This document is a preprint of: L. Zhou, A. Wolf and M. Motani, "On Lossy Multi-Connectivity: Finite Blocklength Performance and Second-Order Asymptotics," in IEEE Journal on Selected Areas in Communications (JSAC), Feb 2019. DOI:10.1109/JSAC.2019.2898755

On Lossy Multi-Connectivity: Finite Blocklength Performance and Second-Order Asymptotics

Lin Zhou, *Member, IEEE*, Albrecht Wolf, and Mehul Motani, *Fellow, IEEE*

***Abstract*—We consider the lossy transmission of a single source over parallel additive white Gaussian noise channels with independent quasi-static fading, which we term the lossy multi-connectivity problem. We assume that only the decoder has access to the channel state information. Motivated by ultra-reliable and low latency communication requirements, we are interested in the finite blocklength performance of the problem, i.e., the minimal excess-distortion probability of transmitting k source symbols over n channel uses. By generalizing non-asymptotic bounds by Kostina and Verdú for the lossy joint source-channel coding problem, we derive non-asymptotic achievability and converse bounds for the lossy multi-connectivity problem. Using these non-asymptotic bounds and under mild conditions on the fading distribution, we derive approximations for the finite blocklength performance in the spirit of second-order asymptotics for any discrete memoryless source under an arbitrary bounded distortion measure. Furthermore, in the achievability part, we analyze the performance of a universal coding scheme by modifying the universal joint source-channel coding scheme by Csiszár and using a generalized minimum distance decoder. Our results demonstrate that the asymptotic notions of outage probability and outage capacity are in fact reasonable criteria even in the finite blocklength regime. Finally, we illustrate our results via numerical examples.**

***Index Terms*—Parallel channels, Dispersion, 5G, Ultra-reliable and low latency communication, Finite blocklength analysis, Quasi-static fading, Joint source-channel coding**

I. Introduction

In cellular communication for 5G systems, one is interested in ultra-reliable and low latency communication (URLLC). To gain insight for the wireless communication system for 5G, in this paper, using tools from finite blocklength information theory, we characterize the fundamental limits of lossy data transmission over parallel additive white Gaussian noise (AWGN) channels with independent quasi-static fading. This problem is also known as the multi-connectivity problem in its lossless form [2]. Thus, in this paper, we term our problem the lossy multi-connectivity problem.

Lin Zhou is with the Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, MI, USA, 48109 (Email: linzhou@umich.edu). He was with the Department of Electrical and Computer Engineering, National University of Singapore, Singapore, 117583 (Email: lzhou@u.nus.edu).

Albrecht Wolf is with Vodafone Chair Mobile Communication Systems, Technical University Dresden, Dresden, Germany, 01069 (Email: albrecht.wolf@tu-dresden.de). The work of A. Wolf was supported by the Federal Ministry of Education and Research within the programme "Twenty20Partnership for Innovation" under contract 03ZZ0505B–"fast wireless".

Mehul Motani is with the Department of Electrical and Computer Engineering, National University of Singapore, Singapore, 117583 (Email: motani@nus.edu.sg).

The material in this work has been presented in part at the IEEE GLOBECOM, Abu Dhabi, UAE, Dec. 2018 [1].

The main reason for considering quasi-static fading is that for low-latency communication systems, the length of the data packet is rather small and usually smaller than the channel coherence time of fading channels. This observation was made previously by Yang *et al.* [3]. The motivation for multi-connectivity is mainly for the reliable communication over fading channels, especially for the quasi-static fading channels suitable for low-latency communication. If only one channel is available and unfortunately, the channel is in deep fade condition, then it is clear that nothing can be transmitted reliably over the channel. However, when multiple channels are available, if at least one channel is not in deep fade condition, then some information can always be transmitted reliably. Hence, multi-connectivity provides a flexible communication framework that can trade diversity for multiplexing via the multiple routes to the destination. Furthermore, multi-connectivity can use different carrier frequencies, such that the multiple copies of the same information can, in the best case, be delivered in a single time slot.

To the best of our knowledge, existing works on quasi-static fading channels mainly focus on outage analysis with infinite signal to noise ratio (SNR) via the diversity-multiplexing tradeoff (DMT) ([4], [5], [6]). For the lossless multi-connectivity problem, the outage analysis was established by Wolf *et al.* [2]. However, as pointed by Yang *et al.* [3], the outage analysis is not necessarily a valid criterion for low-latency communication without careful verification. The main reason why one adopts the notion of outage analysis for quasi-static fading channel is that the classical capacity with vanishing error probability proposed by Shannon [7] is zero for most commonly encountered quasi-static fading channels [8, Chapter 5]. However, capacity itself is an asymptotic notion requiring infinite blocklength and thus contradicts the need for low-latency communication. To fully understand the finite blocklength performance of channel coding over quasi-static fading channels, Yang *et al.* [3] adopted tools from finite blocklength information theory for point-to-point channels [9], [10], [11] and adapted the ideas to quasi-static MIMO channels.

In this paper, in the spirit of [3], we provide a detailed analysis for the lossy multi-connectivity problem using tools from finite blocklength information theory [10], [12], [13] and thus provide insights for the designers of future 5G URLLC systems. Our lossy multi-connectivity problem is essentially a lossy joint source channel coding (JSCC) problem where the channel consists of parallel AWGN channels with independent quasi-static fading.

This document is a preprint of: L. Zhou, A. Wolf and M. Motani, "On Lossy Multi-Connectivity: Finite Blocklength Performance and Second-Order Asymptotics," in IEEE Journal on Selected Areas in Communications (JSAC), Feb 2019. DOI:10.1109/JSAC.2019.2898755

A. Related Work

For the AWGN channel, the finite blocklength performance was characterized by Polyanskiy, Poor and Verdú [10], by Hayashi [9] and by Tan and Tomamichel [14]. Subsequently, the result was generalized to block fading channels by Polyanskiy and Verdú in [15]. Furthermore, for MIMO channel with quasi-static fading, under mild conditions on the fading distributions, the finite blocklength analysis was performed by Yang *et al.* in [3]. In general, it was shown in [3] that for quasi-static fading channels, the maximum number of messages per channel use which can transmitted over n uses of the quasi-static channel allowing average error probability ε is given by the outage capacity [8, Chapter 5] plus a remainder term which scales in the order of $O(\log n/n)$. Yang *et al.* proved the interesting result by generalizing the $\kappa\beta$-bound [10, Theorem 25] in the achievability part and the meta-converse theorem [10, Theorem 30] in the converse part. For the single input multiple output (SIMO) channel with quasi-static Rayleigh fading, MolavianJazi and Laneman [16] derived a similar result using information spectrum method for both directions. For other works on finite blocklength analysis for quasi-static fading channels, see [17] and the references therein.

We also review existing works on the lossy source coding (rate-distortion) problem dating back to Shannon [18]. The error exponent, which characterizes the speed of exponential decay of the excess-distortion probability for the rate-distortion problem, was characterized by Marton [19] for any discrete memoryless source (DMS) under any bounded distortion measure and by Ihara and Kubo [20] for any Gaussian memoryless source (GMS) under the quadratic distortion measure. In terms of second-order asymptotics, for any DMS under any bounded distortion measure, the result was derived by Ingber and Kochman [21] using method of types [22] and a refined version of the type covering lemma [19]. Similar results were also derived by Kostina and Verdú [12] for any DMS under any bounded distortion measure and any GMS under the quadratic distortion measure. Furthermore, Kostina and Verdú [12] derived non-asymptotic achievability and converse bounds using the so-called distortion-tilted information density. Subsequently, Kostina and Verdú [13] also generalized the results in [12] to derive the finite blocklength performance of the lossy joint source-channel coding problem [18].

For the lossy multi-connectivity problem with two parallel AWGN channels and quasi-static fading, Laneman *et al.* [23] considered the source and channel diversity for multiple settings of separate source-channel coding and joint source-channel coding under the average quadratic distortion criterion in the limit of *high* SNR. Note that, however, in this paper, we are interested in the finite blocklength performance for the lossy multi-connectivity problem for *any* value of SNR and *any* finite number of parallel AWGN channels with independent quasi-static fading under the excess-distortion probability criterion.

B. Main Contributions

First, we derive non-asymptotic converse and achievability bounds for the lossy multi-connectivity problem by generalizing the corresponding results for the lossy joint source-channel coding problem by Kostina and Verdú [13]. The non-asymptotic bounds hold for any value of SNR and any source distribution, discrete or continuous. In the expression of the non-asymptotic bounds, we make use of the distortion-tilted information density [12] and the fading information density [24].

Second, we derive the second-order asymptotics of the lossy multi-connectivity problem for discrete memoryless sources under bounded distortion measures. These results provide tight approximation for the finite blocklength performance. Both the achievability and converse parts follow by applying the Berry-Esseen theorem to our non-asymptotic bounds and analyzing the remainder term judiciously. We remark that the JSCC scheme used to prove our non-asymptotic achievability bound is non-universal in the sense that the decoder requires the source distribution, the channel lawand the channel state information (CSI) to make a decision. Motivated by practical universal communication systems, we adapt a universal joint source-channel coding scheme dating back to Csiszár [25] and used also by Wang, Ingber and Kochman [26] to our lossy multi-connectivity setting. Inspired by [27], our JSCC scheme combines unequal error protection [28] and modified minimum distance decoding [29]. We show that our universal scheme achieves the same second-order asymptotics as the non-universal one and thus demonstrates that the knowledge of the source distribution and the channel law at the decoder is not necessary to achieve the optimal second-order performance in the lossy multi-connectivity problem.

Finally, we demonstrate the benefit of multi-connectivity numerically by comparing our results as the number of parallel channels varies. We show that with multiple parallel channels, the achievable excess-distortion probability $\mathrm{P}^*_{\mathrm{e},k,n}(P,D)$ (see (7)) is significantly decreased.

C. Organization of the Rest of the Paper

The rest of the paper is organized as follows. In Section II, we set up the notation and formulate our lossy multi-connectivity problem. Subsequently, in Section III, we present our main results: non-asymptotic bounds and second-order asymptotics. In Section IV, we present numerical simulations to illustrate our results. The proofs of second-order asymptotics are given in Section V. Finally, in Section VI, we conclude the paper and discuss future research directions. For the smooth presentation of our main results, we defer all other proofs to the appendices.

II. Problem Formulation

A. Notation

Random variables and their realizations are denoted in capital letters (e.g., X) and lower case letters (e.g., x), respectively. All sets (e.g., alphabets of random variables) are denoted in calligraphic font (e.g., $\mathcal{X}$). Let $X^n := (X_1, \ldots, X_n)$ be a

This document is a preprint of: L. Zhou, A. Wolf and M. Motani, "On Lossy Multi-Connectivity: Finite Blocklength Performance and Second-Order Asymptotics," in IEEE Journal on Selected Areas in Communications (JSAC), Feb 2019. DOI:10.1109/JSAC.2019.2898755

random vector of length n and x^n a particular realization. We use $\|x^n\|$ to denote the ℓ_2 norm $\sqrt{\sum_i x_i^2}$. The set of all probability distribution on $\mathcal{X}$ is denoted as $\mathcal{P}(\mathcal{X})$ and the set of all conditional probability distributions with the input alphabet $\mathcal{X}$ and the output alphabet $\mathcal{Y}$ is denoted as $\mathcal{P}(\mathcal{Y}|\mathcal{X})$. We use $\mathcal{R}$, $\mathcal{R}_+$ and $\mathcal{N}$ to denote the sets of real numbers, non-negative real numbers, and natural numbers, respectively. Given any two integers $a, b \in \mathcal{N}^2$, we use $[a:b]$ to denote the inclusive collection of all integers between a and b, i.e., $[a:b] := \{c : c \in \mathcal{N}\ , a \leq c \leq b\}$. Furthermore, we use $[a]$ to denote the set $[1:a]$ for any integer a. We use $\mathbf{I}_n$ to denote the $n \times n$ identity matrix. Given a real number $a \in \mathcal{R}$, we use $|a|^+$ to denote $\max\{a, 0\}$. Finally, all logarithms are natural logarithms and we use standard asymptotic notation such as $O(\cdot)$, $o(\cdot)$ and $\Theta(\cdot)$ [30]. To be specific, for any two sequences a_n and b_n, we say $a_n = O(b_n)$ if $\limsup_{n\to\infty} \left|\frac{a_n}{b_n}\right| < \infty$, we say $a_n = o(b_n)$ if $\lim_{n\to\infty} \frac{a_n}{b_n} = 0$ and we say $a_n = \Theta(b_n)$ if $c_1 \leq \lim_{n\to\infty} \frac{a_n}{b_n} \leq c_2$ where c_1 and c_2 are two constants.

B. System Model

In this paper, as shown in Figure 1, we consider the lossy transmission of a single memoryless source over parallel AWGN channels with independent quasi-static fading. We assume that the memoryless source is generated i.i.d. according to a distribution P_S defined on the alphabet $\mathcal{S}$, i.e., for any integer $k \in \mathbb{N}$ and any $s^k \in \mathcal{S}^k$, $P_{S^k}(s^k) = \prod_{i\in[k]} P_S(s_i)$.

Furthermore, we assume that there are in total Ψ parallel AWGN channels with independent quasi-static fading. For each channel indexed by $t \in [\Psi]$, the additive noise Z_t^n is generated i.i.d. according to the normal distribution, i.e., $\mathcal{N}(0,1)$. Furthermore, the fading coefficient $A_t \in \mathcal{A}$ for each channel ($t \in [\Psi]$) remains unchanged in n channel uses and the fading parameters $\{A_t\}_{t\in[\Psi]}$ are distributed i.i.d. according to a fading distribution P_A. Thus, the channel law for each independent channel indexed by $t \in [\Psi]$ is given by

$$Y_t^n = A_t X_t^n + Z_t^n,\ t \in [\Psi]. \tag{1}$$

Throughout the paper, we assume a maximal power constraint P, i.e., for any codewords $\{x_t^n\}_{t\in[\Psi]}$ for parallel channels, $\frac{1}{n}\sum_{t\in[\Psi]} \|x_t^n\|^2 \leq P$. Furthermore, we assume that only the receiver has access to CSI, i.e., $\{A_t\}_{t\in[\Psi]}$. Given the channel outputs $\{Y^n(t)\}_{t\in[\Psi]}$ and CSI $\{A_t\}_{t\in[\Psi]}$, the decoder estimates the source sequence as $\hat{S}^k$, which takes values in the alphabet $\hat{\mathcal{S}}^k$. To measure the performance, we define the one-shot distortion measure $d : \mathcal{S} \times \hat{\mathcal{S}} \to [0,\infty)$ and its multi-letter version $d(s^k, \hat{s}^k) := \frac{1}{k}\sum_{i\in[k]} d(s_i, \hat{s}_i)$ for any pair of source sequence $s^k \in \mathcal{S}^k$ and the reproduced sequence $\hat{s}^k \in \hat{\mathcal{S}}^k$.

To formally define a code for our problem, for each $t \in [\Psi]$, given any $P_t \in \mathcal{R}_+$, let $\mathbb{R}(n, P_t)$ be the collection of all length-n sequences of real numbers with power P_t, i.e.,

$$\mathbb{R}(n, P_t) := \{x^n \in \mathcal{R}^n : \|x^n\|^2 = nP_t\}. \tag{2}$$

Definition 1. *An (k, n, P)-code consists of*

- Ψ *encoders*

$$f_t : \mathcal{S}^k \to \mathbb{R}(n, P_t),\ \forall\ t \in [\Psi],\ \text{and} \tag{3}$$

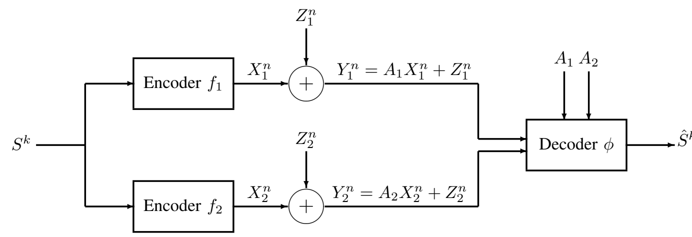


Fig. 1: System model for the lossy multi-connectivity problem with two parallel channels (i.e., $\Psi = 2$). We are interested in the finite blocklength performance of lossy transmission over parallel AWGN channels with independent quasi-static fading. We assume that only the decoder has access to the channel state information, i.e., the fading parameters $\{A_t\}_{t\in[\Psi]}$.

- *a decoder*

$$\phi : \mathcal{R}^{n\Psi} \times \mathcal{A}^\Psi \to \hat{\mathcal{S}}^k, \tag{4}$$

for some power allocation vector $\{P_t\}_{t\in[\Psi]}$ satisfying

$$\sum_{t\in[\Psi]} P_t \leq P. \tag{5}$$

Given an (k, n, P)-code, the source estimate can be expressed as $\hat{S}^k = \phi(\{Y_t^n, A_t\}_{t\in[\Psi]})$ and the excess-distortion probability with respect to a distortion level D is defined as follows:

$$\mathrm{P}_{\mathrm{e},k,n}(P, D) := \Pr\left\{d(S^k, \hat{S}^k) > D\right\}. \tag{6}$$

Given (k, n, P) and D, the optimal performance is evaluated by the minimal excess-distortion probability of any (k, n, P)-code with respect to the distortion level D, i.e.,

$$\mathrm{P}^*_{\mathrm{e},k,n}(P, D) := \inf\{\varepsilon : \exists \text{ an } (k, n, P)\text{−code s.t. } \mathrm{P}_{\mathrm{e},k,n}(P, D) \leq \varepsilon\}. \tag{7}$$

Symmetrically, we let $k^*(n, \varepsilon, P, D)$ denote the maximum number of source symbols that can be transmitted over n uses of parallel AWGN channels with maximum power constraint P and independent quasi-static fading so that the excess-distortion probability $\mathrm{P}_{\mathrm{e},k,n}(P, D)$ is bounded above by ε, i.e.,

$$k^*(n, \varepsilon, P, D) := \sup\{k : \ \exists \text{ an } (k, n, P)\text{−code s.t. } \mathrm{P}_{\mathrm{e},k,n}(P, D) \leq \varepsilon\}. \tag{8}$$

In this paper, we will present non-asymptotic bounds on $\mathrm{P}^*_{\mathrm{e},k,n}(P, D)$ and second-order asymptotics for $\mathrm{P}^*_{\mathrm{e},k,n}(P, D)$ and $k^*(n, \varepsilon, P, D)$.

III. Main Results

A. Preliminaries

Recall the definitions of the rate-distortion function $R(P_S, D)$ [18], [19] and the distortion-tilted information density [31], [12], i.e.,

$$R(P_S, D) := \inf_{P_{\hat{S}|S} : \mathbb{E}[d(S,\hat{S})] \leq D} I(S; \hat{S}), \tag{9}$$

$$\jmath(s, D) := -\log \mathbb{E}[\exp(\lambda^* D - d(s, \hat{S}))], \tag{10}$$

$$\lambda^* := -\frac{\partial R(P_S, D)}{\partial D}, \tag{11}$$

This document is a preprint of: L. Zhou, A. Wolf and M. Motani, "On Lossy Multi-Connectivity: Finite Blocklength Performance and Second-Order Asymptotics," in IEEE Journal on Selected Areas in Communications (JSAC), Feb 2019. DOI:10.1109/JSAC.2019.2898755

where the expectation in (10) is with respect to the unconditional distribution $P_{\hat{S}}^*$, which is induced by the optimal test channel $P_{\hat{S}|S}^*$ for (9) and the source distribution P_S.

The definition of the distortion-tilted information density in (10) can also be extended to multi-letter case and thus we can define $\jmath(s^k, D)$ for any length-k source sequence s^k. It was shown in [12] that for any memoryless source, we have

$$\jmath(s^k, D) = \sum_{i \in [k]} \jmath(s_i, D). \tag{12}$$

Recall the channel law in (1). For each $t \in [\Psi]$, the conditional distribution of the channel output Y_t^n given the input X_t^n and the fading parameter A_t is a product of normal distributions, i.e., for any (x_t^n, a_t, y_t^n),

$$P_{Y_t^n|X_t^n A_t}(y_t^n|x_t^n, a_t) = \prod_{i \in [n]} \frac{1}{\sqrt{2\pi}} \exp\bigg(-\frac{(y_{t,i} - a_t x_{t,i})^2}{2}\bigg). \tag{13}$$

Now, let $\{Q_{Y_t^n|A_t}\}_{t \in [\Psi]}$ be arbitrary conditional distributions and define the fading information densities [16, Eq. (9)] as

$$\tilde{\imath}(x_t^n; y_t^n|a_t) := \log \frac{P_{Y_t^n|X_t^n A_t}(y_t^n|x_t^n, a_t)}{Q_{Y_t^n|A_t}(y_t^n|a_t)}. \tag{14}$$

For simplicity, throughout the paper, we let $\mathbf{P} := (P_1, \ldots, P_\Psi)$ denote the power allocation vector and use $\mathbf{A} := (A_1, \ldots, A_\Psi)$ to denote fading parameters. Furthermore, we let $\mathcal{P}_{\max}(P) := \{\mathbf{P} : \sum_{t \in [\Psi]} P_t \le P\}$ and $\mathcal{P}_{\mathrm{eq}}(P) := \{\mathbf{P} : \sum_{t \in [\Psi]} P_t = P\}$ be the set of power allocation vectors satisfying the maximum and equal power constraints respectively.

B. Non-Asymptotic Bounds

Our first result is a lower bound on the excess-distortion probability for any (k, n, P)-code.

Theorem 1. *Given a distortion level D, any (k, n, P)-code satisfies that*

$$\begin{aligned}
&\mathrm{P}_{\mathrm{e},k,n}(P, D) \\
&\ge \inf_{\substack{(\mathbf{P}, \{P_{X_t^n|S^k}\}_{t \in [\Psi]}): \\ \mathbf{P} \in \mathcal{P}_{\max}(P),\ \text{and}\ \forall\ t \in [\Psi] \\ P_{X_t^n|S^k} \in \mathcal{P}(\mathbb{R}(n, P_t)|\mathcal{S}^k)}} \sup_{\gamma > 0} \bigg(-\exp(-n\gamma) \\
&\quad + \Pr\bigg\{\sum_{t \in [\Psi]} \tilde{\imath}(X_t^n; Y_t^n|A_t) \le \jmath(S^k, D) - n\gamma\bigg\}\bigg).
\end{aligned} \tag{15}$$

We remark that the proof of Theorem 1 is similar to [13, Theorem 1] and omitted for simplicity. Note that in the right hand side of (15), the infimum over all stochastic encoders $P_{X_t^n|S^k}$, $t \in [\Psi]$ makes the lower bound difficult to calculate. However, the following lemma states that for any AWGN channel with quasi-static fading, the right hand side of (15) does not depend on encoders under particular choices of auxiliary distributions $\{Q_{Y_t^n|A_t}\}_{t \in [\Psi]}$.

To present the lemma, for each $t \in [\Psi]$, given the power P_t and the fading coefficient A_t, define

$$\begin{aligned}
L_t^n(P_t, A_t, Z_t^n) &:= \frac{n}{2} \log(1 + P_t A_t^2) \\
&\quad + \frac{\sum_{j \in [n]} P_t A_t^2 (1 - Z_{t,j}^2) + 2A_t \sqrt{P_t} Z_{t,j}}{2(1 + P_t A_t^2)}.
\end{aligned} \tag{16}$$

Lemma 2. *For each $t \in [\Psi]$, given P_t and a_t, choose the distribution $Q_{Y_t^n|A_t}$ such that*

$$Q_{Y_t^n|A_t}(\cdot|a_t) \sim \mathcal{N}(0, (1 + P_t a_t^2)\mathbf{I}_n). \tag{17}$$

Then, for any $t \in [\Psi]$, channel input x_t^n and fading parameter a_t, under the distribution of $P_{Y_t^n|X_t^n A_t}(\cdot|x_t^n, a_t)$ (see (13)), the distribution of the fading information density $\tilde{\imath}(x_t^n; Y_t^n|a_t)$ depends on x_t^n only through its power $P_t = \frac{\|x_t^n\|^2}{n}$. Furthermore, $\tilde{\imath}(x_t^n; Y_t^n|A_t)$ has the same distribution as $L_t^n(P_t, A_t, Z_t^n)$.

The proof of Lemma 2 follows from spherical symmetry and is given in Appendix A.

Invoking Theorem 1 and Lemma 2, we can now obtain the following non-asymptotic converse bound.

Corollary 3. *Given a distortion level D, any (k, n, P)-code satisfies that*

$$\begin{aligned}
\mathrm{P}_{\mathrm{e},k,n}(P, D) &\ge \inf_{\mathbf{P} \in \mathcal{P}_{\max}(P)} \sup_{\gamma > 0} \bigg(-\exp(-n\gamma) \\
&\quad + \Pr\bigg\{\sum_{t \in [\Psi]} L_t^n(P_t, A_t, Z_t^n) \le \jmath(S^k, D) - n\gamma\bigg\}\bigg).
\end{aligned} \tag{18}$$

In the following, we present a non-asymptotic achievability bound by generalizing [13, Theorems 7 and 8] to our lossy multi-connectivity setting. Recall the definition of $L_t^n(\cdot)$ in (16) and that P_S^* is the distribution induced by the optimal test channel for the rate-distortion function $R(P_S, D)$ (see (9)). Given any source sequence s^k, define the non-excess-distortion probability

$$\Phi(s^k, D) := \Pr_{P_{\hat{S}}^{*k}}\{d(s^k, \hat{S}^k) \le D\}. \tag{19}$$

Finally, for simplicity, let

$$\tilde{L}_t^n(P_t, A_t, Z_t^n) := L_t^n(P_t, A_t, Z_t^n) - \log \frac{1 + P_t A_t^2}{\sqrt{1 + 2P_t A_t^2}}. \tag{20}$$

Theorem 4. *Given any distortion level D, there exists an (k, n, P)-code such that*

$$\begin{aligned}
\mathrm{P}_{\mathrm{e},k,n}(P, D) &\le \inf_{\mathbf{P} \in \mathcal{P}_{\max}(P)} \inf_{\gamma > 0} \bigg\{\exp(-n\gamma + 1) \\
&\quad + \mathbb{E}\bigg[\exp\bigg(-\bigg|\sum_{t \in [\Psi]} \tilde{L}_t^n(P_t, A_t, Z_t^n) - \log \frac{n\gamma}{\Phi(S^k, D)}\bigg|^+\bigg)\bigg]\bigg\}.
\end{aligned} \tag{21}$$

The proof of Theorem 4 is given in Appendix B. We adapt the joint source-channel coding scheme in [13] which consists of the concatenation of source and channel codes. We remark that the coding scheme to prove Theorem 4 is non-universal since the decoder needs to know the exact channel law and the source distribution.

We remark that the result in Theorem 4 holds for arbitrary source distributions (discrete and continuous). As we will show later, using non-asymptotic bounds in Corollary 3 and Theorem 4, we can derive the second-order asymptotics for our lossy multi-connectivity problem, which provides tight approximation for $\mathrm{P}^*_{\mathrm{e},k,n}(P,D)$ (see (7)) and $k^*(n,\varepsilon,P,D)$ (see (8)). For the case of lossless transmission, the results in Corollary 3 and Theorem 4 holds with $\jmath(S^k,D)$ and $\Phi(S^k,D)$ replaced by $\log P_S^k(S^k)$.

C. Second-Order Asymptotics for a DMS

In this subsection, we consider any discrete memoryless source under arbitrary bounded distortion measure, i.e., $\mathcal{S}$ and $\hat{\mathcal{S}}$ are finite and $\max_{(s,\hat{s})\in\mathcal{S}\times\hat{\mathcal{S}}} d(s,\hat{s})<\infty$. Before presenting the main theorem, we first recall and define necessary quantities. Recall the definitions of the distortion-dispersion function and the third absolute moment of the distortion-tilted information density [12], i.e.,

$$\mathrm{V_s}(P_S,D):=\mathrm{Var}[\jmath(S,D)], \tag{22}$$
$$T(P_S,D):=\mathbb{E}[|\jmath(S,D)^3|]. \tag{23}$$

Furthermore, given power allocation vector $\mathbf{P}$ and fading parameters $\mathbf{A}$, define

$$U_1(\mathbf{P},\mathbf{A}):=\sum_{t\in[\Psi]}\frac{1}{2}\log(1+P_tA_t^2), \tag{24}$$
$$U_2(\mathbf{P},\mathbf{A}):=\sum_{t\in[\Psi]}\frac{P_tA_t^2(P_tA_t^2+2)}{2(1+P_tA_t^2)}, \tag{25}$$
$$U_3(\mathbf{P},\mathbf{A}):=\sum_{t\in[\Psi]}\frac{\sqrt{2}P_t^2A_t^4(7P_tA_t^2+12)}{(\sqrt{2(1+P_tA_t^2)})^3}. \tag{26}$$

For our result to hold, we need the following assumptions:

(i) Given source distribution P_S and distortion level D,

$$\mathrm{V_s}(P_S,D)>0,\ T(P_S,D)<\infty; \tag{27}$$

(ii) Given any $\mathbf{P}\in\mathcal{P}_{\mathrm{eq}}(P)$,

$$\mathbb{E}_{\mathbf{A}}\left[\frac{U_3(\mathbf{P},\mathbf{A})}{U_2(\mathbf{P},\mathbf{A})}\right]<\infty,\ \mathbb{E}_{\mathbf{A}}\left[\left(\frac{U_3(\mathbf{P},\mathbf{A})}{U_2(\mathbf{P},\mathbf{A})}\right)^2\right]<\infty. \tag{28}$$

Theorem 5. *Suppose $k=\Theta(n)$. Under the assumptions in* (27) *and* (28)*, the optimal excess-distortion probability satisfies that*

$$\mathrm{P}^*_{\mathrm{e},k,n}(P,D)-O\left(\frac{\log n}{\sqrt{n}}\right) = \inf_{\mathbf{P}\in\mathcal{P}_{\mathrm{eq}}(P)}\mathbb{E}_{\mathbf{A}}\left[\mathrm{Q}\left(\frac{nU_1(\mathbf{P},\mathbf{A})-kR(P_S,D)}{\sqrt{k\mathrm{V_s}(P_S,D)+nU_2(\mathbf{P},\mathbf{A})}}\right)\right]. \tag{29}$$

The proof of Theorem 5 is given in Sections V-A to V-C. Several remarks are in order.

First, the result in Theorem 5 follows by applying the Berry-Esseen theorem to our non-asymptotic bounds in Corollary 3 and Theorem 4. Since the result in Theorem 4 is based on a non-universal JSCC scheme, we also prove the achievability part using a universal JSCC scheme in Section V-D, which uses unequal error protection [28] and modified minimum distance decoding [27]. We remark that the additional requirement $\mathbb{E}_{P_A}\left[\frac{1\{A>0\}}{A}\right]<\infty$ and $\mathbb{E}_{P_A}[A]<\infty$ is needed in order for our universal JSCC scheme to achieve the performance in Theorem 5.

Second, as can be seen in Theorem 5, the excess-distortion probability is dominated by the first term, which can be understood as the outage probability in the second-order sense. Actually, using [3, Lemma 17], we have

$$\inf_{\mathbf{P}\in\mathcal{P}_{\mathrm{eq}}(P)}\mathbb{E}_{\mathbf{A}}\left[\mathrm{Q}\left(\frac{nU_1(\mathbf{P},\mathbf{A})-kR(P_S,D)}{\sqrt{k\mathrm{V_s}(P_S,D)+nU_2(\mathbf{P},\mathbf{A})}}\right)\right] = \mathrm{P_{out}}(k,n,D)+O\left(\frac{1}{n}\right), \tag{30}$$

where the outage probability is defined as

$$\mathrm{P_{out}}(k,n,D):=\inf_{\substack{\mathbf{P}\in\\ \mathcal{P}_{\mathrm{eq}}(P)}}\Pr\left\{nU_1(\mathbf{P},\mathbf{A})\le kR(P_S,D)\right\}. \tag{31}$$

Some interesting observations can be made. Theorem 5 indicates that the optimal excess-distortion probability is dominated by the second-order outage probability expressed using the $\mathrm{Q}(\cdot)$ function and the linear combination of dispersion functions $k\mathrm{V_s}(P_S,D)+nU_2(\mathbf{P},\mathbf{A})$. However, (30) shows that with negligible loss of performance (for relatively large blocklength), it suffices to use the outage probability directly. Therefore, our results indicate that the outage probability is still a valid criterion even in finite blocklength. In general, the optimal power allocation to minimize the outage probability is unknown. However, for parallel fading channels with i.i.d. Rayleigh fading, from [8, Exercise 5.17], we know that the outage probability $\mathrm{P_{out}}(k,n,D)$ is achieved by equal power allocation, i.e., $P_t=\frac{P}{\Psi}$ for all $t\in[\Psi]$.

Finally, we comment a bit on the assumptions for Theorem 5. The assumption in (28) is mild since it is satisfied by many common fading distributions, such as the Rayleigh fading [24]. Furthermore, the assumption that $k=\Theta(n)$ is valid since we can show that by allowing any non-vanishing excess-distortion probability, the asymptotic ratio of $\frac{k}{n}$ is finite and positive.

Using Theorem 5 and [3, Lemma 17], we obtain an approximation for $k^*(n,\varepsilon,P,D)$ in (8). To present the result, given any maximal power constraint $P\in\mathcal{R}_+$ and any excess-distortion probability $\varepsilon\in(0,1)$, define the outage capacity for our lossy multi-connectivity problem as

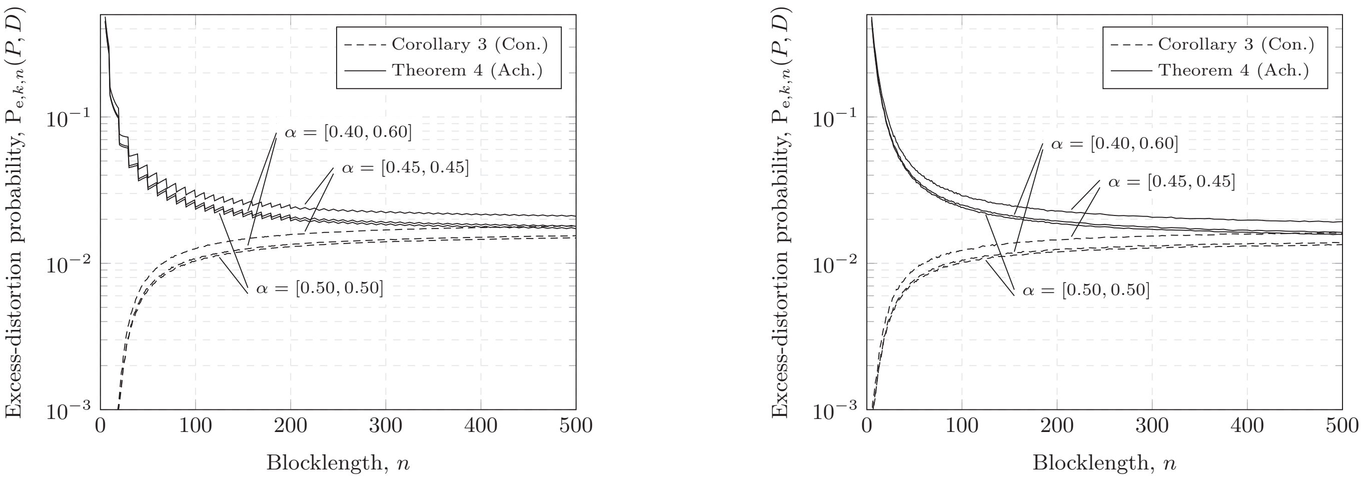
$$\mathrm{C_{out}^{MC}}(P,D,\varepsilon):=\sup\left\{r\in\mathcal{R}_+:\mathrm{P_{out}}(r,1,D)\le\varepsilon\right\}. \tag{32}$$

Corollary 6. *Under conditions of Theorem 5, for any continuously differentiable fading distribution P_A, under any maximal power constraint $P\in\mathcal{R}_+$ and any excess-distortion probability $\varepsilon\in(0,1)$, we have*

$$k^*(n,\varepsilon,P,D)=n\mathrm{C_{out}^{MC}}(P,D,\varepsilon)+O(\log n). \tag{33}$$

The proof of Corollary 6 is provided in Section V-E. We remark that Corollary 6 implies that the dispersion for our lossy multi-connectivity problem is zero, which is consistent with existing literatures [3], [16]. Furthermore, our result in Corollary 6 justifies that the asymptotic notion of outage capacity, used extensively in wireless communication systems

This document is a preprint of: L. Zhou, A. Wolf and M. Motani, “On Lossy Multi-Connectivity: Finite Blocklength Performance and Second-Order Asymptotics,” in IEEE Journal on Selected Areas in Communications (JSAC), Feb 2019. DOI:10.1109/JSAC.2019.2898755



(a) BMS with distortion level $D = 0.1$

(b) GMS with distortion level $D = 0.5$

Fig. 2: Power allocation: excess-distortion probability for lossy transmission over $\Psi = 2$ parallel AWGN channels with independent quasi-static Rayleigh fading for different power allocations $\alpha = [\alpha_1, \alpha_2]$ with power constraint $P = 5$.

with quasi-static fading, is actually an accurate performance criterion even in the finite blocklength setting.

We remark that the results in Theorem 5 and Corollary 6 can also be established for Gaussian memoryless sources under quadratic distortion measures by using our non-asymptotic converse bounds. The converse proof remains the same and the achievability part can be done by adapting the universal coding scheme in [27] to the lossy multi-connectivity setting considered in the present paper.

IV. Numerical Simulation

In this section, we illustrate our results in Corollary 3 and Theorem 4 via numerical simulations. We consider Rayleigh fading with scale parameter one, i.e., each fading parameter follows the same Rayleigh distribution $P_A(a_t) = a_t \exp\left(-a_t^2/2\right)$ for any $a_t \in \mathcal{R}_+$. We are interested in the finite blocklength performance of the lossy transmission of the following two memoryless sources over parallel AWGN channels with independent quasi-static Rayleigh fading:

- Binary memoryless source (BMS) with bias p under Hamming distortion measure. The source alphabet is $\mathcal{S} = \{0, 1\}$ and the source distribution is $P_S(0) = p$ and $P_S(1) = 1 - p$. For any two sequences s^k and $\hat{s}^k$, the Hamming distortion is defined as $d(s^k, \hat{s}^k) = \frac{1}{k}\sum_{i\in[k]} 1\{s_i \neq \hat{s}_i\}$.
- GMS under the quadratic distortion measure. The source distribution is $\mathcal{N}(0, \sigma_S^2)$ and the quadratic distortion measure for any two sequences s^k and $\hat{s}^k$ is defined as $d(s^k, \hat{s}^k) = \frac{1}{k}\sum_{i\in[k]}(s_i - \hat{s}_i)^2$.

In the following, we assume a maximal power constraint of $P = 5$, a fixed rate of source symbols per channel use of $k/n = 1$ and up to three parallel AWGN channels, i.e., $\Psi \leq 3$.

In Figure 2, we plot the non-asymptotic bounds in Corollary 3 (dashed lines) and Theorem 4 (solid lines) via Monte-Carlo simulations for the case of $\Psi = 2$ and different power allocations $\mathbf{P} = [P_1, P_2] = [\alpha_1, \alpha_2]P$ where $(\alpha_1, \alpha_2) \in \mathcal{R}_+$ satisfying $\alpha_1 + \alpha_2 \leq 1$. It is shown that the equal power allocation (i.e., $\alpha_1 = \alpha_2 = 0.5$) achieves the smallest excess-distortion probabilities in both the achievability and converse parts. We have verified that this is true for any possible values of $(\alpha_1, \alpha_2) \in \mathcal{R}_+$ such that $\alpha_1 + \alpha_2 \leq 1$ and thus conclude that the equal power allocation is optimal in the finite blocklength regime[1]. Hence, in further numerical simulations, we plot our results for the equal power allocation only.

In Figure 3, for the case of $\Psi \in [3]$, we plot the non-asymptotic achievability bound in Theorem 4 (solid lines), the second-order asymptotics in Theorem 5 (dotted lines), and outage probability in (31) (dash-dotted lines) versus the blocklength. It can be observed that the achievable excess-distortion probability decreases significantly as the number of parallel channels increases and thus demonstrates the benefits of multi-connectivity in the finite blocklength setting. The same result is true for the converse result (see Corollary 3).

V. Proof of Second-Order Asymptotics

A. Preliminaries

Using the definitions in (22), and (24) to (26), define the average dispersion function and the third-absolute moment as

$$\mathrm{V}_{k+n\Psi}^{(\mathbf{P},\mathbf{A})} := \frac{k\mathrm{V}_{\mathrm{s}}(P_S, D) + nU_2(\mathbf{P}, \mathbf{A})}{k + n\Psi}, \tag{34}$$

$$T_{k+n\Psi}^{(\mathbf{P},\mathbf{A})} := \frac{kT(P_S, D) + nU_3(\mathbf{P}, \mathbf{A})}{k + n\Psi}. \tag{35}$$

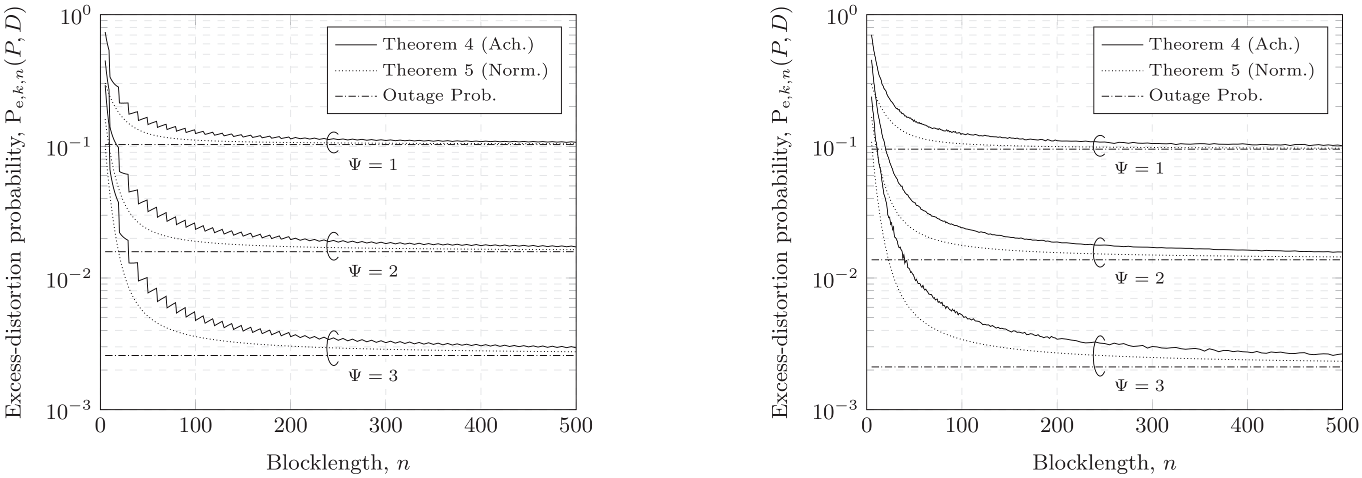
Furthermore, define the following three quantities

$$\Omega_1(k, n, P) := \sup_{\mathbf{P}\in\mathcal{P}_{\mathrm{eq}}(P)} \mathbb{E}_{\mathbf{A}}\left[6T_{k+n\Psi}^{(\mathbf{P},\mathbf{A})}\left(\sqrt{\mathrm{V}_{k+n\Psi}^{(\mathbf{P},\mathbf{A})}}\right)^{-3}\right], \tag{36}$$

$$\Omega_2(k, n, P) := \sup_{\mathbf{P}\in\mathcal{P}_{\mathrm{eq}}(P)} \mathbb{E}_{\mathbf{A}}\left[\left(\sqrt{\mathrm{V}_{k+n\Psi}^{(\mathbf{P},\mathbf{A})}}\right)^{-1}\right], \tag{37}$$

[1]The simulations included a great number of power allocations, which are omitted in Fig. 2 for clarity.

This document is a preprint of: L. Zhou, A. Wolf and M. Motani, “On Lossy Multi-Connectivity: Finite Blocklength Performance and Second-Order Asymptotics,” in IEEE Journal on Selected Areas in Communications (JSAC), Feb 2019. DOI:10.1109/JSAC.2019.2898755



(a) BMS with distortion level $D = 0.1$

(b) GMS with distortion level $D = 0.5$

Fig. 3: Multiple-links: excess-distortion probability for lossy transmission over parallel AWGN channels with independent quasi-static Rayleigh fading channel for equal power allocation with $P = 5$ and $\Psi \in [3]$.

$$\Omega_3(k,n,P) := \sup_{\substack{\mathbf{P}\in\\ \mathcal{P}_{\mathrm{eq}}(P)}} \mathbb{E}_{\mathbf{A}}\left[U_1(\mathbf{P},\mathbf{A})\left(\sqrt{\mathrm{V}_{k+n\Psi}^{(\mathbf{P},\mathbf{A})}}\right)^{-1}\right]. \tag{38}$$

B. Converse Proof

In this section, we present the converse proof of our main results. Given a distortion level D, similar to the definition of $\mathrm{P}^*_{\mathrm{e},k,n}(P,D)$ in (7), we let $\mathrm{P}^{*,\mathrm{eq}}_{\mathrm{e},k,n}(P,D)$ denote the corresponding minimal excess-distortion probability for any (k,n,P)-code with equal power constraint. It can be easily shown that

$$\mathrm{P}^*_{\mathrm{e},k,n}(P,D) \geq \mathrm{P}^{*,\mathrm{eq}}_{\mathrm{e},k,n+1}(P,D) \tag{39}$$

since for any (k,n,P)-code with strict inequality in (5), we can always construct a $(k,n+1,P)$-code with equality in (5). In this section, we will first derive a lower bound on $\mathrm{P}^{*,\mathrm{eq}}_{\mathrm{e},k,n}(P,D)$ and then use (39) to establish a lower bound for $\mathrm{P}^*_{\mathrm{e},k,n}(P,D)$. Furthermore, the results in Theorem 1 and Corollary 3 hold also with equal power constraint as can be gleaned in their proofs. For ease of notation, let

$$\underline{L}(P_t,A_t,Z_t^n) := \frac{\sum_{j\in[n]} P_t A_t^2(1-Z_{t,j}^2) + 2A_t\sqrt{P_t}Z_{t,j}}{2(1+P_tA_t^2)}. \tag{40}$$

In the following, we first bound the dominant term in Corollary 3. Using the definitions of $L_t^n(P_t,A_t,Z_t^n)$ in (16) and $\jmath(S^k,D)$ in (12), for any power allocation vector $\mathbf{P}\in\mathcal{P}_{\mathrm{eq}}(P)$ and $\gamma > 0$, we have that

$$\begin{aligned}
&\Pr\left\{\sum_{t\in[\Psi]} L_t^n(P_t,A_t,Z_t^n) \leq \jmath(S^k,D) - n\gamma\right\}\\
&= \Pr\left\{\sum_{t\in[\Psi]}\left(\frac{n}{2}\log(1+P_tA_t^2) + \underline{L}(P_t,A_t,Z_t^n)\right)\right.\\
&\qquad\qquad \left.\leq \sum_{i\in[k]} \jmath(S_i,D) - n\gamma\right\} \qquad (41)\\
&= \Pr\left\{\sum_{i\in[k]}\big(\jmath(S_i,D) - R(P_S,D)\big) - \sum_{t\in[\Psi]}\underline{L}(P_t,A_t,Z_t^n)\right.\\
&\qquad\qquad \left.\geq \sum_{t\in[\Psi]}\frac{n}{2}\log(1+P_tA_t^2) - kR(P_S,D) + n\gamma\right\}. \qquad (42)
\end{aligned}$$

Using the fact that $\mathbb{E}[\jmath(S_i,D)] = R(P_S,D)$ [12, Property 1], we conclude that $\{\jmath(S_i,D) - R(P_S,D)\}_{i\in[k]}$ and $\left\{\frac{P_tA_t^2(1-Z_{t,j}^2+2A_t\sqrt{P_t}Z_{t,j})}{2(1+P_tA_t^2)}\right\}_{t\in[\Psi],j\in[n]}$ forms a sequence of $(k+n\Psi)$ independent random variables with zero mean. Applying the Berry-Esseen theorem [32], [33] (see also [34, Chapter 1]) to the result in (42), we obtain that

$$\begin{aligned}
&\mathbb{E}_{\mathbf{A}}\left[\left|\Pr\left\{\sum_{t\in[\Psi]} L_t^n(P_t,A_t,Z_t^n) \leq \jmath(S^k,D) - n\gamma\right\}\right.\right.\\
&\qquad \left.\left. - \mathrm{Q}\left(\frac{\sum_{t\in[\Psi]}\frac{n}{2}\log(1+P_tA_t^2) - kR(P_S,D) + n\gamma}{\sqrt{(k+n\Psi)\mathrm{V}_{k+n\Psi}^{(\mathbf{P},\mathbf{A})}}}\right)\right|\right]\\
&\leq \mathbb{E}_{\mathbf{A}}\left[\frac{6T_{k+n\Psi}^{(\mathbf{P},\mathbf{A})}\left(\mathrm{V}_{k+n\Psi}^{(\mathbf{P},\mathbf{A})}\right)^{-3/2}}{\sqrt{(k+n\Psi)}}\right]. \qquad (43)
\end{aligned}$$

Now, using Corollary 3 and the result in (43), by choosing $\gamma = \frac{\log n}{n}$, we have

$$\begin{aligned}
&\mathrm{P}^{*,\mathrm{eq}}_{\mathrm{e},k,n}(P,D) + \frac{1}{n}\\
&\geq \inf_{\mathbf{P}\in\mathcal{P}_{\mathrm{eq}}(P)}\left\{\mathbb{E}_{\mathbf{A}}\left[\mathrm{Q}\left(\frac{nU_1(\mathbf{P},\mathbf{A}) - kR(P_S,D) + \log n}{\sqrt{(k+n\Psi)\mathrm{V}_{k+n\Psi}^{(\mathbf{P},\mathbf{A})}}}\right)\right.\right.\\
&\qquad\qquad \left.\left. - \frac{6T_{k+n\Psi}^{(\mathbf{P},\mathbf{A})}\left(\mathrm{V}_{k+n\Psi}^{(\mathbf{P},\mathbf{A})}\right)^{-3/2}}{\sqrt{(k+n\Psi)}}\right]\right\} \qquad (44)\\
&\geq \inf_{\mathbf{P}\in\mathcal{P}_{\mathrm{eq}}(P)}\mathbb{E}_{\mathbf{A}}\left[\mathrm{Q}\left(\frac{nU_1(\mathbf{P},\mathbf{A}) - kR(P_S,D) + \log n}{\sqrt{(k+n\Psi)\mathrm{V}_{k+n\Psi}^{(\mathbf{P},\mathbf{A})}}}\right)\right]\\
&\qquad - \sup_{\mathbf{P}\in\mathcal{P}_{\mathrm{eq}}(P)}\mathbb{E}_{\mathbf{A}}\left[\frac{6T_{k+n\Psi}^{(\mathbf{P},\mathbf{A})}\left(\mathrm{V}_{k+n\Psi}^{(\mathbf{P},\mathbf{A})}\right)^{-3/2}}{\sqrt{(k+n\Psi)}}\right] \qquad (45)
\end{aligned}$$

$$= \inf_{\mathbf{P}\in\mathcal{P}_{\mathrm{eq}}(P)} \mathbb{E}_{\mathbf{A}}\left[\mathrm{Q}\left(\frac{nU_1(\mathbf{P},\mathbf{A})-kR(P_S,D)+\log n}{\sqrt{(k+n\Psi)\mathrm{V}_{k+n\Psi}^{(\mathbf{P},\mathbf{A})}}}\right)\right] - \frac{\Omega_1(k,n,P)}{\sqrt{(k+n\Psi)}} \quad (46)$$

$$= \inf_{\mathbf{P}\in\mathcal{P}_{\mathrm{eq}}(P)} \mathbb{E}_{\mathbf{A}}\left[\mathrm{Q}\left(\frac{nU_1(\mathbf{P},\mathbf{A})-kR(P_S,D)+\log n}{\sqrt{(k+n\Psi)\mathrm{V}_{k+n\Psi}^{(\mathbf{P},\mathbf{A})}}}\right)\right] + O\Big(\frac{1}{\sqrt{n}}\Big), \quad (47)$$

where (46) follows from the definition of $\Omega_1(k,n,P)$ in (36) and (47) follows since $\Omega_1(k,n,P)=O(1)$ from the assumptions of Theorem 5.

Combining (39) and (46), we obtain that

$$\mathrm{P}_{\mathrm{e},k,n}^*(P,D)$$
$$\geq \mathrm{P}_{\mathrm{e},k,n+1}^{*,\mathrm{eq}}(P,D) \quad (48)$$
$$\geq O\Big(\frac{1}{\sqrt{n+1}}\Big) + \inf_{\mathbf{P}\in\mathcal{P}_{\mathrm{eq}}(P)} \mathbb{E}_{\mathbf{A}}\left[\mathrm{Q}\left(\frac{(n+1)U_1(\mathbf{P},\mathbf{A})-kR(P_S,D)+\log(n+1)}{\sqrt{(k+(n+1)\Psi)\mathrm{V}_{k+(n+1)\Psi}^{(\mathbf{P},\mathbf{A})}}}\right)\right] \quad (49)$$
$$\geq \inf_{\mathbf{P}\in\mathcal{P}_{\mathrm{eq}}(P)} \mathbb{E}_{\mathbf{A}}\left[\mathrm{Q}\left(\frac{nU_1(\mathbf{P},\mathbf{A})-kR(P_S,D)+\log n}{\sqrt{(k+n\Psi)\mathrm{V}_{k+n\Psi}^{(\mathbf{P},\mathbf{A})}}}\right) - \frac{U_1(\mathbf{P},\mathbf{A})+\log(1+1/n)}{\sqrt{(k+n\Psi)\mathrm{V}_{k+n\Psi}^{(\mathbf{P},\mathbf{A})}}}\right] + O\Big(\frac{1}{\sqrt{n}}\Big) \quad (50)$$
$$= \inf_{\mathbf{P}\in\mathcal{P}_{\mathrm{eq}}(P)} \mathbb{E}_{\mathbf{A}}\left[\mathrm{Q}\left(\frac{nU_1(\mathbf{P},\mathbf{A})-kR(P_S,D)+\log n}{\sqrt{(k+n\Psi)\mathrm{V}_{k+n\Psi}^{(\mathbf{P},\mathbf{A})}}}\right)\right] - \frac{\Omega_3(k,n,P)}{\sqrt{k+n\Psi}} + O\Big(\frac{1}{\sqrt{n}}\Big) \quad (51)$$
$$= \inf_{\mathbf{P}\in\mathcal{P}_{\mathrm{eq}}(P)} \mathbb{E}_{\mathbf{A}}\left[\mathrm{Q}\left(\frac{nU_1(\mathbf{P},\mathbf{A})-kR(P_S,D)+\log n}{\sqrt{(k+n\Psi)\mathrm{V}_{k+n\Psi}^{(\mathbf{P},\mathbf{A})}}}\right)\right] + O\Big(\frac{1}{\sqrt{n}}\Big), \quad (52)$$
$$\geq \inf_{\mathbf{P}\in\mathcal{P}_{\mathrm{eq}}(P)} \mathbb{E}_{\mathbf{A}}\left[\mathrm{Q}\left(\frac{nU_1(\mathbf{P},\mathbf{A})-kR(P_S,D)}{\sqrt{(k+n\Psi)\mathrm{V}_{k+n\Psi}^{(\mathbf{P},\mathbf{A})}}}\right)\right] - \frac{\Omega_2(k,n,P)\log n}{\sqrt{k+n\Psi}} + O\Big(\frac{1}{\sqrt{n}}\Big) \quad (53)$$
$$= \inf_{\mathbf{P}\in\mathcal{P}_{\mathrm{eq}}(P)} \mathbb{E}_{\mathbf{A}}\left[\mathrm{Q}\left(\frac{nU_1(\mathbf{P},\mathbf{A})-kR(P_S,D)}{\sqrt{(k+n\Psi)\mathrm{V}_{k+n\Psi}^{(\mathbf{P},\mathbf{A})}}}\right)\right] + O\Big(\frac{\log n}{\sqrt{n}}\Big), \quad (54)$$

where (50) follows from [13, Eq. (255)] which states that $\mathrm{Q}(x+a)\geq \mathrm{Q}(x)-\frac{a}{\sqrt{2\pi}}\geq \mathrm{Q}(x)-a$ for any x and any $a\geq 0$, (51) follows from the definition of $\Omega_3(k,n,P)$ in (38) and the fact that $\frac{\log(1+1/n)}{\sqrt{(k+n\Psi)\mathrm{V}_{k+n\Psi}^{(\mathbf{P},\mathbf{A})}}}\leq \frac{\log 2}{\sqrt{k\mathrm{V}(P_S)}}=O(\frac{1}{\sqrt{n}})$ for any $n\geq 1$ since $k=\Theta(n)$, and (52) follows since $\Omega_3(k,n,P)=O(1)$ similarly to (47); (53) follows similarly to (50) and using the definition of $\Omega_2(\cdot)$ in (37), and (54) follows since $\Omega_2(k,n,P)=O(1)$ according to the assumption of Theorem 5.

C. Achievability Proof

For simplicity, in the following, we use $\mathbf{Z}$ to denote $\{Z_t^n\}_{t\in[\Psi]}$ and similarly $\mathbf{z}$. In the achievability proof, the following lemma [12, Lemma 2] is important. Recall the definitions of $\jmath(s,D)$ in (10) and $\Phi(s^k,D)$ in (19).

Lemma 7. *For any DMS under any bounded distortion measure, there exists constants c_1, c_2, c_3 such that*

$$\Pr\left\{\log\frac{1}{\Phi(S^k,D)}\leq \sum_{i\in[k]}\jmath(S_i,D)+(c_1-0.5)\log k+c_2\right\} \geq 1-\frac{c_3}{\sqrt{k}}. \quad (55)$$

Recall the definition of $\tilde{L}_t^n(\cdot)$ in (20). For simplicity, let

$$\Upsilon_1(\mathbf{P},\mathbf{A},\mathbf{Z},S^k,D):=\sum_{t\in[\Psi]}\tilde{L}_t^n(P_t,A_t,Z_t^n)-\log n - \log\gamma-\log\frac{1}{\Phi(S^k,D)}, \quad (56)$$

$$\Upsilon(\mathbf{P},\mathbf{A},\mathbf{Z},S^k,D):=\Upsilon_1(\mathbf{P},\mathbf{A},\mathbf{Z},S^k,D)+\log\frac{1}{\Phi(S^k,D)} - \Upsilon_2(S^k,D), \quad (57)$$

where

$$\Upsilon_2(S^k,D):=\sum_{i\in[k]}\jmath(S_i,D)+(c_1-0.5)\log k+c_2. \quad (58)$$

Furthermore, for any power vector $\mathbf{P}$, define the set:

$$\mathcal{T}_{k,n}(\mathbf{P}):=\Big\{(\mathbf{a},\mathbf{z},s^k):\Upsilon(\mathbf{P},\mathbf{a},\mathbf{z},s^k,D)\geq\log n\Big\}. \quad (59)$$

Using the result in Lemma 7, and weakening the result in Theorem 4 by letting $\gamma=\frac{\log n+1}{n}$, we obtain that there exists an (k,n,P)-code such that for any $\mathbf{P}\in\mathcal{P}_{\max}(P)$,

$$\mathrm{P}_{\mathrm{e},k,n}(P,D)$$
$$\leq \mathbb{E}\left[\exp\left(-\left|\Upsilon_1(\mathbf{P},\mathbf{A},\mathbf{Z},S^k,D)\right|^+\right)\right]+\frac{1}{n} \quad (60)$$
$$\leq \mathbb{E}\Bigg[\exp\left(-\left|\Upsilon(\mathbf{P},\mathbf{A},\mathbf{Z},S^k,D)\right|^+\right) \times 1\left\{\log\frac{1}{\Phi(S^k,D)}\leq\Upsilon_2(S^k,D)\right\}\Bigg] + \Pr\left\{\log\frac{1}{\Phi(S^k,D)}>\Upsilon_2(S^k,D)\right\}+\frac{1}{n} \quad (61)$$
$$\leq \mathbb{E}\left[\exp\left(-\left|\Upsilon(\mathbf{P},\mathbf{A},\mathbf{Z},S^k,D)\right|^+\right)\right]+\frac{1}{n}+\frac{c_3}{\sqrt{k}} \quad (62)$$
$$\leq \frac{1}{n}\Pr\{(\mathbf{A},\mathbf{Z},S^k)\in\mathcal{T}_{k,n}(\mathbf{P})\} + \Pr\Big\{(\mathbf{A},\mathbf{Z},S^k)\notin\mathcal{T}_{k,n}(\mathbf{P})\Big\}+\frac{1}{n}+\frac{c_3}{\sqrt{k}} \quad (63)$$
$$\leq \frac{2}{n}+\frac{c_3}{\sqrt{k}}+\Pr\Big\{(\mathbf{A},\mathbf{Z},S^k)\notin\mathcal{T}_{k,n}(\mathbf{P})\Big\}, \quad (64)$$

where (61) follows since i) $\Upsilon(\mathbf{P},\mathbf{A},\mathbf{Z},S^k,D)\leq\Upsilon_1(\mathbf{P},\mathbf{A},\mathbf{Z},S^k,D)$ if $\log\frac{1}{\Phi(S^k,D)}\leq\Upsilon_2(S^k,D)$

This document is a preprint of: L. Zhou, A. Wolf and M. Motani, “On Lossy Multi-Connectivity: Finite Blocklength Performance and Second-Order Asymptotics,” in IEEE Journal on Selected Areas in Communications (JSAC), Feb 2019. DOI:10.1109/JSAC.2019.2898755

and $\exp(-|x|^+)$ is non-increasing in x, and ii) $\exp(-|x|^+) \le 1$ for any x (recall that $|x|^+ = \max\{0,x\}$), (62) follows from Lemma 7, and (63) follows since $\mathbb{E}[\exp(-|X|^+)] \le \frac{1}{a}\Pr\{X \ge \log a\} + \Pr\{X \le \log a\}$ for any $a \ge 1$.

In the remaining part of this subsection, we will upper bound the third term in (64). For simplicity, define the following function of power allocation vector $\mathbf{P}$ and the fading vector $\mathbf{A}$,

$$B_1(\mathbf{P},\mathbf{A}) := \sum_{t\in[\Psi]} \log \frac{1+P_tA_t^2}{\sqrt{1+2P_tA_t^2}} + (c_1-0.5)\log k + c_2 + 2\log n + \log\gamma. \tag{65}$$

For ease of following analysis, let

$$C(n,k) := B_1(\mathbf{P},\mathbf{A}) - \sum_{t\in[\Psi]} \log \frac{1+P_tA_t^2}{\sqrt{1+2P_tA_t^2}}. \tag{66}$$

Then, using the definitions of $U_i(\mathbf{P},\mathbf{A})$ for $i\in[3]$ in (24) to (26) and applying the Berry-Esseen theorem similarly to (43), we have that

$$\begin{aligned}
&\Pr\Big\{(\mathbf{A},\mathbf{Z},S^k)\notin\mathcal{T}_{k,n}(\mathbf{P})\Big\}\\
&= \inf_{\mathbf{P}\in\mathcal{P}_{\max}(P)} \Pr\Bigg\{\sum_{i\in[k]} \jmath(S_i,D) - \sum_{t\in[\Psi]} L_t^n(P_t,A_t,Z_t^n) - B_1(\mathbf{P},\mathbf{A}) \ge 0\Bigg\} &(67)\\
&\le \inf_{\mathbf{P}\in\mathcal{P}_{\rm eq}(P)} \Pr\Bigg\{\bigg(\sum_{i\in[k]}\big(\jmath(S_i,D)-R(P_S,D)\big) - \sum_{t\in[\Psi]}\Big(L_t^n(P_t,A_t,Z_t^n)-\frac{n}{2}\log(1+P_tA_t^2)\Big)\bigg) \ge nU_1(\mathbf{P},\mathbf{A}) - kR(P_S,D) - B_1(\mathbf{P},\mathbf{A})\Bigg\} &(68)\\
&\le \inf_{\substack{\mathbf{P}\in\\ \mathcal{P}_{\rm eq}(P)}} \mathbb{E}_{\mathbf{A}}\Bigg[\mathrm{Q}\Bigg(\frac{nU_1(\mathbf{P},\mathbf{A}) - kR(P_S,D) - B_1(\mathbf{P},\mathbf{A})}{\sqrt{(k+n\Psi)\mathrm{V}_{k+n\Psi}^{(\mathbf{P},\mathbf{A})}}}\Bigg)\Bigg] + \mathbb{E}_{\mathbf{A}}\Bigg[\frac{6T_{k+n\Psi}^{(\mathbf{P},\mathbf{A})}\big(\mathrm{V}_{k+n\Psi}^{(\mathbf{P},\mathbf{A})}\big)^{-3/2}}{\sqrt{k+n\Psi}}\Bigg] &(69)\\
&\le \inf_{\substack{\mathbf{P}\in\\ \mathcal{P}_{\rm eq}(P)}} \mathbb{E}_{\mathbf{A}}\Bigg[\mathrm{Q}\Bigg(\frac{nU_1(\mathbf{P},\mathbf{A}) - kR(P_S,D) - B_1(\mathbf{P},\mathbf{A})}{\sqrt{(k+n\Psi)\mathrm{V}_{k+n\Psi}^{(\mathbf{P},\mathbf{A})}}}\Bigg)\Bigg] + \sup_{\mathbf{P}\in\mathcal{P}_{\rm eq}(P)} \mathbb{E}_{\mathbf{A}}\Bigg[\frac{6T_{k+n\Psi}^{(\mathbf{P},\mathbf{A})}\big(\mathrm{V}_{k+n\Psi}^{(\mathbf{P},\mathbf{A})}\big)^{-3/2}}{\sqrt{k+n\Psi}}\Bigg] &(70)\\
&\le \inf_{\substack{\mathbf{P}\in\\ \mathcal{P}_{\rm eq}(P)}} \mathbb{E}_{\mathbf{A}}\Bigg[\mathrm{Q}\Bigg(\frac{nU_1(\mathbf{P},\mathbf{A}) - kR(P_S,D) - C(n,k)}{\sqrt{(k+n\Psi)\mathrm{V}_{k+n\Psi}^{(\mathbf{P},\mathbf{A})}}} + \frac{\sum_{t\in[\Psi]}\log\frac{1+P_tA_t^2}{\sqrt{1+2P_tA_t^2}}}{\sqrt{(k+n\Psi)\mathrm{V}_{k+n\Psi}^{(\mathbf{P},\mathbf{A})}}}\Bigg)\Bigg] + \frac{\Omega_1(k,n,P)}{\sqrt{k+n\Psi}} &(71)\\
&\le \inf_{\substack{\mathbf{P}\in\\ \mathcal{P}_{\rm eq}(P)}} \mathbb{E}_{\mathbf{A}}\Bigg[\mathrm{Q}\Bigg(\frac{nU_1(\mathbf{P},\mathbf{A}) - kR(P_S,D) - C(n,k)}{\sqrt{(k+n\Psi)\mathrm{V}_{k+n\Psi}^{(\mathbf{P},\mathbf{A})}}}\Bigg)\Bigg] + \frac{\Omega_1(k,n,P)+\Omega_3(k,n,P)}{\sqrt{k+n\Psi}} &(72)\\
&= \inf_{\substack{\mathbf{P}\in\\ \mathcal{P}_{\rm eq}(P)}} \mathbb{E}_{\mathbf{A}}\Bigg[\mathrm{Q}\Bigg(\frac{nU_1(\mathbf{P},\mathbf{A}) - kR(P_S,D) + O(\log n)}{\sqrt{(k+n\Psi)\mathrm{V}_{k+n\Psi}^{(\mathbf{P},\mathbf{A})}}}\Bigg)\Bigg] + O\left(\frac{1}{\sqrt{n}}\right) &(73)\\
&\le \inf_{\substack{\mathbf{P}\in\\ \mathcal{P}_{\rm eq}(P)}} \mathbb{E}_{\mathbf{A}}\Bigg[\mathrm{Q}\Bigg(\frac{nU_1(\mathbf{P},\mathbf{A}) - kR(P_S,D)}{\sqrt{(k+n\Psi)\mathrm{V}_{k+n\Psi}^{(\mathbf{P},\mathbf{A})}}}\Bigg)\Bigg] + O\left(\frac{\log n}{\sqrt{n}}\right), &(74)
\end{aligned}$$

where (71) follows from the definitions of $\Omega_1(k,n,P)$ in (36), $C(n,k)$ in (66) and by applying [13, Eq. (255)], (72) follows from the fact that $\log\frac{1+P_tA_t^2}{\sqrt{1+2P_tA_t^2}} \le \frac{1}{2}\log(1+P_tA_t^2)$ for any $t\in[\Psi]$ and the definition of $\Omega_3(k,n,P)$ in (38), (73) follows since $\Omega_1(k,n,P)=O(1)$ and $\Omega_3(k,n,P)=O(1)$ according to the assumptions in Theorem 5, and (74) follows by applying [13, Eq. (255)] similarly to (53) and (54).

D. Universal Achievability Coding Scheme

In this subsection, we present a universal JSCC scheme, which can achieve the same second-order asymptotics as our non-universal counterpart presented in Section V-C. The JSCC scheme considered here dates back to Csiszár [25] who proposed a JSCC scheme consisting of unequal error protection and method of types to derive the error exponent for the JSCC problem. We adapted the idea to our setting by using a modified minimum distance decoder.

1) Notation for Method of Types: Given a length-k sequence S^k, we use $\hat{P}_{S^k}$ to denote its type (empirical distribution). The collection of all types on $\mathcal{S}^k$ is denoted as $\mathcal{P}_k(\mathcal{S})$. For any type $Q\in\mathcal{P}_k(\mathcal{S})$, we use $\mathcal{T}_Q$ to denote the type class, i.e., all the length-k sequences such that $\hat{P}_{S^k}=Q$. For simplicity, we assume that the types in $\mathcal{P}_k(\mathcal{S})$ are ordered, i.e., $\mathcal{P}_k(\mathcal{S}) = \{Q_i\}_{i\in|\mathcal{P}_k(\mathcal{S})|}$ and each index $i\in[|\mathcal{P}_k(\mathcal{S})|]$ is associated with a type Q_i. It is known that [22] $|\mathcal{P}_k(\mathcal{S})| \le (k+1)^{|\mathcal{S}|}$.

2) JSCC Scheme: Let $\{M_i\}_{i\in[|\mathcal{P}_k(\mathcal{S})|]}$ be a sequence of integers to be determined. Furthermore, define the set

$$\mathcal{D} := \{(r_1,r_2)\in\mathcal{R}_+^2 : r_1\in[|\mathcal{P}_k(\mathcal{S})|], r_2\in[M_{r_1}]\}. \tag{75}$$

Codebook generation: For each $i\in[|\mathcal{P}_k(\mathcal{S})|]$, generate M_i independent source codewords $\hat{\mathbf{S}}_i := \{\hat{S}^k(i,1),\ldots,\hat{S}^k(i,M_i)\}$, each i.i.d. according to $P_{\hat{S}}^*$. Furthermore, for each $i\in[|\mathcal{P}_k(\mathcal{S})|]$ and $t\in[\Psi]$, generate M_i independent channel codewords $\{X_t^n(i,1),\ldots,X_t^n(i,M_{t_i})\}$, each according to the uniform distribution over a sphere with radius $\sqrt{nP_t}$.

Encoding: Given S^k, for each $t\in[\Psi]$, the encoder f_t transmits $X_t^n(I,J)$ if the type index of the source sequence is I (i.e., $S^k\in\mathcal{T}_{Q_I}$) and the index minimizing the distortion between the source sequence and the source codewords in the i-th subcodebook $\mathbf{S}_i$ is J, i.e.,

$$J = \operatorname*{arg\,min}_{\tilde{j}\in[1:M_I]} d(S^k,\hat{S}^k(I,\tilde{j})). \tag{76}$$

This document is a preprint of: L. Zhou, A. Wolf and M. Motani, “On Lossy Multi-Connectivity: Finite Blocklength Performance and Second-Order Asymptotics,” in IEEE Journal on Selected Areas in Communications (JSAC), Feb 2019. DOI:10.1109/JSAC.2019.2898755

Decoding: Fix a threshold $\gamma_{k,n}$ to be specified. Given the channel outputs $\{Y_t^n\}_{t\in[\Psi]}$, the decoder ϕ outputs source estimate $\hat{S}^k(\hat{I},\hat{J})$ if $(\hat{I},\hat{J})$ are the unique pair such that

$$\sum_{t\in[\Psi]}\tilde{\imath}(X_t^n(\hat{I},\hat{J});Y_t^n|A_t)-\log M_{\hat{I}}>\log\gamma_{k,n}. \tag{77}$$

Otherwise, the decoder declares an error.

Recalling the definition of $\tilde{\imath}(X_t^n;Y_t^n|A_t)$ in (14), we obtain that

$$\begin{aligned}\sum_{t\in[\Psi]}\tilde{\imath}(X_t^n(\hat{I},\hat{J});Y_t^n|A_t)&=-\frac{\sum_{t\in[\Psi]}\|Y_t^n-A_tX_t^n(\hat{I},\hat{J})\|_2^2}{2}\\&\quad+\sum_{t\in[\Psi]}\left(\frac{n}{2}\log(1+P_tA_t^2)+\frac{\|Y_t^n\|_2^2}{2(1+P_tA_t^2)}\right).\end{aligned} \tag{78}$$

Hence, our decoder in (77) is a universal decoder which requires knowledge of only the channel outputs $\{Y_t^n\}_{t\in[\Psi]}$, the fading parameters $\{A_t\}_{t\in[\Psi]}$, the channel codebook $\{X_t^n(i,j)\}_{(i,j)\in\mathcal{D}}$ for each $t\in[\Psi]$ and the integers $\{M_i\}_{i\in[|\mathcal{P}_k(\mathcal{S})|]}$.

3) Analysis of Excess-Distortion Probability: Given our coding scheme, following the analyses in [26], [27], we can upper bound the excess-distortion probability by

$$\begin{aligned}&\mathrm{P}_{\mathrm{e},k,n}(P,D)\\&\le\sum_{i\in[|\mathcal{P}_k(\mathcal{S})|]}\Pr\{S^k\in\mathcal{T}_{Q_i},\ \min_{\tilde{j}\in[M_i]}d(S^k,\hat{S}^k(i,\tilde{j}))>D\}\\&\quad+\sum_{i\in[|\mathcal{P}_k(\mathcal{S})|]}\Pr\{S^k\in\mathcal{T}_{Q_i},\ (\hat{I},\hat{J})\neq(i,J)\}.\end{aligned} \tag{79}$$

For each $i\in[|\mathcal{P}_k(\mathcal{S})|]$, let

$$\log M_i:=kR(Q_i,D)+C\log k, \tag{80}$$

where $C=4|\mathcal{S}||\hat{\mathcal{S}}|+9$ (see [35, Eq. (74)]). With this choice of M_i, invoking the type covering lemma for the lossy source coding problem [19], [21], [36], we conclude that for n sufficiently large and any $i\in[|\mathcal{P}_k(\mathcal{S})|$,

$$\Pr\{S^k\in\mathcal{T}_{Q_i},\ \min_{\tilde{j}\in[M_i]}d(S^k,\hat{S}^k(i,\tilde{j}))>D\}=0. \tag{81}$$

For simplicity, given $s^k\in\mathcal{T}_{Q_i}$ and the i-th subcodebook $\hat{\mathbf{s}}_i$, we use $j(s^k,\hat{\mathbf{s}}_i)$ to denote the index of the codeword in $\mathbf{s}_i$ which minimizes the distortion with respect to s^k. In the following lemma, we upper bound the error probability of decoding $(i,j(s^k,\hat{\mathbf{s}}_i))$ conditioning on s^k and $\mathbf{s}_i$. To present the lemma, define

$$\begin{aligned}B_2(\mathbf{P},\mathbf{A}):=\sum_{t\in[\Psi]}\frac{1+P_tA_t^2}{\sqrt{1+2P_tA_t^2}}\Bigg(&\frac{2(1+P_tA_t^2)}{\sqrt{\pi}\sqrt{P_tA_t^2(P_tA_t^2+2)}}\\&+\frac{12\sqrt{2}P_tA_t^2(7P_tA_t^2+12)}{(P_tA_t^2+2)\sqrt{P_tA_t^2(P_tA_t^2+2)}}\Bigg).\end{aligned} \tag{82}$$

It can be verified that under the conditions of Theorem 5, we have that there exists a constant c_6 such that

$$\sup_{\mathbf{P}\in\mathcal{P}_{\mathrm{eq}}(P)}\mathbb{E}[B_2(\mathbf{P},\mathbf{A})]\le c_6. \tag{83}$$

Lemma 8. *For any $i\in[|\mathcal{P}_k(\mathcal{S})|]$, given $s^k\in\mathcal{T}_{Q_i}$ and $\hat{\mathbf{s}}_i$, we obtain that*

$$\begin{aligned}&\Pr\{(\hat{I},\hat{J})\neq(i,j(s^k,\hat{\mathbf{s}}_i))|S^k=s^k,\ \hat{\mathbf{S}}_i=\hat{\mathbf{s}}_i\}\\&\le\Pr\left\{\sum_{t\in[\Psi]}\tilde{\imath}(X_t^n;Y_t^n|A_t)\le\log\gamma_{k,n}+\log M_i\right\}\\&\qquad+\frac{|\mathcal{P}_k(\mathcal{S})|\mathbb{E}_{\mathbf{A}}[B_2(\mathbf{P},\mathbf{A})]}{\gamma_{k,n}\sqrt{n}}.\end{aligned} \tag{84}$$

The proof of Lemma 8 is inspired by the dependence testing bound [10, Theorem 18] (see also [16, Theorem 1]) and the analysis of its variant to quasi-static fading channel in [24, Chapter 4]. The proof of Lemma 8 is available in Appendix C for completeness.

Using Lemma 8, choosing $\gamma_{k,n}=|\mathcal{P}_k(\mathcal{S})|\le(k+1)^{|\mathcal{S}|}$ and applying the Berry-Esseen theorem, we obtain that for any $\mathbf{P}\in\mathcal{P}_{\mathrm{eq}}(P)$,

$$\begin{aligned}&\sum_{i\in[|\mathcal{P}_k(\mathcal{S})|]}\Pr\{S^k\in\mathcal{T}_{Q_i},\ (\hat{I},\hat{J})\neq(i,J)\}\\&=\sum_{i\in[|\mathcal{P}_k(\mathcal{S})|]}\sum_{s^k\in\mathcal{T}_{Q_i}}P_S^k(s^k)\sum_{\hat{\mathbf{s}}_i}P_{\hat{S}|S}^k(\hat{s}^k|s^k)\\&\qquad\times\Pr\{(\hat{I},\hat{J})\neq(i,j(s^k,\hat{\mathbf{s}}_i))|S^k=s^k,\ \hat{\mathbf{S}}_i=\hat{\mathbf{s}}_i\} &(85)\\&\le\sum_{i\in[|\mathcal{P}_k(\mathcal{S})|]}\sum_{s^k\in\mathcal{T}_{Q_i}}P_S^k(s^k)\Pr\Bigg\{\sum_{t\in[\Psi]}\tilde{\imath}(X_t^n;Y_t^n|A_t)\\&\qquad\le\log\gamma_{k,n}+\log M_i\Bigg\}+\frac{|\mathcal{P}_k(\mathcal{S})|\mathbb{E}_{\mathbf{A}}[B_2(\mathbf{P},\mathbf{A})]}{\gamma_{k,n}\sqrt{n}} &(86)\\&=\Pr\Bigg\{\sum_{t\in[\Psi]}\tilde{\imath}(X_t^n;Y_t^n|A_t)\le\log\gamma_{k,n}+nR(\hat{P}_{S^k},D)\\&\qquad\qquad+C\log k\Bigg\}+\frac{|\mathcal{P}_k(\mathcal{S})|\mathbb{E}_{\mathbf{A}}[B_2(\mathbf{P},\mathbf{A})]}{\gamma_{k,n}\sqrt{n}} &(87)\\&\le\Pr\Bigg\{\sum_{t\in[\Psi]}\tilde{\imath}(X_t^n;Y_t^n|A_t)\le|\mathcal{S}|\log(k+1)+nR(\hat{P}_{S^k},D)\\&\qquad+C\log k\Bigg\}+\frac{\mathbb{E}_{\mathbf{A}}[B_2(\mathbf{P},\mathbf{A})]}{\sqrt{n}} &(88)\\&\le\Pr\Bigg\{S^k\in\mathcal{T}^k,\sum_{t\in[\Psi]}\tilde{\imath}(X_t^n;Y_t^n|A_t)\le|\mathcal{S}|\log(k+1)\\&\qquad\quad+kR(\hat{P}_{S^k},D)+C\log k\Bigg\}+\Pr\{S^k\notin\mathcal{T}^k\}\\&\qquad+\frac{\mathbb{E}_{\mathbf{A}}[B_2(\mathbf{P},\mathbf{A})]}{\sqrt{n}} &(89)\\&\le\Pr\Bigg\{\sum_{t\in[\Psi]}L_t^n(P_t,A_t,Z_t^n)\le\sum_{i\in[k]}\jmath(S_i,D)+O(\log k)\Bigg\}\\&\qquad+\frac{2|\mathcal{S}|}{k^2}+\frac{\mathbb{E}_{\mathbf{A}}[B_2(\mathbf{P},\mathbf{A})]}{\sqrt{n}} &(90)\\&=\mathbb{E}_{\mathbf{A}}\Bigg[\Pr\Bigg\{\sum_{i\in[k]}(\jmath(S_i,D)-R(P_X,D))-\sum_{t\in[\Psi]}\tilde{L}_t^n(P_t,A_t,Z^n)\\&\qquad\ge\Bigg(\sum_{t\in[\Psi]}\frac{n}{2}\log(1+P_tA_t^2)-kR(P_S,D)\\&\qquad\quad+O(\log k)\Bigg)\Bigg\}\Bigg|\mathbf{A}\Bigg]+\frac{2|\mathcal{S}|}{k^2}+\frac{\mathbb{E}_{\mathbf{A}}[B_2(\mathbf{P},\mathbf{A})]}{\sqrt{n}} &(91)\end{aligned}$$

$$\leq \mathbb{E}_{\mathbf{A}}\left[\mathrm{Q}\left(\frac{nU_1(\mathbf{P},\mathbf{A})-kR(P_S,D)+O(\log n)}{\sqrt{k\mathrm{V}_{\mathrm{s}}(P_S,D)+nU_2(\mathbf{P},\mathbf{A})}}\right)\right] + \frac{\Omega_1(k,n,P)}{\sqrt{(k+n\Psi)}}+\frac{2|\mathcal{S}|}{k^2}+\frac{\mathbb{E}_{\mathbf{A}}[B_2(\mathbf{P},\mathbf{A})]}{\sqrt{n}} \quad (92)$$

$$= \mathbb{E}_{\mathbf{A}}\left[\mathrm{Q}\left(\frac{nU_1(\mathbf{P},\mathbf{A})-kR(P_S,D)+O(\log n)}{\sqrt{k\mathrm{V}_{\mathrm{s}}(P_S,D)+nU_2(\mathbf{P},\mathbf{A})}}\right)\right] + O\Big(\frac{1}{\sqrt{n}}\Big) \quad (93)$$

$$\leq \mathbb{E}_{\mathbf{A}}\left[\mathrm{Q}\left(\frac{nU_1(\mathbf{P},\mathbf{A})-kR(P_S,D)}{\sqrt{k\mathrm{V}_{\mathrm{s}}(P_S,D)+nU_2(\mathbf{P},\mathbf{A})}}\right)\right]+O\left(\frac{\log n}{\sqrt{n}}\right), \quad (94)$$

where (87) follows from the choice of M_i in (87); (90) follows from Taylor expansion of $R(Q_i,D)$ at P_S (see [35, Eq. (91)]), the result in Lemma 2 and [37, Lemma 22]; (91) is similar to (42); (92) follows by applying the Berry-Esseen theorem as in (43) and using the definition of $\Omega_1(k,n,P)$ in (36) and assumption that $k=O(n)$, (93) follows from the result in (83) and the result that $\Omega_1(k,n,P)=O(1)$ according to the assumptions of Theorem 5, and (94) follows similarly to (53) and (54).

E. Proof of Corollary 6

Combining (52), (73) (or (93)), applying [3, Lemma 17] and using the definition of outage probability in (31), we have that the optimal excess-distortion probability $\mathrm{P}^*_{\mathrm{e},k,n}$ in (7) satisfies that

$$\mathrm{P}^*_{\mathrm{e},k,n} \geq \inf_{\mathbf{P}\in\mathcal{P}_{\mathrm{eq}}(P)} \Pr\{nU_1(\mathbf{P},\mathbf{A})-kR(P_S,D)+\log n\leq 0\} + O\left(\frac{1}{\sqrt{n}}\right), \quad (95)$$

$$= \mathrm{P}_{\mathrm{out}}(k+O(\log n),n,D)+O\left(\frac{1}{\sqrt{n}}\right), \quad (96)$$

$$\mathrm{P}^*_{\mathrm{e},k,n} \leq \inf_{\mathbf{P}\in\mathcal{P}_{\mathrm{eq}}(P)} \Pr\{nU_1(\mathbf{P},\mathbf{A})-kR(P_S,D)\leq O(\log n)\} + O\left(\frac{1}{\sqrt{n}}\right) \quad (97)$$

$$= \mathrm{P}_{\mathrm{out}}(k+O(\log n),n,D)+O\left(\frac{1}{\sqrt{n}}\right). \quad (98)$$

Using the definition of $k^*(n,\varepsilon,P,D)$ in (8) and combining (96), (98), we obtain that when n is sufficient large,

$$k^*(n,\varepsilon,P,D)=n\mathrm{C}_{\mathrm{out}}^{\mathrm{MC}}(P,D,\varepsilon)+O(\log n). \quad (99)$$

VI. Conclusion

In this paper, we considered the lossy multi-connectivity problem, which is a lossy joint source-channel coding problem over parallel AWGN channels with independent quasi-static fading. We derived non-asymptotic achievability and converse bounds on the excess-distortion probability for optimal codes and illustrated our results using numerical examples. Furthermore, for discrete memoryless sources under bounded distortion measures, we derived second-order asymptotics for the optimal excess-distortion probability and the optimal coding rates. Our results imply that the asymptotic notions of outage capacity and outage probability are still valid criteria even in the finite blocklength scenario with negligible loss of performance. In a nutshell, we demonstrate that in the lossy multi-connectivity problem, the optimal performance of URLLC can be well approximated by simple outage analysis. In future, one can nail down the exact coefficients for the remainder term scaling in the order of $\frac{\log n}{n}$ in Corollary 6. It is also interesting to derive the finite blocklength performance when CSI is not provided at the decoder and explore whether there is a loss of performance compared with the present setting. Finally, one can also adapt the coding scheme in [27] to propose a universal JSCC scheme to transmit an arbitrary memoryless source over parallel additive arbitrary noise channels with independent quasi-static fading.

Appendix

A. Proof of Lemma 2

Recall the distributions $P_{Y_t^n|X_t^nA_t}$ in (13) and $Q_{Y_t^n|A_t}$ in (17). Given any $t\in[\Psi]$, channel input $x_t^n\in\mathbb{R}(n,P_t)$ (see (2)) and fading parameter a_t, using the definition of $\tilde{\imath}(X_t^n;Y_t^n|A_t)$ and the fact $Z_t^n=Y_t^n-a_tx_t^n$ due to the channel law (see (1)), under the distribution of $P_{Y_t^n|X_t^nA_t}(\cdot|x_t^n,a_t)$, we have

$$\tilde{\imath}(x_t^n;Y_t^n|a_t)=\frac{n}{2}\log(1+P_ta_t^2)-\frac{\|Z_t^n\|^2}{2}+\frac{\|Z_t^n+a_tx_t^n\|^2}{2(1+P_ta_t^2)}. \quad (100)$$

Hence, the distribution of $\tilde{\imath}(x_t^n;Y_t^n|a_t)$ depends on x_t^n only through $\|Z_t^n+a_tx_t^n\|^2$. Noting that

$$\|Z_t^n+a_tx_t^n\|^2=\|Z_t^n\|^2+a_t^2\|x_t^n\|^2+2\sum_{j\in[n]}a_tx_{t,j}Z_{t,j} \quad (101)$$

$$=\|Z_t^n\|^2+nP_ta_t^2+2\sum_{j\in[n]}a_tx_{t,j}Z_{t,j}, \quad (102)$$

we conclude that the distribution of $\tilde{\imath}(x_t^n;Y_t^n|a_t)$ depends on x_t^n only through $\sum_{j\in[n]}A_tx_{t,j}Z_{t,j}$. Since Z_t^n is i.i.d. according to $\mathcal{N}(0,1)$, we conclude that

$$\sum_{j\in[n]}a_tx_{t,j}Z_{t,j}\sim\mathcal{N}(0,nP_ta_t^2), \quad (103)$$

independent of the choice of $x_t^n\in\mathbb{R}(n,P_t)$. Therefore, we conclude that the distribution of $\tilde{\imath}(x_t^n;Y_t^n|a_t)$ depends on x_t^n only through its power P_t under $P_{Y_t^n|X_t^nA_t}(\cdot|x_t^n,a_t)$. It remains to show that $\tilde{\imath}(x_t^n;Y_t^n|a_t)$ is distributed as $L_t^n(P_t,A_t,Z_t^n)$. We can now choose $x_t^n=x_*^n=(\sqrt{P_t},\ldots,\sqrt{P_t})$ for simplicity. With this choice, under the distribution of $P_{Y_t^n|X_t^nA_t}(\cdot|x_t^n,a_t)$, using the result in (100) and the definition of $L_t^n(P_t,A_t,Z_t^n)$ in (16), we obtain that $\tilde{\imath}(x_t^n;Y_t^n|a_t)$ for any $x_t^n\in\mathbb{R}(n,P_t)$ and any fading parameter a_t has the same distribution as $\tilde{\imath}(x_*^n;Y_t^n|A_t)=L_t^n(P_t,a_t,Z_t^n)$.

B. Proof of Theorem 4

Since the proof of Theorem 4 is similar to [13], we only emphasize the differences here. Fix an arbitrary real number

$\gamma > 0$. Given any source sequence s^k, recalling the definition of $\Phi(s^k, D)$ in (19), define

$$W(s^k, D) := \left\lfloor \frac{n\gamma}{\Phi(s^k, D)} \right\rfloor. \tag{104}$$

Codebook generation: For each $i \in [M]$, generate a random sequence $\hat{S}^k(i)$ i.i.d. according to $P^*_{\hat{S}}$. The collection of these M recovery sequences form the random source codebook. Furthermore, for each $t \in [\Psi]$ and each $i \in [M]$, generate a channel codeword $X_t^n(i)$ according to a uniform distribution over the sphere with radius $\sqrt{nP_t}$, i.e.,

$$P_{X_t^n}(x_t^n) = \frac{1\{\|x_t^n\|^2 - nP_t\}}{A_n(\sqrt{nP_t})}, \tag{105}$$

where $1\{\cdot\}$ is indicator function and $A_n(r)$ is the surface area of an n-dimensional sphere with radius r. The collection of the $n\Psi$ channel codewords form the channel codebook. We assume that the source and channel codebooks are known by both the encoder and the decoder. For simplicity, we will use $\hat{\mathbf{S}}$ to denote the random source codebook and $\hat{\mathbf{s}}$ a particular realization. Similarly we will use $\mathbf{X}$ and $\mathbf{x}$ to denote the channel codebooks. Recall that we use $\mathbf{A}$ to denote the fading parameters. We will also use $\mathbf{a}$ to denote a particular realization of $\mathbf{A}$.

Encoding: Fix an integer M s.t. $W(s^k, D) \le M$ for all s^k. Given a source sequence S^k, for each $t \in [\Psi]$, the encoder f_t outputs the channel codeword $X_t^n(t, j(S^k))$ where $j(S^k) := \min\{m, W(S^k, D)\}$ if $d(S^k, \hat{S}^k(m)) \le D < \min_{i \in [m-1]} d(S^k, \hat{S}^k(i))$ and $j(S^k) = M$ otherwise. The index $j(S^k)$ can be essentially understood as the minimum of $W(S^k, D)$ and the index of the source codeword which first satisfies the distortion requirement D with respect to S^k.

Decoding: Prior to describing the decoding rule, define the random variable U taking values in $[M+1]$ as follows:

$$U := \begin{cases} j(S^k) & \text{if } d(S^k, \hat{S}^k(j(S^k))) \le D \\ M+1 & \text{otherwise.} \end{cases} \tag{106}$$

Given channel outputs $\{Y_t^n\}_{t \in [\Psi]}$ and fading parameters $\{A_t\}_{t \in \Psi}$, the decoder declares $\hat{S}^k(\hat{J})$ as the source estimate if

$$\hat{J} = \operatorname*{arg\,max}_{\bar{j} \in [M]} P_{U|\hat{\mathbf{S}}}(\bar{j}|\hat{\mathbf{S}}) \prod_{t=1}^{\Psi} P_{Y_t^n|X_t^n A_t}(Y_t^n | X_t^n(\bar{j}), A_t). \tag{107}$$

Note that the decoder in (107) is similar to a MAP decoder with the only difference that $P_{U|\mathbf{S}}(\bar{j}|\hat{\mathbf{S}})$ should be replaced by $\prod_{t \in \Psi} P_{X_t^n}(X_t^n(\bar{j}))$.

Conditioning on the fading parameters $\mathbf{a}$, the source codebook $\hat{\mathbf{s}}$ and the channel codebook $\mathbf{x}$, similarly to [13, Eq. (90)], we can upper bound the excess-distortion probability by

$$\begin{aligned}
&\mathrm{P}_{\mathrm{e},k,n}(P, D|\mathbf{A} = \mathbf{a}, \mathbf{X} = \mathbf{x}, \hat{\mathbf{S}} = \hat{\mathbf{s}}) \\
&\le \sum_{u \in [M]} P_{U|\hat{\mathbf{S}}}(u|\hat{\mathbf{s}}) \Pr\Bigg\{ \bigcup_{\bar{j} \in [M]: \bar{j} \neq u} \Bigg\{ \frac{P_{U|\hat{\mathbf{S}}}(\bar{j}|\hat{\mathbf{s}})}{P_{U|\hat{\mathbf{S}}}(u|\hat{\mathbf{s}})} \\
&\qquad \times \frac{\prod_{t=1}^{\Psi} P_{Y_t^n|X_t^n A_t}(Y_t^n | x_t^n(\bar{j}), a_t)}{\prod_{t=1}^{\Psi} P_{Y_t^n|X_t^n A_t}(Y_t^n | x_t^n(u), a_t)} \ge 1 \Bigg\} \Bigg\} \\
&\quad + \Pr_{P_{U|\hat{\mathbf{S}}}} \Big\{ U > W(S^k, D) | \hat{\mathbf{S}} = \hat{\mathbf{s}} \Big\}.
\end{aligned} \tag{108}$$

For each $t \in [\Psi]$, let $P_{Y_t^n|A_t}$ be induced by the channel law $P_{Y_t^n|X_t^n A_t}$, the input distribution $P_{X_t^n}$ and the fading distribution P_A. Using the definition of $Q_{Y_t^n|A_t}$ in (17) and [24, Eq. (4.91)], we obtain that for any $t \in [\Psi]$ and any fading parameter a_t,

$$\max_{y^n} \frac{P_{Y_t^n|A_t}(y_t^n|a_t)}{Q_{Y_t^n|A_t}(y_t^n|a_t)} \le \frac{1 + P_t a_t^2}{\sqrt{1 + 2P_t a_t^2}}. \tag{109}$$

In the remaining part of this subsection, wherever we use $\mathbb{E}$, we mean the expectation with respect to the following distribution

$$\Big(\prod_{i \in [M]} P_S^k(s^k)\Big)\Big(\prod_{t \in [\Psi]} \prod_{j \in [M]} P_{X_t^n}(X_t^n(j))\Big)\Big(\prod_{t \in [\Psi]} P_{Y_t^n|X_t^n A_t}\Big). \tag{110}$$

Using the definition of $W(S^k, D)$ in (104), combing (108), (109) and following the analyses in [13, Eq. (91)-(106)], we can obtain that

$$\begin{aligned}
&\mathrm{P}_{\mathrm{e},k,n}(P, D) \\
&\le \mathbb{E}\Bigg[\min\Bigg\{ 1, W(S^k, D) \prod_{t \in [\Psi]} \frac{P_{Y_t^n|A_t}(Y_t^n|A_t)}{P_{Y_t^n|X_t^n A_t}(Y_t^n|X_t^n, A_t)} \Bigg\} \Bigg] \\
&\qquad + \mathbb{E}\Big[(1 - \Phi(S^k, D))^{W(S^k, D)} \Big] \qquad (111) \\
&\le \mathbb{E}\Bigg[\min\Bigg\{ 1, W(S^k, D) \prod_{t \in [\Psi]} \frac{1 + P_t A_t^2}{\sqrt{1 + 2P_t A_t^2}} \\
&\qquad\qquad \times \frac{Q_{Y_t^n|A_t}(Y_t^n|A_t)}{P_{Y_t^n|X_t^n A_t}(Y_t^n|X_t^n, A_t)} \Bigg\} \Bigg] \\
&\qquad + \mathbb{E}\Big[(1 - \Phi(S^k, D))^{\frac{\gamma}{\Phi(S^k, D)} - 1} \Big] \qquad (112) \\
&\le \mathbb{E}\Bigg[\exp\Bigg(- \Bigg| \sum_{t \in [\Psi]} \Bigg(\tilde{\imath}(X_t^n; Y_t^n | A_t) - \log \frac{1 + P_t A_t^2}{\sqrt{1 + 2P_t A_t^2}} \Bigg) \\
&\qquad\qquad - \log W(S^k, D) \Bigg|^+ \Bigg) \Bigg] \\
&\qquad + \mathbb{E}\Bigg[\exp\Bigg\{ - \Phi(S^k, D) \Bigg(\frac{n\gamma}{\Phi(S^k, D)} - 1 \Bigg) \Bigg\} \Bigg] \qquad (113) \\
&= \mathbb{E}\Bigg[\exp\Bigg(- \Bigg| \sum_{t \in [\Psi]} \Bigg(\tilde{\imath}(X_t^n; Y_t^n | A_t) - \log \frac{1 + P_t A_t^2}{\sqrt{1 + 2P_t A_t^2}} \Bigg) \\
&\qquad\qquad - \log \frac{n\gamma}{\Phi(S^k, D)} \Bigg|^+ \Bigg) \Bigg] \\
&\qquad + \exp(-n\gamma + 1) \qquad (114) \\
&= \mathbb{E}\Bigg[\exp\Bigg(- \Bigg| \sum_{t \in [\Psi]} \Bigg(L_t^n(P_t, A_t, Z_t^n) - \log \frac{1 + P_t A_t^2}{\sqrt{1 + 2P_t A_t^2}} \Bigg) \\
&\qquad\qquad - \log \frac{n\gamma}{\Phi(S^k, D)} \Bigg|^+ \Bigg) \Bigg] \qquad (115) \\
&\qquad + \exp(-n\gamma + 1), \qquad (116)
\end{aligned}$$

where (113) follows by using the inequality $(1-a)^M \le \exp(-Ma)$ for any $a \in [0, 1]$, and (116) follows from $\Phi(s^k, D) \le 1$ for any s^k and the result in Lemma 2 which states that under the distribution $P_{Y_t^n|X_t^n A_t}$, the fading information density $\tilde{\imath}(X_t^n;Y_t^n|A_t)$ has the same distribution as $L_t^n(P_t,A_t,Z_t^n)$.

The proof of Theorem 4 is now completed.

C. Proof of Lemma 8

Recall the definition of $\mathcal{D}$ in (75). Given $s^k\in\mathcal{T}_{Q_i}$ and $\hat{\mathbf{s}}_i$, we obtain that

$$\begin{aligned}
&\Pr\{(\hat{I},\hat{J})\neq(i,j(S^k,\hat{\mathbf{S}}_i))|S^k=s^k,\ \hat{\mathbf{S}}_i=\hat{\mathbf{s}}_i\}\\
&=\Pr\left\{\sum_{t\in[\Psi]}\tilde{\imath}(X_t^n(i,j(s^k,\hat{\mathbf{s}}_i));Y_t^n|A_t)\le\log\gamma_{k,n}+\log M_i\right\}\\
&\quad+\Pr\Bigg\{\exists\ (\bar{i},\bar{j})\in\mathcal{D}:(\bar{i},\bar{j})\neq(i,j(s^k,\hat{\mathbf{s}}_i)\ \ \text{s.t.}\\
&\qquad\sum_{t\in[\Psi]}\tilde{\imath}(X_t^n(\bar{i},\bar{j});Y_t^n|A_t)>\gamma_{k,n}+\log M_{\bar{i}}\Bigg\}, \quad (117)\\
&\le\Pr\left\{\sum_{t\in[\Psi]}\tilde{\imath}(X_t^n(i,j(s^k,\hat{\mathbf{s}}_i));Y_t^n|A_t)\le\log\gamma_{k,n}+\log M_i\right\}\\
&\quad+\sum_{(\bar{i},\bar{j})\in\mathcal{D}:(\bar{i},\bar{j})\neq(i,j(s^k,\hat{\mathbf{s}}_i)}\Pr\Bigg\{\sum_{t\in[\Psi]}\tilde{\imath}(X_t^n(\bar{i},\bar{j});Y_t^n|A_t)\\
&\qquad>\log\gamma_{k,n}+\log M_{\bar{i}}\Bigg\} \quad (118)\\
&\le\Pr\left\{\sum_{t\in[\Psi]}\tilde{\imath}(X_t^n;Y_t^n|A_t)\le\log\gamma_{k,n}+\log M_i\right\}\\
&\quad+\sum_{\bar{i}=1}^{|\mathcal{P}_k(\mathcal{S})|}M_{\bar{i}}\Pr\Bigg\{\sum_{t\in[\Psi]}\tilde{\imath}(X_t^n;\bar{Y}_t^n|A_t)\\
&\qquad>\log\gamma_{k,n}+\log M_{\bar{i}}\Bigg\} \quad (119)\\
&\le\Pr\left\{\sum_{t\in[\Psi]}\tilde{\imath}(X_t^n;Y_t^n|A_t)\le\log\gamma_{k,n}+\log M_i\right\}\\
&\quad+\frac{|\mathcal{P}_k(\mathcal{S})|\mathbb{E}_{\mathbf{A}}[B_2(\mathbf{P},\mathbf{A})]}{\gamma_{k,n}\sqrt{n}}, \quad (120)
\end{aligned}$$

where in (119), for each $t\in[\Psi]$, $(X_t^n,A_t,Y_t^n,\bar{Y}_t^n)\sim P_{X_t^n}P_{A_t}P_{Y_t^n|X_t^nA_t}P_{Y_t^n|A_t}$, (119) follows since the channel output $Y_t^n=X_t^n(i,j(s^k,\hat{\mathbf{s}}_i))+Z_t^n$ is independent of all codewords $X_t^n(\bar{i},\bar{j})$ where $(\bar{i},\bar{j})\neq(i,j(s^k,\hat{\mathbf{s}}_i))$, and (120) follows from using the result in [24, Eq. (4.93)] and the definition of $B_2(\mathbf{P},\mathbf{A})$ in (82).

Acknowledge

The authors acknowledge Prof. Vincent Y. F. Tan at the National University of Singapore for helpful discussions and comments.

References

[1] L. Zhou, A. Wolf, and M. Motani, “On the finite blocklength performance of lossy multi-connectivity,” in *IEEE Globecom*, 2018.
[2] A. Wolf, P. Schulz, M. Dörpinghaus, J. C. S. Santos Filho, and G. Fettweis, “How reliable and capable is multi-connectivity?” *To appear in IEEE Trans. Commun.*, 2019.
[3] W. Yang, G. Durisi, T. Koch, and Y. Polyanskiy, “Quasi-static multiple-antenna fading channels at finite blocklength,” *IEEE Trans. Inf. Theory*, vol. 60, no. 7, pp. 4232–4265, 2014.
[4] L. Zheng and D. N. C. Tse, “Diversity and multiplexing: A fundamental tradeoff in multiple-antenna channels,” *IEEE Trans. Inf. Theory*, vol. 49, no. 5, pp. 1073–1096, 2003.
[5] Y. Yona and M. Feder, “On the diversity-multiplexing tradeoff of unconstrained multiple-access channels,” *IEEE Trans. Inf. Theory*, vol. 61, no. 9, pp. 4630–4662, 2015.
[6] A. Khisti and S. C. Draper, “The streaming-DMT of fading channels,” *IEEE Trans. Inf. Theory*, vol. 60, no. 11, pp. 7058–7072, 2014.
[7] C. E. Shannon, “A mathematical theory of communication,” *Bell Syst. Tech. J.*, vol. 27, no. 1, pp. 379–423, 1948.
[8] D. Tse and P. Viswanath, *Fundamentals of wireless communication*. Cambridge university press, 2005.
[9] M. Hayashi, “Information spectrum approach to second-order coding rate in channel coding,” *IEEE Trans. Inf. Theory*, vol. 55, no. 11, pp. 4947–4966, 2009.
[10] Y. Polyanskiy, H. V. Poor, and S. Verdú, “Channel coding rate in the finite blocklength regime,” *IEEE Trans. Inf. Theory*, vol. 56, no. 5, pp. 2307–2359, 2010.
[11] V. Strassen, “Asymptotische abschätzungen in shannons informationstheorie,” in *Trans. Third Prague Conf. Information Theory*, 1962, pp. 689–723.
[12] V. Kostina and S. Verdú, “Fixed-length lossy compression in the finite blocklength regime,” *IEEE Trans. Inf. Theory*, vol. 58, no. 6, pp. 3309–3338, 2012.
[13] V. Kostina and S. Verdú, “Lossy joint source-channel coding in the finite blocklength regime,” *IEEE Trans. Inf. Theory*, vol. 59, no. 5, pp. 2545–2575, 2013.
[14] 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, no. 5, pp. 2430–2438, May 2015.
[15] Y. Polyanskiy and S. Verdú, “Scalar coherent fading channel: Dispersion analysis,” in *IEEE ISIT*, 2011, pp. 2959–2963.
[16] E. MolavianJazi and J. N. Laneman, “On the second-order coding rate of non-ergodic fading channels,” in *Proc. 51st Annu. Allerton Conf.*, 2013, pp. 583–587.
[17] W. Yang, “Fading channels: Capacity and channel coding rate in the finite-blocklength regime,” Ph.D. dissertation, Department of Signals and Systems, Chalmers University of Technology, 2015.
[18] C. E. Shannon, “Coding theorems for a discrete source with a fidelity criterion,” *IRE Nat. Conv. Rec*, vol. 4, no. 142-163, p. 1, 1959.
[19] K. Marton, “Error exponent for source coding with a fidelity criterion,” *IEEE Trans. Inf. Theory*, vol. 20, no. 2, pp. 197–199, 1974.
[20] S. Ihara and M. Kubo, “Error exponent for coding of memoryless Gaussian sources with a fidelity criterion,” *IEICE Trans. Fundamentals*, vol. 83, no. 10, pp. 1891–1897, 2000.
[21] A. Ingber and Y. Kochman, “The dispersion of lossy source coding,” in *IEEE DCC*, 2011, pp. 53–62.
[22] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*. Cambridge University Press, 2011.
[23] J. N. Laneman, E. Martinian, G. W. Wornell, and J. G. Apostolopoulos, “Source-channel diversity for parallel channels,” *IEEE Trans. Inf. Theory*, vol. 51, no. 10, pp. 3518–3539, 2005.
[24] E. MolavianJazi, “A unified approach to Gaussian channels with finite blocklength,” Ph.D. dissertation, Dept. Electrical Engineering, University of Notre Dame, 2014.
[25] I. Csiszár, “Joint source-channel error exponent,” *Prob. Contr. & Info. Theory*, vol. 9, no. 5, pp. 315–328, 1980.
[26] D. Wang, A. Ingber, and Y. Kochman, “The dispersion of joint source-channel coding,” in *Proc. 49th Annu. Allerton Conf.* IEEE, 2011, pp. 180–187.
[27] L. Zhou, V. Y. F. Tan, and M. Motani, “The dispersion of mismatched joint-source channel coding for arbitrary sources and additive channels,” *To appear in IEEE Trans. Inf. Theory*, 2018.
[28] Y. Y. Shkel, V. Y. F. Tan, and S. C. Draper, “Unequal message protection: Asymptotic and non-asymptotic tradeoffs,” *IEEE Trans. Inf. Theory*, vol. 61, no. 10, pp. 5396–5416, 2015.
[29] A. Lapidoth, “Nearest neighbor decoding for additive non-Gaussian noise channels,” *IEEE Trans. Inf. Theory*, vol. 42, no. 5, pp. 1520–1529, 1996.
[30] T. Cormen, C. Leiserson, R. Rivest, and C. Stein, *Introduction to Algorithms*, 2nd ed. McGraw-Hill Science/Engineering/Math, 2003.
[31] V. Kostina and S. Verdú, “A new converse in rate-distortion theory,” in *CISS*, 2012, pp. 1–6.
[32] A. C. Berry, “The accuracy of the Gaussian approximation to the sum of independent variates,” *Transactions of the American mathematical society*, vol. 49, no. 1, pp. 122–136, 1941.
[33] C.-G. Esseen, *On the Liapounoff limit of error in the theory of probability*. Almqvist & Wiksell, 1942.

This document is a preprint of: L. Zhou, A. Wolf and M. Motani, “On Lossy Multi-Connectivity: Finite Blocklength Performance and Second-Order Asymptotics,” in IEEE Journal on Selected Areas in Communications (JSAC), Feb 2019. DOI:10.1109/JSAC.2019.2898755

[34] V. Y. F. Tan, “Asymptotic estimates in information theory with non-vanishing error probabilities,” *Foundations and Trends® in Communications and Information Theory*, vol. 11, no. 1–2, pp. 1–184, 2014.

[35] L. Zhou, V. Y. F. Tan, and M. Motani, “Second-order and moderate deviation asymptotics for successive refinement,” *IEEE Trans. Inf. Theory*, vol. 63, no. 5, pp. 2896–2921, 2017.

[36] A. No, A. Ingber, and T. Weissman, “Strong successive refinability and rate-distortion-complexity tradeoff,” *IEEE Trans. Inf. Theory*, vol. 62, no. 6, pp. 3618–3635, 2016.

[37] M. Tomamichel and V. Y. F. Tan, “Second-order coding rates for channels with state,” *IEEE Trans. Inf. Theory*, vol. 60, no. 8, pp. 4427–4448, 2014.

Lin Zhou (S’15-M’18) is currently a Research Fellow in the Department of Electrical Engineering and Computer Science at the University of Michigan, Ann Arbor. From Mar. 2018 to Dec. 2018, he was first a Research Engineer and then a Research Fellow in the Department of Electrical and Computer Engineering (ECE) at the National University of Singapore (NUS). He received the Ph.D. degree in ECE from NUS in 2018 and the B.E. degree in Information Engineering from Shanghai Jiao Tong University (SJTU) in 2014. His research interests include information theory and its applications in wireless communication, signal processing and machine learning.

Albrecht Wolf received the Dipl.-Ing. degree in electrical engineering from Technical University Dresden, Dresden, Germany, in 2014, where he is currently pursuing the Ph.D. degree with the Vodafone Chair Mobile Communication Systems. During his studies, he focused on mobile communications systems and communication theory. He completed his internship with Qualcomm Inc., Frankfurt, where he performed field testing and log analyzing of modem chip sets. His research interests include network information theory and cooperative wireless communications. He was a recipient of a Best Student Paper Award at the 2018 IEEE Wireless Communications and Networking Conference (WCNC), Barcelona, Spain.

Mehul Motani (S’92-M’01-SM’15-F’18) received the B.E. degree from Cooper Union, New York, NY, the M.S. degree from Syracuse University, Syracuse, NY, and the Ph.D. degree from Cornell University, Ithaca, NY, all in Electrical and Computer Engineering.

Dr. Motani is currently an Associate Professor in the Electrical and Computer Engineering Department at the National University of Singapore (NUS) and a Visiting Research Collaborator at Princeton University. Previously, he was a Visiting Fellow at Princeton University. He was also a Research Scientist at the Institute for Infocomm Research in Singapore, for three years, and a Systems Engineer at Lockheed Martin in Syracuse, NY for over four years. His research interests include information theory, machine learning, wireless and sensor networks, and energy harvesting communications.

Dr. Motani was the recipient of the Intel Foundation Fellowship for his Ph.D. research, the NUS Annual Teaching Excellence Award, the NUS Faculty of Engineering Innovative Teaching Award, and the NUS Faculty of Engineering Teaching Honours List Award. He actively participates in the Institute of Electrical and Electronics Engineers (IEEE) and the Association for Computing Machinery (ACM). He is a Fellow of the IEEE and has served as the Secretary of the IEEE Information Theory Society Board of Governors. He has served as an Associate Editor for both the IEEE Transactions on Information Theory and the IEEE Transactions on Communications. He has also served on the Organizing and Technical Program Committees of numerous IEEE and ACM conferences.