_0\frac{\mathsf{T}\rho((1+\mathsf{T}\rho)Z_1+Z_2)}{1+\mathsf{T}\rho}\right)\right] \\ = 0. \tag{B.73}$$

Indeed, the dominated convergence theorem can be applied in (B.72) because

$$\left|\log\left(Z_1+\frac{Z_2}{1+\mathsf{T}\rho}\right)\right| \leq |\log(Z_1+Z_2)| + |\log(Z_1)| \tag{B.74}$$

and $\mathsf{E}[|\log(Z_1+Z_2)|+|\log(Z_1)|] < \infty$. Likewise, the dominated convergence theorem can be applied in (B.73) because

$$\left|\log\tilde{\gamma}\left(\mathsf{T}-1, \frac{\mathsf{T}\rho((1+\mathsf{T}\rho)Z_1+Z_2)}{1+\mathsf{T}\rho}\right)\right| \leq (\mathsf{T}-1)\log\left(1+\frac{\beta(\rho)}{(1+\mathsf{T}\rho)Z_1+Z_2}\right) \\ \leq (\mathsf{T}-1)\log\left(1+\frac{\beta(\rho_0)}{Z_1+Z_2}\right), \quad \rho\geq\rho_0 \tag{B.75}$$

and because the expected value of the RHS of (B.75) is finite. (In (B.75), we define $\beta(\rho) \triangleq \Gamma(\mathsf{T})^{\frac{1}{\mathsf{T}-1}} \frac{1+\mathsf{T}\rho}{\mathsf{T}\rho}$.) Combining (B.71)–(B.73) with (B.70) yields (B.67). ■

## B.4 Second Derivative of CGF Bounded Away from Zero

Let $X_\theta$ be a zero-mean random variable parametrized by $\theta$, whose MGF and CGF are defined in (6.1) and (6.2), respectively. We have the following result.

**Lemma B.3** *Assume that there exists a $\zeta_0 > 0$ such that*

$$\sup_{\substack{\theta\in\Theta,\\ |\zeta|<\zeta_0}} \left|m_\theta^{(k)}(\zeta)\right| < \infty, \quad k = 0,1,2,3,4 \tag{B.76}$$

*and that*

$$\inf_{\theta\in\Theta} \psi_\theta''(0) > 0. \tag{B.77}$$

*Then*

$$\inf_{\substack{\theta\in\Theta,\\ |\zeta|<\zeta_0}} \psi_\theta''(\zeta) > 0. \tag{B.78}$$

*Proof:* The LHS of (B.78) can be lower-bounded as

$$\begin{aligned}\inf_{\substack{\theta\in\Theta,\\ |\zeta|<\zeta_0}} |\psi_\theta''(\zeta)| &= \inf_{\substack{\theta\in\Theta,\\ |\zeta|<\zeta_0}} \frac{1}{m_\theta(\zeta)^2}\left\{\mathsf{E}\big[X_\theta^2 e^{\zeta X_\theta}\big]\,\mathsf{E}\big[e^{\zeta X_\theta}\big] - \mathsf{E}\big[X_\theta e^{\zeta X_\theta}\big]^2\right\} \\ &\geq \frac{1}{\sup\limits_{\substack{\theta\in\Theta,\\ |\zeta|<\zeta_0}} m_\theta(\zeta)^2} \inf_{\substack{\theta\in\Theta,\\ |\zeta|<\zeta_0}} \left\{\mathsf{E}\big[X_\theta^2 e^{\zeta X_\theta}\big]\,\mathsf{E}\big[e^{\zeta X_\theta}\big] - \mathsf{E}\big[X_\theta e^{\zeta X_\theta}\big]^2\right\}. \end{aligned} \tag{B.79}$$

By (B.76), the first term in (B.79) is bounded away from zero. Thus, in order to show (B.78), it suffices to show that

$$\inf_{\substack{\theta\in\Theta,\\ |\zeta|<\zeta_0}} \left\{\mathsf{E}\big[X_\theta^2 e^{\zeta X_\theta}\big]\,\mathsf{E}\big[e^{\zeta X_\theta}\big] - \mathsf{E}\big[X_\theta e^{\zeta X_\theta}\big]^2\right\} > 0. \tag{B.80}$$

To shorten notation, we next define

$$A \triangleq X_\theta e^{\frac{\zeta}{2} X_\theta} \tag{B.81a}$$

$$B \triangleq e^{\frac{\zeta}{2} X_\theta} \tag{B.81b}$$

as well as $\sigma_A^2 \triangleq \mathsf{E}\big[A^2\big]$ and $\sigma_B^2 \triangleq \mathsf{E}\big[B^2\big]$. Hence, (B.80) can be written as

$$\inf_{\substack{\theta\in\Theta,\\ |\tau|<\zeta}} \left\{\sigma_A^2\sigma_B^2 - \mathsf{E}[AB]^2\right\} > 0. \tag{B.82}$$

By following the proof of the Cauchy-Schwarz inequality [46, Th. 3.3.1], it can be shown that

$$\mathsf{E}[AB] \le \sigma_A\sigma_B \left(1 - \frac{1}{2}\inf_{\substack{\theta\in\Theta,\\ |\zeta|<\zeta_0}} \mathsf{E}\left[\left(\frac{A}{\sigma_A} - \frac{B}{\sigma_B}\right)^2\right]\right)^+ \tag{B.83a}$$

$$\mathsf{E}[AB] \ge -\sigma_A\sigma_B \left(1 - \frac{1}{2}\inf_{\substack{\theta\in\Theta,\\ |\zeta|<\zeta_0}} \mathsf{E}\left[\left(\frac{A}{\sigma_A} + \frac{B}{\sigma_B}\right)^2\right]\right)^+ . \tag{B.83b}$$

Consequently,

$$|\mathsf{E}[AB]| \le \sigma_A\sigma_B \max\left\{ \left(1 - \frac{1}{2}\inf_{\substack{\theta\in\Theta,\\ |\zeta|<\zeta_0}} \mathsf{E}\left[\left(\frac{A}{\sigma_A} - \frac{B}{\sigma_B}\right)^2\right]\right)^+, \left(1 - \frac{1}{2}\inf_{\substack{\theta\in\Theta,\\ |\zeta|<\zeta_0}} \mathsf{E}\left[\left(\frac{A}{\sigma_A} + \frac{B}{\sigma_B}\right)^2\right]\right)^+ \right\}. \tag{B.84}$$

Using (B.84), we can lower-bound the LHS of (B.82) as

$$\begin{aligned}\inf_{\substack{\theta\in\Theta,\\ |\zeta|<\zeta_0}} \left\{\sigma_A^2\sigma_B^2 - \mathsf{E}[AB]^2\right\} &\ge \inf_{\substack{\theta\in\Theta,\\ |\zeta|<\zeta_0}} \mathsf{E}\big[A^2\big] \inf_{\substack{\theta\in\Theta,\\ |\zeta|<\zeta_0}} \mathsf{E}\big[B^2\big] \\ &\times \left(1 - \max\left\{ \left[\left(1 - \frac{1}{2}\inf_{\substack{\theta\in\Theta,\\ |\zeta|<\zeta_0}} \mathsf{E}\left[\left(\frac{A}{\sigma_A} - \frac{B}{\sigma_B}\right)^2\right]\right)^+\right]^2, \left[\left(1 - \frac{1}{2}\inf_{\substack{\theta\in\Theta,\\ |\zeta|<\zeta_0}} \mathsf{E}\left[\left(\frac{A}{\sigma_A} + \frac{B}{\sigma_B}\right)^2\right]\right)^+\right]^2 \right\}\right).\end{aligned} \tag{B.85}$$

Thus, in order to show (B.82), it suffices to prove that

$$\inf_{\substack{\theta\in\Theta,\\ |\zeta|<\zeta_0}} \mathsf{E}\big[A^2\big] > 0 \tag{B.86a}$$

$$\inf_{\substack{\theta\in\Theta,\\ |\zeta|<\zeta_0}} \mathsf{E}\big[B^2\big] > 0 \tag{B.86b}$$

$$\inf_{\substack{\theta\in\Theta,\\ |\zeta|<\zeta_0}} \mathsf{E}\left[\left(\frac{A}{\sigma_A} - \frac{B}{\sigma_B}\right)^2\right] > 0 \tag{B.86c}$$

$$\inf_{\substack{\theta\in\Theta,\\ |\zeta|<\zeta_0}} \mathsf{E}\left[\left(\frac{A}{\sigma_A} + \frac{B}{\sigma_B}\right)^2\right] > 0. \tag{B.86d}$$

To prove (B.86a), recall that, by (B.81a),

$$\inf_{\substack{\theta\in\Theta,\\ |\zeta|<\zeta_0}} \mathsf{E}\big[A^2\big] = \inf_{\substack{\theta\in\Theta,\\ |\zeta|<\zeta_0}} \mathsf{E}\big[X_\theta^2 e^{\zeta X_\theta}\big]. \tag{B.87}$$

We next show that

$$\inf_{\substack{\theta\in\Theta,\\ |\zeta|<\zeta_0}} \mathsf{E}\big[X_\theta^2 e^{\zeta X_\theta}\big] = 0 \quad \Longrightarrow \quad \inf_{\theta\in\Theta} \mathsf{E}\big[X_\theta^2\big] = 0. \tag{B.88}$$

Since $\mathsf{E}\big[X_\theta^2\big] = \psi_\theta''(0)$, it follows by assumption (B.77) that $\inf_{\theta\in\Theta} \mathsf{E}\big[X_\theta^2\big] > 0$. Hence, by reverse logic, (B.88) implies that $\inf_{\substack{\theta\in\Theta,\\ |\zeta|<\zeta_0}} \mathsf{E}\big[X_\theta^2 e^{\zeta X_\theta}\big] > 0$, which is (B.86a).

To prove (B.88), we first note that, for every arbitrary $\delta > 0$,

$$\begin{aligned}\mathsf{E}\big[X_\theta^2 e^{\zeta X_\theta}\big] &= \mathsf{E}\big[X_\theta^2 e^{\zeta X_\theta}\mathbb{I}\{|X_\theta|\le\delta\}\big] + \mathsf{E}\big[X_\theta^2 e^{\zeta X_\theta}\mathbb{I}\{|X_\theta|>\delta\}\big] \\ &\ge \mathsf{E}\big[X_\theta^2\mathbb{I}\{|X_\theta|\le\delta\}\big]\, e^{-\zeta\delta}.\end{aligned} \tag{B.89}$$

We further have that

$$\begin{aligned}\mathsf{E}\big[X_\theta^2\big] &= \mathsf{E}\big[X_\theta^2\mathbb{I}\{|X_\theta|\le\delta\}\big] + \mathsf{E}\big[X_\theta^2\mathbb{I}\{|X_\theta|>\delta\}\big] \\ &\le \mathsf{E}\big[X_\theta^2\mathbb{I}\{|X_\theta|\le\delta\}\big] + \frac{\sqrt{\sup_{\theta\in\Theta}\mathsf{E}[X_\theta^4]\sup_{\theta\in\Theta}\mathsf{E}[X_\theta^2]}}{\delta}\end{aligned} \tag{B.90}$$

by the Cauchy-Schwarz and the Chebyshev inequality. Using (B.89) and (B.90), it follows that

$$\begin{aligned}\inf_{\theta\in\Theta}\mathsf{E}\big[X_\theta^2\big] &\le \inf_{\theta\in\Theta}\mathsf{E}\big[X_\theta^2\mathbb{I}\{|X_\theta|\le\delta\}\big] + \frac{\sqrt{\sup_{\theta\in\Theta}\mathsf{E}[X_\theta^4]\sup_{\theta\in\Theta}\mathsf{E}[X_\theta^2]}}{\delta} \\ &\le e^{\zeta\delta}\inf_{\substack{\theta\in\Theta,\\ |\zeta|<\zeta_0}}\mathsf{E}\big[X_\theta^2 e^{\zeta X_\theta}\big] + \frac{\sqrt{\sup_{\theta\in\Theta}\mathsf{E}[X_\theta^4]\sup_{\theta\in\Theta}\mathsf{E}[X_\theta^2]}}{\delta}.\end{aligned} \tag{B.91}$$

Thus, if $\inf_{\substack{\theta\in\Theta,\\ |\zeta|<\zeta}}\mathsf{E}\big[X_\theta^2 e^{\zeta X_\theta}\big] = 0$ then, for every arbitrary $\delta > 0$,

$$\inf_{\theta\in\Theta}\mathsf{E}\big[X_\theta^2\big] \le \frac{\sqrt{\sup_{\theta\in\Theta}\mathsf{E}[X_\theta^4]\sup_{\theta\in\Theta}\mathsf{E}[X_\theta^2]}}{\delta}. \tag{B.92}$$

Since the suprema on the RHS of (B.92) are bounded by assumption, we obtain that $\inf_{\theta\in\Theta}\mathsf{E}\big[X_\theta^2\big] \le 0$ upon letting $\delta\to\infty$. Since $X_\theta^2$ is nonnegative, the claim (B.88) follows.

To prove (B.86b), recall that $\mathsf{E}[B] = \mathsf{E}\big[e^{\zeta X_\theta}\big]$. Since $X_\theta$ is zero-mean by assumption, it follows by Jensen's inequality that

$$\mathsf{E}\big[e^{\zeta X_\theta}\big] \ge 1. \tag{B.93}$$

Hence, the claim follows.

We next show (B.86c). Using (B.81a) and (B.81b), we can lower-bound the LHS of (B.86c) by

$$\frac{1}{\sup_{\substack{\theta\in\Theta,\\ |\zeta|<\zeta_0}}\mathsf{E}[X_\theta^2 e^{\zeta X_\theta}]}\inf_{\substack{\theta\in\Theta,\\ |\zeta|<\zeta_0}}\mathsf{E}\left[e^{\zeta X_\theta}\left(X_\theta - \sqrt{\frac{\mathsf{E}[X_\theta^2 e^{\zeta X_\theta}]}{\mathsf{E}[e^{\zeta X_\theta}]}}\right)^2\right]. \tag{B.94}$$

The first term is bounded away from zero by assumption (B.76). We next show that

$$\inf_{\substack{\theta\in\Theta,\\ |\zeta|<\zeta_0}} \mathsf{E}\left[e^{\zeta X_\theta}\left(X_\theta - \sqrt{\frac{\mathsf{E}[X_\theta^2 e^{\zeta X_\theta}]}{\mathsf{E}[e^{\zeta X_\theta}]}}\right)^2\right] > 0. \tag{B.95}$$

To this end, we follow along the steps (B.88)–(B.92) used to show (B.86a), but replacing $X_\theta^2$ by

$$\left(X_\theta - \eta_{\theta,\zeta}\right)^2 \tag{B.96}$$

where

$$\eta_{\theta,\zeta} \triangleq \sqrt{\frac{\mathsf{E}[X_\theta^2 e^{\zeta X_\theta}]}{\mathsf{E}[e^{\zeta X_\theta}]}}. \tag{B.97}$$

Specifically, we shall show that

$$\inf_{\substack{\theta\in\Theta,\\ |\zeta|<\zeta_0}} \mathsf{E}\left[\left(X_\theta - \eta_{\theta,\zeta}\right)^2 e^{\zeta X_\theta}\right] = 0 \quad \Longrightarrow \quad \inf_{\substack{\theta\in\Theta,\\ |\zeta|<\zeta_0}} \mathsf{E}\left[\left(X_\theta - \eta_{\theta,\zeta}\right)^2\right] = 0. \tag{B.98}$$

Since $X_\theta$ is zero mean, we have that $\mathsf{E}\left[\left(X_\theta - \eta_{\theta,\zeta}\right)^2\right] \geq \mathsf{E}\left[X_\theta^2\right]$, so if $\inf_{\substack{\theta\in\Theta,\\ |\zeta|<\zeta_0}} \mathsf{E}\left[\left(X_\theta - \eta_{\theta,\zeta}\right)^2\right] = 0$, then $\inf_{\theta\in\Theta} \mathsf{E}\left[X_\theta^2\right] = 0$, too. Furthermore, by assumption (B.77), $\inf_{\theta\in\Theta} \mathsf{E}\left[X_\theta^2\right] > 0$. Hence, by reverse logic, (B.98) implies that $\inf_{\substack{\theta\in\Theta,\\ |\zeta|<\zeta_0}} \mathsf{E}\left[\left(X_\theta - \eta_{\theta,\zeta}\right)^2 e^{\zeta X_\theta}\right] > 0$, which is (B.86c).

It remains to prove (B.98). Indeed, following the steps (B.89)–(B.92) but with $X_\theta^2$ replaced by $\left(X_\theta - \eta_{\theta,\zeta}\right)^2$, we obtain that, if $\inf_{\substack{\theta\in\Theta,\\ |\zeta|<\zeta_0}} \mathsf{E}\left[\left(X_\theta - \eta_{\theta,\zeta}\right)^2 e^{\zeta X_\theta}\right] = 0$, then

$$\inf_{\substack{\theta\in\Theta,\\ |\zeta|<\zeta_0}} \mathsf{E}\left[\left(X_\theta - \eta_{\theta,\zeta}\right)^2\right] \leq \frac{\sqrt{\sup_{\substack{\theta\in\Theta,\\ |\zeta|<\zeta_0}} \mathsf{E}\left[\left(X_\theta - \eta_{\theta,\zeta}\right)^4\right] \sup_{\substack{\theta\in\Theta,\\ |\zeta|<\zeta_0}} \mathsf{E}\left[\left(X_\theta - \eta_{\theta,\zeta}\right)^2\right]}}{\delta}. \tag{B.99}$$

The suprema in (B.99) are bounded. Indeed, using that

$$|a_1 + \cdots + a_\eta|^\nu \leq c_{\eta,\nu}(|a_1|^\nu + \cdots + |a_\eta|^\nu), \quad \eta,\nu \in \mathbb{Z}^+ < \infty \tag{B.100}$$

for some positive constant $c_{\eta,\nu}$ that only depends on $\eta$ and $\nu$, we can upper-bound $\left(X_\theta - \eta_{\theta,\zeta}\right)^k$, $k = \{2,4\}$ as

$$\left(X_\theta - \eta_{\theta,\zeta}\right)^k \leq c_{2,k} X_\theta^k + c_{2,k}\left(\frac{\mathsf{E}\left[X_\theta^2 e^{\zeta X_\theta}\right]}{\mathsf{E}[e^{\zeta X_\theta}]}\right)^k, \quad k = \{2,4\} \tag{B.101}$$

where $c_{2,k}$ is a positive constant that only depends on $k$. Hence, the claim follows by (B.76) and (B.93). We thus obtain (B.98) from (B.99) upon letting $\delta$ tend to infinity.

Finally, (B.86d) follows by the same steps as the ones used to show (B.86c), but with $\eta_{\theta,\zeta}$ replaced by $-\eta_{\theta,\zeta}$. ■

## B.5 Proof of Lemma 4.2

By [45, App. A.9], proving Parts 1) and 2) of Lemma 4.2 is equivalent to proving that, for $0 < s_0 < 1$, $\rho_0 > 0$, and $0 < a < 1/(\mathsf{T}-1)$

$$\sup_{\substack{-a(\mathsf{T}-1)\le\tau\le a,\\ s\in[s_0,1],\\ \rho\ge\rho_0}} \mathsf{E}\Big[(I_s(\rho)-i_{\ell,s}(\rho))^k e^{\tau(i_{\ell,s}(\rho)-I_s(\rho))}\Big] < \infty, \quad k\in\mathbb{Z}_0^+ \tag{B.102}$$

and that for $0 < s_0 < s_{\max}$, $0 < \rho_0 < \rho_{\max}$, $0 < a < 1$, and $0 < b < \min\Big\{\frac{\mathsf{T}}{\mathsf{T}-1}, \frac{1+\mathsf{T}\rho_{\max}}{\mathsf{T}\rho_{\max}s_{\max}}\Big\}$

$$\sup_{\substack{-a\le\tau\le b,\\ s\in[s_0,s_{\max}],\\ \rho\in[\rho_0,\rho_{\max}]}} \mathsf{E}\Big[(I_s(\rho)-i_{\ell,s}(\rho))^k e^{\tau(i_{\ell,s}(\rho)-I_s(\rho))}\Big] < \infty, \quad k\in\mathbb{Z}_0^+. \tag{B.103}$$

### B.5.1 Proof of Part 1)

To prove Part 1) of Lemma 4.2, we need to show that (B.102) holds. By Hölder's inequality, for any arbitrary $\delta\in(0,1-a(\mathsf{T}-1))$ such that $k/\delta$ is an integer, the LHS of (B.102) can be upper-bounded as

$$\sup_{\substack{s\in[s_0,1],\\ \rho\ge\rho_0}} \mathsf{E}\Big[(I_s(\rho)-i_{\ell,s}(\rho))^{k/\delta}\Big]^{\delta} \sup_{\substack{-a(\mathsf{T}-1)\le\tau\le a,\\ s\in[s_0,1],\\ \rho\ge\rho_0}} \mathsf{E}\Big[e^{\frac{\tau}{1-\delta}(I_s(\rho)-i_{\ell,s}(\rho))}\Big]^{1-\delta}, \quad k\in\mathbb{Z}_0^+. \tag{B.104}$$

By following along similar lines as in the proof of Lemma B.9 (Appendix B.9), it can be shown that the first supremum in (B.104) is bounded. We next show that the second supremum in (B.104) is bounded by proving that, for every $0<\xi<1$,

$$\sup_{\substack{s\in[s_0,1],\\ \rho\ge\rho_0}} \mathsf{E}\Big[e^{-\xi(I_s(\rho)-i_{\ell,s}(\rho))}\Big] < \infty \tag{B.105}$$

and that, for every $0<\xi<1/(\mathsf{T}-1)$,

$$\sup_{\substack{s\in[s_0,1],\\ \rho\ge\rho_0}} \mathsf{E}\Big[e^{\xi(I_s(\rho)-i_{\ell,s}(\rho))}\Big] < \infty. \tag{B.106}$$

Part 1) of Lemma 4.2 follows then by the convexity of the MGF [47, Lemma 2.2.5].

The LHS of (B.105) can be written as

$$\mathsf{E}\Big[e^{\xi(i_{\ell,s}(\rho)-I_s(\rho))}\Big] = \mathsf{E}\Bigg[\exp\Bigg\{\xi\Bigg(-s\frac{\mathsf{T}\rho}{1+\mathsf{T}\rho}Z_2+(\mathsf{T}-1)\log\Bigg(Z_1+\frac{Z_2}{1+\mathsf{T}\rho}\Bigg)$$

$$
\begin{aligned}
&\qquad\qquad - \log \tilde{\gamma}\Big(\mathsf{T}-1, s\frac{\mathsf{T}\rho((1+\mathsf{T}\rho)Z_1+Z_2)}{1+\mathsf{T}\rho}\Big)\Big)\Big\}\Big] \\
&\times \exp\Big\{-\xi\Big(-s\frac{\mathsf{T}\rho}{1+\mathsf{T}\rho}(\mathsf{T}-1)+(\mathsf{T}-1)\mathsf{E}\Big[\log\Big(Z_1+\frac{Z_2}{1+\mathsf{T}\rho}\Big)\Big] \\
&\qquad -\mathsf{E}\Big[\log\tilde{\gamma}\Big(\mathsf{T}-1, s\frac{\mathsf{T}\rho((1+\mathsf{T}\rho)Z_1+Z_2)}{1+\mathsf{T}\rho}\Big)\Big]\Big)\Big\}. \qquad \text{(B.107)}
\end{aligned}
$$

We next upper-bound the first expected value on the RHS of (B.107) as follows. Define $\beta_{\rho,s} \triangleq \Gamma(\mathsf{T})^{\frac{1}{\mathsf{T}-1}}\frac{1+\mathsf{T}\rho}{s\mathsf{T}\rho}$. For every $Z_1 \geq 0$ and $Z_2 \geq 0$, the exponent inside this expected value can be upper-bounded as

$$
\begin{aligned}
-s\frac{\mathsf{T}\rho}{1+\mathsf{T}\rho}Z_2 &+ (\mathsf{T}-1)\log\Big(Z_1+\frac{Z_2}{1+\mathsf{T}\rho}\Big) - \log\tilde{\gamma}\Big(\mathsf{T}-1, s\frac{\mathsf{T}\rho((1+\mathsf{T}\rho)Z_1+Z_2)}{1+\mathsf{T}\rho}\Big) \\
&\leq (\mathsf{T}-1)\log(Z_1+Z_2) + (\mathsf{T}-1)\log\Big(1+\frac{\beta_{s_0,\rho_0}}{Z_1+Z_2}\Big) \\
&= (\mathsf{T}-1)\log\Big(\frac{\beta_{s_0,\rho_0}+Z_1+Z_2}{Z_1+Z_2}\Big) \qquad \text{(B.108)}
\end{aligned}
$$

where we have used (4.1) to bound the regularized lower incomplete gamma function. Hence, we obtain

$$
\begin{aligned}
\mathsf{E}\Big[\exp\Big\{\xi\Big(-s\frac{\mathsf{T}\rho}{1+\mathsf{T}\rho}Z_2 &+ (\mathsf{T}-1)\log\Big(Z_1+\frac{Z_2}{1+\mathsf{T}\rho}\Big) \\
&- \log\tilde{\gamma}\Big(\mathsf{T}-1, s\frac{\mathsf{T}\rho((1+\mathsf{T}\rho)Z_1+Z_2)}{1+\mathsf{T}\rho}\Big)\Big)\Big\}\Big] \\
&\leq \mathsf{E}\Big[\exp\Big\{\xi(\mathsf{T}-1)\log\Big(\frac{\beta_{s_0,\rho_0}+Z_1+Z_2}{Z_1+Z_2}\Big)\Big\}\Big] \\
&= \mathsf{E}\Big[\Big(1+\frac{\beta_{s_0,\rho_0}}{Z_{1,2}}\Big)^{\xi(\mathsf{T}-1)}\Big] \qquad \text{(B.109)}
\end{aligned}
$$

where the last equality follows by defining $Z_{1,2} \triangleq Z_1+Z_2$, which is Gamma-distributed with parameters $(\mathsf{T}, 1)$. We next show that

$$
\mathsf{E}\Big[\Big(1+\frac{\beta_{s_0,\rho_0}}{Z_{1,2}}\Big)^{\xi(\mathsf{T}-1)}\Big] < \infty. \qquad \text{(B.110)}
$$

To this end, we first use (B.100) to establish the upper bound

$$
\Big(1+\frac{\beta_{s_0,\rho_0}}{Z_{1,2}}\Big)^{\xi(\mathsf{T}-1)} \leq c_{2,\lceil\xi(\mathsf{T}-1)\rceil}\Big(1+\Big(\frac{\beta_{s_0,\rho_0}}{Z_{1,2}}\Big)^{\lceil\xi(\mathsf{T}-1)\rceil}\Big) \qquad \text{(B.111)}
$$

where $c_{2,\lceil\xi(\mathsf{T}-1)\rceil}$ is a positive constant that only depends on $\lceil\xi(\mathsf{T}-1)\rceil$. The expected value of the RHS of (B.111) can be evaluated as

$$
c_{2,\lceil\xi(\mathsf{T}-1)\rceil} + c_{2,\lceil\xi(\mathsf{T}-1)\rceil}\mathsf{E}\Big[\Big(\frac{\beta_{s_0,\rho_0}}{Z_{1,2}}\Big)^{\lceil\xi(\mathsf{T}-1)\rceil}\Big]
$$

$$
\begin{aligned}
&= c_{2,\lceil\xi(\mathsf{T}-1)\rceil} + \frac{c_{2,\lceil\xi(\mathsf{T}-1)\rceil}\beta_{s_0,\rho_0}^{\lceil\xi(\mathsf{T}-1)\rceil}}{\Gamma(\mathsf{T})}\int_0^\infty z^{(\mathsf{T}-1)-\lceil\xi(\mathsf{T}-1)\rceil}e^{-z}dz \\
&= c_{2,\lceil\xi(\mathsf{T}-1)\rceil} + \frac{c_{2,\lceil\xi(\mathsf{T}-1)\rceil}\beta_{s_0,\rho_0}^{\lceil\xi(\mathsf{T}-1)\rceil}\Gamma(\mathsf{T}-\lceil\xi(\mathsf{T}-1)\rceil)}{\Gamma(\mathsf{T})} \qquad \text{(B.112)}
\end{aligned}
$$

where to solve the integral we have used [27, Sec. 3.381-4].

The remaining terms in (B.107) can be bounded for every $Z_1 \geq 0$ and $Z_2 \geq 0$ as follows:

$$
\xi\frac{s\mathsf{T}\rho}{1+\mathsf{T}\rho}(\mathsf{T}-1) \leq \xi s_{\max}(\mathsf{T}-1), \quad s \leq s_{\max} \qquad \text{(B.113a)}
$$

$$
-\xi(\mathsf{T}-1)\mathsf{E}\left[\log\left(Z_1+\frac{Z_2}{1+\mathsf{T}\rho}\right)\right] \leq -\xi(\mathsf{T}-1)\mathsf{E}\big[\log(Z_1)\big] = \xi(\mathsf{T}-1)\gamma \qquad \text{(B.113b)}
$$

$$
\xi\mathsf{E}\left[\log\tilde{\gamma}\left(\mathsf{T}-1, s\frac{\mathsf{T}\rho((1+\mathsf{T}\rho)Z_1+Z_2)}{1+\mathsf{T}\rho}\right)\right] \leq 0. \qquad \text{(B.113c)}
$$

Applying (B.113a)–(B.113c) to the remaining terms in (B.107) yields

$$
\begin{aligned}
&\exp\left\{-\xi\left(-s\frac{\mathsf{T}\rho}{1+\mathsf{T}\rho}(\mathsf{T}-1) + (\mathsf{T}-1)\mathsf{E}\left[\log\left(Z_1+\frac{Z_2}{1+\mathsf{T}\rho}\right)\right]\right.\right. \\
&\qquad\left.\left. - \mathsf{E}\left[\log\tilde{\gamma}\left(\mathsf{T}-1, s\frac{\mathsf{T}\rho((1+\mathsf{T}\rho)Z_1+Z_2)}{1+\mathsf{T}\rho}\right)\right]\right)\right\} \\
&\quad \leq e^{\xi(\mathsf{T}-1)(s_{\max}+\gamma)}. \qquad \text{(B.114)}
\end{aligned}
$$

Using (B.114) with $s_{\max} = 1$ and (B.112), it follows that (B.107) is bounded in $s \in (s_0, 1]$ and $\rho \geq \rho_0$ for every $0 < \xi < 1$. This proves (B.105).

To prove (B.106), we follow along similar lines. We have

$$
\begin{aligned}
\mathsf{E}\big[e^{\xi(I_s(\rho)-i_{\ell,s}(\rho))}\big] &= \mathsf{E}\left[\exp\left\{-\xi\left(-s\frac{\mathsf{T}\rho}{1+\mathsf{T}\rho}Z_2 + (\mathsf{T}-1)\log\left(Z_1+\frac{Z_2}{1+\mathsf{T}\rho}\right)\right.\right.\right. \\
&\qquad\left.\left.\left. - \log\tilde{\gamma}\left(\mathsf{T}-1, s\frac{\mathsf{T}\rho((1+\mathsf{T}\rho)Z_1+Z_2)}{1+\mathsf{T}\rho}\right)\right)\right\}\right] \\
&\quad\times\exp\left\{\xi\left(-s\frac{\mathsf{T}\rho}{1+\mathsf{T}\rho}(\mathsf{T}-1) + (\mathsf{T}-1)\mathsf{E}\left[\log\left(Z_1+\frac{Z_2}{1+\mathsf{T}\rho}\right)\right]\right.\right. \\
&\qquad\left.\left. - \mathsf{E}\left[\log\tilde{\gamma}\left(\mathsf{T}-1, s\frac{\mathsf{T}\rho((1+\mathsf{T}\rho)Z_1+Z_2)}{1+\mathsf{T}\rho}\right)\right]\right)\right\}. \qquad \text{(B.115)}
\end{aligned}
$$

By applying similar bounds as in (B.113), for every $Z_1 \geq 0$ and $Z_2 \geq 0$, the first expected value on the RHS of (B.115) can be upper-bounded as

$$
\begin{aligned}
&\mathsf{E}\left[\exp\left\{-\xi\left(-s\frac{\mathsf{T}\rho}{1+\mathsf{T}\rho}Z_2 + (\mathsf{T}-1)\log\left(Z_1+\frac{Z_2}{1+\mathsf{T}\rho}\right)\right.\right.\right. \\
&\qquad\left.\left.\left. - \log\tilde{\gamma}\left(\mathsf{T}-1, s\frac{\mathsf{T}\rho((1+\mathsf{T}\rho)Z_1+Z_2)}{1+\mathsf{T}\rho}\right)\right)\right\}\right]
\end{aligned}
$$

$$
\begin{aligned}
&\leq \mathsf{E}\Big[\exp\Big\{-\xi\Big(-Z_2+(\mathsf{T}-1)\log(Z_1)\Big)\Big\}\Big] \\
&= \mathsf{E}\Big[e^{\xi Z_2} Z_1^{-\xi(\mathsf{T}-1)}\Big] \\
&= \frac{\Gamma(1-\xi(\mathsf{T}-1))}{(1-\xi)^{(\mathsf{T}-1)}}
\end{aligned} \tag{B.116}
$$

where the last expected value in (B.116) has been solved using [27, Sec. 3.381-4].

We next focus on the remaining terms in (B.115). We solve each expected value separately by using the following bounds:

$$
\mathsf{E}\Big[\log\Big(Z_1+\frac{Z_2}{1+\mathsf{T}\rho}\Big)\Big] \leq \log \mathsf{E}\Big[\Big(Z_1+\frac{Z_2}{1+\mathsf{T}\rho}\Big)\Big] \leq (\mathsf{T}-1) \tag{B.117a}
$$

$$
\begin{aligned}
\mathsf{E}\Big[-\log\tilde{\gamma}\Big(\mathsf{T}-1, s\frac{\mathsf{T}\rho((1+\mathsf{T}\rho)Z_1+Z_2)}{1+\mathsf{T}\rho}\Big)\Big]& \\
&\leq \mathsf{E}\Big[-\log\tilde{\gamma}\Big(\mathsf{T}-1, s_0\frac{\mathsf{T}\rho((1+\mathsf{T}\rho)Z_1+Z_2)}{1+\mathsf{T}\rho}\Big)\Big] \\
&\leq (\mathsf{T}-1)\mathsf{E}\Big[\log\Big(1+\frac{\beta_{\rho_0,s_0}}{Z_1+Z_2}\Big)\Big] \\
&\leq (\mathsf{T}-1)\mathsf{E}\Big[\frac{\beta_{\rho_0,s_0}}{Z_1+Z_2}\Big] \\
&= \beta_{\rho_0,s_0}, \quad s>s_0,\ \rho\geq\rho_0
\end{aligned} \tag{B.117b}
$$

$$
\frac{-s\mathsf{T}\rho}{1+\mathsf{T}\rho} \leq \frac{-s_0\mathsf{T}\rho_0}{1+\mathsf{T}\rho_0}, \quad s>s_0,\ \rho\geq\rho_0. \tag{B.117c}
$$

In (B.117b), we define $\beta_{\rho,s} = \Gamma(\mathsf{T})^{\frac{1}{\mathsf{T}-1}}\frac{1+\mathsf{T}\rho}{s\mathsf{T}\rho}$. Combining (B.117) with (B.116), we can upper-bound (B.115) as

$$
\begin{aligned}
&\mathsf{E}\Big[e^{-\xi(i_{\ell,s}(\rho)-I_s(\rho))}\Big] \\
&\quad\leq \exp\Big\{\xi\Big(-\frac{s_0\mathsf{T}\rho_0}{1+\mathsf{T}\rho_0}(\mathsf{T}-1)+(\mathsf{T}-1)+\beta_{\rho_0,s_0}\Big)\Big\}\frac{\Gamma(1-\xi(\mathsf{T}-1))}{(1-\xi)^{(\mathsf{T}-1)}}
\end{aligned} \tag{B.118}
$$

which is bounded in $\rho\geq\rho_0$ and $s\in[s_0,1]$ for every $0<\xi<1/(\mathsf{T}-1)$. This proves (B.106).

### B.5.2 Proof of Part 2)

The proof follows along similar lines as the proof of Part 1). Again, by Hölder's inequality, for any arbitrary $\delta\in(0,1)$ satisfying

$$
1-\delta > \max\left\{a, \frac{b}{\min\Big\{\frac{\mathsf{T}}{\mathsf{T}-1}, \frac{1+\mathsf{T}\rho}{\mathsf{T}\rho s}\Big\}}\right\} \tag{B.119}
$$

such that $k/\delta$ is an integer, the LHS of (B.103) can be upper-bounded as

$$\sup_{\substack{s\in[s_0,s_{\max}],\\ \rho\in[\rho_0,\rho_{\max}]}} \mathsf{E}\Big[(I_s(\rho)-i_{\ell,s}(\rho))^{k/\delta}\Big]^{\delta} \sup_{\substack{-a\le\tau\le b,\\ s\in[s_0,s_{\max}],\\ \rho\in[\rho_0,\rho_{\max}]}} \mathsf{E}\Big[e^{\frac{\tau}{1-\delta}(I_s(\rho)-i_{\ell,s}(\rho))}\Big]^{1-\delta}, \quad k\in\mathbb{Z}_0^+. \tag{B.120}$$

As in Part 1), it can be shown that the first supremum in (B.120) is bounded. It thus suffices to prove that, for every $0<\xi<1$,

$$\sup_{\substack{s\in[s_0,s_{\max}],\\ \rho\in[\rho_0,\rho_{\max}]}} \mathsf{E}\big[e^{-\xi(I_s(\rho)-i_{\ell,s}(\rho))}\big]<\infty \tag{B.121}$$

and that, for every $0<\xi<\min\Big\{\frac{\mathsf{T}}{\mathsf{T}-1},\frac{1+\mathsf{T}\rho}{\mathsf{T}\rho s}\Big\}$,

$$\sup_{\substack{s\in[s_0,s_{\max}],\\ \rho\in[\rho_0,\rho_{\max}]}} \mathsf{E}\big[e^{\xi(I_s(\rho)-i_{\ell,s}(\rho))}\big]<\infty. \tag{B.122}$$

The claim (B.121) follows from (B.107)–(B.114). It remains to prove (B.122). This follows analogously to (B.115)–(B.118), with the only difference that (B.116) needs to be adapted in order to account for the different region of $(\tau,\rho,s)$.

Indeed, using that the logarithm of the regularized lower incomplete gamma function is smaller than or equal to zero, the LHS of (B.116) can be upper-bounded as

$$\begin{aligned}
&\mathsf{E}\bigg[\exp\bigg\{-\xi\bigg(-s\frac{\mathsf{T}\rho}{1+\mathsf{T}\rho}Z_2+(\mathsf{T}-1)\log\Big(Z_1+\frac{Z_2}{1+\mathsf{T}\rho}\Big)\\
&\qquad -\log\tilde{\gamma}\Big(\mathsf{T}-1,s\frac{\mathsf{T}\rho((1+\mathsf{T}\rho)Z_1+Z_2)}{1+\mathsf{T}\rho}\Big)\bigg)\bigg\}\bigg]\\
&\quad\le \mathsf{E}\bigg[\exp\Big\{-\xi\Big(-\frac{s\mathsf{T}\rho Z_2}{1+\mathsf{T}\rho}+(\mathsf{T}-1)\log\Big(Z_1+\frac{Z_2}{1+\mathsf{T}\rho}\Big)\Big\}\bigg]\\
&\quad=\frac{\eta_\rho^{\nu(\xi)}\Gamma(\mathsf{T}-1+\nu(\xi))}{\Gamma(\mathsf{T})(\eta_\rho+\lambda_{\rho,s}(\xi))^{\mathsf{T}-1+\nu(\xi)}}\,{}_2F_1\bigg(1,\mathsf{T}-1+\nu(\xi);\mathsf{T};\frac{\lambda_{\rho,s}(\xi)}{\eta_\rho+\lambda_{\rho,s}(\xi)}\bigg)
\end{aligned} \tag{B.123}$$

where $\eta_\rho=\frac{1}{1+\mathsf{T}\rho}$, $\nu(\xi)=-\xi(\mathsf{T}-1)+1$, and $\lambda_{\rho,s}(\xi)=\frac{\mathsf{T}\rho}{1+\mathsf{T}\rho}(1-\xi s)$. The expected value has been solved using [27, Sec. 3.381-3.[8]] to integrate with respect to $Z_1$, and [27, Sec. 6.455-1] to integrate with respect to $Z_2$. Note that the RHS of (B.123) is well-defined and finite for $0<\xi<\mathsf{T}/(\mathsf{T}-1)$ and $0<\xi<\frac{1+\mathsf{T}\rho}{\mathsf{T}\rho s}$.

It thus follows from (B.115), (B.117), and (B.123) that

$$\begin{aligned}
&\mathsf{E}\big[e^{-\xi(i_{\ell,s}(\rho)-I_s(\rho))}\big]\\
&\quad\le \exp\Big\{\xi\Big(-\frac{s_0\mathsf{T}\rho_0}{1+\mathsf{T}\rho_0}(\mathsf{T}-1)+(\mathsf{T}-1)+\beta_{\rho_0,s_0}\Big)\Big\}\\
&\qquad\times\frac{\eta_\rho^{\nu(\xi)}\Gamma(\mathsf{T}-1+\nu(\xi))}{\Gamma(\mathsf{T})(\eta_\rho+\lambda_{\rho,s}(\xi))^{\mathsf{T}-1+\nu(\xi)}}\,{}_2F_1\bigg(1,\mathsf{T}-1+\nu(\xi);\mathsf{T};\frac{\lambda_{\rho,s}(\xi)}{\eta_\rho+\lambda_{\rho,s}(\xi)}\bigg)
\end{aligned} \tag{B.124}$$

which is a continuous function of $(\rho, s)$, hence it is bounded in $\rho_0 \le \rho \le \rho_{\max}$ and $s_0 \le s \le s_{\max}$. This proves (B.122).

## B.6 Proof of Lemma 6.7

Throughout the proof, we shall assume that $0 \le \tau < a$, $\rho \ge \rho_0$, and $s_0 \le s \le 1$ for some arbitrary $a < 1/(\mathsf{T}-1)$, $s_0 > 0$, and $\rho_0 > 0$ independent of $(L, \rho, s, \tau)$. To prove (6.126a), we shall first show that,

$$\mathsf{E}\left[e^{\tau(I_s(\rho) - i_{\ell,s}(\rho))}\right] = \mathsf{E}\left[e^{\tau(\underline{I}_s(\rho) - \underline{i}_{\ell,s}(\rho))}\right] + o_\rho(1) \tag{B.125}$$

where (cf. (4.15))

$$\underline{i}_{\ell,s}(\rho) \triangleq (\mathsf{T}-1)\log(s\mathsf{T}\rho) - \log\Gamma(\mathsf{T}) - \frac{s\mathsf{T}\rho Z_{2,\ell}}{1+\mathsf{T}\rho} + (\mathsf{T}-1)\log\left(\frac{(1+\mathsf{T}\rho)Z_1 + Z_2}{1+\mathsf{T}\rho}\right) \tag{B.126}$$

and (cf. (4.31a))

$$\underline{I}_s(\rho) \triangleq (\mathsf{T}-1)\log(s\mathsf{T}\rho) - \log\Gamma(\mathsf{T}) - \frac{(\mathsf{T}-1)s\mathsf{T}\rho}{1+\mathsf{T}\rho} + (\mathsf{T}-1)\mathsf{E}\left[\log\left(\frac{(1+\mathsf{T}\rho)Z_1 + Z_2}{1+\mathsf{T}\rho}\right)\right]. \tag{B.127}$$

We further consider $B_\rho$ defined in (B.55).

To show (B.125), we perform the following steps:

$$\begin{aligned}\psi_{\rho,s}(\tau) &= \tau(I_s(\rho) - B_\rho) + \log\left(\mathsf{E}\left[e^{-\tau(i_{\ell,s}(\rho) - B_\rho)}\right]\right) \\ &= \tau(\underline{I}_s(\rho) - B_\rho) + \log\left(\mathsf{E}\left[e^{-\tau\left(\underline{i}_{\ell,s}(\rho) - B_\rho\right)}\right] + o_\rho(1)\right) + o_\rho(1).\end{aligned} \tag{B.128}$$

Indeed, the difference between $I_s(\rho)$ given in (4.27) and $\underline{I}_s(\rho)$ given in (B.127) is

$$\mathsf{E}\left[-\log\tilde{\gamma}\left(\mathsf{T}-1, s\frac{\mathsf{T}\rho((1+\mathsf{T}\rho)Z_1 + Z_2)}{1+\mathsf{T}\rho}\right)\right]. \tag{B.129}$$

By the monotonicity of the regularized lower incomplete gamma function, it thus follows that

$$\begin{aligned}&\sup_{\substack{\tau\in[0,a),\\ s\in(s_0,1]}} \mathsf{E}\left[-\tau\log\tilde{\gamma}\left(\mathsf{T}-1, s\frac{\mathsf{T}\rho((1+\mathsf{T}\rho)Z_1 + Z_2)}{1+\mathsf{T}\rho}\right)\right] \\ &\qquad \le \mathsf{E}\left[-a\log\tilde{\gamma}\left(\mathsf{T}-1, s_0\frac{\mathsf{T}\rho((1+\mathsf{T}\rho)Z_1 + Z_2)}{1+\mathsf{T}\rho}\right)\right]\end{aligned} \tag{B.130}$$

which vanishes as $\rho \to \infty$ by (A.100). Furthermore, we show in Corollary B.7 (Appendix B.9) that

$$\sup_{\substack{\tau\in[0,a),\\ s\in(s_0,1]}} \left|\mathsf{E}\left[e^{-\tau(i_{\ell,s}(\rho) - B_\rho)}\right] - \mathsf{E}\left[e^{-\tau\left(\underline{i}_{\ell,s}(\rho) - B_\rho\right)}\right]\right| = o_\rho(1). \tag{B.131}$$

Hence, (B.128) follows. By applying a Taylor series expansion of the logarithm function, (B.128) can be written as

$$\tau(\underline{I}_s(\rho) - B_\rho) + \log\left(\mathsf{E}\left[e^{-\tau\left(\underline{i}_{\ell,s}(\rho)-B_\rho\right)}\right] + o_\rho(1)\right) + o_\rho(1)$$
$$= \tau(\underline{I}_s(\rho) - B_\rho) + \log\left(\mathsf{E}\left[e^{-\tau\left(\underline{i}_{\ell,s}(\rho)-B_\rho\right)}\right]\right) + o_\rho(1). \quad \text{(B.132)}$$

By Lemma B.8 (Appendix B.9), the expected value inside the logarithm in (B.128) is bounded away from zero in $(\rho, \tau, s)$. It follows that the $o_\rho(1)$ term in (B.132) is uniform in $(\tau, s)$. We conclude the proof of (6.126a) by noting that

$$\bar{\psi}_{\rho,s}(\tau) = \tau(\underline{I}_s(\rho) - B_\rho) + \log\left(\mathsf{E}\left[e^{-\tau\left(\underline{i}_{\ell,s}(\rho)-B_\rho\right)}\right]\right). \quad \text{(B.133)}$$

We next prove (6.126b) by analyzing $\psi'_{\rho,s}(\tau)$ in the limit as $\rho \to \infty$. To this end, we take the derivative of $\tau \mapsto \psi_{\rho,s}(\tau)$ to obtain

$$\begin{aligned}
\psi'_s(\tau) &= \frac{\partial}{\partial \tau} \log \mathsf{E}\left[e^{\tau(I_s(\rho)-i_{\ell,s}(\rho))}\right] \\
&= \frac{\mathsf{E}\left[(I_s(\rho) - i_{\ell,s}(\rho))e^{\tau(I_s(\rho)-i_{\ell,s}(\rho)}\right]}{\mathsf{E}\left[e^{\tau(I_s(\rho)-i_{\ell,s}(\rho)}\right]} \\
&= I_s(\rho) - \frac{\mathsf{E}\left[i_{\ell,s}(\rho)e^{-\tau i_{\ell,s}(\rho)}\right]}{\mathsf{E}\left[e^{-\tau i_{\ell,s}(\rho)}\right]} \\
&= (I_s(\rho) - B_\rho) - \frac{\mathsf{E}\left[(i_{\ell,s}(\rho) - B_\rho)e^{-\tau(i_{\ell,s}(\rho)-B_\rho)}\right]}{\mathsf{E}\left[e^{-\tau(i_{\ell,s}(\rho)-B_\rho)}\right]}
\end{aligned} \quad \text{(B.134)}$$

where the second step follows by swapping derivative and expected value, which can be justified by using [45, App. A.9] together with Lemma B.6 (Appendix B.9). We show in Corollary B.7 (Appendix B.9) that the denominator in (B.134) satisfies

$$\mathsf{E}\left[e^{-\tau(i_{\ell,s}(\rho)-B_\rho)}\right] = \mathsf{E}\left[e^{-\tau(\underline{i}_{\ell,s}(\rho)-B_\rho)}\right] + o_\rho(1). \quad \text{(B.135)}$$

Furthermore, Corollary B.10 (Appendix B.9) particularized for $b = 1$ yields that

$$\mathsf{E}\left[(i_{\ell,s}(\rho) - B_\rho)e^{-\tau(i_{\ell,s}(\rho)-B_\rho))}\right] = \mathsf{E}\left[(\underline{i}_{\ell,s}(\rho) - B_\rho)e^{-\tau(\underline{i}_{\ell,s}(\rho)-B_\rho)}\right] + o_\rho(1). \quad \text{(B.136)}$$

Consequently,

$$\psi'_s(\tau) = (\underline{I}(\rho) - B_\rho) + o_\rho(1) - \frac{\mathsf{E}\left[(\underline{i}_{\ell,s}(\rho) - B_\rho)e^{-\tau(\underline{i}_{\ell,s}(\rho)-B_\rho)}\right] + o_\rho(1)}{\mathsf{E}\left[e^{-\tau(\underline{i}_{\ell,s}(\rho)-B_\rho)}\right] + o_\rho(1)}$$

$$= (\underline{I}(\rho) - B_\rho) - \frac{\mathsf{E}\Big[(\underline{i}_{\ell,s}(\rho) - B_\rho)e^{-\tau(\underline{i}_{\ell,s}(\rho)-B_\rho)}\Big]}{\mathsf{E}\Big[e^{-\tau(\underline{i}_{\ell,s}(\rho)-B_\rho)}\Big]} + o_\rho(1). \tag{B.137}$$

where in the first step we used that $I(\rho) = \underline{I}(\rho) + o_\rho(1)$ (cf. (4.39a)), and in the second step we performed a Taylor series expansion of the fraction which is well defined because by, Lemma B.8 (Appendix B.9),

$$\inf_{\substack{\tau\in[0,a),\\ s\in(s_0,1]}} \mathsf{E}\Big[e^{-\tau(i_{\ell,s}(\rho)-B_\rho)}\Big] > 0.$$

The proof of (6.126b) is concluded by noting that

$$\bar{\psi}'_{\rho,s}(\tau) = (\underline{I}(\rho) - B_\rho) - \frac{\mathsf{E}\Big[(\underline{i}_{\ell,s}(\rho) - B_\rho)e^{-\tau(\underline{i}_{\ell,s}(\rho)-B_\rho)}\Big]}{\mathsf{E}\Big[e^{-\tau(\underline{i}_{\ell,s}(\rho)-B_\rho)}\Big]}. \tag{B.138}$$

To prove (6.126c), we follow along similar lines. Indeed, by swapping derivative and expected value, we obtain

$$\begin{aligned}\psi''_s(\tau) &= \frac{\partial}{\partial\tau}\left\{(I_s(\rho) - B_\rho) - \frac{\mathsf{E}\Big[(i_{\ell,s}(\rho) - B_\rho)e^{-\tau(i_{\ell,s}(\rho)-B_\rho)}\Big]}{\mathsf{E}\Big[e^{-\tau(i_{\ell,s}(\rho)-B_\rho)}\Big]}\right\}\\ &= \frac{\mathsf{E}\Big[(i_{\ell,s}(\rho) - B_\rho)^2e^{-\tau(i_{\ell,s}(\rho)-B_\rho)}\Big]}{\mathsf{E}\Big[e^{-\tau(i_{\ell,s}(\rho)-B_\rho)}\Big]} - \frac{\mathsf{E}\Big[(i_{\ell,s}(\rho) - B_\rho)e^{-\tau(i_{\ell,s}(\rho)-B_\rho)}\Big]^2}{\mathsf{E}\Big[e^{-\tau(i_{\ell,s}(\rho)-B_\rho)}\Big]^2}.\end{aligned} \tag{B.139}$$

We show in Corollary B.10 (Appendix B.9) that the numerator of the first term on the RHS of (B.139) satisfies

$$\begin{aligned}\sup_{\substack{\tau\in[0,a),\\ s\in(s_0,1]}}\Bigg|&\mathsf{E}\Big[(i_{\ell,s}(\rho) - B_\rho)^2e^{-\tau(i_{\ell,s}(\rho)-B_\rho)}\Big]\\ &- \mathsf{E}\Big[\big(\underline{i}_{\ell,s}(\rho) - B_\rho\big)^2 e^{-\tau\left(\underline{i}_{\ell,s}(\rho)-B_\rho\right)}\Big]\Bigg| = o_\rho(1).\end{aligned} \tag{B.140}$$

We next show that the numerator of the second term on the RHS of (B.139) satisfies

$$\begin{aligned}\sup_{\substack{\tau\in[0,a),\\ s\in(s_0,1]}}\Bigg|&\mathsf{E}\Big[(i_{\ell,s}(\rho) - B_\rho)\, e^{-\tau(i_{\ell,s}(\rho)-B_\rho)}\Big]^2\\ &- \mathsf{E}\Big[\big(\underline{i}_{\ell,s}(\rho) - B_\rho\big)\, e^{-\tau\left(\underline{i}_{\ell,s}(\rho)-B_\rho\right)}\Big]^2\Bigg| = o_\rho(1).\end{aligned} \tag{B.141}$$

Indeed, the LHS of (B.141) can be upper-bounded by

$$\begin{aligned}&\sup_{\substack{\tau\in[0,a),\\ s\in(s_0,1]}}\left|\left(\mathsf{E}\left[(i_{\ell,s}(\rho)-B_\rho)\,e^{-\tau(i_{\ell,s}(\rho)-B_\rho)}\right]+\mathsf{E}\left[\left(\underline{i}_{\ell,s}(\rho)-B_\rho\right)e^{-\tau\left(\underline{i}_{\ell,s}(\rho)-B_\rho\right)}\right]\right)\right|\\ &\quad\times\sup_{\substack{\tau\in[0,a),\\ s\in(s_0,1]}}\left|\left(\mathsf{E}\left[(i_{\ell,s}(\rho)-B_\rho)\,e^{-\tau(i_{\ell,s}(\rho)-B_\rho)}\right]-\mathsf{E}\left[\left(\underline{i}_{\ell,s}(\rho)-B_\rho\right)e^{-\tau\left(\underline{i}_{\ell,s}(\rho)-B_\rho\right)}\right]\right)\right|.\end{aligned} \tag{B.142}$$

Using Hölder's inequality, we can upper-bound the first supremum by

$$\begin{aligned}&\sup_{\substack{\tau\in[0,a),\\ s\in(s_0,1]}}\left|\mathsf{E}\left[(i_{\ell,s}(\rho)-B_\rho)^{\frac{1}{\delta}}\right]\right|^{\delta}\sup_{\substack{\tau\in[0,a),\\ s\in(s_0,1]}}\left|\mathsf{E}\left[e^{-\frac{\tau}{1-\delta}(i_{\ell,s}(\rho)-A_\rho)}\right]\right|^{1-\delta}\\ &+\sup_{\substack{\tau\in[0,a),\\ s\in(s_0,1]}}\left|\mathsf{E}\left[\left(\underline{i}_{\ell,s}(\rho)-B_\rho\right)^{\frac{1}{\delta}}\right]\right|^{\delta}\sup_{\substack{\tau\in[0,a),\\ s\in(s_0,1]}}\left|\mathsf{E}\left[e^{-\frac{\tau}{1-\delta}\left(\underline{i}_{\ell,s}(\rho)-A_\rho\right)}\right]\right|^{1-\delta}\end{aligned} \tag{B.143}$$

for any arbitrary $\delta\in(0,1-a(T-1))$ such that $1/\delta$ is integer. Applying Lemmas B.6 and B.9 (both Appendix B.9), we conclude that the first supremum in (B.142) is bounded in $\rho$. Furthermore, Corollary B.10 (Appendix B.9) shows that the second supremum in (B.142) is $o_\rho(1)$. Thus, (B.141) follows.

Back to (B.139), by Corollary B.7 (Appendix B.9), the denominator of the first term on the RHS of (B.139) can be written as

$$\mathsf{E}\left[e^{-\tau(i_{\ell,s}(\rho)-B_\rho)}\right]=\mathsf{E}\left[e^{-\tau(\underline{i}_{\ell,s}(\rho)-B_\rho)}\right]+o_\rho(1). \tag{B.144}$$

We next show that the denominator of the second term on the RHS of (B.139) satisifes

$$\sup_{\substack{\tau\in[0,a),\\ s\in(s_0,1]}}\left|\mathsf{E}\left[e^{-\tau(i_{\ell,s}(\rho)-B_\rho)}\right]^2-\mathsf{E}\left[e^{-\tau\left(\underline{i}_{\ell,s}(\rho)-B_\rho\right)}\right]^2\right|=o_\rho(1). \tag{B.145}$$

To this end, we upper-bound the LHS of (B.145) by

$$\begin{aligned}&\le\sup_{\substack{\tau\in[0,a),\\ s\in(s_0,1]}}\left|\left(\mathsf{E}\left[e^{-\tau(i_{\ell,s}(\rho)-B_\rho)}\right]+\mathsf{E}\left[e^{-\tau\left(\underline{i}_{\ell,s}(\rho)-B_\rho\right)}\right]\right)\right|\\ &\qquad\times\sup_{\substack{\tau\in[0,a),\\ s\in(s_0,1]}}\left|\left(\mathsf{E}\left[e^{-\tau(i_{\ell,s}(\rho)-B_\rho)}\right]-\mathsf{E}\left[e^{-\tau\left(\underline{i}_{\ell,s}(\rho)-B_\rho\right)}\right]\right)\right|.\end{aligned} \tag{B.146}$$

The first supremum is bounded by Corollary B.6 (Appendix B.9). The remaining terms are $o_\rho(1)$ by Lemma B.7 (Appendix B.9). Hence, (B.145) follows.

Combining (B.139) with (B.140), (B.141), (B.144), and (B.145), we obtain that

$$\begin{aligned}
\psi_s''(\tau) &= \frac{\mathsf{E}\Big[(\underline{i}_{\ell,s}(\rho)-B_\rho)^2 e^{-\tau(\underline{i}_{\ell,s}(\rho)-B_\rho)}\Big]+o_\rho(1)}{\mathsf{E}\Big[e^{-\tau(\underline{i}_{\ell,s}(\rho)-B_\rho)}\Big]+o_\rho(1)} \\
&\quad - \frac{\mathsf{E}\Big[(\underline{i}_{\ell,s}(\rho)-B_\rho) e^{-\tau(\underline{i}_{\ell,s}(\rho)-B_\rho)}\Big]^2+o_\rho(1)}{\mathsf{E}\Big[e^{-\tau(\underline{i}_{\ell,s}(\rho)-B_\rho)}\Big]^2+o_\rho(1)} \\
&= \frac{\mathsf{E}\Big[(\underline{i}_{\ell,s}(\rho)-B_\rho)^2 e^{-\tau(\underline{i}_{\ell,s}(\rho)-B_\rho)}\Big]}{\mathsf{E}\Big[e^{-\tau(\underline{i}_{\ell,s}(\rho)-B_\rho)}\Big]} \\
&\quad - \frac{\mathsf{E}\Big[(\underline{i}_{\ell,s}(\rho)-B_\rho) e^{-\tau(\underline{i}_{\ell,s}(\rho)-B_\rho)}\Big]^2}{\mathsf{E}\Big[e^{-\tau(\underline{i}_{\ell,s}(\rho)-B_\rho)}\Big]^2}+o_\rho(1)
\end{aligned} \tag{B.147}$$

where the last equality can be justified by using Taylor series expansions analogously as it was done in (B.137). Identifying the first two terms as $\bar{\psi}_{\rho,s}''(\tau)$, (6.126c) follows.

The last result (6.126d) follows again along similar lines as (6.126b) and (6.126c). Indeed, by swapping derivative and expected value, we obtain

$$\begin{aligned}
\psi_{\rho,s}'''(\tau) &= \frac{\partial^3}{\partial\tau^3}\Big\{\log \mathsf{E}\Big[e^{\tau(I_s(\rho)-i_{\ell,s}(\rho))}\Big]\Big\} \\
&= -\frac{\mathsf{E}\Big[(i_{\ell,s}(\rho)-B_\rho)^3 e^{-\tau(i_{\ell,s}(\rho)-B_\rho)}\Big]}{\mathsf{E}\Big[e^{-\tau(i_{\ell,s}(\rho)-B_\rho)}\Big]} - \frac{2\mathsf{E}\Big[(i_{\ell,s}(\rho)-B_\rho) e^{-\tau(i_{\ell,s}(\rho)-B_\rho)}\Big]^3}{\mathsf{E}\Big[e^{-\tau(i_{\ell,s}(\rho)-B_\rho)}\Big]^3} \\
&\quad + \frac{3\mathsf{E}\Big[(i_{\ell,s}(\rho)-B_\rho) e^{-\tau(i_{\ell,s}(\rho)-B_\rho)}\Big]\mathsf{E}\Big[(i_{\ell,s}(\rho)-B_\rho)^2 e^{-\tau(i_{\ell,s}(\rho)-B_\rho)}\Big]}{\mathsf{E}\Big[e^{-\tau(i_{\ell,s}(\rho)-B_\rho)}\Big]^2}.
\end{aligned} \tag{B.148}$$

The denominators of the first and third term on the RHS of (B.148) have been expanded in (B.144) and (B.145), respectively. We next show that the denominator of the second term on the RHS of (B.148) satisfies

$$\sup_{\substack{\tau\in[0,\tau_*),\\ s\in(s_*,1]}}\left|\mathsf{E}\Big[e^{-\tau(i_{\ell,s}(\rho)-B_\rho)}\Big]^3-\mathsf{E}\Big[e^{-\tau\left(\underline{i}_{\ell,s}(\rho)-B_\rho\right)}\Big]^3\right| = o_\rho(1). \tag{B.149}$$

To this end, we upper-bound the LHS of (B.149) by

$$\begin{aligned}
\sup_{\substack{\tau\in[0,a),\\ s\in(s_0,1]}}\Bigg|&\left(\mathsf{E}\Big[e^{-\tau(i_{\ell,s}(\rho)-B_\rho)}\Big]+\mathsf{E}\Big[e^{-\tau\left(\underline{i}_{\ell,s}(\rho)-B_\rho\right)}\Big]\right)^2 \\
&\times\left(\mathsf{E}\Big[e^{-\tau(i_{\ell,s}(\rho)-B_\rho)}\Big]-\mathsf{E}\Big[e^{-\tau\left(\underline{i}_{\ell,s}(\rho)-B_\rho\right)}\Big]\right)\Bigg|
\end{aligned}$$

$$\leq \sup_{\substack{\tau\in[0,a),\\ s\in(s_0,1]}} \left| 2\left( \mathsf{E}\left[e^{-\tau(i_{\ell,s}(\rho)-B_\rho)}\right]^2 + \mathsf{E}\left[e^{-\tau\left(\underline{i}_{\ell,s}(\rho)-B_\rho\right)}\right]^2 \right) \right|$$

$$\times \sup_{\substack{\tau\in[0,a),\\ s\in(s_0,1]}} \left| \left( \mathsf{E}\left[e^{-\tau(i_{\ell,s}(\rho)-B_\rho)}\right] - \mathsf{E}\left[e^{-\tau\left(\underline{i}_{\ell,s}(\rho)-B_\rho\right)}\right] \right) \right| \tag{B.150}$$

where we have used that $(a+b)^2 \leq 2(a^2+b^2)$ for any $a, b \in \mathbb{R}$. The first supremum is bounded in $\rho$ by Lemma B.6 (Appendix B.9). The second supremum is $o_\rho(1)$ by Corollary B.7 (Appendix B.9). Thus, (B.149) follows.

We continue by noting that, by Corollary B.10 (Appendix B.9), the numerator of the first term on the RHS of (B.148) satisfies

$$\sup_{\substack{\tau\in[0,\tau_*),\\ s\in(s_*,1]}} \left| \mathsf{E}\left[\left(i_{\ell,s}(\rho)-B_\rho\right)^3 e^{-\tau(i_{\ell,s}(\rho)-B_\rho)}\right] \right.$$

$$\left. - \mathsf{E}\left[\left(\underline{i}_{\ell,s}(\rho)-B_\rho\right)^3 e^{-\tau\left(\underline{i}_{\ell,s}(\rho)-B_\rho\right)}\right] \right| = o_\rho(1). \tag{B.151}$$

Similarly, the numerator of the second term on the RHS of (B.148) satisfies

$$\sup_{\substack{\tau\in[0,a),\\ s\in(s_0,1]}} \left| \mathsf{E}\left[\left(i_{\ell,s}(\rho)-B_\rho\right) e^{-\tau(i_{\ell,s}(\rho)-B_\rho)}\right]^3 \right.$$

$$\left. - \mathsf{E}\left[\left(\underline{i}_{\ell,s}(\rho)-B_\rho\right) e^{-\tau\left(\underline{i}_{\ell,s}(\rho)-B_\rho\right)}\right]^3 \right| = o_\rho(1). \tag{B.152}$$

Indeed, the LHS of (B.152) can be upper-bounded as

$$\sup_{\substack{\tau\in[0,a),\\ s\in(s_0,1]}} \left| 2\left( \mathsf{E}\left[\left(i_{\ell,s}(\rho)-B_\rho\right) e^{-\tau(i_{\ell,s}(\rho)-B_\rho)}\right]^2 + \mathsf{E}\left[\left(\underline{i}_{\ell,s}(\rho)-B_\rho\right) e^{-\tau\left(\underline{i}_{\ell,s}(\rho)-B_\rho\right)}\right]^2 \right) \right|$$

$$\times \sup_{\substack{\tau\in[0,a),\\ s\in(s_0,1]}} \left| \left( \mathsf{E}\left[\left(i_{\ell,s}(\rho)-B_\rho\right) e^{-\tau(i_{\ell,s}(\rho)-B_\rho)}\right] - \mathsf{E}\left[\left(\underline{i}_{\ell,s}(\rho)-B_\rho\right) e^{-\tau\left(\underline{i}_{\ell,s}(\rho)-B_\rho\right)}\right] \right) \right|. \tag{B.153}$$

Using Hölder's inequality together with Lemmas B.6 and B.9 (both Appendix B.9), the first supremum is bounded in $\rho$. Furthermore, Corollary B.10 (Appendix B.9) shows that the second supremum is $o_\rho(1)$. Thus, (B.152) follows.

As for the numerator of the third term on the RHS of (B.148), we next show that

$$\sup_{\substack{\tau\in[0,a),\\ s\in(s_0,1]}} \left| \mathsf{E}\left[\left(i_{\ell,s}(\rho)-B_\rho\right)e^{-\tau(i_{\ell,s}(\rho)-B_\rho)}\right] \mathsf{E}\left[\left(i_{\ell,s}(\rho)-B_\rho\right)^2 e^{-\tau(i_{\ell,s}(\rho)-B_\rho)}\right] \right.$$

$$- \mathsf{E}\Big[(\underline{i}_{\ell,s}(\rho) - B_\rho)e^{-\tau(\underline{i}_{\ell,s}(\rho)-B_\rho)}\Big]\mathsf{E}\Big[(\underline{i}_{\ell,s}(\rho) - B_\rho)^2 e^{-\tau(\underline{i}_{\ell,s}(\rho)-B_\rho)}\Big]\Bigg| = o_\rho(1). \quad \text{(B.154)}$$

To this end, we first show that

$$\sup_{\substack{\tau\in[0,a),\\ s\in(s_0,1]}}\Bigg|\mathsf{E}\Big[(i_{\ell,s}(\rho) - B_\rho)e^{-\tau(i_{\ell,s}(\rho)-B_\rho)}\Big]\mathsf{E}\Big[(\underline{i}_{\ell,s}(\rho) - B_\rho)^2 e^{-\tau(\underline{i}_{\ell,s}(\rho)-B_\rho)}\Big]$$
$$- \mathsf{E}\Big[(\underline{i}_{\ell,s}(\rho) - B_\rho)e^{-\tau(\underline{i}_{\ell,s}(\rho)-B_\rho)}\Big]\mathsf{E}\Big[(\underline{i}_{\ell,s}(\rho) - B_\rho)^2 e^{-\tau(\underline{i}_{\ell,s}(\rho)-B_\rho)}\Big]\Bigg| = o_\rho(1) \quad \text{(B.155)}$$

and then that

$$\sup_{\substack{\tau\in[0,a),\\ s\in(s_0,1]}}\Bigg|\mathsf{E}\Big[(i_{\ell,s}(\rho) - B_\rho)e^{-\tau(i_{\ell,s}(\rho)-B_\rho)}\Big]\mathsf{E}\Big[(i_{\ell,s}(\rho) - B_\rho)^2 e^{-\tau(i_{\ell,s}(\rho)-B_\rho)}\Big]$$
$$- \mathsf{E}\Big[(i_{\ell,s}(\rho) - B_\rho)e^{-\tau(i_{\ell,s}(\rho)-B_\rho)}\Big]\mathsf{E}\Big[(\underline{i}_{\ell,s}(\rho) - B_\rho)^2 e^{-\tau(\underline{i}_{\ell,s}(\rho)-B_\rho)}\Big]\Bigg| = o_\rho(1). \quad \text{(B.156)}$$

The identity (B.154) follows then from the triangle inequality.

To prove (B.155), we note that its LHS can be upper-bounded as

$$\sup_{\substack{\tau\in[0,a),\\ s\in(s_0,1]}}\Bigg|\mathsf{E}\Big[(i_{\ell,s}(\rho) - B_\rho)e^{-\tau(i_{\ell,s}(\rho)-B_\rho)}\Big] - \mathsf{E}\Big[(\underline{i}_{\ell,s}(\rho) - B_\rho)e^{-\tau(\underline{i}_{\ell,s}(\rho)-B_\rho)}\Big]\Bigg|$$
$$\times \sup_{\substack{\tau\in[0,a),\\ s\in(s_0,1]}}\Bigg|\mathsf{E}\Big[(\underline{i}_{\ell,s}(\rho) - B_\rho)^2 e^{-\tau(\underline{i}_{\ell,s}(\rho)-B_\rho)}\Big]\Bigg|. \quad \text{(B.157)}$$

The first supremum in (B.157) is $o_\rho(1)$ by Corollary B.10 (Appendix B.9). The second supremum in (B.157) is finite by Hölder's inequality and Lemmas B.6 and B.9 (Appendix B.9). Hence, (B.155) follows.

To prove (B.156), we note that its LHS can be upper-bounded as

$$\sup_{\substack{\tau\in[0,a),\\ s\in(s_0,1]}}\Bigg|\mathsf{E}\Big[(i_{\ell,s}(\rho) - B_\rho)e^{-\tau(i_{\ell,s}(\rho)-B_\rho)}\Big]\Bigg|$$
$$\times \sup_{\substack{\tau\in[0,a),\\ s\in(s_0,1]}}\Bigg|\mathsf{E}\Big[(i_{\ell,s}(\rho) - B_\rho)^2 e^{-\tau(i_{\ell,s}(\rho)-B_\rho)}\Big] - \mathsf{E}\Big[(\underline{i}_{\ell,s}(\rho) - B_\rho)^2 e^{-\tau(\underline{i}_{\ell,s}(\rho)-B_\rho)}\Big]\Bigg|. \quad \text{(B.158)}$$

The first supremum in (B.158) is finite by Hölder's inequality and Lemmas B.6 and B.9 (both Appendix B.9). The second supremum in (B.158) is $o_\rho(1)$ by Corollary B.10 (Appendix B.9). This proves (B.156).

Back to (B.148), combining (B.144), (B.145), (B.149), (B.151), (B.152), and (B.154) with (B.148) yields

$$\begin{aligned}
\psi'''_{\rho,s}(\tau) &= -\frac{\mathsf{E}\Big[(\underline{i}_{\ell,s}(\rho)-B_\rho)^3 e^{-\tau(\underline{i}_{\ell,s}(\rho)-B_\rho)}\Big]+o_\rho(1)}{\mathsf{E}\Big[e^{-\tau(\underline{i}_{\ell,s}(\rho)-B_\rho)}\Big]+o_\rho(1)} \\
&\quad - \frac{2\mathsf{E}\Big[(\underline{i}_{\ell,s}(\rho)-B_\rho) e^{-\tau(\underline{i}_{\ell,s}(\rho)-B_\rho)}\Big]^3+o_\rho(1)}{\mathsf{E}\Big[e^{-\tau(\underline{i}_{\ell,s}(\rho)-B_\rho)}\Big]^3+o_\rho(1)} \\
&\quad + \frac{3\mathsf{E}\Big[(\underline{i}_{\ell,s}(\rho)-B_\rho) e^{-\tau(\underline{i}_{\ell,s}(\rho)-B_\rho)}\Big]\mathsf{E}\Big[(\underline{i}_{\ell,s}(\rho)-B_\rho)^2 e^{-\tau(\underline{i}_{\ell,s}(\rho)-B_\rho)}\Big]+o_\rho(1)}{\mathsf{E}\Big[e^{-\tau(\underline{i}_{\ell,s}(\rho)-B_\rho)}\Big]^2+o_\rho(1)} \\
&= -\frac{\mathsf{E}\Big[(\underline{i}_{\ell,s}(\rho)-B_\rho)^3 e^{-\tau(\underline{i}_{\ell,s}(\rho)-B_\rho)}\Big]}{\mathsf{E}\Big[e^{-\tau(\underline{i}_{\ell,s}(\rho)-B_\rho)}\Big]} - \frac{2\mathsf{E}\Big[(\underline{i}_{\ell,s}(\rho)-B_\rho) e^{-\tau(\underline{i}_{\ell,s}(\rho)-B_\rho)}\Big]^3}{\mathsf{E}\Big[e^{-\tau(\underline{i}_{\ell,s}(\rho)-B_\rho)}\Big]^3} \\
&\quad + \frac{3\mathsf{E}\Big[(\underline{i}_{\ell,s}(\rho)-B_\rho) e^{-\tau(\underline{i}_{\ell,s}(\rho)-B_\rho)}\Big]\mathsf{E}\Big[(\underline{i}_{\ell,s}(\rho)-B_\rho)^2 e^{-\tau(\underline{i}_{\ell,s}(\rho)-B_\rho)}\Big]}{\mathsf{E}\Big[e^{-\tau(\underline{i}_{\ell,s}(\rho)-B_\rho)}\Big]^2} + o_\rho(1)
\end{aligned} \tag{B.159}$$

where the last equality follows by Taylor series expansions analogously as it was done for the first and second derivatives of $\tau \mapsto \psi_{\rho,s}(\tau)$. Identifying the first three terms as $\bar{\psi}'''_{\rho,s}(\tau)$, we obtain (6.126d). This concludes the proof of Lemma 6.7.

## B.7 Analysis of $K_{\rho,s}(\tau, L)$

Recall that $K_{\rho,s}(\tau, L)$ was defined in (6.9b) as

$$K_{\rho,s}(\tau, L) \triangleq \frac{\psi'''_{\rho,s}(\tau)}{6\psi''_{\rho,s}(\tau)^{3/2}}\left(-\frac{1}{\sqrt{2\pi}} + \frac{\tau^2\psi''_{\rho,s}(\tau)L}{\sqrt{2\pi}} - \tau^3\psi''_{\rho,s}(\tau)^{3/2}L^{3/2}f_{\rho,s}(\tau,\tau)\right). \tag{B.160}$$

Further recall the definition of $\Psi_{\rho,s}(\tau)$ in (6.106)

$$\Psi_{\rho,s}(\tau) \triangleq \tau^2\psi''_{\rho,s}(\tau). \tag{B.161}$$

Using that

$$\begin{aligned}
f_{\rho,s}(\tau,\tau) &= e^{L\frac{\Psi_{\rho,s}(\tau)}{2}} Q\left(\sqrt{L\Psi_{\rho,s}(\tau)}\right) \\
&\leq \frac{1}{\sqrt{2\pi L\Psi_{\rho,s}(\tau)}}\left(1 - \frac{1}{L\Psi_{\rho,s}(\tau)} + \frac{3}{(L\Psi_{\rho,s}(\tau))^2}\right)
\end{aligned} \tag{B.162}$$

together with (B.161), we can lower-bound the bracketed term in (B.160) by

$$
\begin{aligned}
&-\frac{1}{\sqrt{2\pi}} + \frac{\Psi_{\rho,s}(\tau)L}{\sqrt{2\pi}} - \frac{\Psi_{\rho,s}(\tau)L}{\sqrt{2\pi}}\left(1 - \frac{1}{L\Psi_{\rho,s}(\tau)} + \frac{3}{(L\Psi_{\rho,s}(\tau))^2}\right) \\
&\quad = -\frac{3}{\sqrt{2\pi}L\Psi_{\rho,s}(\tau)}.
\end{aligned}
\tag{B.163}
$$

Likewise, using that

$$
f_{\rho,s}(\tau,\tau) \geq \frac{1}{\sqrt{2\pi L\Psi_{\rho,s}(\tau)}}\left(1 - \frac{1}{L\Psi_{\rho,s}(\tau)} + \frac{3}{(L\Psi_{\rho,s}(\tau))^2} - \frac{15}{(L\Psi_{\rho,s}(\tau))^3}\right) \tag{B.164}
$$

we can upper-bound the bracketed term in (B.160) by

$$
\begin{aligned}
&-\frac{1}{\sqrt{2\pi}} + \frac{\Psi_{\rho,s}(\tau)L}{\sqrt{2\pi}} - \frac{\Psi_{\rho,s}(\tau)L}{\sqrt{2\pi}}\left(1 - \frac{1}{L\Psi_{\rho,s}(\tau)} + \frac{3}{(L\Psi_{\rho,s}(\tau))^2} - \frac{15}{(L\Psi_{\rho,s}(\tau))^3}\right) \\
&\quad = -\frac{3}{\sqrt{2\pi}L\Psi_{\rho,s}(\tau)}\left(1 - \frac{5}{L\Psi_{\rho,s}(\tau)}\right).
\end{aligned}
\tag{B.165}
$$

It follows that

$$
|K_{\rho,s}(\tau,L)| \leq \frac{|\psi'''_{\rho,s}(\tau)|}{6\psi''_{\rho,s}(\tau)^{3/2}}\frac{3}{\sqrt{2\pi}L\Psi_{\rho,s}(\tau)}\left(1 + \frac{5}{L\Psi_{\rho,s}(\tau)}\right). \tag{B.166}
$$

By Part 2) of Lemma 4.2, we have that $\tau \mapsto \left|\psi'''_{\rho,s}(\tau)\right|$ is bounded in $\underline{\tau} < \tau < \bar{\tau}$, $s_0 < s < s_{\max}$, and $\rho_0 < \rho < \rho_{\max}$ for some arbitrary $0 < \underline{\tau} < \bar{\tau} < 1$, $0 < s_0 < s_{\max} < \infty$ and $0 < \rho_0 < \rho_{\max} < \infty$. Furthermore, by (6.108) and (6.109), both $\tau \mapsto \psi''_{\rho,s}(\tau)$ and $\tau \mapsto \Psi_{\rho,s}(\tau)$ are bounded away from zero in $(\tau, s, \rho)$. We thus conclude that

$$
\sup_{\rho_0 \leq \rho \leq \rho_{\max}} \sup_{\substack{\underline{\tau} < \tau < \bar{\tau}, \\ s \in (s_0, s_{\max}]}} \frac{|\psi'''_{\rho,s}(\tau)|}{6\psi''_{\rho,s}(\tau)^{3/2}}\frac{3}{\sqrt{2\pi}L\Psi_{\rho,s}(\tau)}\left(1 + \frac{5}{L\Psi_{\rho,s}(\tau)}\right) = \mathcal{O}_L\left(\frac{1}{L}\right). \tag{B.167}
$$

## B.8 Derivatives $E_{0,\rho}$-function

Let

$$
\psi_\rho(\tau) = \log \mathsf{E}\left[\exp\left\{\tau\left(I_{\frac{1}{1+\tau}}(\rho) - i_{\ell,\frac{1}{1+\tau}}(\rho)\right)\right\}\right] \tag{B.168}
$$

and

$$
E_{0,\rho}(\tau) = -\log \mathsf{E}\left[\exp\left\{-\tau i_{\ell,\frac{1}{1+\tau}}(\rho)\right\}\right] \tag{B.169}
$$

where $i_{\ell,\frac{1}{1+\tau}}(\rho)$ and $I_{\frac{1}{1+\tau}}(\rho)$ are given in (4.12) and (4.27), respectively. Using (B.168) and (B.169), it follows that

$$
\begin{aligned}
E_{0,\rho}(\tau) &= \tau I_{\frac{1}{1+\tau}}(\rho) - \psi_\rho(\tau) && \text{(B.170a)} \\
E'_{0,\rho}(\tau) &= I_{\frac{1}{1+\tau}}(\rho) + \tau I'_{\frac{1}{1+\tau}}(\rho) - \psi'_\rho(\tau) && \text{(B.170b)} \\
E''_{0,\rho}(\tau) &= 2I'_{\frac{1}{1+\tau}}(\rho) + \tau I''_{\frac{1}{1+\tau}}(\rho) - \psi''_\rho(\tau) && \text{(B.170c)} \\
E'''_{0,\rho}(\tau) &= 3I''_{\frac{1}{1+\tau}}(\rho) + \tau I'''_{\frac{1}{1+\tau}}(\rho) - \psi'''_\rho(\tau) && \text{(B.170d)}
\end{aligned}
$$

where we slightly abuse notation and write $I'_{\frac{1}{1+\tau}}(\rho)$, $I''_{\frac{1}{1+\tau}}(\rho)$ and $I'''_{\frac{1}{1+\tau}}(\rho)$ to denote the first three derivatives of $I_{\frac{1}{1+\tau}}(\rho)$ with respect to $\tau$. The following lemma shows that the second and third derivatives of $E_{0,\rho}(\tau)$ are bounded in $\rho_0 \leq \rho \leq \rho_{\max}$ for every $0 < \rho_0 < \rho_{\max}$ and $\tau \in (0,1)$.

**Lemma B.4** *For every* $0 < \rho_0 < \rho_{\max}$*, and* $\tau \in (0,1)$*, we have*

$$\sup_{\rho_0 \leq \rho \leq \rho_{\max}} \left| E''_{0,\rho}(\tau) \right| < \infty \tag{B.171a}$$

$$\sup_{\rho_0 \leq \rho \leq \rho_{\max}} \left| E'''_{0,\rho}(\tau) \right| < \infty. \tag{B.171b}$$

*Proof:* In view of (B.170c) and (B.170d), in order to prove (B.171a) and (B.171b), it suffices to show that, for every $0 < \rho_0 < \rho_{\max}$ and $\tau \in (0,1)$,

$$\sup_{\rho_0 \leq \rho \leq \rho_{\max}} \left| \psi''_{\rho}(\tau) \right| < \infty \tag{B.172a}$$

$$\sup_{\rho_0 \leq \rho \leq \rho_{\max}} \left| \psi'''_{\rho}(\tau) \right| < \infty \tag{B.172b}$$

and that

$$\sup_{\rho_0 \leq \rho \leq \rho_{\max}} \left| I'_{\frac{1}{1+\tau}}(\rho) \right| < \infty \tag{B.173a}$$

$$\sup_{\rho_0 \leq \rho \leq \rho_{\max}} \left| I''_{\frac{1}{1+\tau}}(\rho) \right| < \infty \tag{B.173b}$$

$$\sup_{\rho_0 \leq \rho \leq \rho_{\max}} \left| I'''_{\frac{1}{1+\tau}}(\rho) \right| < \infty. \tag{B.173c}$$

We start by analyzing $\psi''_{\rho}(\tau)$ and $\psi'''_{\rho}(\tau)$. To this end, we define

$$g_{\rho}(\tau) \triangleq I_{\frac{1}{1+\tau}}(\rho) - i_{\ell,\frac{1}{1+\tau}}(\rho). \tag{B.174}$$

Hence, we can write

$$\psi_{\rho}(\tau) = \log \mathsf{E}\left[e^{\tau g_{\rho}(\tau)}\right]. \tag{B.175}$$

Using [45, App. A.9], we can swap derivative and expected value, so the following identities follow:

$$\psi'_{\rho}(\tau) = \frac{\mathsf{E}\left[g_{\rho}(\tau)e^{\tau g_{\rho}(\tau)}\right] + \mathsf{E}\left[\tau g'_{\rho}(\tau)e^{\tau g_{\rho}(\tau)}\right]}{\mathsf{E}\left[e^{\tau g_{\rho}(\tau)}\right]} \tag{B.176a}$$

$$\begin{aligned}\psi''_{\rho}(\tau) = \frac{1}{\mathsf{E}\left[e^{\tau g_{\rho}(\tau)}\right]^2}\Bigg\{\Bigg(&\mathsf{E}\left[g_{\rho}(\tau)^2 e^{\tau g_{\rho}(\tau)}\right] + 2\mathsf{E}\left[g'_{\rho}(\tau)e^{\tau g_{\rho}(\tau)}\right] \\ &+ 2\mathsf{E}\left[\tau g_{\rho}(\tau)g'_{\rho}(\tau)e^{\tau g_{\rho}(\tau)}\right] + \mathsf{E}\left[\tau^2 g'_{\rho}(\tau)^2 e^{\tau g_{\rho}(\tau)}\right] \\ &+ \mathsf{E}\left[\tau g''_{\rho}(\tau)e^{\tau g_{\rho}(\tau)}\right]\Bigg)\mathsf{E}\left[e^{\tau g_{\rho}(\tau)}\right]\end{aligned}$$

$$
\begin{aligned}
&- \mathsf{E}\Big[g_\rho(\tau)e^{\tau g_\rho(\tau)}\Big]^2 - \mathsf{E}\Big[\tau g_\rho'(\tau)e^{\tau g_\rho(\tau)}\Big]^2 \\
&- 2\mathsf{E}\Big[g_\rho(\tau)e^{\tau g_\rho(\tau)}\Big]\mathsf{E}\Big[\tau g_\rho'(\tau)e^{\tau g_\rho(\tau)}\Big]\Bigg\}
\end{aligned} \tag{B.176b}
$$

$$
\begin{aligned}
\psi_\rho'''(\tau) = \frac{1}{\mathsf{E}\big[e^{\tau g_\rho(\tau)}\big]^3}\Bigg\{\Bigg(&\mathsf{E}\Big[g_\rho(\tau)^3 e^{\tau g_\rho(\tau)}\Big] + 6\mathsf{E}\Big[g_\rho(\tau)g_\rho'(\tau)e^{\tau g_\rho(\tau)}\Big] \\
&+ 3\mathsf{E}\Big[\tau g_\rho(\tau)^2 g_\rho'(\tau)e^{\tau g_\rho(\tau)}\Big] + 6\mathsf{E}\Big[\tau g_\rho'(\tau)^2 e^{\tau g_\rho(\tau)}\Big] \\
&+ 3\mathsf{E}\Big[g_\rho''(\tau)e^{\tau g_\rho(\tau)}\Big] + 3\mathsf{E}\Big[\tau^2 g_\rho(\tau)g_\rho'(\tau)^2 e^{\tau g_\rho(\tau)}\Big] \\
&+ 3\mathsf{E}\Big[\tau g_\rho(\tau)g_\rho''(\tau)e^{\tau g_\rho(\tau)}\Big] + 3\mathsf{E}\Big[\tau^2 g_\rho'(\tau)g_\rho''(\tau)e^{\tau g_\rho(\tau)}\Big] \\
&+ \mathsf{E}\Big[\tau^2 g_\rho'(\tau)^3 e^{\tau g_\rho(\tau)}\Big] + \mathsf{E}\Big[\tau g_\rho'''(\tau)e^{\tau g_\rho(\tau)}\Big]\Bigg)\mathsf{E}\Big[e^{\tau g_\rho(\tau)}\Big]^2 \\
&- 3\mathsf{E}\Big[e^{\tau g_\rho(\tau)}\Big]\Bigg(\mathsf{E}\Big[g_\rho(\tau)e^{\tau g_\rho(\tau)}\Big] + \mathsf{E}\Big[\tau g_\rho'(\tau)e^{\tau g_\rho(\tau)}\Big]\Bigg) \\
&\times \Bigg(\mathsf{E}\Big[g_\rho(\tau)^2 e^{\tau g_\rho(\tau)}\Big] + 2\mathsf{E}\Big[g_\rho'(\tau)e^{\tau g_\rho(\tau)}\Big] \\
&\quad + 2\mathsf{E}\Big[\tau g_\rho(\tau)g_\rho'(\tau)e^{\tau g_\rho(\tau)}\Big] \\
&\quad + \mathsf{E}\Big[\tau^2 g_\rho'(\tau)^2 e^{\tau g_\rho(\tau)}\Big] + \mathsf{E}\Big[\tau g_\rho''(\tau)e^{\tau g_\rho(\tau)}\Big]\Bigg) \\
&+ 2\Bigg(\mathsf{E}\Big[g_\rho(\tau)e^{\tau g_\rho(\tau)}\Big] + \mathsf{E}\Big[\tau g_\rho'(\tau)e^{\tau g_\rho(\tau)}\Big]\Bigg)^3\Bigg\}.
\end{aligned} \tag{B.176c}
$$

We can use Hölder's inequality over all the expected values in (B.176b) and (B.176c), similarly as we did, for instance, in (B.142)–(B.143). Then, (B.172a) and (B.172b) follow by showing that, for any arbitrary $\delta \in (0, 1-\tau)$

$$\sup_{\rho_0 \le \rho \le \rho_{\max}} \mathsf{E}\Big[e^{\frac{\tau}{1-\delta}g_\rho(\tau)}\Big] < \infty \tag{B.177a}$$

$$\sup_{\rho_0 \le \rho \le \rho_{\max}} \mathsf{E}\Big[|g_\rho(\tau)|^k\Big] < \infty, \quad k \in \mathbb{Z}^+ \tag{B.177b}$$

$$\sup_{\rho_0 \le \rho \le \rho_{\max}} \mathsf{E}\Big[\big|g_\rho'(\tau)\big|^k\Big] < \infty, \quad k \in \mathbb{Z}^+ \tag{B.177c}$$

$$\sup_{\rho_0 \le \rho \le \rho_{\max}} \mathsf{E}\Big[\big|g_\rho''(\tau)\big|^k\Big] < \infty, \quad k \in \mathbb{Z}^+ \tag{B.177d}$$

$$\sup_{\rho_0 \le \rho \le \rho_{\max}} \mathsf{E}\Big[\big|g_\rho'''(\tau)\big|^k\Big] < \infty, \quad k \in \mathbb{Z}^+. \tag{B.177e}$$

The first inequality (B.177a) follows by Part 2) of Lemma 4.2. The second inequality (B.177b) can be obtained by following along similar lines as in Lemma B.9 (Appendix B.9). We next prove (B.177c)–(B.177e).

We start with (B.177c). Let $s_\tau = 1/(1+\tau)$, and let

$$\tilde{g}_\rho(s) \triangleq \log\tilde{\gamma}\left(\mathsf{T}-1, \frac{s\mathsf{T}\rho((1+\mathsf{T}\rho)Z_1 + Z_2)}{1+\mathsf{T}\rho}\right), \quad s > 0. \tag{B.178}$$

Using the definitions of $i_{\ell,s}(\rho)$ and $I_s(\rho)$ in (4.12) and (4.27), and applying (B.100), one can show that

$$\begin{aligned}
&\mathsf{E}\Big[\big|g'_\rho(\tau)\big|^k\Big] \\
&\quad= \mathsf{E}\left[\left|s'_\tau\left(\frac{\mathsf{T}-1}{s_\tau} - \frac{\mathsf{T}\rho(\mathsf{T}-1)}{1+\mathsf{T}\rho} - \mathsf{E}\big[\tilde{g}'_\rho(s_\tau)\big] - \frac{\mathsf{T}-1}{s_\tau} + \frac{\mathsf{T}\rho Z_2}{1+\mathsf{T}\rho} + \tilde{g}'_\rho(s_\tau)\right)\right|^k\right] \\
&\quad\leq c_{4,k}|s'_\tau|^k\left\{\left(\frac{\mathsf{T}\rho}{1+\mathsf{T}\rho}\right)^k\left((\mathsf{T}-1)^k + \mathsf{E}\big[Z_2^k\big]\right) + \big|\mathsf{E}\big[\tilde{g}'_\rho(s_\tau)\big]\big|^k + \mathsf{E}\Big[\big|\tilde{g}'_\rho(s_\tau)\big|^k\Big]\right\}
\end{aligned} \tag{B.179}$$

where $s'_\tau$ denotes the derivative of $\tau \mapsto 1/(1+\tau)$ evaluated at $\tau$, i.e., $s'_\tau = -1/(1+\tau)^2$, and $\tilde{g}'_\rho$ denotes the derivative of $s \mapsto \tilde{g}_\rho(s)$ with respect to $s$. By Lemma B.11 (Appendix B.10), the last two terms on the RHS of (B.179) are bounded in $\rho$. Since the first two terms on the RHS of (B.179) are bounded in $\rho$, too, (B.177c) follows.

We next show (B.177d) following along similar lines. Using (B.100) and the definition of $g_\rho$ in (B.174), we obtain the upper bound

$$\begin{aligned}
&\mathsf{E}\Big[\big|g''_\rho(\tau)\big|^k\Big] \\
&\quad= \mathsf{E}\left[\left|s''_\tau\left(-\frac{\mathsf{T}\rho(\mathsf{T}-1)}{1+\mathsf{T}\rho} - \mathsf{E}\big[\tilde{g}'_\rho(s_\tau)\big] + \frac{\mathsf{T}\rho Z_2}{1+\mathsf{T}\rho} + \tilde{g}'_\rho(s_\tau)\right)\right.\right. \\
&\qquad\qquad \left.\left. + (s'_\tau)^2\big(-\mathsf{E}\big[\tilde{g}''_\rho(s_\tau)\big] + \tilde{g}''_\rho(s_\tau)\big)\right|^k\right] \\
&\quad\leq c_{6,k}|s''_\tau|^k\left\{\left(\frac{\mathsf{T}\rho}{1+\mathsf{T}\rho}\right)^k\left((\mathsf{T}-1)^k + \mathsf{E}\big[Z_2^k\big]\right) + \big|\mathsf{E}\big[\tilde{g}'_\rho(s_\tau)\big]\big|^k + \mathsf{E}\Big[\big|\tilde{g}'_\rho(s_\tau)\big|^k\Big]\right\} \\
&\qquad + c_{6,k}|s'_\tau|^{2k}\left\{\Big|\mathsf{E}\Big[\tilde{g}''_{\rho(s_\tau)}\Big]\Big|^k + \mathsf{E}\Big[\big|\tilde{g}''_\rho(s_\tau)\big|^k\Big]\right\}
\end{aligned} \tag{B.180}$$

where $s''_\tau$ denotes the second derivative of $\tau \mapsto 1/(1+\tau)$ evaluated at $\tau$, i.e., $s''_\tau = 2/(1+\tau)^3$, and $\tilde{g}''_\rho$ denotes the second derivative of $s \mapsto \tilde{g}_\rho(s)$ with respect to $s$. The terms that are multiplying $c_{6,k}|s''_\tau|^k$ were shown to be bounded before, and the terms that are multiplying $c_{6,k}|s'_\tau|^{2k}$ are bounded in $\rho$ by Lemma B.11 (Appendix B.10). The claim (B.177d) thus follows.

We finally show (B.177e). Using (B.100) and the definition of $g_\rho$ in (B.174), we establish the upper bound

$$\begin{aligned}
\mathsf{E}\Big[\big|g'''_\rho(\tau)\big|^k\Big] = \mathsf{E}\Bigg[\Bigg|&s'''_\tau\left(-\frac{\mathsf{T}\rho(\mathsf{T}-1)}{1+\mathsf{T}\rho} - \mathsf{E}\big[\tilde{g}'_\rho(s_\tau)\big] + \frac{\mathsf{T}\rho Z_2}{1+\mathsf{T}\rho} + \tilde{g}'_\rho(s_\tau)\right) \\
&+ 3s'_\tau s''_\tau\left(-\mathsf{E}\big[\tilde{g}''_\rho(s_\tau)\big] + \tilde{g}''_\rho(s_\tau)\right)
\end{aligned}$$

$$
\begin{aligned}
&\quad + (s'_\tau)^3\big(-\mathsf{E}\big[\tilde{g}'''_\rho(s_\tau)\big] + \tilde{g}'''_\rho(s_\tau)\big)\Big|^k\Big] \\
&\le c_{8,k}|s'''_\tau|^k\Bigg\{\bigg(\frac{\mathsf{T}\rho}{1+\mathsf{T}\rho}\bigg)^k\big((\mathsf{T}-1)^k + \mathsf{E}\big[Z_2^k\big]\big) \\
&\qquad + \big|\mathsf{E}\big[\tilde{g}'_\rho(s_\tau)\big]\big|^k + \mathsf{E}\Big[\big|\tilde{g}'_\rho(s_\tau)\big|^k\Big]\Bigg\} \\
&\quad + 3c_{8,k}|s'_\tau s''_\tau|^k\Big\{\big|\mathsf{E}\big[\tilde{g}''_\rho(s_\tau)\big]\big|^k + \mathsf{E}\Big[\big|\tilde{g}''_\rho(s_\tau)\big|^k\Big]\Big\} \\
&\quad + c_{8,k}|s'_\tau|^{3k}\Big\{\big|\mathsf{E}\big[\tilde{g}'''_\rho(s_\tau)\big]\big|^k + \mathsf{E}\Big[\big|\tilde{g}'''_\rho(s_\tau)\big|^k\Big]\Big\}
\end{aligned}
\tag{B.181}
$$

where $s'''_\tau$ denotes the third derivative of $\tau \mapsto 1/(1+\tau)$ evaluated at $\tau$, i.e., $s'''_\tau = -6/(1+\tau)^4$, and $\tilde{g}'''_\rho$ denotes the third derivative of $s \mapsto \tilde{g}_\rho(s)$ with respect to $s$. By the same arguments as above, we can conclude that the RHS of (B.181) is bounded in $\rho$. The claim (B.177e) thus follows.

To prove (B.173), we first note that the derivatives $I'_{\frac{1}{1+\tau}}(\rho)$, $I''_{\frac{1}{1+\tau}}(\rho)$ and $I'''_{\frac{1}{1+\tau}}(\rho)$ are given by

$$
I'_{\frac{1}{1+\tau}}(\rho) = -\frac{1}{(1+\tau)^2}\bigg\{\frac{\mathsf{T}-1}{s_\tau} - \frac{\mathsf{T}\rho(\mathsf{T}-1)}{1+\mathsf{T}\rho} - \mathsf{E}\big[\tilde{g}'_\rho(s_\tau)\big]\bigg\}
\tag{B.182a}
$$

$$
\begin{aligned}
I''_{\frac{1}{1+\tau}}(\rho) = {}& \frac{2}{(1+\tau)^3}\bigg\{\frac{\mathsf{T}-1}{s_\tau} - \frac{\mathsf{T}\rho(\mathsf{T}-1)}{1+\mathsf{T}\rho} - \mathsf{E}\big[\tilde{g}'_\rho(s_\tau)\big]\bigg\} \\
& - \frac{1}{(1+\tau)^4}\bigg\{\frac{\mathsf{T}-1}{s_\tau^2} + \mathsf{E}\big[\tilde{g}''_\rho(s_\tau)\big]\bigg\}
\end{aligned}
\tag{B.182b}
$$

$$
\begin{aligned}
I'''_{\frac{1}{1+\tau}}(\rho) = {}& -\frac{6}{(1+\tau)^4}\bigg\{\frac{\mathsf{T}-1}{s_\tau} - \frac{\mathsf{T}\rho(\mathsf{T}-1)}{1+\mathsf{T}\rho} - \mathsf{E}\big[\tilde{g}'_\rho(s_\tau)\big]\bigg\} \\
& + \frac{6}{(1+\tau)^5}\bigg\{\frac{\mathsf{T}-1}{s_\tau^2} + \mathsf{E}\big[\tilde{g}''_\rho(s_\tau)\big]\bigg\} \\
& - \frac{1}{(1+\tau)^6}\bigg\{\frac{2(\mathsf{T}-1)}{s_\tau^3} - \mathsf{E}\big[\tilde{g}'''_\rho(s_\tau)\big]\bigg\}.
\end{aligned}
\tag{B.182c}
$$

Note that the terms $|(\mathsf{T}-1)/s_\tau|$, $\big|(\mathsf{T}-1)/s_\tau^2\big|$ and $\big|(\mathsf{T}-1)/s_\tau^3\big|$ are bounded for $\tau \in (0,1)$. Furthermore,

$$
\bigg|\frac{\mathsf{T}\rho(\mathsf{T}-1)}{1+\mathsf{T}\rho}\bigg| \le (\mathsf{T}-1), \quad \rho \ge 0.
\tag{B.183}
$$

Finally, the derivatives of the logarithm of the regularized lower incomplete gamma function are bounded by Lemma B.11 (Appendix B.10). Thus, all the terms in (B.182) are bounded in $\rho$, so (B.173) follows. ■

## B.9 Auxiliary Results for MGF and CGF Analyses

In this appendix, we present auxiliary lemmas and corollaries that are used throughout the proof of Lemma 6.7 (Appendix B.6), the proof of Lemma 4.2 (Appendix B.5), and the proof of Lemma B.4 (Appendix B.8).

**Lemma B.5** *Let $0 \leq a < 1/(\mathsf{T}-1)$, $\rho_0 > 0$, and $0 < s_0 \leq 1$. For every $\rho \geq \rho_0$ and every $\delta \in (0, 1 - a(\mathsf{T}-1))$, we have*

$$\sup_{\substack{\tau\in[0,a),\\ s\in(s_0,1]}} \mathsf{E}\left[\left(1-\exp\left\{-\tau\left(-\log\tilde{\gamma}\left(\mathsf{T}-1, \frac{s\mathsf{T}\rho((1+\mathsf{T}\rho)Z_1+Z_2)}{1+\mathsf{T}\rho}\right)\right)\right\}\right)^{\frac{1}{\delta}}\right] = \mathcal{O}_\rho\left(\frac{1}{\rho}\right). \tag{B.184}$$

*Proof:* By(4.1),

$$\begin{aligned}
&\sup_{\substack{\tau\in[0,a),\\ s\in(s_0,1]}} \left|\mathsf{E}\left[\left(1-\exp\left\{-\tau\left(-\log\tilde{\gamma}\left(\mathsf{T}-1, s\frac{\mathsf{T}\rho((1+\mathsf{T}\rho)Z_1+Z_2)}{1+\mathsf{T}\rho}\right)\right)\right\}\right)^{\frac{1}{\delta}}\right]\right| \\
&\quad\leq \sup_{\substack{\tau\in[0,a),\\ s\in(s_0,1]}} \left|\mathsf{E}\left[\left(1-\left(1-\exp\left\{-\Gamma(\mathsf{T})^{-\frac{1}{\mathsf{T}-1}}\frac{s\mathsf{T}\rho((1+\mathsf{T}\rho)Z_1+Z_2)}{1+\mathsf{T}\rho}\right\}\right)^{\tau(\mathsf{T}-1)}\right)^{\frac{1}{\delta}}\right]\right|.
\end{aligned} \tag{B.185}$$

The function inside the expected value on the RHS of (B.185) can be upper-bounded by replacing $\tau$ by 1 and $s$ by $s_0$. Hence, we obtain that

$$\begin{aligned}
&\sup_{\substack{\tau\in[0,a),\\ s\in(s_0,1]}} \left|\mathsf{E}\left[\left(1-\left(1-\exp\left\{-\Gamma(\mathsf{T})^{-\frac{1}{\mathsf{T}-1}}\frac{s\mathsf{T}\rho((1+\mathsf{T}\rho)Z_1+Z_2)}{1+\mathsf{T}\rho}\right\}\right)^{\tau(\mathsf{T}-1)}\right)^{\frac{1}{\delta}}\right]\right| \\
&\quad\leq \mathsf{E}\left[\left(1-\left(1-\exp\left\{-\Gamma(\mathsf{T})^{-\frac{1}{\mathsf{T}-1}}\frac{s_0\mathsf{T}\rho((1+\mathsf{T}\rho)Z_1+Z_2)}{1+\mathsf{T}\rho}\right\}\right)^{\mathsf{T}-1}\right)^{\frac{1}{\delta}}\right].
\end{aligned} \tag{B.186}$$

Let $\beta_{\rho,s} \triangleq \Gamma(\mathsf{T})^{-\frac{1}{\mathsf{T}-1}}\frac{1+\mathsf{T}\rho}{s\mathsf{T}\rho}$. Using that, for every $x \geq 0$, $(1-e^x)^{(\mathsf{T}-1)} \geq 1-(\mathsf{T}-1)e^x$, and that $\beta_{\rho,s_0} \geq \beta_{\rho_0,s_0}$ for every $\rho \geq \rho_0 > 0$, the RHS of (B.186) can be further upper-bounded by

$$\begin{aligned}
&\mathsf{E}\left[(\mathsf{T}-1)^{\frac{1}{\delta}}\exp\left\{-\frac{\beta_{\rho_0,s_0}}{\delta}((1+\mathsf{T}\rho)Z_1+Z_2)\right\}\right] \\
&\qquad\qquad = \frac{(\mathsf{T}-1)^{\frac{1}{\delta}}}{\left(1+\frac{\beta_{\rho_0,s_0}}{\delta}(1+\mathsf{T}\rho)\right)\left(1+\frac{\beta_{\rho_0,s_*}}{\delta}\right)^{\mathsf{T}-1}}.
\end{aligned} \tag{B.187}$$

We conclude by noting that the RHS of (B.187) is of order $1/\rho$. ∎

**Lemma B.6** *Let $0 \le a < 1/(\mathsf{T}-1)$ and $0 < s_0 \le 1$. For every $\delta \in (0, 1-a(\mathsf{T}-1))$, we have*

$$\sup_{\rho>0} \sup_{\substack{\tau\in[0,a),\\ s\in(s_0,1]}} \mathsf{E}\Big[e^{-\frac{\tau}{1-\delta}\left(\underline{i}_{\ell,s}(\rho)-B_\rho\right)}\Big] < \infty. \tag{B.188a}$$

*Since $\underline{i}_{\ell,s}(\rho) \le i_{\ell,s}(\rho)$, this implies that*

$$\sup_{\rho>0} \sup_{\substack{\tau\in[0,a),\\ s\in(s_0,1]}} \mathsf{E}\Big[e^{-\frac{\tau}{1-\delta}(i_{\ell,s}(\rho)-B_\rho)}\Big] < \infty. \tag{B.188b}$$

*Proof:* We first lower-bound $\underline{i}_{\ell,s}(\rho)$ using that, for every $Z_1 \ge 0$ and every $Z_2 \ge 0$,

$$(\mathsf{T}-1)\log\left(Z_1 + \frac{Z_2}{1+\mathsf{T}\rho}\right) \ge (\mathsf{T}-1)\log(Z_1) \tag{B.189a}$$

$$-s\frac{\mathsf{T}\rho}{1+\mathsf{T}\rho}Z_2 \ge -Z_2, \quad s_0 < s \le 1. \tag{B.189b}$$

Hence, we can upper-bound the LHS of (B.188a) by

$$\begin{aligned}
&\sup_{\tau\in[0,a)}\left|\mathsf{E}\Big[e^{-\frac{\tau}{1-\delta}(-Z_2+(\mathsf{T}-1)\log Z_1)}\Big]\right| \\
&= \sup_{\tau\in[0,a)}\left|\frac{1}{\Gamma(\mathsf{T}-1)}\int_0^\infty z_1^{-\frac{\tau(\mathsf{T}-1)}{1-\delta}} e^{-z_1}dz_1 \int_0^\infty z_2^{\mathsf{T}-2}e^{-z_2(1-\tau/(1-\delta))}dz_2\right| \\
&= \sup_{\tau\in[0,a)}\left|\frac{\Gamma\left(1-\frac{\tau(T-1)}{1-\delta}\right)}{(1-\frac{\tau}{1-\delta})^{\mathsf{T}-1}}\right|.
\end{aligned} \tag{B.190}$$

Here, the integrals have been computed using [27, Sec. 3.381-4]. We next show that the RHS of (B.190) is finite provided that $a < \frac{1-\delta}{\mathsf{T}-1}$. Indeed, we have

$$\sup_{\tau\in[0,a)} \frac{\Gamma\left(1-\frac{\tau(T-1)}{1-\delta}\right)}{(1-\frac{\tau}{1-\delta})^{\mathsf{T}-1}} \le \frac{\Gamma\left(1-\frac{a(T-1)}{1-\delta}\right)}{(1-\frac{a}{1-\delta})^{\mathsf{T}-1}} < \infty \tag{B.191}$$

where the last step follows because $x \mapsto \Gamma(x)$ is a continuous convex function in $x > 0$. ■

**Corollary B.7** *Let $0 \le a < 1/(\mathsf{T}-1)$ and $0 < s_0 \le 1$. Then,*

$$\sup_{\substack{\tau\in[0,a),\\ s\in(s_0,1]}}\left|\mathsf{E}\Big[e^{-\tau(i_{\ell,s}(\rho)-B_\rho)}\Big] - \mathsf{E}\Big[e^{-\tau\left(\underline{i}_{\ell,s}(\rho)-B_\rho\right)}\Big]\right| = o_\rho(1). \tag{B.192}$$

*Proof:* The LHS of (B.192) can be written as

$$\sup_{\substack{\tau\in[0,a),\\ s\in(s_0,1]}}\left|\mathsf{E}\Bigg[e^{-\tau\left(\underline{i}_{\ell,s}(\rho)-B_\rho\right)} \times\left(1-\exp\left\{-\tau\left(-\log\tilde{\gamma}\left(\mathsf{T}-1, s\frac{\mathsf{T}\rho((1+\mathsf{T}\rho)Z_1+Z_2)}{1+\mathsf{T}\rho}\right)\right)\right\}\right)\Bigg]\right|. \tag{B.193}$$

By Hölder's inequality, for any $\delta \in (0, 1 - a(\mathsf{T}-1))$, this can be upper-bounded by

$$\sup_{\substack{\tau\in[0,a),\\ s\in(s_0,1]}} \mathsf{E}\left[e^{-\frac{\tau}{1-\delta}\left(\underline{i}_{\ell,s}(\rho)-B_\rho\right)}\right]^{1-\delta}$$
$$\times \sup_{\substack{\tau\in[0,a),\\ s\in(s_0,1]}} \mathsf{E}\left[\left(1-\exp\left\{-\tau\left(-\log\tilde{\gamma}\left(\mathsf{T}-1, \frac{s\mathsf{T}\rho((1+\mathsf{T}\rho)Z_1+Z_2)}{1+\mathsf{T}\rho}\right)\right)\right\}\right)^{\frac{1}{\delta}}\right]^{\delta}. \tag{B.194}$$

The first supremum in (B.194) is finite by Lemma B.6. The second supremum in (B.194) is of order $\rho^{-\delta}$ by Lemma B.5. ■

**Lemma B.8** *Let* $0 \le a < 1/(\mathsf{T}-1)$ *and* $0 < s_0 \le 1$. *Then*

$$\inf_{\substack{\tau\in[0,a),\\ s\in(s_0,1]}} \mathsf{E}\left[e^{-\tau\left(\underline{i}_{\ell,s}(\rho)-B_\rho\right)}\right] \ge \tilde{\gamma}(\mathsf{T},1) + (1-\tilde{\gamma}(\mathsf{T},1))e^{-1}. \tag{B.195}$$

*Proof:* Note that, for every $Z_1 \ge 0$ and $Z_2 \ge 0$,

$$\underline{i}_{\ell,s}(\rho) - B_\rho = -s\frac{\mathsf{T}\rho}{1+\mathsf{T}\rho}Z_2 + (\mathsf{T}-1)\log\left(Z_1 + \frac{Z_2}{1+\mathsf{T}\rho}\right) \le (\mathsf{T}-1)\log(Z_1+Z_2). \tag{B.196}$$

Consequently, we have that

$$\mathsf{E}\left[e^{-\tau\left(\underline{i}_{\ell,s}(\rho)-B_\rho\right)}\right] \ge \mathsf{E}\left[e^{-\tau(\mathsf{T}-1)\log(Z_1+Z_2)}\right], \quad 0 \le \tau < a,\ s_0 < s \le 1. \tag{B.197}$$

The RHS of (B.197) can be further lower-bounded as

$$\begin{aligned}
\mathsf{E}\left[e^{-\tau(\mathsf{T}-1)\log(Z_1+Z_2)}\right] &= \mathsf{E}\left[e^{-\tau(\mathsf{T}-1)\log(Z_1+Z_2)}\middle| Z_1+Z_2<1\right]\mathsf{P}\left[Z_1+Z_2<1\right] \\
&\quad + \mathsf{E}\left[e^{-\tau(\mathsf{T}-1)\log(Z_1+Z_2)}\middle| Z_1+Z_2\ge 1\right]\mathsf{P}\left[Z_1+Z_2\ge 1\right] \\
&\ge \tilde{\gamma}(\mathsf{T},1) + \mathsf{E}\left[(Z_1+Z_2)^{-(\mathsf{T}-1)}\middle| Z_1+Z_2\ge 1\right]\mathsf{P}\left[Z_1+Z_2\ge 1\right] \\
&= \tilde{\gamma}(\mathsf{T},1) + (1-\tilde{\gamma}(\mathsf{T},1))e^{-1}.
\end{aligned} \tag{B.198}$$

In (B.198), the inequality follows by substituting $\tau$ by 0 in the first expected value and $\tau$ by 1 in the second expected value. To solve the remaining expectations and probability terms, we have used that $Z_1 + Z_2$ is Gamma$(\mathsf{T},1)$-distributed—so $\mathsf{P}[Z_1+Z_2<1] = \tilde{\gamma}(\mathsf{T},1)$—and [27, Eq. 3.381-3.[8]]. ■

**Lemma B.9** *For every* $b \in \mathbb{Z}^+$, $0 < s_0 < s_{\max}$, *and* $\rho_0 > 0$, *we have*

$$\sup_{\rho\ge\rho_0}\ \sup_{s\in(s_0,s_{\max}]} \mathsf{E}\left[\left|(i_{\ell,s}(\rho)-B_\rho)\right|^b\right] < \infty. \tag{B.199}$$

*Proof:* For every $s \in (s_0, s_{\max}]$, we have that

$$\begin{aligned}
\mathsf{E}\left[\left|(i_{\ell,s}(\rho) - B_\rho)\right|^b\right] &= \mathsf{E}\left[\left|\left(-s\frac{\mathsf{T}\rho}{1+\mathsf{T}\rho}Z_2 + (\mathsf{T}-1)\log\left(Z_1 + \frac{Z_2}{1+\mathsf{T}\rho}\right)\right.\right.\right. \\
&\qquad \left.\left.\left. - \log\tilde{\gamma}\left(\mathsf{T}-1, s\frac{\mathsf{T}\rho((1+\mathsf{T}\rho)Z_1 + Z_2)}{1+\mathsf{T}\rho}\right)\right)\right|^b\right] \\
&\le c_{3,b}\left(\mathsf{E}\left[\left|\frac{s\mathsf{T}\rho}{1+\mathsf{T}\rho}Z_2\right|^b\right] + \mathsf{E}\left[\left|(\mathsf{T}-1)\log\left(Z_1 + \frac{Z_2}{1+\mathsf{T}\rho}\right)\right|^b\right]\right. \\
&\qquad \left. + \mathsf{E}\left[\left|\log\tilde{\gamma}\left(\mathsf{T}-1, s_0\frac{\mathsf{T}\rho((1+\mathsf{T}\rho)Z_1 + Z_2)}{1+\mathsf{T}\rho}\right)\right|^b\right]\right) \qquad \text{(B.200)}
\end{aligned}$$

where $c_{3,b}$ is a positive constant that only depends on $k$. Indeed, the inequality follows by (B.100). The first term on the RHS of (B.200) can be upper-bounded as

$$\mathsf{E}\left[\left|\frac{s\mathsf{T}\rho}{1+\mathsf{T}\rho}Z_2\right|^b\right] \le s_{\max}^b \mathsf{E}\left[Z_2^b\right], \quad s_0 < s \le s_{\max}. \qquad \text{(B.201)}$$

The second term on the RHS of (B.200) can be upper-bounded as

$$\begin{aligned}
\mathsf{E}\left[\left|(\mathsf{T}-1)\log\left(Z_1 + \frac{Z_2}{1+\mathsf{T}\rho}\right)\right|^b\right] &\le c_{2,b}\left((\mathsf{T}-1)^b\mathsf{E}\left[|\log(Z_1)|^b\right]\right. \\
&\qquad \left. + (\mathsf{T}-1)^b\mathsf{E}\left[|\log(Z_1+Z_2)|^b\right]\right) \\
&< \infty \qquad \text{(B.202)}
\end{aligned}$$

where we have used (B.100) and that

$$\left|(\mathsf{T}-1)\log\left(Z_1 + \frac{Z_2}{1+\mathsf{T}\rho}\right)\right| \le (\mathsf{T}-1)(|\log(Z_1)| + |\log(Z_1+Z_2)|). \qquad \text{(B.203)}$$

Finally, using (4.1), the third term on the RHS of (B.200) can be upper-bounded as

$$\begin{aligned}
\mathsf{E}\left[\left|\log\tilde{\gamma}\left(\mathsf{T}-1, s_0\frac{\mathsf{T}\rho((1+\mathsf{T}\rho)Z_1+Z_2)}{1+\mathsf{T}\rho}\right)\right|^b\right] &\le (\mathsf{T}-1)^b\mathsf{E}\left[\log^b\left(1 + \frac{\beta_{\rho_0,s_0}}{Z_1+Z_2}\right)\right] \\
&< \infty, \quad \rho \ge \rho_0 \qquad \text{(B.204)}
\end{aligned}$$

where $\beta_{\rho,s} \triangleq \Gamma(\mathsf{T})^{\frac{1}{\mathsf{T}-1}}\frac{1+\mathsf{T}\rho}{s\mathsf{T}\rho}$. ■

**Corollary B.10** *Let* $0 \le a < 1/(\mathsf{T}-1)$ *and* $0 < s_0 \le 1$. *For* $b \in \{1,2,3\}$,

$$\begin{aligned}
\sup_{\substack{\tau\in[0,a),\\ s\in(s_0,1]}}&\left|\left(\mathsf{E}\left[(i_{\ell,s}(\rho) - B_\rho)^b e^{-\tau(i_{\ell,s}(\rho)-B_\rho)}\right]\right.\right. \\
&\qquad \left.\left. - \mathsf{E}\left[(\underline{i}_{\ell,s}(\rho) - B_\rho)^b e^{-\tau\left(\underline{i}_{\ell,s}(\rho)-B_\rho\right)}\right]\right)\right| = o_\rho(1). \qquad \text{(B.205)}
\end{aligned}$$

*Proof:* To show (B.205), we proceed in two steps. We first show that

$$\lim_{\rho\to\infty} \sup_{\substack{\tau\in[0,a),\\ s\in(s_0,1]}} \Bigg| \Bigg( \mathsf{E}\Big[ (i_{\ell,s}(\rho) - B_\rho)^b e^{-\tau(i_{\ell,s}(\rho)-B_\rho)} \Big] - \mathsf{E}\Big[ (i_{\ell,s}(\rho) - B_\rho)^b e^{-\tau\left(\underline{i}_{\ell,s}(\rho)-B_\rho\right)} \Big] \Bigg) \Bigg| = 0. \quad \text{(B.206)}$$

We then show that

$$\lim_{\rho\to\infty} \sup_{\substack{\tau\in[0,a),\\ s\in(s_0,1]}} \Bigg| \Bigg( \mathsf{E}\Big[ (i_{\ell,s}(\rho) - B_\rho)^b e^{-\tau\left(\underline{i}_{\ell,s}(\rho)-B_\rho\right)} \Big] - \mathsf{E}\Big[ (\underline{i}_{\ell,s}(\rho) - B_\rho)^b e^{-\tau\left(\underline{i}_{\ell,s}(\rho)-B_\rho\right)} \Big] \Bigg) \Bigg| = 0. \quad \text{(B.207)}$$

Corollary B.10 follows then by the triangle inequality.

The LHS of (B.206) can be written as

$$\sup_{\substack{\tau\in[0,a),\\ s\in(s_0,1]}} \left| \mathsf{E}\left[ (i_{\ell,s}(\rho) - B_\rho)^b e^{-\tau\left(\underline{i}_{\ell,s}(\rho)-B_\rho\right)} \left( 1 - \tilde{\gamma}\Big(\mathsf{T}-1, s\frac{\mathsf{T}\rho((1+\mathsf{T}\rho)Z_1+Z_2)}{1+\mathsf{T}\rho}\Big)^\tau \right) \right] \right|. \quad \text{(B.208)}$$

Applying Hölder's inequality, this can be upper-bounded by

$$\sup_{s\in(s_0,1]} \mathsf{E}\left[ \Big| (i_{\ell,s}(\rho) - B_\rho) \Big|^{\frac{2b}{\delta}} \right]^{\frac{\delta}{2}} \sup_{\substack{\tau\in[0,a),\\ s\in(s_0,1]}} \mathsf{E}\left[ e^{-\frac{\tau}{1-\delta}\left(\underline{i}_{\ell,s}(\rho)-B_\rho\right)} \right]^{1-\delta} \times \sup_{\substack{\tau\in[0,a),\\ s\in(s_0,1]}} \mathsf{E}\left[ \left( 1 - \tilde{\gamma}\Big(\mathsf{T}-1, s\frac{\mathsf{T}\rho((1+\mathsf{T}\rho)Z_1+Z_2)}{1+\mathsf{T}\rho}\Big)^\tau \right)^{\frac{2}{\delta}} \right]^{\frac{\delta}{2}} \quad \text{(B.209)}$$

for some arbitrary $\delta \in (0, 1-a(\mathsf{T}-1))$ such that $2/\delta$ is an integer. The first supremum on the RHS of (B.209) is bounded in $\rho$ by Lemma B.9. The second supremum is bounded in $\rho$ by Lemma B.6. The third supremum is $\mathcal{O}_\rho(1/\rho)$ by Lemma B.5. Consequently, (B.208) is $o_\rho(1)$, which proves (B.206).

We next prove (B.207). Since $b \in \{1,2,3\}$, we first focus on the case $b=1$, where it suffices to show that

$$\sup_{\substack{\tau\in[0,a),\\ s\in(s_0,1]}} \left| \mathsf{E}\left[ \left( -\log\tilde{\gamma}\Big(\mathsf{T}-1, s\frac{\mathsf{T}\rho((1+\mathsf{T}\rho)Z_1+Z_2)}{1+\mathsf{T}\rho}\Big) \right) e^{-\tau\left(\underline{i}_{\ell,s}(\rho)-B_\rho\right)} \right] \right| = o_\rho(1). \quad \text{(B.210)}$$

By Hölder's inequality, the LHS of (B.210) can be upper-bounded by

$$\sup_{s\in(s_0,1]} \mathsf{E}\left[\left(-\log\tilde{\gamma}\left(\mathsf{T}-1, s\frac{\mathsf{T}\rho((1+\mathsf{T}\rho)Z_1+Z_2)}{1+\mathsf{T}\rho}\right)\right)^{\frac{1}{\delta}}\right]^{\delta} \times \sup_{\substack{\tau\in[0,a),\\ s\in(s_0,1]}} \mathsf{E}\left[e^{-\frac{\tau}{1-\delta}\left(\underline{i}_{\ell,s}(\rho)-B_\rho\right)}\right]^{1-\delta}. \quad \text{(B.211)}$$

for every $\delta\in(0,1-a(\mathsf{T}-1))$. By Lemma B.6, the second supremum on the RHS of (B.211) is bounded in $\rho$. The first supremum is achieved at $s=s_0$. It thus remains to show that

$$\mathsf{E}\left[\left(-\log\tilde{\gamma}\left(\mathsf{T}-1, s_0\frac{\mathsf{T}\rho((1+\mathsf{T}\rho)Z_1+Z_2)}{1+\mathsf{T}\rho}\right)\right)^{\frac{1}{\delta}}\right] = o_\rho(1). \quad \text{(B.212)}$$

To this end, we use that, by (A.100) and (A.103), we can apply the dominated convergence theorem to obtain

$$\lim_{\rho\to\infty} \mathsf{E}\left[\left(-\log\tilde{\gamma}\left(\mathsf{T}-1, s_0\frac{\mathsf{T}\rho((1+\mathsf{T}\rho)Z_1+Z_2)}{1+\mathsf{T}\rho}\right)\right)^{\frac{1}{\delta}}\right]^{\delta} = \mathsf{E}\left[\lim_{\rho\to\infty}\left(-\log\tilde{\gamma}\left(\mathsf{T}-1, s_0\frac{\mathsf{T}\rho((1+\mathsf{T}\rho)Z_1+Z_2)}{1+\mathsf{T}\rho}\right)\right)^{\frac{1}{\delta}}\right]^{\delta}. \quad \text{(B.213)}$$

Since $\tilde{\gamma}(\mathsf{T}-1,x)\to 1$ as $x\to\infty$, it follows that the term inside the expected value on the RHS of (B.213) is zero almost surely, hence the RHS of (B.213) is zero. This proves (B.212), which together with (B.211) proves (B.210).

We next focus on the case $b=2$, where it suffices to show

$$\sup_{\substack{\tau\in[0,a),\\ s\in(s_0,1]}} \Bigg| \mathsf{E}\left[\left(-\log\tilde{\gamma}\left(\mathsf{T}-1, s\frac{\mathsf{T}\rho((1+\mathsf{T}\rho)Z_1+Z_2)}{1+\mathsf{T}\rho}\right)\right)^2 e^{-\tau\left(\underline{i}_{\ell,s}(\rho)-B_\rho\right)}\right] + \mathsf{E}\left[\left(-2\left(\underline{i}_{\ell,s}(\rho)-B_\rho\right)\log\tilde{\gamma}\left(\mathsf{T}-1, s\frac{\mathsf{T}\rho((1+\mathsf{T}\rho)Z_1+Z_2)}{1+\mathsf{T}\rho}\right)\right) e^{-\tau\left(\underline{i}_{\ell,s}(\rho)-B_\rho\right)}\right]\Bigg| = o_\rho(1). \quad \text{(B.214)}$$

The LHS of (B.214) can be upper-bounded by

$$\sup_{\substack{\tau\in[0,a),\\ s\in(s_0,1]}} \left| \mathsf{E}\left[\left(-\log\tilde{\gamma}\left(\mathsf{T}-1, s\frac{\mathsf{T}\rho((1+\mathsf{T}\rho)Z_1+Z_2)}{1+\mathsf{T}\rho}\right)\right)^2 e^{-\tau\left(\underline{i}_{\ell,s}(\rho)-B_\rho\right)}\right]\right| + \sup_{\substack{\tau\in[0,a),\\ s\in(s_0,1]}} \left| \mathsf{E}\left[\left(-2\left(\underline{i}_{\ell,s}(\rho)-B_\rho\right)\log\tilde{\gamma}\left(\mathsf{T}-1, s\frac{\mathsf{T}\rho((1+\mathsf{T}\rho)Z_1+Z_2)}{1+\mathsf{T}\rho}\right)\right) \times e^{-\tau\left(\underline{i}_{\ell,s}(\rho)-B_\rho\right)}\right]\right|. \quad \text{(B.215)}$$

The first supremum is $o_\rho(1)$ by following similar steps as the ones used to show (B.210). For the second supremum, Hölder's inequality yields for any arbitrary $\delta \in (0, 1 - a(\mathsf{T}-1))$ such that $2/\delta$ is an integer

$$
\begin{aligned}
&\sup_{\substack{\tau\in[0,a),\\ s\in(s_0,1]}} \left| \mathsf{E}\left[ \left( -2\big(\underline{i}_{\ell,s}(\rho) - B_\rho\big) \log \tilde{\gamma}\Big(\mathsf{T}-1, s\frac{\mathsf{T}\rho((1+\mathsf{T}\rho)Z_1 + Z_2)}{1+\mathsf{T}\rho}\Big)\right) \right.\right. \\
&\qquad\qquad \left.\left. \times\, e^{-\tau\left(\underline{i}_{\ell,s}(\rho) - B_\rho\right)} \right] \right| \\
&\quad \leq \sup_{\substack{\tau\in[0,a),\\ s\in(s_0,1]}} \left| \mathsf{E}\left[ e^{-\frac{\tau}{1-\delta}\left(\underline{i}_{\ell,s}(\rho) - B_\rho\right)} \right] \right|^{1-\delta} \sup_{s\in(s_0,1]} \left| \mathsf{E}\left[ \left(2\big(\underline{i}_{\ell,s}(\rho) - B_\rho\big)\right)^{\frac{2}{\delta}} \right] \right|^{\frac{\delta}{2}} \\
&\quad\ \times \sup_{s\in(s_0,1]} \mathsf{E}\left[ \left( -\log \tilde{\gamma}\Big(\mathsf{T}-1, s\frac{\mathsf{T}\rho((1+\mathsf{T}\rho)Z_1 + Z_2)}{1+\mathsf{T}\rho}\Big)\right)^{\frac{2}{\delta}} \right]^{\frac{\delta}{2}}. \qquad \text{(B.216)}
\end{aligned}
$$

The first supremum on the RHS of (B.216) is bounded in $\rho$ by Lemma B.6. The second supremum on the RHS of (B.216) is bounded in $\rho$ by Lemma B.9. The third supremum is $o_\rho(1)$ by following similar steps as the ones used to prove (B.212). This proves (B.214).

Finally, for the case $b = 3$, it suffices to show

$$
\begin{aligned}
&\sup_{\substack{\tau\in[0,a),\\ s\in(s_0,1]}} \left| \mathsf{E}\left[ \left( -\log \tilde{\gamma}\Big(\mathsf{T}-1, s\frac{\mathsf{T}\rho((1+\mathsf{T}\rho)Z_1 + Z_2)}{1+\mathsf{T}\rho}\Big)\right)^3 e^{-\tau\left(\underline{i}_{\ell,s}(\rho) - B_\rho\right)} \right] \right. \\
&\quad - \mathsf{E}\left[ \left( 3\big(\underline{i}_{\ell,s}(\rho) - B_\rho\big)^2 \log \tilde{\gamma}\Big(\mathsf{T}-1, s\frac{\mathsf{T}\rho((1+\mathsf{T}\rho)Z_1 + Z_2)}{1+\mathsf{T}\rho}\Big)\right) e^{-\tau\left(\underline{i}_{\ell,s}(\rho) - B_\rho\right)} \right] \\
&\quad \left. + \mathsf{E}\left[ \left( 3\big(\underline{i}_{\ell,s}(\rho) - B_\rho\big) \log^2 \tilde{\gamma}\Big(\mathsf{T}-1, s\frac{\mathsf{T}\rho((1+\mathsf{T}\rho)Z_1 + Z_2)}{1+\mathsf{T}\rho}\Big)\right) e^{-\tau\left(\underline{i}_{\ell,s}(\rho) - B_\rho\right)} \right] \right| \\
&\quad = o_\rho(1). \qquad \text{(B.217)}
\end{aligned}
$$

The LHS of (B.217) can be upper-bounded by

$$
\begin{aligned}
&\sup_{\substack{\tau\in[0,a),\\ s\in(s_0,1]}} \left| \mathsf{E}\left[ \left( -\log \tilde{\gamma}\Big(\mathsf{T}-1, s\frac{\mathsf{T}\rho((1+\mathsf{T}\rho)Z_1 + Z_2)}{1+\mathsf{T}\rho}\Big)\right)^3 e^{-\tau\left(\underline{i}_{\ell,s}(\rho) - B_\rho\right)} \right] \right| \\
&\quad + \sup_{\substack{\tau\in[0,a),\\ s\in(s_0,1]}} \left| \mathsf{E}\left[ -\left( 3\big(\underline{i}_{\ell,s}(\rho) - B_\rho\big)^2 \log \tilde{\gamma}\Big(\mathsf{T}-1, s\frac{\mathsf{T}\rho((1+\mathsf{T}\rho)Z_1 + Z_2)}{1+\mathsf{T}\rho}\Big)\right) \right.\right. \\
&\qquad\qquad \left.\left. \times e^{-\tau\left(\underline{i}_{\ell,s}(\rho) - B_\rho\right)} \right] \right| \\
&\quad + \sup_{\substack{\tau\in[0,a),\\ s\in(s_0,1]}} \left| \mathsf{E}\left[ \left( 3\big(\underline{i}_{\ell,s}(\rho) - B_\rho\big) \log^2 \tilde{\gamma}\Big(\mathsf{T}-1, s\frac{\mathsf{T}\rho((1+\mathsf{T}\rho)Z_1 + Z_2)}{1+\mathsf{T}\rho}\Big)\right) \right.\right.
\end{aligned}
$$

$$\times e^{-\tau\left(\underline{i}_{\ell,s}(\rho)-B_\rho\right)}\Bigg]\Bigg|. \tag{B.218}$$

By following the same steps used to prove (B.210) and (B.214), it can be shown that (B.218) is $o_\rho(1)$. This concludes the proof of (B.207). ■

## B.10 Bounds on the Derivative of the Regularized Lower Incomplete Gamma Function

**Lemma B.11** *Assume that $a > 1$, $k \in \mathbb{Z}^+$, and $1/2 < s \leq 1$. Then, for every $x > 0$,*

$$\left|\frac{\partial^\ell}{\partial s^\ell}\log\tilde{\gamma}(a,sx)\right|^k \leq c(a,k,s), \quad k \in \mathbb{Z}^+,\ \ell \in \{1,2,3\} \tag{B.219}$$

*where $c(a,b,s)$ only depends on $a$, $k$ and $s$, but not on $x$.*

*Proof:* We start by showing (B.219) for the first derivative, namely,

$$\left|\frac{\tilde{\gamma}'(a,sx)}{\tilde{\gamma}(a,sx)}\right|^k \leq c(a,k,s), \quad k \in \mathbb{Z}^+. \tag{B.220}$$

Here,

$$\tilde{\gamma}'(a,sx) = \frac{\partial}{\partial s}\tilde{\gamma}(a,sx) = \frac{1}{\Gamma(a)}x(sx)^{a-1}e^{-sx} \tag{B.221}$$

which is nonnegative since $a > 1$, $x > 0$ and $1/2 < s \leq 1$. Furthermore, we have [48, Sec. 8.10]

$$\tilde{\gamma}(a,sx) \geq \left(1-e^{-dsx}\right)^a \tag{B.222}$$

where

$$d \triangleq \Gamma(a+1)^{-\frac{1}{a}}. \tag{B.223}$$

The RHS of (B.222) is between 0 and 1 for $x > 0$. It then follows that

$$\begin{aligned}\left|\frac{\tilde{\gamma}'(a,sx)}{\tilde{\gamma}(a,sx)}\right|^k &\leq \left(\frac{1}{\Gamma(a)}\frac{x(sx)^{a-1}e^{-sx}}{(1-e^{-dsx})^a}\right)^k\\ &\leq \left(\frac{1}{\Gamma(a)}\frac{x(sx)^{a-1}}{\left(e^{sx/a}-1\right)^a}\right)^k\\ &\leq \left(\frac{1}{\Gamma(a)}\frac{x(sx)^{a-1}}{\left(\frac{sx}{a}\right)^a}\right)^k\\ &= \left(\frac{a^a}{s\Gamma(a)}\right)^k\end{aligned} \tag{B.224}$$

where the second inequality follows because

$$\frac{1}{d} = \Gamma(a+1)^{\frac{1}{a}} = (a!)^{\frac{1}{a}} \leq a \tag{B.225}$$

and the third inequality in (B.224) follows because $e^x \geq 1 + x$. This proves (B.219) for $\ell = 1$.

While the proof of (B.219) for $\ell = 2$ and $\ell = 3$ follows along similar lines, it requires a more careful analysis. The second and third derivatives of $\tilde{\gamma}(a, sx)$ with respect to $s$ are given by the respective

$$\tilde{\gamma}''(a, sx) = \frac{1}{\Gamma(a)} x^2 (sx)^{a-2} e^{-sx}(a - sx - 1) \tag{B.226a}$$

$$\tilde{\gamma}'''(a, sx) = \frac{1}{\Gamma(a)} x^3 (sx)^{a-3} e^{-sx}(a^2 - a(2sx + 3) + 2). \tag{B.226b}$$

For simplicity, and since the steps involved to show (B.219) for $\ell = 2$ and $\ell = 3$ are analogous, we will only explain the case for $\ell = 2$. First, note that the LHS of (B.219) for $\ell = 2$ can be upper-bounded as

$$\left| \frac{\partial^2}{\partial s^2} \log \tilde{\gamma}(a, sx) \right|^k \leq c_{2,k} \left| \frac{\tilde{\gamma}''(a, sx)}{\tilde{\gamma}(a, sx)} \right|^k + c_{2,k} \left| \frac{\tilde{\gamma}'(a, sx)}{\tilde{\gamma}(a, sx)} \right|^{2k} \tag{B.227}$$

where the inequality follows by (B.100). The second term on the RHS of (B.227) can be analyzed by following the same steps as in (B.224). For the first term on the RHS of (B.227), it follows that

$$\begin{aligned} \left| \frac{\tilde{\gamma}''(a, sx)}{\tilde{\gamma}(a, sx)} \right|^k \leq c_{2,k} &\left( \frac{1}{\Gamma(a)} \frac{x^2 (sx)^{a-2} e^{-sx}(a-1)}{\tilde{\gamma}(a, sx)} \right)^k \\ &+ c_{2,k} \left( \frac{1}{\Gamma(a)} \frac{x^2 (sx)^{a-2} e^{-sx}(sx)}{\tilde{\gamma}(a, sx)} \right)^k \end{aligned} \tag{B.228}$$

where the inequality follows by using (B.226a) and (B.100). Note that both terms in (B.228) are nonnegative. Thus, the first term on the RHS of (B.228) can be analyzed following the same steps as in (B.224):

$$\begin{aligned} \left( \frac{x^2 (sx)^{a-2} e^{-sx}(a-1)}{\Gamma(a) \tilde{\gamma}(a, sx)} \right)^k &\leq \left( \frac{x^2 (sx)^{a-2}(a-1)}{\Gamma(a) \big(e^{sx/a} - 1\big)^a} \right)^k \\ &\leq \left( \frac{x^2 (sx)^{a-2}(a-1)}{\Gamma(a) \big(\frac{sx}{a}\big)^a} \right)^k \\ &= \left( \frac{a^a (a-1)}{s^2 \Gamma(a)} \right)^k. \end{aligned} \tag{B.229}$$

To analyze the second term on the RHS of (B.228), let $x_*(a, s)$ be the maximizer of the numerator, i.e.,

$$x_*(a, s) \triangleq \underset{x \geq 0}{\operatorname{argmax}} \big\{ x^{a+1} s^{a-1} e^{-sx} \big\} = \frac{a+1}{s}. \tag{B.230}$$

For the case $0 < x < x_*(a,s)$, we obtain the upper bound

$$\left(\frac{x^2(sx)^{a-2}e^{-sx}(sx)}{\Gamma(a)\tilde{\gamma}(a,sx)}\right)^k \le \left(\frac{x^2(sx)^{a-1}}{\Gamma(a)\left(\frac{sx}{a}\right)^a}\right)^k = \left(\frac{xa^a}{s\Gamma(a)}\right)^k \le \left(\frac{x_*(a,s)a^a}{s\Gamma(a)}\right)^k \tag{B.231}$$

where the result of the last inequality only depends on $a$, $s$, and $k$.

For $x \ge x_*(a,s)$, we have that

$$\left(1-e^{-dsx}\right)^a \ge \left(1-e^{-dsx_*(a,s)}\right)^a. \tag{B.232}$$

Then, using (B.222), (B.230) and (B.232), the second term on the RHS of (B.228) can be upper-bounded as

$$\left(\frac{x^2(sx)^{a-2}e^{-sx}(sx)}{\Gamma(a)\tilde{\gamma}(a,sx)}\right)^k \le \left(\frac{x_*(a,s)^2(sx_*(a,s))^{a-2}e^{-sx_*(a,s)}(sx_*(a,s))}{\Gamma(a)\left(1-e^{-dsx_*(a,s)}\right)^a}\right)^k \tag{B.233}$$

which, again, only depends on $a$, $s$, and $k$.

Combining (B.231) and (B.233), we obtain that, for $x > 0$,

$$\begin{aligned}&\left(\frac{1}{\Gamma(a)}\frac{x^2(sx)^{a-2}e^{-sx}(sx)}{\tilde{\gamma}(a,sx)}\right)^k \\ &\le \max\left\{\left(\frac{x_*(a,s)a^a}{s\Gamma(a)}\right)^k, \left(\frac{x_*(a,s)^2(sx_*(a,s))^{a-2}e^{-sx_*(a,s)}(sx_*(a,s))}{\Gamma(a)\left(1-e^{-dsx_*(a,s)}\right)^a}\right)^k\right\}\end{aligned} \tag{B.234}$$

which only depends on $a$, $s$, and $k$, but not on $x$. Combining (B.228), (B.229) and (B.234) thus yields (B.219) for $\ell = 2$. ∎

# References


[1] Y. Polyanskiy, H. V. Poor, and S. Verdú, "Channel coding rate in the finite blocklength regime," *IEEE Trans. Inf. Theory*, vol. 56, pp. 2307–2359, May 2010.

[2] A. Martinez and A. Guillén i Fàbregas, "Saddlepoint approximation of random-coding bounds," in *2011 Inf. Theory and Appl. Workshop*, pp. 1–6, Feb. 2011.

[3] G. Durisi, T. Koch, and P. Popovski, "Towards massive, ultra-reliable, and low-latency wireless communication with short packets," *Proc. IEEE*, vol. 104, pp. 1711–1726, Sept. 2016.

[4] P. Popovski, K. F. Trillingsgaard, O. Simeone, and G. Durisi, "5G wireless network slicing for eMBB, URLLC, and mMTC: A communication-theoretic view," *IEEE Access*, vol. 6, pp. 55765–55779, 2018.

[5] Y. Polyanskiy and S. Verdú, "Scalar coherent fading channel: Dispersion analysis," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, pp. 2959–2963, July 2011.

[6] A. Collins and Y. Polyanskiy, "Orthogonal designs optimize achievable dispersion for coherent MISO channels," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, pp. 2524–2528, Jun. 2014.

[7] A. Collins and Y. Polyanskiy, "Dispersion of the coherent MIMO block-fading channel," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, pp. 1068–1072, July 2016.

[8] A. Collins and Y. Polyanskiy, "Coherent multiple-antenna block-fading channels at finite blocklength," *IEEE Trans. Inf. Theory*, vol. 65, pp. 380–405, Jan. 2019.

[9] W. Yang, G. Caire, G. Durisi, and Y. Polyanskiy, "Optimum power control at finite blocklength," *IEEE Trans. Inf. Theory*, vol. 61, pp. 4598–4615, Sept. 2015.

[10] W. Yang, G. Durisi, T. Koch, and Y. Polyanskiy, "Quasi-static multiple-antenna fading channels at finite blocklength," *IEEE Trans. Inf. Theory*, vol. 60, pp. 4232–4265, July 2014.

[11] E. MolavianJazi and J. N. Laneman, "Dispersion of the coherent MIMO block-fading channel," in *Proc. Allerton Conf. Commun. Control Comput.*, (Monticello, IL, USA), Oct. 2013.

[12] J. Hoydis, R. Couillet, and P. Piantanida, "The second-order coding rate of the MIMO quasi-static Rayleigh fading channel," *IEEE Trans. Inf. Theory*, vol. 61, pp. 6591–6622, Dec. 2015.

[13] W. Yang, G. Durisi, T. Koch, and Y. Polyanskiy, "Diversity versus channel knowledge at finite block-length," in *Proc. IEEE Inf. Theory Workshop (ITW)*, pp. 572–576, Sept. 2012.

[14] J. Östman, W. Yang, G. Durisi, and T. Koch, "Diversity versus multiplexing at finite blocklength," in *Proc. IEEE Int. Symp. on Wireless Commun. Syst. (ISWCS)*, pp. 702–706, Aug. 2014.

[15] G. Durisi, T. Koch, J. Östman, Y. Polyanskiy, and W. Yang, "Short-packet communications over multiple-antenna Rayleigh-fading channels," *IEEE Trans. Commun.*, vol. 64, pp. 618–629, Feb. 2016.

[16] B. M. Hochwald and T. L. Marzetta, "Unitary space-time modulation for multiple-antenna communications in Rayleigh flat fading," *IEEE Trans. Inf. Theory*, vol. 46, pp. 543–564, Mar. 2000.

[17] L. Zheng and D. Tse, "Communication on the Grassmann manifold: A geometric approach to the noncoherent multiple-antenna channel," *IEEE Trans. Inf. Theory*, vol. 48, pp. 359–383, Feb. 2002.

[18] W. Yang, G. Durisi, and E. Riegler, "On the capacity of large-MIMO block-fading channels," *IEEE J. Sel. Areas Commun.*, vol. 31, pp. 117–132, Feb. 2013.

[19] R. G. Gallager, *Information Theory and Reliable Communication.* New York, NY, USA: John Wiley & Sons, first ed., 1968.

[20] I. Abou-Faycal and B. M. Hochwald, "Coding requirements for multiple-antenna channels with unknown Rayleigh fading," tech. rep., Bell Labs., Lucent Technologies, 1999.

[21] J. Östman, G. Durisi, E. G. Ström, M. C. Coşkun, and G. Liva, "Short packets over block-memoryless fading channels: Pilot-assisted or noncoherent transmission?," *IEEE Trans. on Commun.*, vol. 67, pp. 1521–1536, Feb. 2019.

[22] W. Feller, *An Introduction To Probability Theory And Its Applications*, vol. II. New York, NY, USA: Wiley, second ed., 1971.

[23] J. Scarlett, A. Martinez, and A. Guillén i Fàbregas, “Mismatched decoding: Error exponents, second-order rates and saddlepoint approximations,” *IEEE Trans. Inf. Theory*, vol. 60, pp. 2647–2666, May 2014.

[24] J. Font-Segura, G. Vazquez-Vilar, A. Guillén i Fàbregas, and A. Lancho, “Saddlepoint approximations in random coding,” in *Proc. Conference on Inf. Sci. and Sys. (CISS)*, (Princeton, NJ, USA), Mar. 2018.

[25] G. Vazquez-Vilar, A. Guillén i Fàbregas, T. Koch, and A. Lancho, “Saddlepoint approximation of the error probability of binary hypothesis testing,” in *Proc. IEEE Int. Symp. on Inf. Theory (ISIT)*, (Vail, CO, USA), Jun. 2018.

[26] M. Abramowitz, *Handbook of Mathematical Functions, With Formulas, Graphs, and Mathematical Tables,*. New York, NY, USA: Dover Publications, 1974.

[27] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*. Elsevier/Academic Press, Amsterdam, seventh ed., 2007.

[28] S. Verdu and T. S. Han, “A general formula for channel capacity,” *IEEE Trans. Inf. Theory*, vol. 40, pp. 1147–1157, July 1994.

[29] M. Hayashi and H. Nagaoka, “General formulas for capacity of classical-quantum channels,” *IEEE Trans. Inf. Theory*, vol. 49, pp. 1753–1768, July 2003.

[30] E. Biglieri, J. Proakis, and S. Shamai, “Fading channels: information-theoretic and communications aspects,” *IEEE Trans. Inf. Theory*, vol. 44, pp. 2619–2692, Oct. 1998.

[31] A. Lapidoth and S. M. Moser, “Capacity bounds via duality with applications to multiple-antenna systems on flat-fading channels,” *IEEE Trans. Inf. Theory*, vol. 49, pp. 2426–2467, Oct. 2003.

[32] T. L. Marzetta and B. M. Hochwald, “Capacity of a mobile multiple-antenna communication link in Rayleigh flat fading,” *IEEE Trans. Inf. Theory*, vol. 45, pp. 139–157, Jan. 1999.

[33] T. Ericson, “A Gaussian channel with slow fading,” *IEEE Trans. Inf. Theory*, vol. 16, pp. 353–355, May 1970.

[34] J. Östman, R. Devassy, G. C. Ferrante, and G. Durisi, “Low-latency short-packet transmissions: Fixed length or HARQ?.” `arXiv:1809.06560 [cs.IT]`, Sept. 2018.

[35] M. Anedda, A. Meloni, and M. Murroni, "64-apsk constellation and mapping optimization for satellite broadcasting using genetic algorithms," *IEEE Trans. Broadcasting*, vol. 62, pp. 1–9, March 2016.

[36] M. C. Coşkun, G. Durisi, T. Jerkovits, G. Liva, W. Ryan, B. Stein, and F. Steiner, "Efficient error-correcting codes in the short blocklength regime," *Physical Communication*, vol. 34, pp. 66 – 79, 2019.

[37] M. C. Coşkun, G. Liva, J. Östman, and G. Durisi, "Low-complexity joint channel estimation and list decoding of short codes," in *Proc. Int. ITG Conf. Sys. Commun. Coding (SCC)*, (Rostock, Germany), Feb. 2019.

[38] V. Y. F. Tan and M. Tomamichel, "The third-order term in the normal approximation for the AWGN channel," *IEEE Trans. Inf. Theory*, vol. 61, pp. 2430–2438, May 2015.

[39] J. F. Kurose and K. W. Ross, *Computer Networking: A Top-Down Approach.* New Jersey, USA: Pearson, 6th ed., 2012.

[40] A. Gut, *Probability: A Graduate Course.* New York, U.S.: Springer, second ed., 2013.

[41] H. Alzer, "On some inequalities for the incomplete gamma function," *Math Comp.*, vol. 66, pp. 771–778, Apr. 1997.

[42] W. Gautschi, "The incomplete gamma functions since tricomi," in *Atti dei Convegni Lincei, 147*, (Accademia Nazionale dei Lincei, Rome), pp. 203–237, 1998.

[43] W. Rudin, *Real and Complex Analysis.* New York, NY, USA: McGraw-Hill, third ed., 1987.

[44] W. Rudin, *Principles of Mathematical Analysis.* New York, NY, USA: McGraw-Hill, third ed., 1976.

[45] R. Durrett, *Probability: Theory and Examples.* Belmont, USA: Duxbury Advances Series, third ed., 2005.

[46] A. Lapidoth, *A Foundation in Digital Communication.* New York: Cambridge University Press, 2009.

[47] A. Dembo and O. Zeitouni, *Large Deviations Techniques and Applications.* New York, U.S.: Springer, second ed., 2010.

[48] F. W. Olver, D. W. Lozier, R. F. Boisvert, and C. W. Clark, *NIST Handbook of Mathematical Functions.* New York, NY, USA: Cambridge University Press, 1st ed., 2010.