



HAL
open science

Broadcast Channels with Heterogeneous Arrival and Decoding Deadlines: Second-Order Achievability

Homa Nikbakht, Malcolm Egan, Jean-Marie Gorce, H. Vincent Poor

► **To cite this version:**

Homa Nikbakht, Malcolm Egan, Jean-Marie Gorce, H. Vincent Poor. Broadcast Channels with Heterogeneous Arrival and Decoding Deadlines: Second-Order Achievability. 2023. hal-04360147

HAL Id: hal-04360147

<https://inria.hal.science/hal-04360147>

Preprint submitted on 21 Dec 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

Broadcast Channels with Heterogeneous Arrival and Decoding Deadlines: Second-Order Achievability

Homa Nikbakht, *Member, IEEE*, Malcolm Egan, Jean-Marie Gorce, *Senior Member, IEEE*,
and H. Vincent Poor, *Life Fellow, IEEE*

Abstract

A standard assumption in the design of ultra-reliable low-latency communication systems is that the duration between message arrivals is larger than the number of channel uses before the decoding deadline. Nevertheless, this assumption fails when messages arrive rapidly and reliability constraints require that the number of channel uses exceed the time between arrivals. In this paper, we consider a broadcast setting in which a transmitter wishes to send two different messages to two receivers over Gaussian channels. Messages have different arrival times and decoding deadlines such that their transmission windows overlap. For this setting, we propose a coding scheme that exploits Marton's coding strategy. We derive rigorous bounds on the achievable rate regions. Those bounds can be easily employed in point-to-point settings with one or multiple parallel channels. In the point-to-point setting with one or multiple parallel channels, the proposed achievability scheme outperforms the Normal Approximation, especially when the number of channel uses is smaller than 200. In the broadcast setting, our scheme agrees with Marton's strategy for sufficiently large numbers of channel uses and shows significant performance improvements over standard approaches based on time sharing for transmission of short packets.

I. INTRODUCTION

Mobile wireless networks in 5G and in 6G proposals are increasingly intended for use in latency-critical and high-reliability systems, notably in industrial control applications, autonomous vehicles and remote surgery [1]–[6]. In such ultra-reliable low-latency communications (URLLC), packets are typically short. As a consequence, data transmission cannot be made reliable by increasing channel code blocklength arbitrarily.

A key challenge is, therefore, to design coding schemes that support high reliability requirements under finite blocklength. In recent years, a number of channel coding schemes have been proposed to address such requirements including short LDPC and polar codes [7], [8]. At the same time, new characterizations of fundamental tradeoffs among the size of the message set, the probability of error, and the length of the code have been obtained via achievability and converse bounds, thereby building on the work of [9]–[11].

A standard assumption in the design of coding schemes, even for URLLC, is that consecutive messages have distinct arrival times and decoding deadlines. As such, there is no choice but to encode the messages independently. However, this assumption is violated when a message arrives before the decoding deadline of a prior message. For example, a sensor in an unstable control system may send rapid measurements in order to stabilize the system [12]. In order to ensure reliability of the sensor observations, the channel uses allocated to each observation of the speed may partially overlap.

It is therefore desirable to consider joint encoding of multiple sensor observations, albeit with heterogeneous decoding deadlines. That is, if the channel uses for two separate observations overlap, it is not possible to wait until the entire transmission for both sensor observations is received before decoding. In this situation, code design must account for two issues:

- (i) messages with close arrival times; and
- (ii) messages with heterogeneous decoding deadlines.

In addition to heterogeneous arrival times and decoding deadlines, a transmitter may seek to send messages to distinct receivers. While this setting has clear relevance for URLLC applications, there are currently no known designs for appropriate coding schemes.

A. Related Work

1) *Finite Blocklength Regime*: To capture both reliability and latency, investigation of coding bounds in the finite blocklength regime is a requirement that dates back to the work of Shannon, Gallager, and Berlekamp [13]. Most of the existing literature focuses on identifying the limits of communication between a single transmitter and a single receiver for a coding block of

H. Nikbakht and H. V. Poor are with the Department of Electrical and Computer Engineering, Princeton University, NJ, USA (email: homa.poor@princeton.edu). M. Egan and J.-M. Gorce are with the INRIA and University of Lyon, CITI, INSA Lyon, 69100 Villeurbanne, France (e-mail: jean-marie.gorce@inria.fr, malcolm.egan@inria.fr). The work of H. Nikbakht and H. V. Poor has been supported by the U.S. National Science Foundation (NSF) under Grants CNS-2128448 and ECCS-2335876. The work of J.-M. Gorce and M. Egan has been partly supported by the French National Agency for Research within the PEPR-5 program.

size n . Among the proposed bounds, the widely used bound is the *Normal Approximation* that for a given n approximates the point-to-point transmission rate R by [9], [14]:

$$R \approx C - \sqrt{\frac{V}{n}} \log(e) Q^{-1}(\epsilon) + \frac{\log n}{2n}, \quad (1)$$

where C is the channel capacity, V is the channel dispersion coefficient, ϵ is the average error probability and $Q^{-1}(\cdot)$ is the inverse of the Gaussian cumulative distribution function. However this approximation has been proved to be a valid $O(n^{-1})$ asymptotic approximation for converse and achievability bounds [15], but an $O(n^{-1})$ bound is not necessarily reliable for small values of n corresponding to URLLC applications. Therefore, exact bounds are of considerable interest.

2) *Heterogeneous Decoding Deadlines*: The problem of designing a code to handle heterogeneous decoding deadlines was first considered in the context of static broadcasting [16], where a single message is decoded at multiple receivers under different relative decoding delay constraints. The work in [16] was recently generalized to multi-source and multi-terminal networks by Langberg and Effros in [17]. In particular, the notion of a time-rate region was introduced, which accounted for different decoding delay constraints for each message at each receiver.

The work in both [16] and [17] focused on the asymptotic regime. In the finite blocklength regime, a coding scheme for the Gaussian broadcast channel with heterogeneous blocklength constraints was introduced in [18], which decodes the messages at time-instances that depend on the realizations of the random channel fading. By employing an early decoding scheme, the authors showed that significant improvements over standard successive interference cancellation are possible. In [19] achievable rates and latency of the early-decoding scheme in [18] are improved by introducing *concatenated shell codes*.

3) *Heterogeneous Arrival Times*: The work in [16]–[19] focuses on the case where both messages are available at the time of encoding. In our previous works [20], [21], we introduced a coding scheme for the Gaussian point-to-point channel that encodes the first message before the second message arrives. The scheme proposed in [20] exploited power sharing for symbols between the arrival time of the second message and the decoding deadline of the first message. Under a Gaussian interference assumption, bounds on the error probabilities for each message were established based on the message set size and finite decoding deadline constraints. In [21], a coding scheme was proposed that exploits dirty-paper coding (DPC) [22]–[24]. We further derived rigorous bounds on the achievable error probabilities of the messages.

4) *Broadcast Channels in the Finite Blocklength Regime*: Finite blocklength analysis of broadcast channels (BCs) was studied in [25]–[28]. The second-order Gaussian BC setting was investigated in [26] where the authors studied the concatenate-and-code protocol [29] in which the transmitter concatenates the users' message bits into a single data packet and each user decodes the entire packet to extract its own bits. This scheme was shown to outperform superposition coding and time division multiplexing (TDM) schemes. The work in [28] is an extension of [26] to K -user BCs. The work in [27] considered a two-user static BC and showed that under per-user reliability constraint, superposition coding combined with a rate splitting technique in which the message intended for the user with the lowest signal-to-noise ratio (SNR) (the cloud center message) is allocated to either users gives the largest second-order rate region.

B. Main Contributions

While coding schemes for heterogeneous arrival and decoding deadlines have been developed for point-to-point channels, adapting these codes to broadcast channels remains an open problem. In this paper, we address this question by developing and analyzing a coding scheme tailored to broadcast channels with heterogeneous arrival times and decoding deadlines.

The main contributions of this work are:

- We introduce a coding scheme for two-user Gaussian BCs with heterogeneous arrival times and decoding deadlines in [20], which exploits Marton's coding strategy [30]. Accounting for finite decoding deadline constraints (corresponding to fixed blocklengths), we first derive rigorous bounds on the achievable transmission rate for each of the messages. This is achieved by combining techniques to analyze the Gel'fand-Pinsker channel in the finite blocklength regime [23] and multiple parallel channels [15].
- With the developed bounds, we obtain further rigorous bounds for point-to-point settings with one or multiple parallel channels. In the point-to-point setting with one channel, we show that our achievability scheme outperforms the Normal Approximation in (1) proposed by Polyanskiy, Poor and Verdú in [9] especially when the number of channel uses (the size of coding block) is smaller than 200. In the point-to-point setting with multiple parallel channels our achievability bound outperforms the Normal Approximation bound proposed by Erseghe in [15].
- In the broadcast setup, we provide a second-order analysis for the achievable rate regions. We show that our scheme agrees with Marton's bound in [30] for a sufficiently large number of channel uses.
- Finally, we show that our scheme outperforms the time-sharing scheme that transmits each message independently but over fewer number of channel uses.

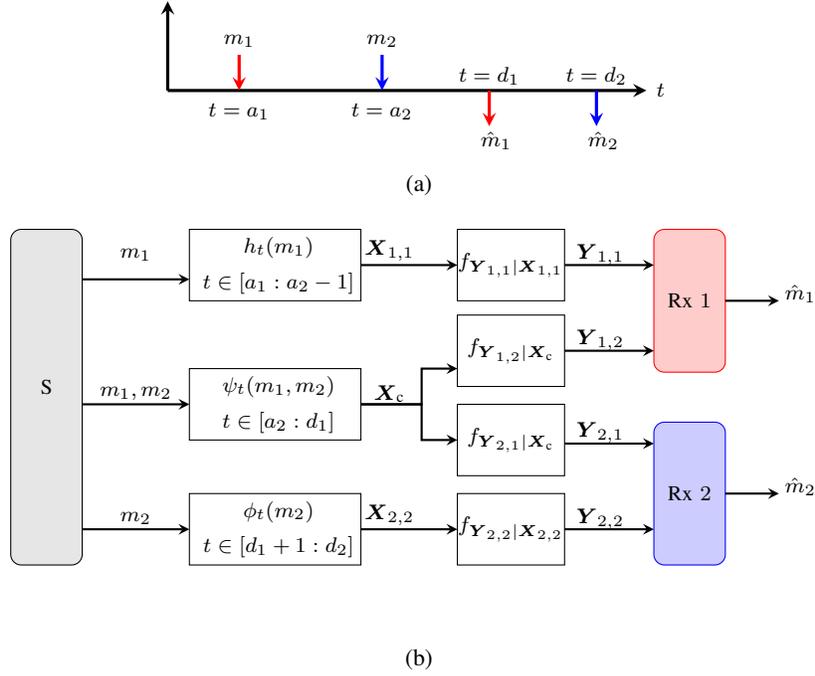


Fig. 1: System model. (a) Arrival times and decoding deadlines of m_1 and m_2 , (b) Problem setup with one transmitter and two receivers.

C. Organization

The rest of this paper is organized as follows. We end this section with some remarks on notation. Sections II and III describe the problem setup and the proposed coding scheme. Section IV and V present our main results and discussions on the related works. Section VI concludes the paper. Some technical proofs are referred to in appendices.

D. Notation

The set of all integers is denoted by \mathbb{Z} , the set of positive integers by \mathbb{Z}^+ and the set of real numbers by \mathbb{R} . For other sets we use calligraphic letters, e.g., \mathcal{X} . Random variables are denoted by uppercase letters, e.g., X , and their realizations by lowercase letters, e.g., x . For vectors we use boldface notation, i.e., upper case boldface letters such as \mathbf{X} for random vectors and lower case boldface letters such as \mathbf{x} for deterministic vectors. Matrices are depicted with sans serif font, e.g., \mathbf{H} . We also write X^n for the tuple of random variables (X_1, \dots, X_n) and \mathbf{X}^n for the tuple of random vectors $(\mathbf{X}_1, \dots, \mathbf{X}_n)$.

II. PROBLEM SETUP

Consider a transmitter S that seeks to communicate with two receivers Rx 1 and Rx 2. It wishes to transmit message m_1 to receiver Rx 1 and message m_2 to receiver Rx 2. At time $t = a_1$, transmission commences for the first message m_1 . At time $t = a_2$, transmission commences for the second message m_2 . The two messages m_1, m_2 are assumed to be drawn independently and uniformly on $\{1, \dots, M_1\}$ and $\{1, \dots, M_2\}$, respectively.

Each message is subject to different decoding delay constraints.¹ In particular, at time d_1 , Rx 1 attempts to reconstruct the message m_1 . Similarly, at time $d_2 > d_1$, Rx 2 attempts to reconstruct the message m_2 . See Fig. 1a.

Under the assumption that $a_1 < a_2$ and $a_2 < d_1 < d_2$, the encoder outputs symbols at time $t \in \{a_1, \dots, d_2\}$ as

$$X_t = \begin{cases} h_t(m_1), & t \in \{a_1, \dots, a_2 - 1\} \\ \psi_t(m_1, m_2), & t \in \{a_2, \dots, d_1\} \\ \phi_t(m_2), & t \in \{d_1 + 1, \dots, d_2\}, \end{cases} \quad (2)$$

where h, ψ, ϕ are the encoding functions corresponding to the channel uses where only message m_1 has arrived but not m_2 , where both m_1, m_2 are present, and m_1 has been decoded at Rx 1. We highlight that m_2 is not known before time $t = a_2$; i.e., encoding is causal. Define

$$n := d_2 - a_1 + 1, \quad n_{1,1} := a_2 - a_1, \quad n_{1,2} := d_1 - a_2 + 1 \quad \text{and} \quad n_{2,2} := d_2 - d_1. \quad (3)$$

¹In this work we do not consider the encoding/decoding time delays, but rather the delay due to transmission, i.e., in terms of number of channel uses.

Choose $\beta_{1,1}$, β_c and $\beta_{2,2}$ in $[0, 1]$ such that

$$n_{1,1}\beta_{1,1} + n_{1,2}\beta_c + n_{2,2}\beta_{2,2} = n. \quad (4)$$

We assume that the encoding functions satisfy an average block power constraint; namely,

$$\frac{1}{n} \sum_{i=a_1}^{d_2} X_i^2 \leq P \quad (5)$$

and consequently

$$\frac{1}{n_{1,1}} \sum_{i=a_1}^{a_2-1} X_i^2 \leq \beta_{1,1}P, \quad (6)$$

$$\frac{1}{n_{1,2}} \sum_{i=a_2}^{d_1} X_i^2 \leq \beta_c P, \quad (7)$$

$$\frac{1}{n_{2,2}} \sum_{i=d_1+1}^{d_2} X_i^2 \leq \beta_{2,2}P. \quad (8)$$

Denote the channel inputs by

$$\begin{aligned} \mathbf{X}_{1,1} &= \{X_{a_1}, \dots, X_{a_2-1}\}, \\ \mathbf{X}_c &= \{X_{a_2}, \dots, X_{d_1}\}, \\ \mathbf{X}_{2,2} &= \{X_{d_1+1}, \dots, X_{d_2}\}, \end{aligned} \quad (9)$$

and the corresponding channel outputs at Rx 1 by $\mathbf{Y}_{1,1}$ and $\mathbf{Y}_{1,2}$ and the channel outputs at Rx 2 by $\mathbf{Y}_{2,1}$ and $\mathbf{Y}_{2,2}$. The conditional distributions governing the four channels are then denoted by $f_{\mathbf{Y}_{1,1}|\mathbf{X}_{1,1}}$, $f_{\mathbf{Y}_{1,2}|\mathbf{X}_c}$, $f_{\mathbf{Y}_{2,1}|\mathbf{X}_c}$ and $f_{\mathbf{Y}_{2,2}|\mathbf{X}_{2,2}}$. We assume that each channel is additive, memoryless, stationary, and Gaussian; that is,

$$\mathbf{Y}_{1,1} = \mathbf{X}_{1,1} + \mathbf{Z}_{1,1}, \quad (10)$$

$$\mathbf{Y}_{1,2} = \mathbf{X}_c + \mathbf{Z}_{1,2}, \quad (11)$$

$$\mathbf{Y}_{2,1} = \mathbf{X}_c + \mathbf{Z}_{2,1}, \quad (12)$$

$$\mathbf{Y}_{2,2} = \mathbf{X}_{2,2} + \mathbf{Z}_{2,2}, \quad (13)$$

where $\mathbf{Z}_{i,j} \sim \mathcal{N}(\mathbf{0}, \sigma_{i,j}^2 \mathbf{I})$ for $i = 1, 2$ and $j = 1, 2$.

This setup is illustrated in Fig. 1b.

Receiver Rx 1 attempts to reconstruct message m_1 based on the channel outputs $\mathbf{Y}_{1,1}$ and $\mathbf{Y}_{1,2}$ via the decoding function g_1 ; i.e.,

$$\hat{m}_1 = g_1(\mathbf{Y}_{1,1}, \mathbf{Y}_{1,2}). \quad (14)$$

Receiver Rx 2 attempts to reconstruct message m_2 based on the channel outputs $\mathbf{Y}_{2,1}$ and $\mathbf{Y}_{2,2}$ via the decoding function g_2 ; i.e.,

$$\hat{m}_2 = g_2(\mathbf{Y}_{2,1}, \mathbf{Y}_{2,2}). \quad (15)$$

Observe that both receivers are causal.

The average probability of error for each of the messages is then

$$\epsilon_1 = \mathbb{P}[\hat{m}_1 \neq m_1], \quad \epsilon_2 = \mathbb{P}[\hat{m}_2 \neq m_2]. \quad (16)$$

The focus of the remainder of this paper is on characterizing the tradeoff among the size of the message sets M_1 and M_2 , the error probabilities ϵ_1, ϵ_2 , and the decoding deadlines d_1, d_2 . Formally, we study the achievable region defined as follows.

Definition 1: Given the power constraint P , a tuple $(a_1, a_2, d_1, d_2, M_1, M_2, \epsilon_1, \epsilon_2)$ is achievable if the messages m_1 and m_2 of cardinality M_1 and M_2 , respectively, arriving at the a_1 -th and a_2 -th channel uses can be decoded by the d_1 -th and d_2 -th channel uses with an average probability of error not exceeding ϵ_1 and ϵ_2 , respectively.

III. RANDOM CODING SCHEME

In this section, we introduce our random coding scheme. Notice that instead of employing Gaussian codebooks, our analysis relies on the use of power-shell codebooks. A power-shell codebook of length n consists of codewords that are uniformly distributed on the centered $(n-1)$ -dimensional sphere with radius \sqrt{nP} where P is the average input power constraint.

A. Encoding

The encoding process consists of three phases.

1) *Transmitting only m_1* : In the first channel, consisting of $n_{1,1}$ channel uses, only m_1 is known to the encoder. The channel input $\mathbf{X}_{1,1}$ corresponding to message m_1 is a codeword $\mathbf{X}_{1,1}(m_1) \in \mathbb{R}^{n_{1,1}}$, which is independently distributed on the sphere $\mathbb{S}^{n_{1,1}-1}$ with power $n_{1,1}\beta_{1,1}\mathsf{P}$. That is, the probability density function of $\mathbf{X}_{1,1}$ is given by

$$f_{\mathbf{X}_{1,1}}(\mathbf{x}_{1,1}) = \frac{\delta(\|\mathbf{x}_{1,1}\|^2 - n_{1,1}\beta_{1,1}\mathsf{P})}{S_{n_{1,1}}(\sqrt{n_{1,1}\beta_{1,1}\mathsf{P}})}, \quad (17)$$

where $\delta(\cdot)$ is the Dirac delta function, and $S_n(r)$ is the surface area of a sphere of radius r in n -dimensional space.

2) *Transmitting both m_1 and m_2* : Over the next $n_{1,2}$ channel uses, the encoder exploits Marton's coding strategy [30] to jointly encode m_1 and m_2 . To this end, we choose $\beta_{2,1} \in [0, 1]$ and $\beta_{1,2} \in [0, 1]$ such that for a given $\rho \in [0, 1]$,

$$\beta_{1,2} + \beta_{2,1} + \rho\sqrt{\beta_{1,2}\beta_{2,1}} = \beta_c \quad (18)$$

The parameter ρ is defined shortly.

The following two codebooks are then generated.

Codebook Generation:

- Denote by L_1 the random coding parameter illustrating the number of auxiliary codewords for each message $m_1 \in \{1, \dots, M_1\}$. A random codebook $C_{\mathbf{X}_{1,2}}$ containing $M_1 L_1$ auxiliary codewords $\{\mathbf{X}_{1,2}(m_1, \ell_1)\}$ with $\ell_1 \in \{1, \dots, L_1\}$ and $m_1 \in \{1, \dots, M_1\}$ is generated where each codeword is independently distributed on the sphere $\mathbb{S}^{n_{1,2}-1}$ with power $n_{1,2}\beta_{1,2}\mathsf{P}$.
- Denote by L_2 the random coding parameter illustrating the number of auxiliary codewords for each message $m_2 \in \{1, \dots, M_2\}$. A random codebook $C_{\mathbf{X}_{2,1}}$ containing $M_2 L_2$ auxiliary codewords $\{\mathbf{X}_{2,1}(m_2, \ell_2)\}$ with $\ell_2 \in \{1, \dots, L_2\}$ and $m_2 \in \{1, \dots, M_2\}$ is generated where each codeword is independently distributed on the sphere $\mathbb{S}^{n_{1,2}-1}$ with power $n_{1,2}\beta_{2,1}\mathsf{P}$.

The codebooks are revealed to all terminals.

The transmitter chooses the pair (ℓ_1, ℓ_2) such that $\mathbf{X}_{1,2}(m_1, \ell_1) \in C_{\mathbf{X}_{1,2}}$ and $\mathbf{X}_{2,1}(m_2, \ell_2) \in C_{\mathbf{X}_{2,1}}$ satisfy

$$\langle \mathbf{X}_{1,2}(m_1, \ell_1), \mathbf{X}_{2,1}(m_2, \ell_2) \rangle \in \mathcal{D} \quad (19)$$

where

$$\mathcal{D} \triangleq \left[n_{1,2}\sqrt{\beta_{1,2}\beta_{2,1}}\mathsf{P}\rho : n_{1,2}\sqrt{\beta_{1,2}\beta_{2,1}}\mathsf{P} \right] \quad (20)$$

and ρ is the correlation parameter. If more than one such a pair exists, then one pair is selected arbitrarily.

The channel input over the $n_{1,2}$ symbols allocated for the joint transmission of m_1 and m_2 is given by

$$\mathbf{X}_c = \alpha(\mathbf{X}_{1,2}(m_1, \ell_1) + \mathbf{X}_{2,1}(m_2, \ell_2)), \quad (21)$$

where

$$\alpha := \sqrt{\frac{\beta_c}{\beta_c^*}}. \quad (22)$$

with

$$\beta_c^* := \beta_{1,2} + \beta_{2,1} + \rho^*\sqrt{\beta_{1,2}\beta_{2,1}}, \quad (23)$$

and $\rho^* \in [\rho, 1]$ is the correlation coefficient between the chosen codewords $\mathbf{X}_{1,2}(m_1, \ell_1)$ and $\mathbf{X}_{2,1}(m_2, \ell_2)$. Notice that α is a power normalization coefficient that ensures that the transmit signal satisfies the power constraint in (5).

We have the encoding error event:

$$\mathcal{E}_{1,2} \triangleq \{\text{no } (\ell_1, \ell_2) \text{ exists such that (19) is satisfied}\}. \quad (24)$$

In our analysis, we set $\epsilon_{1,2} \in [0, 1]$ as the threshold on $\mathbb{P}[\mathcal{E}_{1,2}]$, i.e., given $n_{1,2}, \beta_{1,2}, \beta_{2,1}, \mathsf{P}$ and ρ we choose L_1 and L_2 such that $\mathbb{P}[\mathcal{E}_{1,2}] \leq \epsilon_{1,2}$.

3) *Transmitting only m_2* : Over the last $n_{2,2}$ channel uses, the transmitter encodes only m_2 with a codeword $\mathbf{X}_{2,2}(m_2)$ which is independently distributed on the sphere $\mathbb{S}^{n_{2,2}-1}$ with power $n_{2,2}\beta_{2,2}\mathsf{P}$.

B. Decoding

Given the structure of the encoding functions, the output sequences at each receiver can be viewed as arising from two parallel channels: Rx 1 observes the two channel outputs $\mathbf{Y}_{1,1}$ and $\mathbf{Y}_{1,2}$ of $n_{1,1}$ and $n_{1,2}$ blocks, respectively; and Rx 2 observes the two channel outputs $\mathbf{Y}_{2,1}$ and $\mathbf{Y}_{2,2}$ of $n_{1,2}$ and $n_{2,2}$ blocks, respectively.

1) *Decoding m_1* : Given observations $\mathbf{Y}_{1,1}$ and $\mathbf{Y}_{1,2}$, Rx 1 estimates m_1 according to the pair $(\hat{m}_1, \hat{\ell}_1)$, such that the corresponding sequences $\mathbf{X}_{1,2}(\hat{m}_1, \hat{\ell}_1)$ and $\mathbf{X}_{1,1}(\hat{m}_1)$ maximize

$$i_1(\{\mathbf{x}_{1,j}\}_{j \in \{1,2\}}; \{\mathbf{y}_{1,j}\}_{j \in \{1,2\}}) := \log \prod_{j=1,2} \frac{f_{\mathbf{Y}_{1,j}|\mathbf{X}_{1,j}}(\mathbf{y}_{1,j}|\mathbf{x}_{1,j})}{f_{\mathbf{Y}_{1,j}}(\mathbf{y}_{1,j})} \quad (25)$$

over all pairs of $\mathbf{x}_{1,1}$ and $\mathbf{x}_{1,2} \in C_{\mathbf{X}_{1,2}}$. We have the following error event while decoding m_1 :

$$\mathcal{E}_1 \triangleq \{\text{Rx 1 chooses a message } \hat{m}_1 \neq m_1\}. \quad (26)$$

Thus the average error of decoding m_1 is bounded by

$$\epsilon_1 \leq \mathbb{P}[\mathcal{E}_{1,2}] + \mathbb{P}[\mathcal{E}_1 | \mathcal{E}_{1,2}^c]. \quad (27)$$

2) *Decoding m_2* : Given observations $\mathbf{Y}_{2,1}$ and $\mathbf{Y}_{2,2}$, Rx 2 estimates m_2 according to the pair $(\hat{m}_2, \hat{\ell}_2)$, such that the corresponding sequences $\mathbf{X}_{2,1}(\hat{m}_2, \hat{\ell}_2)$ and $\mathbf{X}_{2,2}(\hat{m}_2)$ maximize

$$i_2(\{\mathbf{x}_{2,j}\}_{j \in \{1,2\}}; \{\mathbf{y}_{2,j}\}_{j \in \{1,2\}}) := \log \prod_{j=1,2} \frac{f_{\mathbf{Y}_{2,j}|\mathbf{X}_{2,j}}(\mathbf{y}_{2,j}|\mathbf{x}_{2,j})}{f_{\mathbf{Y}_{2,j}}(\mathbf{y}_{2,j})} \quad (28)$$

over all pairs of $\mathbf{x}_{2,1} \in C_{\mathbf{X}_{2,1}}$ and $\mathbf{x}_{2,2}$. We have the following error event while decoding m_2 :

$$\mathcal{E}_2 \triangleq \{\text{Rx 2 chooses a message } \hat{m}_2 \neq m_2\}. \quad (29)$$

Thus the average error of decoding m_2 is bounded by

$$\epsilon_2 \leq \mathbb{P}[\mathcal{E}_{1,2}] + \mathbb{P}[\mathcal{E}_2 | \mathcal{E}_{1,2}^c]. \quad (30)$$

Remark 1: To improve the finite blocklength performance, all the codewords are uniformly distributed on the *power shell*. According to Shannon's observation, the optimal decay of the probability of error near capacity of the point-to-point Gaussian channel is achieved by codewords on the power-shell [31]. As a result of this code construction, the induced output distributions $f_{\mathbf{Y}_{i,j}}(\mathbf{y}_{i,j})$, with $i = 1, 2$, $j = 1, 2$, $f_{\mathbf{Y}_{1,2}|\mathbf{X}_{1,2}}(\mathbf{y}_{1,2}|\mathbf{x}_{1,2})$ and $f_{\mathbf{Y}_{2,1}|\mathbf{X}_{2,1}}(\mathbf{y}_{2,1}|\mathbf{x}_{2,1})$ are non-i.i.d.; thus we propose to bound the corresponding information density measure $i_i(\{\mathbf{x}_{i,j}\}_{j \in \{1,2\}}; \{\mathbf{y}_{i,j}\}_{j \in \{1,2\}})$, for each $i \in \{1, 2\}$, by

$$\tilde{i}_i(\{\mathbf{x}_{i,j}\}_{j \in \{1,2\}}; \{\mathbf{y}_{i,j}\}_{j \in \{1,2\}}) := \log \frac{f_{\mathbf{Y}_{i,i}|\mathbf{X}_{i,i}}(\mathbf{y}_{i,i}|\mathbf{x}_{i,i}) Q_{\mathbf{Y}_{i,j \neq i}|\mathbf{X}_{i,j \neq i}}(\mathbf{y}_{i,j \neq i}|\mathbf{x}_{i,j \neq i})}{\prod_{j=1}^2 Q_{\mathbf{Y}_{i,j}}(\mathbf{y}_{i,j})}, \quad (31)$$

where the Q s are i.i.d Gaussian distributions. For each $i \in \{1, 2\}$, we thus introduce a parameter J_i , chosen such that

$$i_i(\{\mathbf{x}_{i,j}\}_{j \in \{1,2\}}; \{\mathbf{y}_{i,j}\}_{j \in \{1,2\}}) \geq \tilde{i}_i(\{\mathbf{x}_{i,j}\}_{j \in \{1,2\}}; \{\mathbf{y}_{i,j}\}_{j \in \{1,2\}}) + \log J_i. \quad (32)$$

IV. MAIN RESULTS

Define $n_{2,1} = n_{1,2}$. For each $i \in \{1, 2\}$ let $j \in \{1, 2\} \setminus i$, and define

$$\Omega_{i,i} := \frac{\beta_{i,i} \mathbf{P}}{\sigma_{i,i}^2}, \quad (33a)$$

$$\Omega_{i,j} := \frac{(\beta_{i,j} + \rho^2 \beta_{j,i} + \rho \sqrt{\beta_{i,j} \beta_{j,i}}) \mathbf{P}}{\sigma_{i,j}^2 + (1 - \rho^2) \beta_{j,i} \mathbf{P}}, \quad (33b)$$

$$J_i := e^{\frac{n_{i,i}}{2} (\log 2 - \beta_{j,i} \mathbf{P})} \frac{4 \sqrt{\pi \beta_{i,j} \beta_{j,i}} (1 + 2\Omega_{i,i})}{243 (1 + \Omega_{i,i}) (\beta_{i,j} + \beta_{j,i})}. \quad (33c)$$

Our first result gives upper bounds on the achievable rates of the first and the second message.

Theorem 1: Given $n_{1,1}, n_{1,2}, n_{2,2}$ and \mathbf{P} , for each $i \in \{1, 2\}$, let $j \in \{1, 2\} \setminus i$, $\log L_i := n_{i,j} R_{L_i}$ and $\log M_i := R_i \sum_{j=1}^2 n_{i,j}$. For each $i \in \{1, 2\}$, we then have the following upper bound:

$$\log M_i + \log L_i \leq \sum_{j=1}^2 n_{i,j} C(\Omega_{i,j}) - \sqrt{\sum_{j=1}^2 n_{i,j} V(\Omega_{i,j})} \mathbb{Q}^{-1}(\epsilon_i - \Delta_i) + K_i \log \left(\sum_{j=1}^2 n_{i,j} \right) \quad (34)$$

subject to

$$\sum_{i=1}^2 \log L_i \geq \log \frac{\log(\epsilon_{1,2})}{\log(1 - (1 - \rho^2)^{n_{1,2}-1})}, \quad (35)$$

where $C(x) := \frac{1}{2} \log(1+x)$ and $V(x) := \frac{x(2+x)}{2(1+x)^2} \log e^2$ with

$$\Delta_i := \frac{6T_{\max,i}}{\sqrt{(\sum_{j=1}^2 n_{i,j} V(\Omega_{i,j}))^3}} + \frac{\prod_{j=1}^2 (1 + \Omega_{i,j})^{\frac{n_{i,j}}{2}}}{2^{E_{\min,i}} J_i \left(\sum_{j=1}^2 n_{i,j} \right)^{-K_i}} + \left(1 - (1 - \rho^2)^{n_{i,j}-1} \right)^{L_1 \cdot L_2}, \quad (36)$$

where $T_{\max,i}$ and $E_{\min,i}$ are defined in (37) and $\Phi(\cdot, \cdot, \cdot)$ is the Hurwitz Lerch transcendent, $\Gamma_{i,i} := \Gamma(0.5n_{i,i} + 0.5)/\Gamma(0.5n_{i,i})$, $\Gamma_{i,j} := \Gamma(0.5n_{i,j} + 0.5)/\Gamma(0.5n_{i,j})$, $\tilde{\beta}_c := \beta_{1,2} + \beta_{2,1} + \sqrt{\beta_{1,2}\beta_{2,1}}$, and for any $\kappa_i > 1$, $\zeta_i > 1$ and $\tilde{\zeta}_i > 1$.

Proof: See Appendix A. ■

Proposition 1: For sufficiently large $n_{1,1}, n_{1,2}, n_{2,2}$, and for $\rho > 0$, we have

$$\log M_i + \log L_i \leq \sum_{j=1}^2 n_{i,j} C(\Omega_{i,j}) - \sqrt{\sum_{j=1}^2 n_{i,j} V(\Omega_{i,j})} \mathbb{Q}^{-1}(\epsilon_i) + O\left(\log\left(\sum_{j=1}^2 n_{i,j}\right)\right) \quad (38)$$

subject to

$$\sum_{i=1}^2 \log L_i \geq \log(-\log \epsilon_{1,2}) - n_{1,2} \log(1 - \rho^2). \quad (39)$$

Proof: See Appendix B. ■

V. DISCUSSION ON THE MAIN RESULTS AND RELATED WORKS

In this section, we review related settings that can be covered by Theorem 1.

$$T_{\max,i} := 2^3 \kappa_i + 2^4 \max \left\{ \frac{1}{\Gamma\left(\frac{n_{i,i}}{2}\right)} \zeta_i 2^{\frac{n_{i,i}-2}{4}} e^{-c_i} A(n_{i,i}, k_i, b_i, c_i), \frac{1}{\Gamma\left(\frac{n_{i,j}}{2}\right)} \tilde{\zeta}_i 2^{\frac{n_{i,j}-2}{4}} e^{-\tilde{c}_i} A(n_{i,j}, \tilde{k}_i, \tilde{b}_i, \tilde{c}_i) \right\}, \quad (37a)$$

$$\begin{aligned} E_{\min,i} := & \sum_{j=1}^2 \frac{n_{i,j}}{2} \log(1 + \Omega_{i,j}) + \frac{(1 - \sigma_{i,i}^2) n_{i,i} \Omega_{i,i}}{2(1 + \Omega_{i,i})} - \sqrt{2} \sigma_{i,i}^2 \Gamma_{i,i} k_i b_i - \frac{n_{i,j} \sigma_{i,j}^2 \beta_c \mathbf{P}}{2(\sigma_{i,j}^2 + \beta_c \mathbf{P})} - \frac{\beta_c \mathbf{P}^2 n_{i,j} \beta_{j,i}}{2\sigma_{i,j}^2 (\beta_c \mathbf{P} + \sigma_{i,j}^2)} \\ & + \frac{\beta_c \mathbf{P} n_{i,j} \beta_{j,i}}{2\tilde{\beta}_c \sigma_{i,j}^2 (\beta_c \mathbf{P} + \sigma_{i,j}^2)} \left(\sigma_{i,j}^2 + (\beta_c \mathbf{P} - \sigma_{i,j}^2) \left(\sqrt{\tilde{\beta}_c \beta_c^{-1}} - 1 - \rho \sqrt{\beta_{j,i} \beta_{i,j}^{-1}} \right)^2 - \frac{\sqrt{2\tilde{\mathbf{P}} \tilde{\beta}_c} \Gamma_{i,j}}{\sqrt{n_{i,j} \beta_{j,i}}} \right) \\ & + \sqrt{n_{i,j} \beta_{j,i}} \mathbf{P} \left(\sqrt{\beta_c \tilde{\beta}_c^{-1}} \left(\frac{1}{\sigma_{i,j}^2 + \beta_c \mathbf{P}} + 1 + \rho \sqrt{\beta_{j,i} \beta_{i,j}^{-1}} \right) - \frac{1}{\sigma_{i,j}^2} \right) \left(\rho \sqrt{n_{i,j} \beta_{j,i}} \mathbf{P} - \sigma_{i,j}^2 \sqrt{2\Gamma_{i,j}} \right), \end{aligned} \quad (37b)$$

$$\begin{aligned} A(n, k, b, c) := & \left((\tilde{\kappa} + 1)^3 - \frac{k^3 b^6}{8} - \frac{3}{2} (\tilde{\kappa} + 1) k b^2 (\tilde{\kappa} + 1 - \frac{1}{2} k b^2) \right) \Phi(e^{-1}, -\frac{n}{2} + 1, c) \\ & - 3k (5k^2 b^4 + 4(\tilde{\kappa} + 1)((\tilde{\kappa} + 1) - 3k b^2)) \Phi(e^{-1}, -\frac{n}{2}, c) - 96\sqrt{2} k^3 b \Phi(e^{-1}, -\frac{n}{2} - \frac{3}{2}, c) \\ & - 3kb \left(\frac{k^2 b^4}{\sqrt{2}} + 2\sqrt{2}(\tilde{\kappa} + 1)b((\tilde{\kappa} + 1) - b^2 k) \right) \Phi(e^{-1}, -\frac{n}{2} + \frac{1}{2}, c) - 64k^3 \Phi(e^{-1}, -\frac{n}{2} - 2, c) \\ & + 8\sqrt{2} k^2 b (6(\tilde{\kappa} + 1) - 5k b^2) \Phi(e^{-1}, -\frac{n}{2} - \frac{1}{2}, c) + 24k(2(\tilde{\kappa} + 1)^2 - 5k^2 b^2) \Phi(e^{-1}, -\frac{n}{2} - 1, c), \end{aligned} \quad (37c)$$

$$\begin{aligned} \tilde{\kappa}_i := & \frac{k_i b_i^2}{4} + \frac{\tilde{k}_i \tilde{b}_i^2}{4} - \frac{n_{i,i}}{2} \log(1 + \Omega_{i,i}) - \frac{n_{i,j}}{2} \log \frac{\beta_c \mathbf{P} + \sigma_{i,j}^2}{\sigma_{i,j}^2} - \frac{n_{i,i} \Omega_{i,i}}{2(1 + \Omega_{i,i})} \\ & - \tilde{k}_i \tilde{b}_i - \frac{n_{i,j} \beta_{i,j} \mathbf{P}}{2(\beta_c \mathbf{P} + \sigma_{i,j}^2)} + \frac{\sigma_{i,j}^2}{2} \left(\tilde{k}_i (\tilde{b}_i - \sqrt{n_{i,j} \beta_{j,i}} \mathbf{P}) - \frac{\sqrt{n_{i,j} \beta_{i,j}} \mathbf{P}}{\beta_c \mathbf{P} + \sigma_{i,j}^2} \right)^2, \end{aligned} \quad (37d)$$

$$c_i := \frac{1}{2} \left(\sqrt{\frac{\tilde{\kappa}_i + \kappa_i}{2k_i}} - \frac{b_i}{2} \right)^2, \quad \tilde{c}_i := \frac{1}{2} \left(\sqrt{\frac{\tilde{\kappa}_i + \kappa_i}{2\tilde{k}_i}} - \frac{\tilde{b}_i}{2} \right)^2, \quad k_i := \frac{2 + \Omega_{i,i}}{2\sigma_{i,i}^2 (1 + \Omega_{i,i})}, \quad \tilde{k}_i := \frac{2\sigma_{i,j}^2 + \beta_c \mathbf{P}}{2\sigma_{i,j}^2 (\beta_c \mathbf{P} + \sigma_{i,j}^2)}, \quad (37e)$$

$$b_i := \frac{\sqrt{n_{i,i} \Omega_{i,i}}}{k_i \sigma_{i,i} (1 + \Omega_{i,i})}, \quad \tilde{b}_i := \sqrt{n_{i,j} \beta_{j,i}} \mathbf{P} + \frac{\sqrt{n_{i,j} \beta_{i,j}} \mathbf{P}}{\tilde{k}_i \sigma_{i,j}^2} \left(1 - \sqrt{\frac{\beta_c}{\tilde{\beta}_c}} \left(1 + \rho \sqrt{\frac{\beta_{j,i}}{\beta_{i,j}}} \right) \right) + \frac{\sqrt{n_{i,j} \beta_{i,j}} \mathbf{P}}{\tilde{k}_i (\beta_c \mathbf{P} + \sigma_{i,j}^2)}. \quad (37f)$$

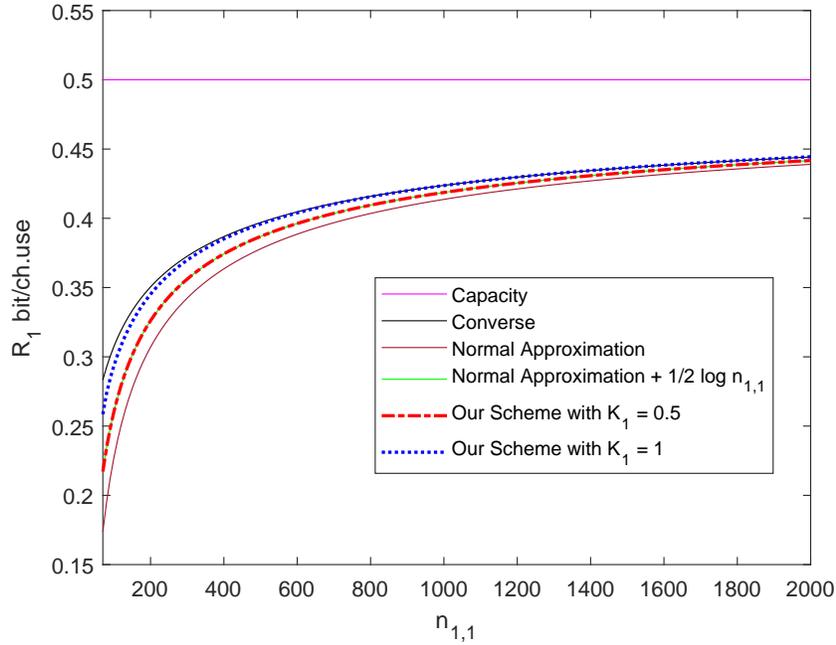


Fig. 2: Converse and achievability bounds on R_1 versus $n_{1,1}$ in the point-to-point case with $\Omega_{1,1} = 0$ dB and $\epsilon_1 = 10^{-3}$. The value of $n_{1,1}$ varies from 70 to 2000 with a step size of 10.

A. Point-to-Point Settings

1) *Transmission of only m_1 over $n_{1,1}$ channel uses:* In this setting, we have only the following channel outputs:

$$\mathbf{Y}_{1,1} = \mathbf{X}_{1,1} + \mathbf{Z}_{1,1} \quad (40)$$

with $\|\mathbf{X}_{1,1}\|^2 = n_{1,1}P$ and $\mathbf{Z}_{1,1} \sim \mathcal{N}(0, \sigma_{1,1}^2 I_{n_{1,1}})$. Set

$$\Omega_{1,1} = \frac{P}{\sigma_{1,1}^2}. \quad (41)$$

Theorem 2 (P2P, Only m_1): Given $n_{1,1}$ and P , $\log M_1 := n_{1,1}R_1$ is upper bounded as

$$\log M_1 \leq n_{1,1}C(\Omega_{1,1}) - \sqrt{n_{1,1}V(\Omega_{1,1})}\mathbb{Q}^{-1}(\epsilon_1 - \Delta_1) + K_1 \log(n_{1,1}), \quad (42)$$

where

$$\Delta_1 := \frac{6T_{\max,1}}{\sqrt{(n_{1,1}V(\Omega_{1,1}))^3}} + \frac{n_{1,1}^{K_1}(1 + \Omega_{1,1})^{\frac{n_{1,1}}{2}}}{J_1 \cdot 2^{E_{\min,1}}}, \quad (43)$$

and by [32, Proposition 2]:

$$J_1 := \frac{\sqrt{8(\sigma_{1,1}^2 + 2P)}}{27\sqrt{\pi}(\sigma_{1,1}^2 + P)} \quad (44)$$

with K_1 a constant. Notice that $T_{\max,1}$ and $E_{\min,1}$ can be easily calculated from (37a) and (37b) by setting $i = 1$ and $n_{i,j} = 0$.

Proof: The proof of this result follows the proof of Theorem 1 by setting

$$n_{1,2} = n_{2,1} = n_{2,2} = 0 \quad \text{and} \quad \beta_{1,1} = 1. \quad (45)$$

Remark 2: By choosing K_1 such that $\Delta_1 = O(n_{1,1}^{-\frac{3}{2}})$, the bound in (42) matches the Normal Approximation bound in [9, Eq. 223] which is for Gaussian channels. ■

We now make a more specific comparison with the work of Polyanskiy, Poor and Verdú in [9]. To this end, note that, the authors in [9] show for both discrete memoryless channels and Gaussian channels, the Normal Approximation achieves

$$\log M_1 = n_{1,1}C(\Omega_{1,1}) - \sqrt{n_{1,1}V(\Omega_{1,1})}\mathbb{Q}^{-1}(\epsilon_1) + O(\log(n_{1,1})). \quad (46)$$

See [9, Eq.223]. The authors then show that for the Gaussian channels with equal-power and maximal-power constraints, the $O(\log(n_{1,1}))$ can be bounded as

$$O(\log(n_{1,1})) \leq \frac{1}{2} \log n_{1,1} + O(1), \quad (47)$$

and with the average power constraint as

$$O(\log(n_{1,1})) \leq \frac{3}{2} \log n_{1,1} + O(1). \quad (48)$$

See [9, Eq.294 and Eq.295]. To compare our achievability bound in (42) with the Normal Approximation, Fig. 2 illustrates the bound in (42) for $K_1 = 0.5$ and $K_1 = 1$ as well as the Normal Approximation with $O(\log(n_{1,1}))$ equal to 0 and $0.5 \log n_{1,1}$. The converse bound is the meta-converse bound proposed in [9, Theorem 41]. As can be seen from this figure, our achievability schemes with $K_1 = 1$ matches the converse bound for large values of $n_{1,1}$ and has an improvement over the Normal approximation especially for values of $n_{1,1}$ smaller than 200.

2) *Transmission of only m_1 over two parallel channels of $n_{1,1}$ and $n_{1,2}$ channel uses:* In this setting, we have the following two channel outputs:

$$\mathbf{Y}_{1,1} = \mathbf{X}_{1,1} + \mathbf{Z}_{1,1}, \quad \mathbf{Y}_{1,2} = \mathbf{X}_{1,2} + \mathbf{Z}_{1,2}, \quad (49)$$

with $\|\mathbf{X}_{1,1}\|^2 = n_{1,1}\beta_{1,1}\mathbf{P}$, $\|\mathbf{X}_{1,2}\|^2 = n_{1,2}\beta_{1,2}\mathbf{P}$, and $\mathbf{Z}_{1,1} \sim \mathcal{N}(0, \sigma_{1,1}^2 I_{n_{1,1}})$ and $\mathbf{Z}_{1,2} \sim \mathcal{N}(0, \sigma_{1,2}^2 I_{n_{1,2}})$. The power sharing parameters $\beta_{1,1} \in [0, 1]$ and $\beta_{1,2} \in [0, 1]$ are chosen such that

$$\beta_{1,1}n_{1,1} + \beta_{1,2}n_{1,2} = n. \quad (50)$$

Set

$$\Omega_{1,1} = \frac{\beta_{1,1}\mathbf{P}}{\sigma_{1,1}^2} \quad \text{and} \quad \Omega_{1,2} = \frac{\beta_{1,2}\mathbf{P}}{\sigma_{1,2}^2}. \quad (51)$$

Theorem 3 (P2P, Parallel Channels, Only m_1): Given $n_{1,1}, n_{1,2}, \beta_{1,1}, \beta_{1,2}$ and \mathbf{P} , $\log M_1 := R_1(n_{1,1} + n_{1,2})$ is upper bounded as

$$\log M_1 \leq \sum_{j=1}^2 n_{1,j}C(\Omega_{1,j}) - \sqrt{\sum_{j=1}^2 n_{1,j}V(\Omega_{1,j})}\mathbb{Q}^{-1}(\epsilon_1 - \Delta_1) + K_1 \log \left(\sum_{j=1}^2 n_{1,j} \right), \quad (52)$$

where

$$\Delta_1 := \frac{6T_{\max,1}}{\sqrt{(\sum_{j=1}^2 n_{1,j}V(\Omega_{1,j}))^3}} + \frac{\prod_{j=1}^2 (1 + \Omega_{1,j})^{\frac{n_{1,j}}{2}}}{2^{E_{\min,1}} J_1 \left(\sum_{j=1}^2 n_{i,j} \right)^{-K_1}} \quad (53)$$

with

$$J_1 := \frac{\sqrt{8(\sigma_{1,1}^2 + 2\beta_{1,1}\mathbf{P})}}{27\sqrt{\pi}(\sigma_{1,1}^2 + \beta_{1,1}\mathbf{P})} \cdot \frac{\sqrt{8(\sigma_{1,2}^2 + 2\beta_{1,2}\mathbf{P})}}{27\sqrt{\pi}(\sigma_{1,2}^2 + \beta_{1,2}\mathbf{P})} \quad (54)$$

and K_1 a constant. Notice that $T_{\max,1}$ and $E_{\min,1}$ can be easily calculated from (37a) and (37b) by setting $i = 1$, $\beta_{2,1} = 0$ and $\beta_c = \beta_{1,2}$.

Proof: The proof follows the proof of Theorem 1 by setting

$$n_{2,2} = 0, \quad \beta_{2,1} = 0, \quad \text{and} \quad \beta_{2,2} = 0. \quad (55)$$

Remark 3: By choosing K_1 such that $\Delta_1 = O(n_{1,1}^{-\frac{3}{2}})$, and $n_{1,1} = n_{1,2}$, the bound in (52) matches the Normal Approximation bound in [15, Eq. 54] that is for a two-parallel additive white Gaussian noise (AWGN) channels. This setting can be easily extended to a setting with more than two parallel channels. ■

We now make a more detailed comparison with the work of Erseghe in [15]. The leading idea of [15] is based on the fact that a probability of the form $\mathbb{P}[\sum_{i=1}^n u_i \geq n\tilde{\lambda}]$ where u_i s are i.i.d continues random variables can be written and numerically evaluated using standard Laplace transforms. The author further employs this idea to propose new asymptotic approximations for the meta-converse and random union coding (RCU) achievability bounds for the parallel AWGN channels. In [15, Theorem 10] the author shows that in a K -parallel AWGN channels with each channel of $\frac{n}{K}$ channel uses, the meta-converse bound, proposed by Polyanskiy, Poor and Verdú in [9, Theorem 41], is consistent with the normal approximation (1) with

$$V = \frac{1}{K} \sum_{k=1}^K \frac{\Omega_k(2 + \Omega_k)}{2(1 + \Omega_k)^2} \quad \text{and} \quad C = \frac{1}{K} \sum_{k=1}^K \frac{1}{2} \log(1 + \Omega_k), \quad (56)$$

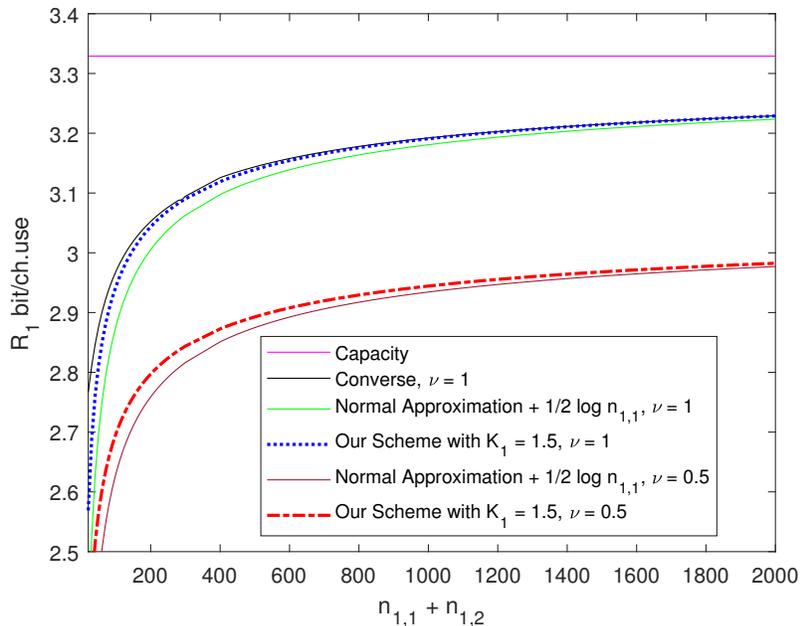


Fig. 3: Converse and achievability bounds on R_1 versus $n_{1,1} + n_{1,2}$ in the two parallel Gaussian channel case with $\Omega_{1,1} = 20$ dB, $\Omega_{1,1} = \nu\Omega_{2,2}$ and $\epsilon_1 = 10^{-6}$. The value of $n = n_{1,1} + n_{1,2}$ varies from 20 to 2000 with a step size of 10.

where Ω_k is the SNR of the k -th channel. We now numerically compare our bound in (52) with the normal approximation bound in [15] for the case of 2-parallel channels. Since in [15] all K channels are of equal blocklength thus we assume $n_{1,1} = n_{1,2}$ and $\Omega_{1,1} = \nu \cdot \Omega_{2,2}$ for some $\nu > 0$. Fig. 3 illustrates this comparison for $\Omega_{1,1} = 20$ dB, $\nu = 0.5$ and $\nu = 1$, $\epsilon_1 = 10^{-6}$ and $K_1 = 1.5$. In the case that channels are of equal SNRs, i.e., $\nu = 1$, the Normal Approximation for K -parallel channels each of blocklength $\frac{n}{K}$ is equivalent to the Normal Approximation for one channel of blocklength n . As can be seen from Fig. 3, for the case of $\nu = 1$, the Normal Approximation is not a good approximation of the meta-converse bound. Also, our achievability bound in (52) with $K_1 = 1.5$ outperforms the normal approximation for $\nu = 1$ and $\nu = 0.5$ specifically at small values of $n = n_{1,1} + n_{1,2}$.

B. Broadcast Setting: Marton's Bounds

In this section, we compare our results with Marton's inner bound in [30] proposed for a general two-receiver broadcast channel. In [30], the Tx wishes to transmit message m_1 to Rx 1 and message m_2 to Rx 2 each over n channel uses. The Tx thus encodes m_1 using the codeword $\mathbf{U}_1^n(m_1)$ and encodes m_2 using the codeword $\mathbf{U}_2^n(m_2)$ where \mathbf{U}_1 and \mathbf{U}_2 are distributed according to $f_{\mathbf{U}_1\mathbf{U}_2}(\mathbf{u}_1, \mathbf{u}_2)$ and are correlated with a correlation parameter ρ such that

$$\langle \mathbf{U}_1, \mathbf{U}_2 \rangle = n\rho\sqrt{\beta_{1,2}\beta_{2,1}}\mathbf{P}. \quad (57)$$

With the codewords \mathbf{U}_1 and \mathbf{U}_2 , the following codeword is formed:

$$\mathbf{X} = \mathbf{U}_1 + \mathbf{U}_2, \quad (58)$$

and is transmitted to Rx 1 over the channel $f_{\mathbf{Y}_1|\mathbf{X}}$ and to Rx 2 over the channel $f_{\mathbf{Y}_2|\mathbf{X}}$. Here, \mathbf{Y}_1 and \mathbf{Y}_2 denote the output sequences of the first and second channels, respectively.

Let $R_1 := \log M_1/n$ and $R_2 := \log M_2/n$, then the following inner bounds hold for this setting:

Theorem 4 (Marton's bound): The capacity region \mathcal{C} of this setting is a set of rate pairs (R_1, R_2) satisfying

$$R_1 \leq I(\mathbf{U}_1; \mathbf{Y}_1), \quad (59a)$$

$$R_2 \leq I(\mathbf{U}_2; \mathbf{Y}_2), \quad (59b)$$

$$R_1 + R_2 \leq I(\mathbf{U}_1; \mathbf{Y}_1) + I(\mathbf{U}_2; \mathbf{Y}_2) - I(\mathbf{U}_1; \mathbf{U}_2). \quad (59c)$$

To compare this setup with ours, let m_1 and m_2 to be jointly sent over the entire n channel uses. Rx 1 observes the channel outputs $\mathbf{Y}_{1,1}$ and Rx 2 observes the channel outputs $\mathbf{Y}_{2,2}$ given by

$$\mathbf{Y}_{1,1} = \alpha(\mathbf{X}_{1,2} + \mathbf{X}_{2,1}) + \mathbf{Z}_{1,1}, \quad \mathbf{Y}_{2,2} = \alpha(\mathbf{X}_{1,2} + \mathbf{X}_{2,1}) + \mathbf{Z}_{2,2} \quad (60)$$

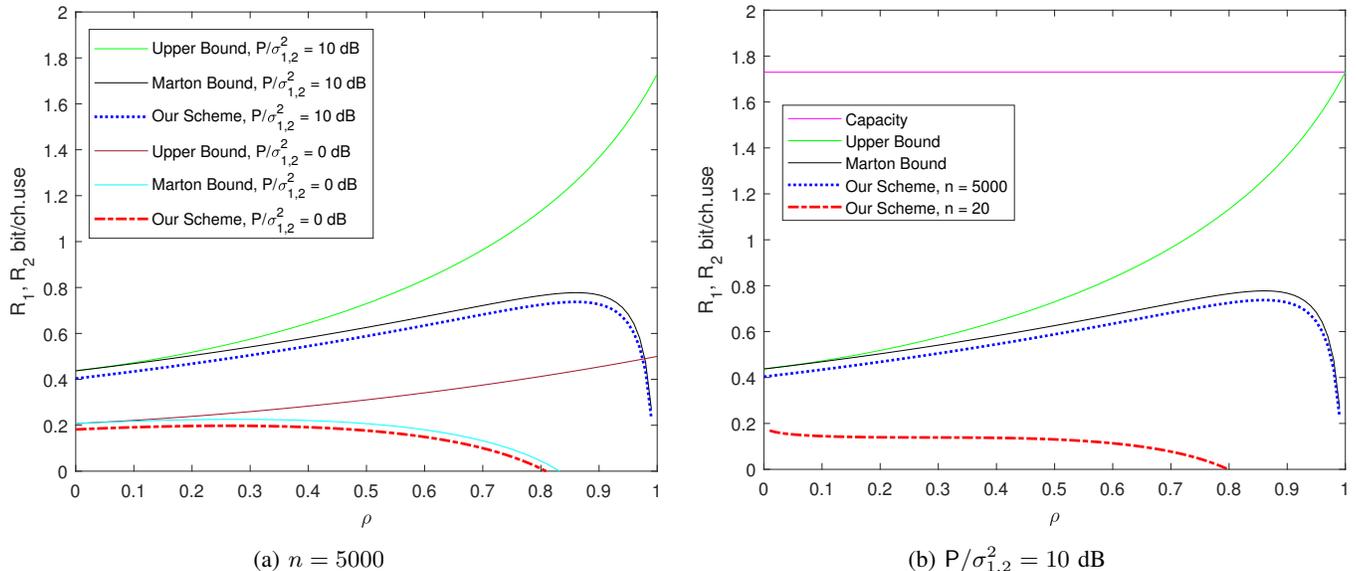


Fig. 4: Effect of the correlation parameter ρ on the bounds in Theorem 1 and Proposition 2.

where $\mathbf{X}_{2,1} \sim \mathcal{N}(0, \beta_{1,2} \mathbf{P} I_n)$ and $\mathbf{X}_{1,2} \sim \mathcal{N}(0, \beta_{2,1} I_n)$ where $\beta_{1,2}$ and $\beta_{2,1}$ are chosen such that

$$\beta_{1,2} + \beta_{2,1} + \rho \sqrt{\beta_{1,2} \beta_{2,1}} = 1. \quad (61)$$

For this Gaussian case, Theorem 4 can be written as the following proposition.

Proposition 2: Let $R_1 := \frac{\log M_1}{n}$, $R_2 := \frac{\log M_2}{n}$, $R_{L_1} := \frac{\log L_1}{n}$, $R_{L_2} := \frac{\log L_2}{n}$, $\mathbf{X}_{1,2} \sim \mathcal{N}(0, \beta_{1,2} \mathbf{P} I_n)$ and $\mathbf{X}_{2,1} \sim \mathcal{N}(0, \beta_{2,1} \mathbf{P} I_n)$, we have the following inner bounds:

$$R_1 + R_{L_1} \leq \frac{1}{2} \log \left(\frac{\sigma_{1,1}^2 + \mathbf{P}}{\sigma_{1,1}^2 + (1 - \rho^2) \beta_{2,1} \mathbf{P}} \right) \quad (62a)$$

$$(62b)$$

and

$$R_2 + R_{L_2} \leq \frac{1}{2} \log \left(\frac{\sigma_{2,2}^2 + \mathbf{P}}{\sigma_{2,2}^2 + (1 - \rho^2) \beta_{1,2} \mathbf{P}} \right) \quad (62c)$$

subject to

$$R_{L_1} + R_{L_2} \geq -\frac{1}{2} \log(1 - \rho^2). \quad (63)$$

Remark 4: By setting $n_{1,1} = n_{2,2} = 0$, $n_{1,2} = n$ and

$$\epsilon_{1,2} = \frac{1}{2^{(1-\rho^2) \frac{n_{1,2}}{2}}}, \quad (64)$$

Proposition 1 matches Proposition 2 in the asymptotic regime (i.e., when $n_{1,2} \rightarrow \infty$) and when all the codewords are i.i.d Gaussian.

Remark 5: If all the codewords are i.i.d Gaussian, by setting either L_1 or L_2 to 0, then in the asymptotic regime Proposition 1 matches the DPC results of Costa [22] and in the finite blocklength regime matches the results of Scarlett [23, Theorem 2].

The correlation parameter ρ is of a vital importance in both our scheme and Marton's scheme. This is due to the fact that increasing ρ increases the right-hand side of (34) and (62) as well as the right-hand side of (35) and (63) of Theorem 1 and Proposition 2, respectively. More specifically, increasing ρ increases the upper bound on $\log M_i + \log L_i$ as well as the lower bound on $\sum_{i=1}^2 \log L_i$. Hence, increasing ρ will not always increase the upper bound on R_i . Fig. 4 illustrates the effects of increasing ρ on upper bounds on R_1 and R_2 in (34) and Marton's bounds in Proposition 2. To make a comparison with Marton's bounds, in this figure, we set $n = n_{1,2}$, $n_{1,1} = n_{2,2} = 0$, $\epsilon_{1,2} = 10^{-3}$, $K_1 = K_2 = 1.5$, $\beta_{1,2} = \beta_{2,1}$, $L_1 = L_2$, $R_1 = R_2$, and $\sigma_{1,2}^2 = \sigma_{2,1}^2$. Fig. 4a is for the case where $n = 5000$ and $P/\sigma_{1,2}^2$ is equal to 10 dB and 0 dB. As can be seen from Fig. 4a, when $P/\sigma_{1,2}^2$ is equal to 10 dB, the transmission rate under our scheme and Marton's scheme maximizes around $\rho = 0.9$ and decays for the values of ρ larger than 0.9. Whereas, for $P/\sigma_{1,2}^2$ equal to 0 dB, the transmission rate under our

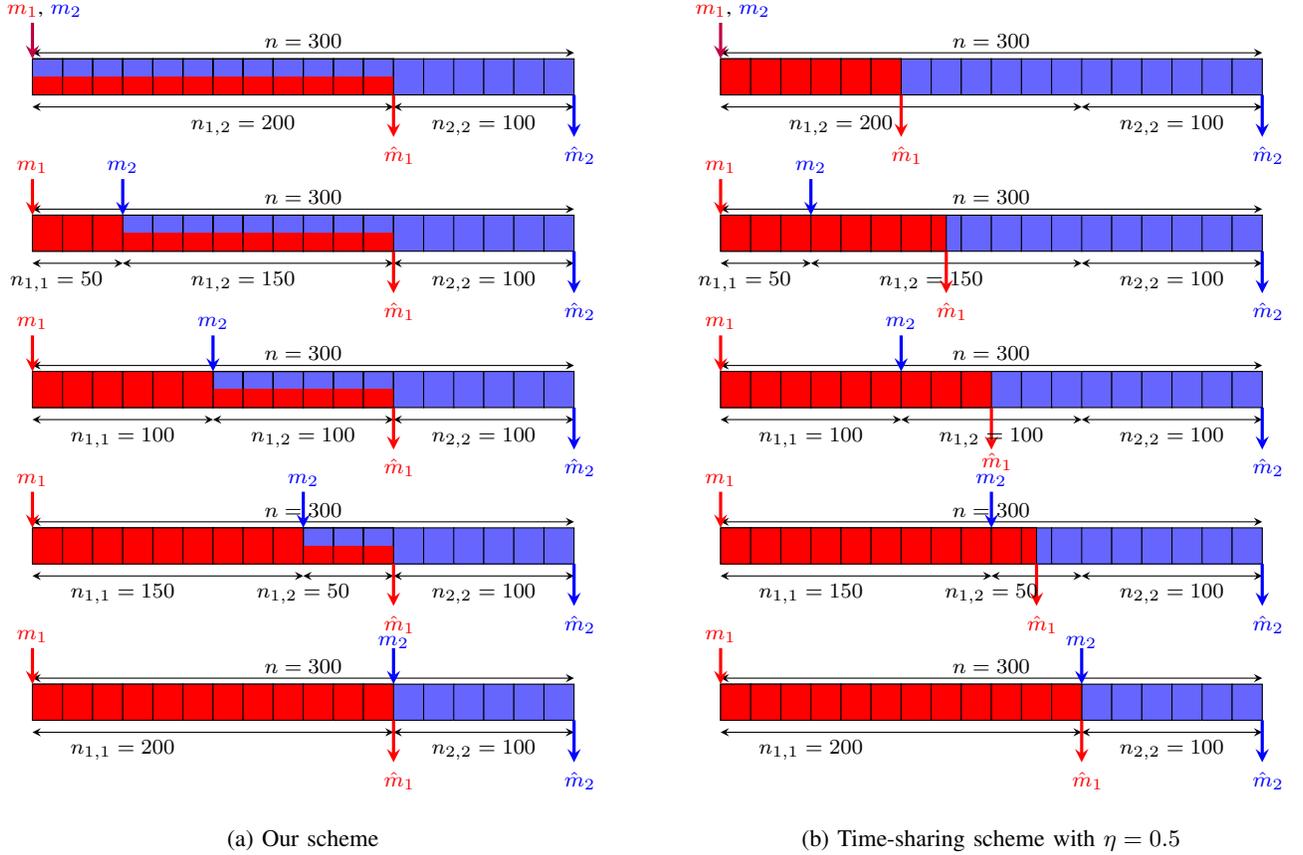


Fig. 5: Example of comparing our scheme with the time-sharing scheme for $n = 300$ and $n_{1,2}$ taking values from 200 to 0 with a step size of 50.

scheme and Marton's scheme maximizes at $\rho = 0$. A similar trend can be seen for very small values of n . See Fig. 4b. Notice that the upper bound is the case where R_{L_1} (or R_{L_2}) is set at 0 thus R_1 (or R_2) can be maximized.

C. Time-sharing: Transmission of both m_1 and m_2

In this setting, we divide the $n_{1,2}$ channel uses into two parts $\eta n_{1,2}$ and $(1 - \eta)n_{1,2}$ for $\eta \in [0, 1]$. Therefore, m_1 is transmitted over $n_{1,1} + \eta n_{1,2}$ channel uses and m_2 is transmitted over $(1 - \eta)n_{1,2} + n_{2,2}$ channel uses. Transmissions of m_1 and m_2 are thus independent. In this setting, we have the following two channel outputs:

$$\mathbf{Y}_{1,1} = \mathbf{X}_{1,1} + \mathbf{Z}_{1,1}, \quad \mathbf{Y}_{2,2} = \mathbf{X}_{2,2} + \mathbf{Z}_{2,2}, \quad (65)$$

with $\|\mathbf{X}_{1,1}\|^2 = (n_{1,1} + \eta n_{1,2})\beta_{1,1}P$, $\|\mathbf{X}_{2,2}\|^2 = ((1 - \eta)n_{1,2} + n_{2,2})\beta_{2,2}P$, and $\mathbf{Z}_{1,1} \sim \mathcal{N}(0, \sigma_{1,1}^2 I_{n_{1,1} + \eta n_{1,2}})$ and $\mathbf{Z}_{2,2} \sim \mathcal{N}(0, \sigma_{2,2}^2 I_{(1 - \eta)n_{1,2} + n_{2,2}})$. The power sharing parameters $\beta_{1,1} \in [0, 1]$ and $\beta_{2,2} \in [0, 1]$ are chosen such that

$$(n_{1,1} + \eta n_{1,2})\beta_{1,1} + ((1 - \eta)n_{1,2} + n_{2,2})\beta_{2,2} = n. \quad (66)$$

Set

$$\Omega_{1,1} = \frac{\beta_{1,1}P}{\sigma_{1,1}^2} \quad \text{and} \quad \Omega_{2,2} = \frac{\beta_{2,2}P}{\sigma_{2,2}^2}. \quad (67)$$

Theorem 5 (Time-Sharing, Both m_1 and m_2): Given $n_{1,1}$, $n_{2,2}$, $\beta_{1,1}$, $\beta_{2,2}$ and P , $\log M_1 := (n_{1,1} + \eta n_{1,2})R_1$ and $\log M_2 := (n_{2,2} + (1 - \eta)n_{1,2})R_2$ are upper bounded as

$$\log M_1 \leq (n_{1,1} + \eta n_{1,2})C(\Omega_{1,1}) - \sqrt{(n_{1,1} + \eta n_{1,2})V(\Omega_{1,1})}Q^{-1}(\epsilon_1 - \Delta_1) + \tilde{K}_1 \log((n_{1,1} + \eta n_{1,2})), \quad (68)$$

and

$$\log M_2 \leq (n_{2,2} + (1 - \eta)n_{1,2})C(\Omega_{2,2}) - \sqrt{(n_{2,2} + (1 - \eta)n_{1,2})V(\Omega_{2,2})}Q^{-1}(\epsilon_2 - \Delta_2) + \tilde{K}_2 \log((n_{2,2} + (1 - \eta)n_{1,2})), \quad (69)$$

where

$$\Delta_1 := \frac{6T_{\max,1}}{\sqrt{(n_{1,1} + \eta n_{1,2})V(\Omega_{1,1})}^3} + \frac{(n_{1,1} + \eta n_{1,2})\tilde{K}_1(1 + \Omega_{1,1})^{\frac{(n_{1,1} + \eta n_{1,2})}{2}}}{J_1 \cdot 2^{E_{\min,1}}} \quad (70)$$

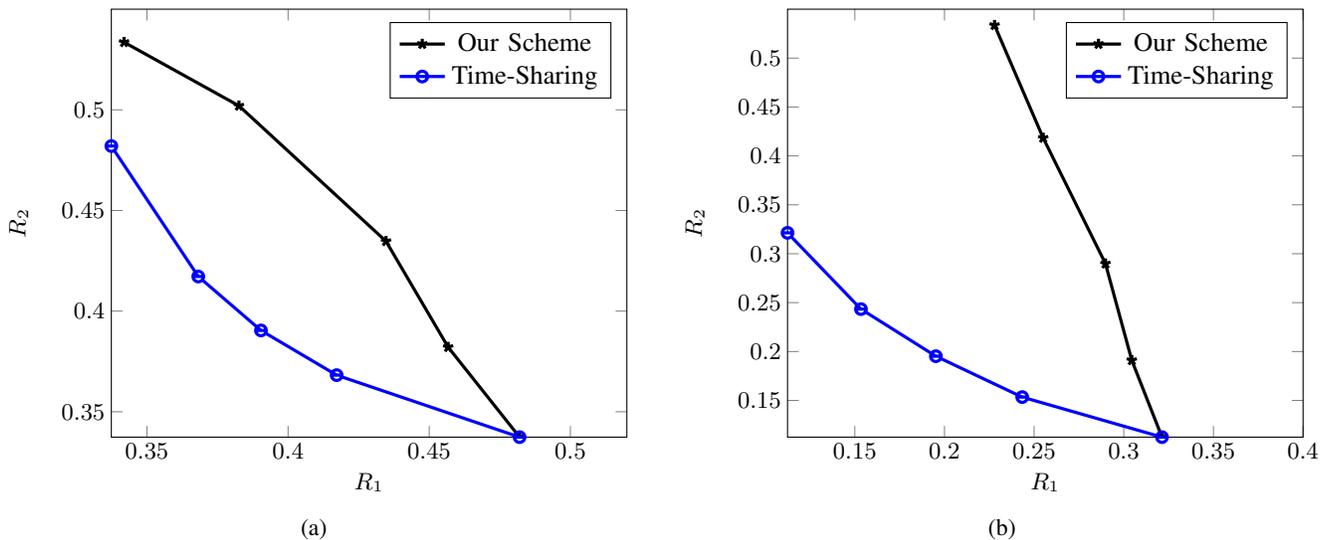


Fig. 6: Comparison between the transmission rate pairs (R_1, R_2) obtained in our scheme and the time-sharing scheme under the example shown in Fig. 5. (a) In our scheme: $R_1 = \log M_1/(n_{1,1} + n_{1,2})$ and $R_2 = \log M_2/(n_{1,1} + n_{2,2})$, in the time-sharing scheme: $R_1 = \log M_1/(n_{1,1} + \eta n_{1,2})$ and $R_2 = \log M_2/(n_{1,1} + (1 - \eta)n_{1,2})$ with $\eta = 0.5$, (b) In both schemes: $R_1 = \log M_1/n$ and $R_2 = \log M_2/n$.

$$\Delta_2 := \frac{6T_{\max,2}}{\sqrt{(n_{2,2} + (1 - \eta)n_{1,2})V(\Omega_{2,2})}^3} + \frac{(n_{2,2} + (1 - \eta)n_{1,2})^{\tilde{K}_2}(1 + \Omega_{2,2})^{\frac{(n_{2,2} + (1 - \eta)n_{1,2})}{2}}}{J_2 \cdot 2E_{\min,2}} \quad (71)$$

$$(72)$$

with

$$J_1 := \frac{\sqrt{8(\sigma_{1,1}^2 + 2\beta_{1,1}\mathsf{P})}}{27\sqrt{\pi}(\sigma_{1,1}^2 + \beta_{1,1}\mathsf{P})}, \quad J_2 := \frac{\sqrt{8(\sigma_{2,2}^2 + 2\beta_{2,2}\mathsf{P})}}{27\sqrt{\pi}(\sigma_{2,2}^2 + \beta_{2,2}\mathsf{P})}, \quad (73)$$

with \tilde{K}_1 and \tilde{K}_2 being constants. Notice that $T_{\max,1}$ and $E_{\min,1}$ can be easily calculated from (37a) and (37b) by setting $i = 1$, $n_{1,2} = 0$ and replacing $n_{1,1}$ by $n_{1,1} + \eta n_{1,2}$. Similarly, $T_{\max,2}$ and $E_{\min,2}$ can be calculated from (37a) and (37b) by setting $i = 2$, $n_{2,1} = 0$ and replacing $n_{2,2}$ by $n_{2,2} + (1 - \eta)n_{1,2}$.

Proof: Follows the proof of Theorem 1 by setting

$$n_{1,2} = n_{2,1} = 0, \quad \beta_{1,2} = 0, \quad \beta_{2,1} = 0, \quad (74)$$

and replacing $n_{1,1}$ by $n_{1,1} + \eta n_{1,2}$ and $n_{2,2}$ by $n_{2,2} + (1 - \eta)n_{1,2}$. ■

To compare our scheme with the time-sharing scheme we consider a scenario where the total number of available channel uses is equal to $n = 300$. The first message m_1 arrives at the beginning of the first channel use (i.e., $a_1 = 1$) and is sent over 200 channel uses (i.e., $d_1 = 200$). The second message m_2 arrives at the time $a_2 \in [1 : 200]$ and has to be decoded at the end of 300 channel uses (i.e., $d_2 = 300$). Fig. 5 illustrates a schematic representation of this example under our scheme (Fig. 5.a) and the time sharing scheme with $\eta = 0.5$ (Fig. 5.b).

In Fig. 6, we plot the transmission rates R_1 and R_2 of our scheme and the time-sharing scheme for this example. Each point corresponds to a different value of $n_{1,2}$ shown in the example starting from $n_{1,2} = 200$ to $n_{1,2} = 0$ with a step size of 50. At $n_{1,2} = 0$, our scheme coincides with the time-sharing scheme. In this figure, $\mathsf{P} = 5$, $\rho = 0.9$, $\sigma_{1,1}^2 = \sigma_{1,2}^2 = \sigma_{2,1}^2 = \sigma_{2,2}^2 = 1$, $K - 1 = K_2 = \tilde{K}_1 = \tilde{K}_2 = 0.5$ and the values of the parameters $\beta_{1,1}, \beta_c, \beta_{2,2}, \beta_{1,2}, \beta_{2,1}$ are optimized to obtain the maximum sum transmission rates. In Fig. 6a, in our scheme $R_1 = \log M_1/(n_{1,1} + n_{1,2})$ and $R_2 = \log M_2/(n_{1,1} + n_{2,2})$ and in the time sharing scheme $R_1 = \log M_1/(n_{1,1} + \eta n_{1,2})$ and $R_2 = \log M_2/(n_{1,1} + (1 - \eta)n_{1,2})$. In Fig. 6b, in both schemes $R_1 = \log M_1/n$ and $R_2 = \log M_2/n$. As can be seen from this figure, our scheme significantly outperforms the time-sharing scheme.

An important difference between our scheme and the time-sharing scheme appears when the channels to one receiver are stronger than the channels to another receiver. Under our scheme, it is possible to adjust the design parameters L_1 and L_2 such that the rate to the weaker receiver increases while keeping the sum-rate approximately constant. In Theorem 1, it is required that parameters L_1 and L_2 be chosen such the condition in (35) is satisfied. It is clear that, large values of L_1 and L_2 reduce the transmission rates R_1 and R_2 and vice versa. In the analysis provided in the previous Section V-B, L_1 and L_2

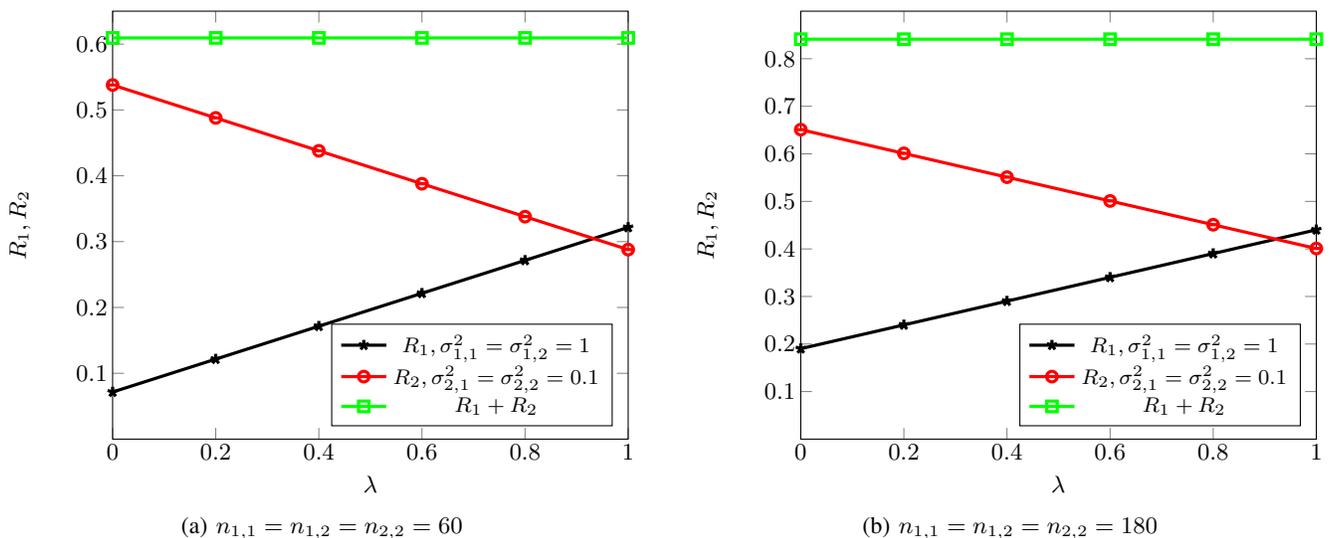


Fig. 7: The figure illustrates the effect of changing $\lambda = \frac{1}{n_{1,2}} \log \left(\frac{L_2}{L_1} \right)$ on the achievable rates R_1 and R_2 in a scenario where $\sigma_{1,1}^2 = \sigma_{1,2}^2 = 1$, and $\sigma_{2,1}^2 = \sigma_{2,2}^2 = 0.1$ and (a) $n_{1,1} = n_{1,2} = n_{2,2} = 60$, (b) $n_{1,1} = n_{1,2} = n_{2,2} = 180$. In this figure, $P = 1$, $\epsilon_1 = \epsilon_2 = 10^{-5}$, $\epsilon_{12} = 10^{-7}$ and $\rho = 0.8$.

are set to be equal. In this section, we focus on the effect of changing the values of L_1 and L_2 on the achievable rates R_1 and R_2 . To this end, we introduce a parameter λ as

$$\lambda := \frac{1}{n_{1,2}} \log \left(\frac{L_2}{L_1} \right). \quad (75)$$

By (35),

$$\log L_1 \geq \frac{1}{2} \log \frac{\log(\epsilon_{1,2})}{\log(1 - (1 - \rho^2)^{n_{1,2}-1})} - \frac{n_{1,2}\lambda}{2}, \quad (76)$$

and

$$\log L_2 \geq \frac{1}{2} \log \frac{\log(\epsilon_{1,2})}{\log(1 - (1 - \rho^2)^{n_{1,2}-1})} + \frac{n_{1,2}\lambda}{2}. \quad (77)$$

Therefore, increasing λ , increases L_2 (decreases L_1) which results in decreasing R_2 (increasing R_1).

To numerically evaluate the effect of λ , we consider a case where the channels of the first receiver are weaker than the channels of the second receiver. In Fig. 7, we assume that $\sigma_{1,1}^2 = \sigma_{1,2}^2 = 1$ and $\sigma_{2,1}^2 = \sigma_{2,2}^2 = 0.1$. In Fig. 7a, we set $n_{1,1} = n_{1,2} = n_{2,2} = 60$ and in Fig. 7b we set $n_{1,1} = n_{1,2} = n_{2,2} = 180$. In this figure, $P = 1$, $\epsilon_1 = \epsilon_2 = 10^{-5}$, $\epsilon_{12} = 10^{-7}$, $K_1 = K_2 = 0.5$, $\rho = 0.8$ and the values of power coefficients $\beta_{1,1}, \beta_c, \beta_{2,2}, \beta_{1,2}, \beta_{2,1}$ are set such that the sum-rate R_1 and R_2 is maximized. We then increase λ from 0 (i.e., $L_1 = L_2$) to 1 with step sizes of 0.2. As can be seen from this figure, by increasing L_2 and consequently decreasing L_1 , it is possible to increase the rate of R_1 while keeping the sum-rate $R_1 + R_2$ constant. For example, under our scheme, it is possible to achieve $R_1 = R_2 = 0.31$ by setting λ at approximately 0.95.

VI. CONCLUSIONS

We have considered a broadcast setting in which a transmitter sends two different messages to two receivers. Messages are considered to have different arrival times and decoding deadlines such that their transmission windows overlap. For this setting, we have proposed a coding scheme that exploits Marton's coding strategy. We have derived rigorous bounds on the achievable rate regions for broadcast setting and point-to-point settings with one or multiple parallel channels. In the point-to-point setup with one channel and more parallel channels, the proposed achievability scheme outperforms the Normal Approximation especially when the number of channel uses is smaller than 200. In the broadcast setup, our scheme agreed with Marton's strategy for sufficiently large numbers of channel uses. Our numerical analysis have shown significant performance improvements over standard approaches based on time sharing for transmission of short packets.

APPENDIX A PROOF OF THEOREM 1

In the following subsections we analyze the probability that events $\mathcal{E}_{1,2}$ and $\mathcal{E}_1 | \mathcal{E}_{1,2}^c$ occur. The analysis related to $\mathbb{P}[\mathcal{E}_2 | \mathcal{E}_{1,2}^c]$ is similar to that of $\mathbb{P}[\mathcal{E}_1 | \mathcal{E}_{1,2}^c]$.

A. Analyzing $\mathbb{P}[\mathcal{E}_{1,2}]$

Recall the definition of the error event $\mathcal{E}_{1,2}$ from (24). Let

$$\cos(\theta) = \frac{\langle \mathbf{X}_{1,2}(m_1, \ell_1), \mathbf{X}_{2,1}(m_2, \ell_2) \rangle}{n_{1,2} \sqrt{\beta_{1,2} \beta_{2,1} \mathbf{P}}}. \quad (78)$$

Then,

$$\mathbb{P}[\mathcal{E}_{1,2}] = \mathbb{P}[\langle \mathbf{X}_{1,2}, \mathbf{X}_{2,1} \rangle \notin \mathcal{D}] \quad (79)$$

$$= \mathbb{P} \left[\langle \mathbf{X}_{1,2}, \mathbf{X}_{2,1} \rangle \notin \left[n_{1,2} \sqrt{\beta_{1,2} \beta_{2,1} \mathbf{P}} \rho : n_{1,2} \sqrt{\beta_{1,2} \beta_{2,1} \mathbf{P}} \right] \right] \quad (80)$$

$$= \prod_{\ell_1=1}^{L_1} \prod_{\ell_2=1}^{L_2} 1 - \mathbb{P} \left[\langle \mathbf{X}_{1,2}(m_1, \ell_1), \mathbf{X}_{2,1}(m_2, \ell_2) \rangle \in \left[n_{1,2} \sqrt{\beta_{1,2} \beta_{2,1} \mathbf{P}} \rho : n_{1,2} \sqrt{\beta_{1,2} \beta_{2,1} \mathbf{P}} \right] \right] \quad (81)$$

$$= \left(1 - \mathbb{P} \left[\langle \mathbf{X}_{1,2}(m_1, 1), \mathbf{X}_{2,1}(m_2, 1) \rangle \in \left[n_{1,2} \sqrt{\beta_{1,2} \beta_{2,1} \mathbf{P}} \rho : n_{1,2} \sqrt{\beta_{1,2} \beta_{2,1} \mathbf{P}} \right] \right] \right)^{L_1 L_2} \quad (82)$$

$$= \left(1 - \mathbb{P} \left[\rho \leq \cos(\theta) \leq 1 \right] \right)^{L_1 L_2} \quad (83)$$

$$= (1 + \mathbb{P}[\cos(\theta) \leq \rho] - \mathbb{P}[\cos(\theta) \leq 1])^{L_1 L_2} \quad (84)$$

$$= (1 + F_{\cos(\theta)}(\rho) - F_{\cos(\theta)}(1))^{L_1 L_2} \quad (85)$$

$$= (F_{\cos^2(\theta)}((\rho)^2))^{L_1 L_2} \quad (86)$$

$$= (I_{(\rho)^2}(1, n_{1,2} - 1))^{L_1 L_2} \quad (87)$$

$$= \left(1 - (1 - \rho^2)^{n_{1,2} - 1} \right)^{L_1 L_2}. \quad (88)$$

Note that in (85), $\cos^2(\theta) \sim \text{Beta}(1, n_{1,2} - 1)$. As such, it follows that

$$F_{\cos^2(\theta)}(x) = I_x(1, n_{1,2} - 1) = 1 - (1 - x)^{n_{1,2} - 1}, \quad (89)$$

where $I_x(\cdot, \cdot)$ is regularized incomplete beta function. It then follows that

$$\mathbb{P}[\mathcal{E}_{1,2}] = \left(1 - (1 - \rho^2)^{n_{1,2} - 1} \right)^{L_1 L_2}. \quad (90)$$

In order to upper bound this error probability by a given threshold $\epsilon_{1,2}$, L_1 and L_2 should be chosen such that

$$L_1 \cdot L_2 \geq \frac{\log(\epsilon_{1,2})}{\log(1 - (1 - \rho^2)^{n_{1,2} - 1})}. \quad (91)$$

This proves the bound in (35).

B. Analyzing $\mathbb{P}[\mathcal{E}_1 | \mathcal{E}_{1,2}^c]$

To analyze $\mathbb{P}[\mathcal{E}_1 | \mathcal{E}_{1,2}^c]$, we use the threshold-based metric bound [9]. Let $\gamma_1 \in \mathbb{R}$, since the first decoder selects among $M_1 L_1$ codewords, thus

$$\mathbb{P}[\mathcal{E}_1 | \mathcal{E}_{1,2}^c] \leq \mathbb{P}[i_1(\mathbf{X}_{1,1}, \mathbf{X}_{1,2}; \mathbf{Y}_{1,1}, \mathbf{Y}_{2,1}) \leq \gamma_1] + M_1 L_1 \mathbb{P}[i_1(\bar{\mathbf{X}}_{1,1}, \bar{\mathbf{X}}_{1,2}; \mathbf{Y}_{1,1}, \mathbf{Y}_{2,1}) \geq \gamma_1] \quad (92)$$

where $\bar{\mathbf{X}}_{1,1} \sim f_{\mathbf{X}_{1,1}}(\mathbf{x}_{1,1})$ and $\bar{\mathbf{X}}_{1,2} \sim f_{\mathbf{X}_{1,2}}(\mathbf{x}_{1,2})$ and are independent of $(\mathbf{X}_{1,1}, \mathbf{X}_{1,2}; \mathbf{Y}_{1,1}, \mathbf{Y}_{2,1})$. Throughout our analysis, we interpret $\mathbb{P}[i_1(\mathbf{X}_{1,1}, \mathbf{X}_{1,2}; \mathbf{Y}_{1,1}, \mathbf{Y}_{2,1}) \leq \gamma_1]$ as the *missed-detection* probability and $\mathbb{P}[i_1(\bar{\mathbf{X}}_{1,1}, \bar{\mathbf{X}}_{1,2}; \mathbf{Y}_{1,1}, \mathbf{Y}_{2,1}) \geq \gamma_1]$ as the *false alarm* probability.

1) *Analyzing the missed-detection probability*: To calculate the missed-detection probability, i.e., $\mathbb{P}[i_1(\mathbf{X}_{1,1}, \mathbf{X}_{1,2}; \mathbf{Y}_{1,1}, \mathbf{Y}_{2,1}) \leq \gamma_1]$, first we define the following Gaussian distributions:

$$Q_{\mathbf{Y}_{1,j}}(\mathbf{y}_{1,j}) \sim \mathcal{N}(\mathbf{y}_{1,j} : \mathbf{0}, \sigma_{y_{1,j}}^2 \mathbf{I}_{n_{1,1}}), \quad \text{for } j = 1, 2 \quad (93)$$

$$Q_{\mathbf{Y}_{2,1} | \mathbf{X}_{1,2}}(\mathbf{y}_{1,2} | \mathbf{x}_{1,2}) \sim \mathcal{N}(\mathbf{Y}_{2,1} : \mu_{1,2}, \sigma_{y_{1,2} | x_{1,2}}^2 \mathbf{I}_{n_{1,2}}), \quad (94)$$

where

$$\sigma_{y_{1,1}}^2 = \beta_{1,1} \mathbf{P} + \sigma_{1,1}^2, \quad (95a)$$

$$\sigma_{y_{1,2}}^2 = \beta_c \mathbf{P} + \sigma_{1,2}^2, \quad (95b)$$

$$\sigma_{1,2}^2 \leq \sigma_{y_{1,2} | x_{1,2}}^2 \leq (1 - \rho^2) \beta_{2,1} \mathbf{P} + \sigma_{1,2}^2, \quad (95c)$$

$$\sqrt{\frac{\beta_c}{\tilde{\beta}_c}} \left(1 + \rho \sqrt{\frac{\beta_{2,1}}{\beta_{1,2}}} \right) \mathbf{X}_{1,2} \leq \mu_{1,2} \leq \left(1 + \sqrt{\frac{\beta_{2,1}}{\beta_{1,2}}} \right) \mathbf{X}_{1,2}. \quad (95d)$$

with $\tilde{\beta}_c := \beta_{1,2} + \beta_{2,1} + \sqrt{\beta_{1,2}\beta_{2,1}}$. We then introduce

$$\tilde{i}_1(\mathbf{X}_{1,1}, \mathbf{X}_{1,2}; \mathbf{Y}_{1,1}, \mathbf{Y}_{2,1}) \triangleq \log \frac{f_{\mathbf{Y}_{1,1}|\mathbf{X}_{1,1}}(\mathbf{y}_{1,1}|\mathbf{x}_{1,1})Q_{\mathbf{Y}_{2,1}|\mathbf{X}_{1,2}}(\mathbf{y}_{1,2}|\mathbf{x}_{1,2})}{\prod_{j=1}^2 Q_{\mathbf{Y}_{1,j}}(\mathbf{y}_{1,j})}. \quad (96)$$

Lemma 1: The following inequality holds:

$$i_1(\mathbf{X}_{1,1}, \mathbf{X}_{1,2}; \mathbf{Y}_{1,1}, \mathbf{Y}_{2,1}) \geq \tilde{i}_1(\mathbf{X}_{1,1}, \mathbf{X}_{1,2}; \mathbf{Y}_{1,1}, \mathbf{Y}_{2,1}) + \log J_1, \quad (97)$$

where J_1 is defined in (33c).

Proof: See Appendix D. ■

Thus

$$\mathbb{P}[i_1(\mathbf{X}_{1,1}, \mathbf{X}_{1,2}; \mathbf{Y}_{1,1}, \mathbf{Y}_{2,1}) \leq \gamma_1] \quad (98)$$

$$\leq \mathbb{P}[\tilde{i}_1(\mathbf{X}_{1,1}, \mathbf{X}_{1,2}; \mathbf{Y}_{1,1}, \mathbf{Y}_{2,1}) \leq \gamma_1 - \log J_1]. \quad (99)$$

Denote by E_1 , V_1 and T_1 the first, second and third central moments of $\tilde{i}_1(\mathbf{X}_{1,1}, \mathbf{X}_{1,2}; \mathbf{Y}_{1,1}, \mathbf{Y}_{2,1})$. By the Berry-Esseen theorem, we have

$$\begin{aligned} & \mathbb{P}[i_1(\mathbf{X}_{1,1}, \mathbf{X}_{1,2}; \mathbf{Y}_{1,1}, \mathbf{Y}_{2,1}) \leq \gamma_1] \\ & \leq \mathbb{P}[\tilde{i}_1(\mathbf{X}_{1,1}, \mathbf{X}_{1,2}; \mathbf{Y}_{1,1}, \mathbf{Y}_{2,1}) \leq \gamma_1 - \log J_1] \\ & \leq \mathbb{Q}(-K) + \frac{6T_1}{\sqrt{V_1^3}}, \end{aligned} \quad (100)$$

where $K > 0$ is chosen such that

$$\gamma_1 = K\sqrt{V_1} + E_1 + \log J_1. \quad (101)$$

We now seek lower bounds on the first and second central moments E_1 and V_1 and an upper bound on the third central moment T_1 since

$$K = \frac{1}{\sqrt{V_1}} (\gamma_1 - \log J_1 - E_1), \quad (102)$$

and

$$\mathbb{P}[\tilde{i}_1(\mathbf{X}_{1,1}, \mathbf{X}_{1,2}; \mathbf{Y}_{1,1}, \mathbf{Y}_{2,1}) - E_1 \leq K\sqrt{V_1}] - \mathbb{Q}(-K) \leq \frac{6T_1}{\sqrt{V_1^3}}. \quad (103)$$

Lemma 2: The following inequalities hold:

$$E_1 \geq E_{\min,1}, \quad (104a)$$

$$V_1 \geq n_{1,1}V(\Omega_{1,1}) + n_{1,2}V(\Omega_{1,2}), \quad (104b)$$

$$T_1 \leq T_{\max,1}, \quad (104c)$$

where $V(x) := \frac{x(2+x)}{2(1+x)^2} \log e^2$ and $T_{\max,1}$, $E_{\min,1}$ are defined in (37a) and (37b), respectively.

Proof: See Appendix C. ■

Set

$$\gamma_1 := \log M_1 + \log L_1 - K_1 \log(n_{1,1} + n_{1,2}) - n_{1,1}C(\Omega_{1,1}) - n_{1,2}C(\Omega_{1,2}) + E_1 + \log J_1 \quad (105)$$

for some K_1 and where $C(x) := \frac{1}{2} \log(1+x)$. Thus

$$K = \frac{\log M_1 + \log L_1 - K_1 \log(n_{1,1} + n_{1,2}) - n_{1,1}C(\Omega_{1,1}) - n_{1,2}C(\Omega_{1,2})}{\sqrt{V_1}}. \quad (106)$$

We then have

$$\begin{aligned} & \mathbb{P}[\tilde{i}_1(\mathbf{X}_{1,1}, \mathbf{X}_{1,2}; \mathbf{Y}_{1,1}, \mathbf{Y}_{2,1}) \leq \gamma_1 - \log J_1] \\ & = \mathbb{P}[\tilde{i}_1(\mathbf{X}_{1,1}, \mathbf{X}_{1,2}; \mathbf{Y}_{1,1}, \mathbf{Y}_{2,1}) - E_1 \leq \log M_1 + \log L_1 - K_1 \log(n_{1,1} + n_{1,2}) - n_{1,1}C(\Omega_{1,1}) - n_{1,2}C(\Omega_{1,2})] \\ & \leq \frac{6T_1}{\sqrt{V_1^3}} + \mathbb{Q}\left(\frac{-\log M_1 - \log L_1 + K_1 \log(n_{1,1} + n_{1,2}) + n_{1,1}C(\Omega_{1,1}) + n_{1,2}C(\Omega_{1,2})}{\sqrt{V_1}}\right) \end{aligned} \quad (107)$$

$$\begin{aligned}
& \stackrel{(a)}{\leq} \frac{6T_{\max,1}}{\sqrt{(n_{1,1}V(\Omega_{1,1}) + n_{1,2}V(\Omega_{1,2}))^3}} \\
& + \mathbb{Q} \left(\frac{-\log M_1 - \log L_1 + K_1 \log(n_{1,1} + n_{1,2}) + n_{1,1}C(\Omega_{1,1}) + n_{1,2}C(\Omega_{1,2})}{\sqrt{n_{1,1}V(\Omega_{1,1}) + n_{1,2}V(\Omega_{1,2})}} \right)
\end{aligned} \tag{108}$$

where step (a) follows by Lemma 2.

2) *Analyzing the false alarm probability:* To bound the false alarm probability, i.e., $\mathbb{P}[i_1(\bar{\mathbf{X}}_{1,1}, \bar{\mathbf{X}}_{1,2}; \mathbf{Y}_{1,1}, \mathbf{Y}_{2,1}) \geq \gamma_1]$, we first use Bayes' formula to write

$$\begin{aligned}
f_{\mathbf{X}_{1,1}}(\bar{\mathbf{x}}_{1,1})f_{\mathbf{X}_{1,2}}(\bar{\mathbf{x}}_{1,2}) &= \frac{f_{\mathbf{Y}_{1,1}}(\mathbf{y}_{1,1})f_{\mathbf{Y}_{1,2}}(\mathbf{y}_{1,2})f_{\mathbf{X}_{1,1}|\mathbf{Y}_{1,1}}(\bar{\mathbf{x}}_{1,1}|\mathbf{y}_{1,1})f_{\mathbf{X}_{1,2}|\mathbf{Y}_{1,2}}(\bar{\mathbf{x}}_{1,2}|\mathbf{y}_{1,2})}{f_{\mathbf{Y}_{1,1}|\mathbf{X}_{1,1}}(\mathbf{y}_{1,1}|\bar{\mathbf{x}}_{1,1})f_{\mathbf{Y}_{1,2}|\mathbf{X}_{1,2}}(\mathbf{y}_{1,2}|\bar{\mathbf{x}}_{1,2})} \\
&= f_{\mathbf{X}_{1,1}|\mathbf{Y}_{1,1}}(\bar{\mathbf{x}}_{1,1}|\mathbf{y}_{1,1})f_{\mathbf{X}_{1,2}|\mathbf{Y}_{1,2}}(\bar{\mathbf{x}}_{1,2}|\mathbf{y}_{1,2})2^{-i_1(\bar{\mathbf{x}}_{1,1}, \bar{\mathbf{x}}_{1,2}; \mathbf{y}_{1,1}, \mathbf{y}_{1,2})}.
\end{aligned} \tag{109}$$

Fix the sequences $\mathbf{y}_{1,1} \in \mathbb{R}^n$ and $\mathbf{y}_{1,2} \in \mathbb{R}^n$. We then have

$$\begin{aligned}
& \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} \mathbf{1} \{i_1(\bar{\mathbf{x}}_{1,1}, \bar{\mathbf{x}}_{1,2}; \mathbf{y}_{1,1}, \mathbf{y}_{1,2}) > \gamma_1\} f_{\mathbf{X}_{1,1}}(\bar{\mathbf{x}}_{1,1})f_{\mathbf{X}_{1,2}}(\bar{\mathbf{x}}_{1,2})d\bar{\mathbf{x}}_{1,1}d\bar{\mathbf{x}}_{1,2} \\
&= \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} 2^{-i_1(\bar{\mathbf{x}}_{1,1}, \bar{\mathbf{x}}_{1,2}; \mathbf{y}_{1,1}, \mathbf{y}_{1,2})} \mathbf{1} \left\{ \prod_{j=1}^2 \frac{f_{\mathbf{Y}_{1,j}|\mathbf{X}_{1,j}}(\mathbf{y}_{1,j}|\bar{\mathbf{x}}_{1,j})}{f_{\mathbf{Y}_{1,j}}(\mathbf{y}_{1,j})} > 2^{\gamma_1} \right\} \\
&\quad \cdot f_{\mathbf{X}_{1,1}|\mathbf{Y}_{1,1}}(\bar{\mathbf{x}}_{1,1}|\mathbf{y}_{1,1})f_{\mathbf{X}_{1,2}|\mathbf{Y}_{1,2}}(\bar{\mathbf{x}}_{1,2}|\mathbf{y}_{1,2})d\bar{\mathbf{x}}_{1,1}d\bar{\mathbf{x}}_{1,2} \\
&\leq \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} 2^{-i_1(\bar{\mathbf{x}}_{1,1}, \bar{\mathbf{x}}_{1,2}; \mathbf{y}_{1,1}, \mathbf{y}_{1,2})} \prod_{j=1}^2 \frac{f_{\mathbf{Y}_{1,j}|\mathbf{X}_{1,j}}(\mathbf{y}_{1,j}|\bar{\mathbf{x}}_{1,j})}{f_{\mathbf{Y}_{1,j}}(\mathbf{y}_{1,j})} 2^{-\gamma_1} \\
&\quad \cdot f_{\mathbf{X}_{1,1}|\mathbf{Y}_{1,1}}(\bar{\mathbf{x}}_{1,1}|\mathbf{y}_{1,1})f_{\mathbf{X}_{1,2}|\mathbf{Y}_{1,2}}(\bar{\mathbf{x}}_{1,2}|\mathbf{y}_{1,2})d\bar{\mathbf{x}}_{1,1}d\bar{\mathbf{x}}_{1,2} \\
&= \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} f_{\mathbf{X}_{1,1}|\mathbf{Y}_{1,1}}(\bar{\mathbf{x}}_{1,1}|\mathbf{y}_{1,1})f_{\mathbf{X}_{1,2}|\mathbf{Y}_{1,2}}(\bar{\mathbf{x}}_{1,2}|\mathbf{y}_{1,2})2^{-\gamma_1} d\bar{\mathbf{x}}_{1,1}d\bar{\mathbf{x}}_{1,2} \\
&= 2^{-\gamma_1}.
\end{aligned} \tag{110}$$

As a consequence,

$$\mathbb{P}[i_1(\bar{\mathbf{X}}_{1,1}, \bar{\mathbf{X}}_{1,2}; \mathbf{Y}_{1,1}, \mathbf{Y}_{2,1}) \geq \gamma_1] \leq 2^{-\gamma_1}. \tag{111}$$

By combining (108) and (111), we can bound the error probability in (92) by

$$\begin{aligned}
\mathbb{P}[\mathcal{E}_1 | \mathcal{E}_{1,2}^c] &\leq \frac{6T_{\max,1}}{\sqrt{(n_{1,1}V(\Omega_{1,1}) + n_{1,2}V(\Omega_{1,2}))^3}} \\
&+ \mathbb{Q} \left(\frac{-\log M_1 - \log L_1 + K_1 \log(n_{1,1} + n_{1,2}) + n_{1,1}C(\Omega_{1,1}) + n_{1,2}C(\Omega_{1,2})}{\sqrt{n_{1,1}V(\Omega_{1,1}) + n_{1,2}V(\Omega_{1,2})}} \right) \\
&+ \frac{(n_{1,1} + n_{1,2})^{K_1} (1 + \Omega_{1,1})^{\frac{n_{1,1}}{2}} (1 + \Omega_{1,2})^{\frac{n_{1,2}}{2}}}{J_1 2^{E_{\min,1}}}.
\end{aligned} \tag{112}$$

As a result

$$\begin{aligned}
\epsilon_1 &\leq \frac{6T_{\max,1}}{\sqrt{(n_{1,1}V(\Omega_{1,1}) + n_{1,2}V(\Omega_{1,2}))^3}} \\
&+ \mathbb{Q} \left(\frac{-\log M_1 - \log L_1 + K_1 \log(n_{1,1} + n_{1,2}) + n_{1,1}C(\Omega_{1,1}) + n_{1,2}C(\Omega_{1,2})}{\sqrt{n_{1,1}V(\Omega_{1,1}) + n_{1,2}V(\Omega_{1,2})}} \right) \\
&+ \frac{(n_{1,1} + n_{1,2})^{K_1} (1 + \Omega_{1,1})^{\frac{n_{1,1}}{2}} (1 + \Omega_{1,2})^{\frac{n_{1,2}}{2}}}{J_1 2^{E_{\min,1}}} + \left(1 - (1 - \rho^2)^{n_{1,2}-1}\right)^{L_1 \cdot L_2}.
\end{aligned} \tag{113}$$

Equivalently,

$$\begin{aligned}
\log M_1 + \log L_1 &\leq n_{1,1}C(\Omega_{1,1}) + n_{1,2}C(\Omega_{1,2}) + K_1 \log(n_{1,1} + n_{1,2}) \\
&\quad - \sqrt{n_{1,1}V(\Omega_{1,1}) + n_{1,2}V(\Omega_{1,2})} \mathbb{Q}^{-1}(\epsilon_1 - \Delta_1)
\end{aligned} \tag{114}$$

where Δ_1 is defined in (36). This concludes the proof of Theorem 1.

APPENDIX B
PROOF OF PROPOSITION 1

For large values of $n_{1,1}$ and $n_{1,2}$, choose K_1 such that

$$\lim_{n_{1,1}, n_{1,2} \rightarrow \infty} \frac{(n_{1,1} + n_{1,2})^{K_1} (1 + \Omega_{1,1})^{\frac{n_{1,1}}{2}} (1 + \Omega_{1,2})^{\frac{n_{1,2}}{2}}}{J_1 2^{E_{\min,1}}} \rightarrow 0. \quad (115)$$

Hence,

$$\Delta_1 = O\left(\frac{1}{\sqrt{(n_{1,1}V(\Omega_{1,1}) + n_{1,2}V(\Omega_{1,2}))^3}}\right) \quad (116)$$

as $n_{1,1}, n_{1,2} \rightarrow \infty$. This proves the bound in (38). To prove the bound in (39), define

$$A := 1 - (1 - \rho^2)^{n_{1,2} - 1}. \quad (117)$$

We then use the approximation $\log(1 - x) \approx -x$ which holds when x is very small. Hence, for large values of $n_{1,2}$

$$\log L_1 + \log L_2 \geq \log(-\log \epsilon_{1,2}) - n_{1,2} \log(1 - \rho^2) \quad (118)$$

which concludes the proof.

APPENDIX C
PROOF OF LEMMA 2

A. *Bounding the first moment*

We start by lower bounding the first moment. To this end, we have

$$\mathbb{E}[\tilde{I}_1(\mathbf{X}_{1,1}, \mathbf{X}_{1,2}; \mathbf{Y}_{1,1}, \mathbf{Y}_{2,1})] \quad (119)$$

$$\stackrel{(i)}{=} \mathbb{E}\left[\frac{n_{1,1}}{2} \log \frac{\sigma_{y_{1,1}}^2}{\sigma_{1,1}^2} + \frac{n_{1,2}}{2} \log \frac{\sigma_{y_{1,2}}^2}{\sigma_{y_{1,2}|x_{1,2}}^2} - \frac{\|\mathbf{Z}_{1,1}\|^2}{2\sigma_{1,1}^2} + \frac{\|\mathbf{X}_{1,1} + \mathbf{Z}_{1,1}\|^2}{2\sigma_{y_{1,1}}^2} - \frac{\|\alpha(\mathbf{X}_{1,2} + \mathbf{X}_{2,1}) + \mathbf{Z}_{1,2} - \mu_{1,2}\|^2}{2\sigma_{y_{1,2}|x_{1,2}}^2} + \frac{\|\alpha(\mathbf{X}_{1,2} + \mathbf{X}_{2,1}) + \mathbf{Z}_{1,2}\|^2}{2\sigma_{y_{1,2}}^2}\right] \quad (120)$$

$$\stackrel{(ii)}{\geq} \frac{n_{1,1}}{2} \log \frac{\sigma_{y_{1,1}}^2}{\sigma_{1,1}^2} + \frac{n_{1,2}}{2} \log \frac{\sigma_{y_{1,2}}^2}{((1 - \rho^2)\beta_{2,1}\mathbf{P} + \sigma_{1,2}^2)} - \frac{\mathbb{E}[\|\mathbf{Z}_{1,1}\|^2]}{2\sigma_{1,1}^2} + \frac{\|\mathbf{X}_{1,1}\|^2 + \mathbb{E}[\|\mathbf{Z}_{1,1}\|^2] + 2\mathbb{E}\langle \mathbf{Z}_{1,1}, \mathbf{X}_{1,1} \rangle}{2\sigma_{y_{1,1}}^2} - \frac{\|\alpha(\mathbf{X}_{1,2} + \mathbf{X}_{2,1}) + \mathbf{Z}_{1,2} - \sqrt{\frac{\beta_c}{\beta_c}}(1 + \rho\sqrt{\frac{\beta_{2,1}}{\beta_{1,2}}})\mathbf{X}_{1,2}\|^2}{2\sigma_{1,2}^2} + \frac{\|\alpha(\mathbf{X}_{1,2} + \mathbf{X}_{2,1}) + \mathbf{Z}_{1,2}\|^2}{2\sigma_{y_{1,2}}^2} \quad (121)$$

$$= \frac{n_{1,1}}{2} \log \frac{\sigma_{y_{1,1}}^2}{\sigma_{1,1}^2} + \frac{n_{1,2}}{2} \log \frac{\sigma_{y_{1,2}}^2}{((1 - \rho^2)\beta_{2,1}\mathbf{P} + \sigma_{1,2}^2)} + \frac{1}{2} \left(\frac{1}{\sigma_{y_{1,1}}^2} - \frac{1}{\sigma_{1,1}^2} \right) \mathbb{E}[\|\mathbf{Z}_{1,1}\|^2] + \frac{\|\mathbf{X}_{1,1}\|^2}{2\sigma_{y_{1,1}}^2} + \frac{\mathbb{E}\langle \mathbf{Z}_{1,1}, \mathbf{X}_{1,1} \rangle}{\sigma_{y_{1,1}}^2} + \frac{1}{2} \left(\frac{\alpha^2}{\sigma_{y_{1,1}}^2} - \frac{\left(\alpha\sqrt{\tilde{\beta}_c} - \sqrt{\beta_c}(1 + \rho\sqrt{\frac{\beta_{2,1}}{\beta_{1,2}}}) \right)^2}{\tilde{\beta}_c\sigma_{1,2}^2} \right) \|\mathbf{X}_{1,2}\|^2 + \left(\frac{\alpha}{\sigma_{y_{1,1}}^2} - \frac{\alpha\sqrt{\tilde{\beta}_c} - \sqrt{\beta_c}(1 + \rho\sqrt{\frac{\beta_{2,1}}{\beta_{1,2}}})}{\sqrt{\tilde{\beta}_c}\sigma_{1,2}^2} \right) (\mathbb{E}\langle \mathbf{X}_{1,2}, \mathbf{X}_{2,1} \rangle + \mathbb{E}\langle \mathbf{X}_{1,2}, \mathbf{Z}_{1,2} \rangle) + \frac{1}{2} \left(\frac{1}{\sigma_{y_{1,2}}^2} - \frac{1}{\sigma_{1,2}^2} \right) (\alpha^2\|\mathbf{X}_{2,1}\|^2 + \mathbb{E}[\|\mathbf{Z}_{1,2}\|^2] + 2\mathbb{E}\langle \mathbf{X}_{2,1}, \mathbf{Z}_{1,2} \rangle) \quad (122)$$

$$\stackrel{(iii)}{\geq} \frac{n_{1,1}}{2} \log \frac{\sigma_{y_{1,1}}^2}{\sigma_{1,1}^2} + \frac{n_{1,2}}{2} \log \frac{\sigma_{y_{1,2}}^2}{((1 - \rho^2)\beta_{2,1}\mathbf{P} + \sigma_{1,2}^2)} + \frac{\sigma_{1,1}^4 n_{1,1}}{2} \left(\frac{1}{\sigma_{y_{1,1}}^2} - \frac{1}{\sigma_{1,1}^2} \right) + \frac{n_{1,1}\beta_{1,1}\mathbf{P}}{2\sigma_{y_{1,1}}^2} - \frac{\sigma_{1,1}^2 \sqrt{2n_{1,1}\beta_{1,1}\mathbf{P}} \Gamma\left(\frac{n_{1,1}+1}{2}\right)}{\sigma_{y_{1,1}}^2 \Gamma\left(\frac{n_{1,1}}{2}\right)} + \frac{1}{2} \left(\frac{\beta_c}{\tilde{\beta}_c\sigma_{y_{1,1}}^2} - \frac{\left(\sqrt{\tilde{\beta}_c} - \sqrt{\beta_c}(1 + \rho\sqrt{\frac{\beta_{2,1}}{\beta_{1,2}}}) \right)^2}{\tilde{\beta}_c\sigma_{1,2}^2} \right) n_{1,2}\beta_{1,2}\mathbf{P}$$

$$\begin{aligned}
& + \left(\frac{\sqrt{\beta_c}}{\sqrt{\tilde{\beta}_c \sigma_{y_{1,1}}^2}} - \frac{\sqrt{\tilde{\beta}_c} - \sqrt{\beta_c}(1 + \rho\sqrt{\frac{\beta_{2,1}}{\beta_{1,2}}})}{\sqrt{\tilde{\beta}_c \sigma_{1,2}^2}} \right) \left(n_{1,2} \sqrt{\beta_{1,2} \beta_{2,1}} \rho \mathbf{P} - \sigma_{1,2}^2 \sqrt{2n_{1,2} \beta_{1,2} \mathbf{P}} \frac{\Gamma\left(\frac{n_{1,2}+1}{2}\right)}{\Gamma\left(\frac{n_{1,2}}{2}\right)} \right) \\
& + \frac{1}{2} \left(\frac{1}{\sigma_{y_{1,2}}^2} - \frac{1}{\sigma_{1,2}^2} \right) \left(n_{1,2} \beta_{2,1} \mathbf{P} + n_{1,2} \sigma_{1,2}^4 - 2\sqrt{2} \sigma_{1,2}^2 \sqrt{n_{1,2} \beta_{2,1} \mathbf{P}} \frac{\Gamma\left(\frac{n_{1,2}+1}{2}\right)}{\Gamma\left(\frac{n_{1,2}}{2}\right)} \right) \tag{123} \\
& = \frac{n_{1,1}}{2} \log(1 + \Omega_{1,1}) + \frac{n_{1,2}}{2} \log(1 + \Omega_{1,2}) + \frac{(1 - \sigma_{1,1}^2) n_{1,1} \Omega_{1,1}}{2(1 + \Omega_{1,1})} - \Gamma_{1,1} \cdot \frac{\sqrt{2n_{1,1} \sigma_{1,1}^2 \Omega_{1,1}}}{1 + \Omega_{1,1}} - \frac{n_{1,2} \sigma_{1,2}^2 \beta_c \mathbf{P}}{2(\sigma_{1,2}^2 + \beta_c \mathbf{P})} \\
& + \sqrt{n_{1,1} \beta_{1,2} \mathbf{P}} \left(\sqrt{\frac{\beta_c}{\tilde{\beta}_c}} \left(\frac{1}{\sigma_{1,2}^2 + \beta_c \mathbf{P}} + 1 + \rho \sqrt{\frac{\beta_{2,1}}{\beta_{1,2}}} \right) - \frac{1}{\sigma_{1,2}^2} \right) \left(\rho \sqrt{n_{1,2} \beta_{2,1} \mathbf{P}} - \sigma_{1,2}^2 \sqrt{2} \Gamma_{1,2} \right) \\
& + \frac{\beta_c n_{1,2} \beta_{1,2} \mathbf{P} \left(\sigma_{1,2}^2 + (\beta_c \mathbf{P} - \sigma_{1,2}^2) \left(\sqrt{\frac{\tilde{\beta}_c}{\beta_c}} - 1 - \rho \sqrt{\frac{\beta_{2,1}}{\beta_{1,2}}} \right)^2 \right)}{2\tilde{\beta}_c \sigma_{1,2}^2 (\beta_c \mathbf{P} + \sigma_{1,2}^2)} - \frac{\beta_c \mathbf{P} \sqrt{n_{1,2} \beta_{2,1} \mathbf{P}}}{2\sigma_{1,2}^2 (\beta_c \mathbf{P} + \sigma_{1,2}^2)} \left(\sqrt{n_{1,2} \beta_{2,1} \mathbf{P}} - 2\sqrt{2} \sigma_{1,2}^2 \Gamma_{1,2} \right) \tag{124}
\end{aligned}$$

where (i) is by (96), (ii) is by (95), (iii) is by (20) and the fact that $\|\mathbf{Z}_{1,1}\|^2/\sigma_{1,1}^4$ and $\|\mathbf{Z}_{1,2}\|^2/\sigma_{1,2}^4$ each follow a central chi-squared distribution with degree of $n_{1,1}$ and $n_{1,2}$, respectively, and $\|\mathbf{Z}_{1,1}\|/\sigma_{1,1}^2$ and $\|\mathbf{Z}_{1,2}\|/\sigma_{1,2}^2$ each follow a central chi-distribution with degree of $n_{1,1}$ and $n_{1,2}$, respectively. The last equality is by (33) and by defining $\Gamma_{1,1} := \Gamma\left(\frac{n_{1,1}+1}{2}\right)/\Gamma\left(\frac{n_{1,1}}{2}\right)$ and $\Gamma_{1,2} := \Gamma\left(\frac{n_{1,2}+1}{2}\right)/\Gamma\left(\frac{n_{1,2}}{2}\right)$.

B. Bounding the second moment

We continue by lower bounding the variance of $\tilde{i}_1(\mathbf{X}_{1,1}, \mathbf{X}_{1,2}; \mathbf{Y}_{1,1}, \mathbf{Y}_{2,1})$. One can write

$$\text{Var}[\tilde{i}_1(\mathbf{X}_{1,1}, \mathbf{X}_{1,2}; \mathbf{Y}_{1,1}, \mathbf{Y}_{2,1})] \tag{125}$$

$$\begin{aligned}
& = \text{Var} \left[\frac{n_{1,1}}{2} \log \frac{\sigma_{y_{1,1}}^2}{\sigma_{1,1}^2} + \frac{n_{1,2}}{2} \log \frac{\sigma_{y_{1,2}}^2}{\sigma_{y_{1,2}|x_{1,2}}^2} - \frac{\|\mathbf{Z}_{1,1}\|^2}{2\sigma_{1,1}^2} + \frac{\|\mathbf{X}_{1,1} + \mathbf{Z}_{1,1}\|^2}{2\sigma_{y_{1,1}}^2} \right. \\
& \quad \left. - \frac{\|\alpha(\mathbf{X}_{1,2} + \mathbf{X}_{2,1}) + \mathbf{Z}_{1,2} - \mu_{1,2}\|^2}{2\sigma_{y_{1,2}|x_{1,2}}^2} + \frac{\|\alpha(\mathbf{X}_{1,2} + \mathbf{X}_{2,1}) + \mathbf{Z}_{1,2}\|^2}{2\sigma_{y_{1,2}}^2} \right] \tag{126}
\end{aligned}$$

$$\begin{aligned}
& \stackrel{(i)}{\geq} \text{Var} \left[\frac{1}{2} \left(\frac{1}{\sigma_{y_{1,1}}^2} - \frac{1}{\sigma_{1,1}^2} \right) \|\mathbf{Z}_{1,1}\|^2 + \frac{1}{2} \left(\frac{1}{\sigma_{y_{1,2}}^2} - \frac{1}{\sigma_{1,2}^2} \right) \|\mathbf{Z}_{1,2}\|^2 + \frac{\langle \mathbf{X}_{1,1}, \mathbf{Z}_{1,1} \rangle}{\sigma_{y_{1,1}}^2} \right. \\
& \quad \left. + \left(\frac{\sqrt{\beta_c}}{\sqrt{\tilde{\beta}_c \sigma_{y_{1,2}}^2}} - \frac{1}{\sigma_{1,2}^2} \right) \langle \mathbf{X}_{2,1}, \mathbf{Z}_{1,2} \rangle + \left(\frac{\sqrt{\beta_c}}{\sqrt{\tilde{\beta}_c \sigma_{y_{1,2}}^2}} - \frac{1 - \sqrt{\frac{\beta_c}{\tilde{\beta}_c}}(1 + \rho\sqrt{\frac{\beta_{2,1}}{\beta_{1,2}}})}{\sigma_{1,2}^2} \right) \langle \mathbf{X}_{1,2}, \mathbf{Z}_{1,2} \rangle \right] \tag{127} \\
& \stackrel{(ii)}{\geq} \frac{\sigma_{1,1}^4 n_{1,1}}{2} \left(\frac{1}{\sigma_{y_{1,1}}^2} - \frac{1}{\sigma_{1,1}^2} \right)^2 + \frac{\sigma_{1,2}^4 n_{1,2}}{2} \left(\frac{1}{\sigma_{y_{1,2}}^2} - \frac{1}{\sigma_{1,2}^2} \right)^2 + \frac{\text{Var}[\langle \mathbf{X}_{1,1}, \mathbf{Z}_{1,1} \rangle]}{(\sigma_{y_{1,1}}^2)^2} \\
& \quad + \left(\frac{\sqrt{\beta_c}}{\sqrt{\tilde{\beta}_c \sigma_{y_{1,2}}^2}} - \frac{1}{\sigma_{1,2}^2} \right)^2 \text{Var}[\langle \mathbf{X}_{2,1}, \mathbf{Z}_{1,2} \rangle] + \left(\frac{\sqrt{\beta_c}}{\sqrt{\tilde{\beta}_c \sigma_{y_{1,2}}^2}} - \frac{1 - \frac{\beta_c}{\tilde{\beta}_c}(1 + \rho\sqrt{\frac{\beta_{2,1}}{\beta_{1,2}}})}{\sigma_{1,2}^2} \right)^2 \text{Var}[\langle \mathbf{X}_{1,2}, \mathbf{Z}_{1,2} \rangle]. \tag{128}
\end{aligned}$$

where (i) is by (95), and (ii) is by the fact that $\text{Var}[\|\mathbf{Z}_{1,1}\|^2] = 2n_{1,1}\sigma_{1,1}^4$ and $\text{Var}[\|\mathbf{Z}_{1,2}\|^2] = 2n_{1,2}\sigma_{1,2}^4$. To bound the remaining variance terms in (128), we use the following Chebyshev inequality.

Theorem 6 (Chebyshev's inequality): Let X be a random variable with expectation μ and variance σ^2 . Then

$$\mathbb{P}[|X - \mu| \geq c] \leq \frac{\sigma^2}{c^2}, \tag{129}$$

for all $c > 0$.

We start by bounding $\text{Var}[\langle \mathbf{X}_{1,1}, \mathbf{Z}_{1,1} \rangle]$. For $c_1 > 0$

$$\text{Var}[\langle \mathbf{X}_{1,1}, \mathbf{Z}_{1,1} \rangle] \geq c_1^2 \mathbb{P}[|\langle \mathbf{X}_{1,1}, \mathbf{Z}_{1,1} \rangle - \mathbb{E}[\langle \mathbf{X}_{1,1}, \mathbf{Z}_{1,1} \rangle]| \geq c_1]. \tag{130}$$

The right-hand side of (141) can be bounded as follows:

$$\mathbb{P}[|\langle \mathbf{X}_{1,1}, \mathbf{Z}_{1,1} \rangle - \mathbb{E}[\langle \mathbf{X}_{1,1}, \mathbf{Z}_{1,1} \rangle]| \geq c_1] \tag{131}$$

$$= \mathbb{P}[\langle \mathbf{X}_{1,1}, \mathbf{Z}_{1,1} \rangle \geq c_1 + \mathbb{E}[\langle \mathbf{X}_{1,1}, \mathbf{Z}_{1,1} \rangle]] + \mathbb{P}[\langle \mathbf{X}_{1,1}, \mathbf{Z}_{1,1} \rangle \leq -c_1 + \mathbb{E}[\langle \mathbf{X}_{1,1}, \mathbf{Z}_{1,1} \rangle]] \quad (132)$$

$$= \mathbb{P}\left[\|\mathbf{Z}_{1,1}\| \cos(\theta) \geq \frac{c_1 + \mathbb{E}[\langle \mathbf{X}_{1,1}, \mathbf{Z}_{1,1} \rangle]}{\|\mathbf{X}_{1,1}\|}\right] + \mathbb{P}\left[\|\mathbf{Z}_{1,1}\| \cos(\theta) \leq \frac{-c_1 + \mathbb{E}[\langle \mathbf{X}_{1,1}, \mathbf{Z}_{1,1} \rangle]}{\|\mathbf{X}_{1,1}\|}\right] \quad (133)$$

$$\geq \mathbb{P}\left[-\|\mathbf{Z}_{1,1}\| \geq \frac{c_1 + \mathbb{E}[\langle \mathbf{X}_{1,1}, \mathbf{Z}_{1,1} \rangle]}{\|\mathbf{X}_{1,1}\|}\right] + \mathbb{P}\left[\|\mathbf{Z}_{1,1}\| \leq \frac{-c_1 + \mathbb{E}[\langle \mathbf{X}_{1,1}, \mathbf{Z}_{1,1} \rangle]}{\|\mathbf{X}_{1,1}\|}\right] \quad (134)$$

$$= \mathbb{P}\left[\|\mathbf{Z}_{1,1}\| \leq -\frac{c_1 + \mathbb{E}[\langle \mathbf{X}_{1,1}, \mathbf{Z}_{1,1} \rangle]}{\|\mathbf{X}_{1,1}\|}\right] + \mathbb{P}\left[\|\mathbf{Z}_{1,1}\| \leq \frac{-c_1 + \mathbb{E}[\langle \mathbf{X}_{1,1}, \mathbf{Z}_{1,1} \rangle]}{\|\mathbf{X}_{1,1}\|}\right] \quad (135)$$

$$\stackrel{(a)}{\geq} \mathbb{P}\left[\|\mathbf{Z}_{1,1}\| \leq -\frac{c_1 - \|\mathbf{X}_{1,1}\| \cdot \mathbb{E}[\|\mathbf{Z}_{1,1}\|]}{\|\mathbf{X}_{1,1}\|}\right] + \mathbb{P}\left[\|\mathbf{Z}_{1,1}\| \leq \frac{-c_1 + \|\mathbf{X}_{1,1}\| \cdot \mathbb{E}[\|\mathbf{Z}_{1,1}\|]}{\|\mathbf{X}_{1,1}\|}\right] \quad (136)$$

$$\stackrel{(b)}{=} 2F_{\|\mathbf{Z}_{1,1}\|/\sigma_{1,1}}\left(-\frac{c_1}{\sigma_{1,1}^2 \sqrt{n_{1,1} \beta_{1,1} \mathbf{P}}} + \frac{\sqrt{2}}{\sigma_{1,1}^2} \frac{\Gamma(\frac{n_{1,1}+1}{2})}{\Gamma(\frac{n_{1,1}}{2})}\right) \quad (137)$$

$$= 2G\left(\frac{n_{1,1}}{2}, \frac{1}{2} \left(-\frac{c_1}{\sigma_{1,1}^2 \sqrt{n_{1,1} \beta_{1,1} \mathbf{P}}} + \frac{\sqrt{2}}{\sigma_{1,1}^2} \frac{\Gamma(\frac{n_{1,1}+1}{2})}{\Gamma(\frac{n_{1,1}}{2})}\right)^2\right) \quad (138)$$

In step (a) we used the following bounds on $\mathbb{E}[\langle \mathbf{X}_{1,1}, \mathbf{Z}_{1,1} \rangle]$:

$$-\|\mathbf{X}_{1,1}\| \cdot \mathbb{E}[\|\mathbf{Z}_{1,1}\|] \leq \mathbb{E}[\langle \mathbf{X}_{1,1}, \mathbf{Z}_{1,1} \rangle] \leq \|\mathbf{X}_{1,1}\| \cdot \mathbb{E}[\|\mathbf{Z}_{1,1}\|] \quad (139)$$

In step (b) we used the fact that $\mathbf{Z}_{1,1}/\sigma_{1,1}$ follows a chi distribution of degree $n_{1,1}$ and $F_{\|\mathbf{Z}_{1,1}\|/\sigma_{1,1}}(\cdot)$ is the CDF of the chi distribution of degree $n_{1,1}$ and is equal to

$$F(x; n) = G\left(\frac{n}{2}, \frac{x^2}{2}\right) \quad (140)$$

where $G(\cdot, \cdot)$ is the regularized gamma function.

Thus

$$\text{Var}[\langle \mathbf{X}_{1,1}, \mathbf{Z}_{1,1} \rangle] \geq 2c_1^2 G\left(\frac{n_{1,1}}{2}, \frac{1}{2} \left(-\frac{c_1}{\sigma_{1,1}^2 \sqrt{n_{1,1} \beta_{1,1} \mathbf{P}}} + \frac{\sqrt{2}}{\sigma_{1,1}^2} \frac{\Gamma(\frac{n_{1,1}+1}{2})}{\Gamma(\frac{n_{1,1}}{2})}\right)^2\right) \quad (141)$$

for all $c_1 > 0$.

Following the same procedure, we prove that

$$\text{Var}[\langle \mathbf{X}_{2,1}, \mathbf{Z}_{1,2} \rangle] \geq 2c_2^2 G\left(\frac{n_{1,2}}{2}, \frac{1}{2} \left(-\frac{c_2}{\sigma_{1,2}^2 \sqrt{n_{1,2} \beta_{2,1} \mathbf{P}}} + \frac{\sqrt{2}}{\sigma_{1,2}^2} \frac{\Gamma(\frac{n_{1,2}+1}{2})}{\Gamma(\frac{n_{1,2}}{2})}\right)^2\right) \quad (142)$$

$$\text{Var}[\langle \mathbf{X}_{1,2}, \mathbf{Z}_{1,2} \rangle] \geq 2c_3^2 G\left(\frac{n_{1,2}}{2}, \frac{1}{2} \left(-\frac{c_3}{\sigma_{1,2}^2 \sqrt{n_{1,2} \beta_{1,2} \mathbf{P}}} + \frac{\sqrt{2}}{\sigma_{1,2}^2} \frac{\Gamma(\frac{n_{1,2}+1}{2})}{\Gamma(\frac{n_{1,2}}{2})}\right)^2\right) \quad (143)$$

for all $c_2 > 0$ and $c_3 > 0$. Thus

$$\text{Var}[\tilde{i}_1(\mathbf{X}_{1,1}, \mathbf{X}_{1,2}; \mathbf{Y}_{1,1}, \mathbf{Y}_{2,1})] \quad (144)$$

$$\begin{aligned} &\geq \frac{\sigma_{1,1}^4 n_{1,1}}{2} \left(\frac{1}{\sigma_{y_{1,1}}^2} - \frac{1}{\sigma_{1,1}^2}\right)^2 + \frac{\sigma_{1,2}^4 n_{1,2}}{2} \left(\frac{1}{\sigma_{y_{1,2}}^2} - \frac{1}{\sigma_{1,2}^2}\right)^2 \\ &\quad + \frac{2c_1^2}{(\sigma_{y_{1,1}}^2)^2} G\left(\frac{n_{1,1}}{2}, \frac{1}{2} \left(-\frac{c_1}{\sigma_{1,1}^2 \sqrt{n_{1,1} \beta_{1,1} \mathbf{P}}} + \frac{\sqrt{2}}{\sigma_{1,1}^2} \frac{\Gamma(\frac{n_{1,1}+1}{2})}{\Gamma(\frac{n_{1,1}}{2})}\right)^2\right) \\ &\quad + 2c_2^2 \left(\frac{\sqrt{\beta_c}}{\sqrt{\tilde{\beta}_c} \sigma_{y_{1,2}}^2} - \frac{1}{\sigma_{1,2}^2}\right)^2 G\left(\frac{n_{1,2}}{2}, \frac{1}{2} \left(-\frac{c_2}{\sigma_{1,2}^2 \sqrt{n_{1,2} \beta_{2,1} \mathbf{P}}} + \frac{\sqrt{2}}{\sigma_{1,2}^2} \frac{\Gamma(\frac{n_{1,2}+1}{2})}{\Gamma(\frac{n_{1,2}}{2})}\right)^2\right) \\ &\quad + 2c_3^2 \left(\frac{\sqrt{\beta_c}}{\sqrt{\tilde{\beta}_c} \sigma_{y_{1,2}}^2} - \frac{1 - \frac{\beta_c}{\tilde{\beta}_c} (1 + \rho \sqrt{\frac{\beta_{2,1}}{\beta_{1,2}}})}{\sigma_{1,2}^2}\right)^2 G\left(\frac{n_{1,2}}{2}, \frac{1}{2} \left(-\frac{c_3}{\sigma_{1,2}^2 \sqrt{n_{1,2} \beta_{1,2} \mathbf{P}}} + \frac{\sqrt{2}}{\sigma_{1,2}^2} \frac{\Gamma(\frac{n_{1,2}+1}{2})}{\Gamma(\frac{n_{1,2}}{2})}\right)^2\right) \\ &= n_{1,1} V(\Omega_{1,1}) + n_{1,2} V(\Omega_{1,2}) - \frac{n_{1,1} \Omega_{1,1}}{(1 + \Omega_{1,1})^2} - \frac{n_{1,2} (2\sigma_{1,2}^2 \beta_c \mathbf{P} - (1 - \rho^2) \beta_{2,1} \mathbf{P} (2\sigma_{1,2}^2 + (1 - \rho^2) \beta_{2,1} \mathbf{P}))}{2(\sigma_{1,2}^2 + \beta_c \mathbf{P})^2} \end{aligned} \quad (145)$$

$$\begin{aligned}
& + \frac{2c_1^2}{(\sigma_{y_{1,1}}^2)^2} G \left(\frac{n_{1,1}}{2}, \frac{1}{2} \left(-\frac{c_1}{\sigma_{1,1}^2 \sqrt{n_{1,1} \beta_{1,1} \mathbf{P}}} + \frac{\sqrt{2}}{\sigma_{1,1}^2} \frac{\Gamma(\frac{n_{1,1}+1}{2})}{\Gamma(\frac{n_{1,1}}{2})} \right)^2 \right) \\
& + 2c_2^2 \left(\frac{\sqrt{\beta_c}}{\sqrt{\tilde{\beta}_c} \sigma_{y_{1,2}}^2} - \frac{1}{\sigma_{1,2}^2} \right)^2 G \left(\frac{n_{1,2}}{2}, \frac{1}{2} \left(-\frac{c_2}{\sigma_{1,2}^2 \sqrt{n_{1,2} \beta_{2,1} \mathbf{P}}} + \frac{\sqrt{2}}{\sigma_{1,2}^2} \frac{\Gamma(\frac{n_{1,2}+1}{2})}{\Gamma(\frac{n_{1,2}}{2})} \right)^2 \right) \\
& + 2c_3^2 \left(\frac{\sqrt{\beta_c}}{\sqrt{\tilde{\beta}_c} \sigma_{y_{1,2}}^2} - \frac{1 - \frac{\beta_c}{\tilde{\beta}_c} (1 + \rho \sqrt{\frac{\beta_{2,1}}{\beta_{1,2}}})}{\sigma_{1,2}^2} \right)^2 G \left(\frac{n_{1,2}}{2}, \frac{1}{2} \left(-\frac{c_3}{\sigma_{1,2}^2 \sqrt{n_{1,2} \beta_{1,2} \mathbf{P}}} + \frac{\sigma_{1,2}^2}{\sqrt{2}} \frac{\Gamma(\frac{n_{1,2}+1}{2})}{\Gamma(\frac{n_{1,2}}{2})} \right)^2 \right) \quad (146)
\end{aligned}$$

for all $c_1 > 0$, $c_2 > 0$ and $c_3 > 0$ with $V(x) := \frac{x(2+x)}{2(1+x)^2} \log e^2$. Given $n_{1,1}, n_{1,2}, \mathbf{P}, \beta_{1,1}, \beta_{1,2}, \beta_{2,1}, \rho$, since $c_1, c_2, c_3 > 0$ are arbitrary, we may choose them such that

$$\begin{aligned}
& \frac{n_{1,1} \Omega_{1,1}}{(1 + \Omega_{1,1})^2} + \frac{n_{1,2} (2\sigma_{1,2}^2 \beta_c \mathbf{P} - (1 - \rho^2) \beta_{2,1} \mathbf{P} (2\sigma_{1,2}^2 + (1 - \rho^2) \beta_{2,1} \mathbf{P}))}{2(\sigma_{1,2}^2 + \beta_c \mathbf{P})^2} \\
& = \frac{2c_1^2}{(\sigma_{y_{1,1}}^2)^2} G \left(\frac{n_{1,1}}{2}, \frac{1}{2} \left(-\frac{c_1}{\sigma_{1,1}^2 \sqrt{n_{1,1} \beta_{1,1} \mathbf{P}}} + \frac{\sqrt{2}}{\sigma_{1,1}^2} \frac{\Gamma(\frac{n_{1,1}+1}{2})}{\Gamma(\frac{n_{1,1}}{2})} \right)^2 \right) \\
& + 2c_2^2 \left(\frac{1}{\sigma_{y_{1,2}}^2} - \frac{1}{\sigma_{1,2}^2} \right)^2 G \left(\frac{n_{1,2}}{2}, \frac{1}{2} \left(-\frac{c_2}{\sigma_{1,2}^2 \sqrt{n_{1,2} \beta_{2,1} \mathbf{P}}} + \frac{\sqrt{2}}{\sigma_{1,2}^2} \frac{\Gamma(\frac{n_{1,2}+1}{2})}{\Gamma(\frac{n_{1,2}}{2})} \right)^2 \right) \\
& + 2c_3^2 \left(\frac{1}{\sigma_{y_{1,2}}^2} + \frac{\frac{\beta_c}{\tilde{\beta}_c} (1 + \rho \sqrt{\frac{\beta_{2,1}}{\beta_{1,2}}})}{(1 - \rho^2) \beta_{2,1} \mathbf{P} + \sigma_{1,2}^2} \right)^2 G \left(\frac{n_{1,2}}{2}, \frac{1}{2} \left(-\frac{c_3}{\sigma_{1,2}^2 \sqrt{n_{1,2} \beta_{1,2} \mathbf{P}}} + \frac{\sigma_{1,2}^2}{\sqrt{2}} \frac{\Gamma(\frac{n_{1,2}+1}{2})}{\Gamma(\frac{n_{1,2}}{2})} \right)^2 \right). \quad (147)
\end{aligned}$$

Note that such $c_1, c_2, c_3 > 0$ always exist since the corresponding factors in (147) are positive and $\Omega_{1,1}, \Omega_{1,2} > 0$. Therefore,

$$\text{Var}[\tilde{i}_1(\mathbf{X}_{1,1}, \mathbf{X}_{1,2}; \mathbf{Y}_{1,1}, \mathbf{Y}_{2,1})] \geq n_{1,1} V(\Omega_{1,1}) + n_{1,2} V(\Omega_{1,2}) \quad (148)$$

with $V(x) := \frac{x(2+x)}{2(1+x)^2} \log e^2$. This proves (104b).

C. Bounding the third moment

In this section, we upper bound the third moment. To this end, we employ the following inequality:

$$\mathbb{E}[|\tilde{i}_1(\mathbf{X}_{1,1}, \mathbf{X}_{1,2}; \mathbf{Y}_{1,1}, \mathbf{Y}_{2,1}) - \mathbb{E}[\tilde{i}_1(\mathbf{X}_{1,1}, \mathbf{X}_{1,2}; \mathbf{Y}_{1,1}, \mathbf{Y}_{2,1})]|^3] \leq 2^3 \mathbb{E}[|\tilde{i}_1(\mathbf{X}_{1,1}, \mathbf{X}_{1,2}; \mathbf{Y}_{1,1}, \mathbf{Y}_{2,1})|^3], \quad (149)$$

which follows from the Minkowski inequality [34].

Let $f_{\tilde{i}}(\cdot)$ be the probability density function of $\tilde{i}_1(\mathbf{X}_{1,1}, \mathbf{X}_{1,2}; \mathbf{Y}_{1,1}, \mathbf{Y}_{2,1})$ with $F_{\tilde{i}}$ as its cumulative density function. For simplicity, define $Z := \tilde{i}_1(\mathbf{X}_{1,1}, \mathbf{X}_{1,2}; \mathbf{Y}_{1,1}, \mathbf{Y}_{2,1})$. For any $\kappa > 1$ we have

$$\begin{aligned}
& \mathbb{E}[|Z|^3] \\
& = \int_{-\infty}^{\infty} |Z|^3 f_{\tilde{i}}(z) dz \\
& \leq \kappa + \int_{\kappa}^{\infty} Z^3 f_{\tilde{i}}(z) dz + \int_{-\infty}^{-\kappa} |Z|^3 f_{\tilde{i}}(z) dz \\
& = \kappa + \int_{\kappa}^{\infty} Z^3 (f_{\tilde{i}}(z) - f_{\tilde{i}}(-z)) dz \\
& = \kappa + \sum_{j=0}^{\infty} \int_{\kappa+j}^{\kappa+j+1} Z^3 (f_{\tilde{i}}(z) - f_{\tilde{i}}(-z)) dz \\
& \leq \kappa + \sum_{j=0}^{\infty} (\kappa + j + 1)^3 \int_{\kappa+j}^{\kappa+j+1} (f_{\tilde{i}}(z) - f_{\tilde{i}}(-z)) dz \\
& = \kappa + \sum_{j=0}^{\infty} (\kappa + j + 1)^3 (F_{\tilde{i}}(\kappa + j + 1) - F_{\tilde{i}}(\kappa + j) + F_{\tilde{i}}(-\kappa - j) - F_{\tilde{i}}(-\kappa - j - 1)) \\
& \leq \kappa + \sum_{j=0}^{\infty} (\kappa + j + 1)^3 (1 - F_{\tilde{i}}(\kappa + j) + F_{\tilde{i}}(-\kappa - j)). \quad (150)
\end{aligned}$$

Notice that

$$1 - F_{\tilde{z}}(\kappa + j) = \mathbb{P}[\tilde{i}_1(\mathbf{X}_{1,1}, \mathbf{X}_{1,2}; \mathbf{Y}_{1,1}, \mathbf{Y}_{2,1}) > \kappa + j] \quad (151)$$

and

$$F_{\tilde{z}}(-\kappa - j) = \mathbb{P}[\tilde{i}_1(\mathbf{X}_{1,1}, \mathbf{X}_{1,2}; \mathbf{Y}_{1,1}, \mathbf{Y}_{2,1}) \leq -\kappa - j]. \quad (152)$$

Hence,

$$\mathbb{E}[|\tilde{i}_1(\mathbf{X}_{1,1}, \mathbf{X}_{1,2}; \mathbf{Y}_{1,1}, \mathbf{Y}_{2,1})|^3] \leq \kappa + 2 \sum_{j=0}^{\infty} (\kappa + j + 1)^3 \mathbb{P}[|\tilde{i}_1(\mathbf{X}_{1,1}, \mathbf{X}_{1,2}; \mathbf{Y}_{1,1}, \mathbf{Y}_{2,1})| > \kappa + j]. \quad (153)$$

Lemma 3: The following inequality holds:

$$\begin{aligned} & \mathbb{P}[|\tilde{i}_1(\mathbf{X}_{1,1}, \mathbf{X}_{1,2}; \mathbf{Y}_{1,1}, \mathbf{Y}_{2,1})| > \kappa + j] \\ & \leq \max \left\{ \frac{2^{1-\frac{n_{1,1}}{2}} \zeta}{\Gamma(\frac{n_{1,1}}{2})} \left(\sqrt{\frac{\tilde{\kappa} + \kappa + j}{2k}} - \frac{b}{2} \right)^{n_{1,1}-2} e^{-\frac{1}{2} \left(\sqrt{\frac{\tilde{\kappa} + \kappa + j}{2k}} - \frac{b}{2} \right)^2}, \right. \\ & \quad \left. \frac{2^{1-\frac{n_{1,2}}{2}} \tilde{\zeta}}{\Gamma(\frac{n_{1,2}}{2})} \left(\sqrt{\frac{\tilde{\kappa} + \kappa + j}{2\tilde{k}}} - \frac{\tilde{b}}{2} \right)^{n_{1,2}-2} e^{-\frac{1}{2} \left(\sqrt{\frac{\tilde{\kappa} + \kappa + j}{2\tilde{k}}} - \frac{\tilde{b}}{2} \right)^2} \right\} \end{aligned} \quad (154)$$

for $\zeta > 1$ and $\tilde{\zeta} > 1$ satisfying

$$\frac{\zeta}{(\zeta - 1)(\frac{n_{1,1}}{2} - 1)} < \frac{1}{2} \left(\sqrt{\frac{\tilde{\kappa} + \kappa + j}{2k}} - \frac{b}{2} \right)^2, \quad (155)$$

$$\frac{\tilde{\zeta}}{(\tilde{\zeta} - 1)(\frac{n_{1,2}}{2} - 1)} < \frac{1}{2} \left(\sqrt{\frac{\tilde{\kappa} + \kappa + j}{2\tilde{k}}} - \frac{\tilde{b}}{2} \right)^2, \quad (156)$$

where $k, \tilde{k}, b, \tilde{b}$ and $\tilde{\kappa}$ are defined in (37).

Proof: See Appendix E. ■

Lemma 4: It holds that

$$\begin{aligned} & \sum_{j=0}^{\infty} (\kappa + j + 1)^3 \mathbb{P}[|\tilde{i}_1(\mathbf{X}_{1,1}, \mathbf{X}_{1,2}; \mathbf{Y}_{1,1}, \mathbf{Y}_{2,1})| > \kappa + j] \\ & \leq \max \left\{ \frac{\zeta \sqrt{2}^{\frac{n_{1,1}}{2}-1}}{\Gamma(\frac{n_{1,1}}{2})} e^{-c} A(n_{1,1}, k, b, c), \frac{\tilde{\zeta} \sqrt{2}^{\frac{n_{1,2}}{2}-1}}{\Gamma(\frac{n_{1,2}}{2})} e^{-\tilde{c}} A(n_{1,2}, \tilde{k}, \tilde{b}, \tilde{c}) \right\} \end{aligned} \quad (157)$$

where c, \tilde{c} and $A(n, k, b, c)$ are defined in (37).

Proof: The proof is based on the following equality [35]:

$$\sum_{j=c}^{\infty} j^n e^{-j} = e^{-c} \Phi(e^{-1}, -n, c). \quad (158)$$

Hence,

$$\mathbb{E}[|\tilde{i}_1(\mathbf{X}_{1,1}, \mathbf{X}_{1,2}; \mathbf{Y}_{1,1}, \mathbf{Y}_{2,1})|^3] \quad (159)$$

$$\leq \kappa + 2 \max \left\{ \frac{\zeta \sqrt{2}^{\frac{n_{1,1}}{2}-1}}{\Gamma(\frac{n_{1,1}}{2})} e^{-c} A(n_{1,1}, k, b, c), \frac{\tilde{\zeta} \sqrt{2}^{\frac{n_{1,2}}{2}-1}}{\Gamma(\frac{n_{1,2}}{2})} e^{-\tilde{c}} A(n_{1,2}, \tilde{k}, \tilde{b}, \tilde{c}) \right\}. \quad (160)$$

As a result

$$\begin{aligned} & \mathbb{E}[|\tilde{i}_1(\mathbf{X}_{1,1}, \mathbf{X}_{1,2}; \mathbf{Y}_{1,1}, \mathbf{Y}_{2,1}) - \mathbb{E}[\tilde{i}_1(\mathbf{X}_{1,1}, \mathbf{X}_{1,2}; \mathbf{Y}_{1,1}, \mathbf{Y}_{2,1})]|^3] \\ & \leq 2^3 \kappa + 2^4 \max \left\{ \frac{\zeta \sqrt{2}^{\frac{n_{1,1}}{2}-1}}{\Gamma(\frac{n_{1,1}}{2})} e^{-c} A(n_{1,1}, k, b, c), \frac{\tilde{\zeta} \sqrt{2}^{\frac{n_{1,2}}{2}-1}}{\Gamma(\frac{n_{1,2}}{2})} e^{-\tilde{c}} A(n_{1,2}, \tilde{k}, \tilde{b}, \tilde{c}) \right\} \end{aligned} \quad (161)$$

for any $\kappa > 1$. This concludes the proof.

APPENDIX D
PROOF OF LEMMA 1

By [32, Proposition 2], we have the following bounds:

$$\frac{f_{\mathbf{Y}_{1,1}}(\mathbf{Y}_{1,1})}{Q_{\mathbf{Y}_{1,1}}(\mathbf{y}_{1,1})} \leq 27\sqrt{\frac{\pi}{8}} \frac{\sigma_{1,1}^2 + \beta_{1,1}\mathbf{P}}{\sqrt{\sigma_{1,1}^2 + 2\beta_{1,1}\mathbf{P}}} \quad (162)$$

$$\frac{f_{\mathbf{Y}_{2,1}}(\mathbf{Y}_{2,1})}{Q_{\mathbf{Y}_{2,1}}(\mathbf{y}_{1,2})} \leq \frac{9(\beta_{1,2} + \beta_{2,1})}{2\pi\sqrt{2\beta_{1,2}\beta_{2,1}}}. \quad (163)$$

By [33, Lemma 5], we have the following bound:

$$\frac{f_{\mathbf{Y}_{2,1}|\mathbf{X}_{1,2}}(\mathbf{y}_{1,2}|\mathbf{x}_{1,2})}{Q_{\mathbf{Y}_{2,1}|\mathbf{X}_{2,1}}(\mathbf{y}_{1,2}|\mathbf{x}_{1,2})} \geq 2^{\frac{n_{1,2}-2}{2}} e^{-\frac{n_{1,2}\beta_{2,1}\mathbf{P}}{2}}. \quad (164)$$

Thus,

$$\log \frac{f_{\mathbf{Y}_{1,1}|\mathbf{X}_{1,1}}(\mathbf{y}_{1,1}|\mathbf{x}_{1,1})f_{\mathbf{Y}_{2,1}|\mathbf{X}_{1,2}}(\mathbf{y}_{1,2}|\mathbf{x}_{1,2})}{f_{\mathbf{Y}_{1,1}}(\mathbf{y}_{1,1})f_{\mathbf{Y}_{2,1}}(\mathbf{y}_{1,2})} \quad (165)$$

$$\geq \log \frac{f_{\mathbf{Y}_{1,1}|\mathbf{X}_{1,1}}(\mathbf{y}_{1,1}|\mathbf{x}_{1,1})Q_{\mathbf{Y}_{2,1}|\mathbf{X}_{1,2}}(\mathbf{y}_{1,2}|\mathbf{x}_{1,2})}{Q_{\mathbf{Y}_{1,1}}(\mathbf{y}_{1,1})Q_{\mathbf{Y}_{2,1}}(\mathbf{y}_{1,2})} + \log \frac{2^{\frac{n_{1,2}+4}{2}} \sqrt{\pi\beta_{1,2}\beta_{2,1}(\sigma_{1,1}^2 + 2\beta_{1,1}\mathbf{P})}}{e^{\frac{n_{1,2}\beta_{2,1}\mathbf{P}}{2}} 243(\sigma_{1,1}^2 + \beta_{1,1}\mathbf{P})(\beta_{1,2} + \beta_{2,1})} \quad (166)$$

which concludes the proof.

APPENDIX E
PROOF OF LEMMA 3

Notice that

$$\begin{aligned} \tilde{i}(\mathbf{X}_{1,1}, \mathbf{X}_{1,2}; \mathbf{Y}_{1,1}, \mathbf{Y}_{2,1}) &= \frac{n_{1,1}}{2} \log \frac{\sigma_{y_{1,1}}^2}{\sigma_{1,1}^2} + \frac{n_{1,2}}{2} \log \frac{\sigma_{y_{1,2}}^2}{\sigma_{y_{1,2}|x_{1,2}}^2} - \frac{\|\mathbf{Z}_{1,1}\|^2}{2\sigma_{1,1}^2} + \frac{\|\mathbf{X}_{1,1} + \mathbf{Z}_{1,1}\|^2}{2\sigma_{y_{1,1}}^2} \\ &\quad - \frac{\|\alpha(\mathbf{X}_{1,2} + \mathbf{X}_{2,1}) + \mathbf{Z}_{1,2} - \mu_{1,2}\|^2}{2\sigma_{y_{1,2}|x_{1,2}}^2} + \frac{\|\alpha(\mathbf{X}_{1,2} + \mathbf{X}_{2,1}) + \mathbf{Z}_{1,2}\|^2}{2\sigma_{y_{1,2}}^2}, \end{aligned} \quad (167)$$

and by (95),

$$\begin{aligned} |\tilde{i}(\mathbf{X}_{1,1}, \mathbf{X}_{1,2}; \mathbf{Y}_{1,1}, \mathbf{Y}_{2,1})| &\leq \frac{n_{1,1}}{2} \log \frac{\sigma_{y_{1,1}}^2}{\sigma_{1,1}^2} + \frac{n_{1,2}}{2} \log \frac{\sigma_{y_{1,2}}^2}{\sigma_{1,2}^2} + \frac{\|\mathbf{Z}_{1,1}\|^2}{2\sigma_{1,1}^2} + \frac{\|\mathbf{X}_{1,1} + \mathbf{Z}_{1,1}\|^2}{2\sigma_{y_{1,1}}^2} \\ &\quad + \frac{\|\mathbf{X}_{2,1} + \mathbf{Z}_{1,2} + \left(1 - \sqrt{\frac{\beta_c}{\beta_c}}(1 + \rho\sqrt{\frac{\beta_{2,1}}{\beta_{1,2}}})\right) \mathbf{X}_{1,2}\|^2}{2\sigma_{1,2}^2} + \frac{\|\mathbf{X}_{1,2} + \mathbf{X}_{2,1} + \mathbf{Z}_{1,2}\|^2}{2\sigma_{y_{1,2}}^2}. \end{aligned} \quad (168)$$

Recall the definitions of $k, \tilde{k}, b, \tilde{b}$ and $\tilde{\kappa}$ from (37). Thus,

$$\mathbb{P}[|\tilde{i}(\mathbf{X}_{1,1}, \mathbf{X}_{1,2}; \mathbf{Y}_{1,1}, \mathbf{Y}_{2,1})| > \kappa + j] \quad (169)$$

$$\begin{aligned} &\leq \mathbb{P}\left[\frac{n_{1,1}}{2} \log \frac{\sigma_{y_{1,1}}^2}{\sigma_{1,1}^2} + \frac{n_{1,2}}{2} \log \frac{\sigma_{y_{1,2}}^2}{\sigma_{1,2}^2} + \frac{\|\mathbf{Z}_{1,1}\|^2}{2\sigma_{1,1}^2} + \frac{\|\mathbf{X}_{1,1} + \mathbf{Z}_{1,1}\|^2}{2\sigma_{y_{1,1}}^2} \right. \\ &\quad \left. + \frac{\|\mathbf{X}_{2,1} + \mathbf{Z}_{1,2} + \left(1 - \sqrt{\frac{\beta_c}{\beta_c}}(1 + \rho\sqrt{\frac{\beta_{2,1}}{\beta_{1,2}}})\right) \mathbf{X}_{1,2}\|^2}{2\sigma_{1,2}^2} + \frac{\|\mathbf{X}_{1,2} + \mathbf{X}_{2,1} + \mathbf{Z}_{1,2}\|^2}{2\sigma_{y_{1,2}}^2} > \kappa + j\right] \\ &\leq \mathbb{P}\left[k\left(\|\mathbf{Z}_{1,1}\| + \frac{b}{2}\right)^2 + \tilde{k}\left(\|\mathbf{Z}_{1,2}\| + \frac{\tilde{b}}{2}\right)^2 > \tilde{\kappa} + \kappa + j\right] \end{aligned} \quad (170)$$

$$\leq \max\left\{\mathbb{P}\left[k\left(\|\mathbf{Z}_{1,1}\| + \frac{b}{2}\right)^2 > \frac{\tilde{\kappa} + \kappa + j}{2}\right], \mathbb{P}\left[\tilde{k}\left(\|\mathbf{Z}_{1,2}\| + \frac{\tilde{b}}{2}\right)^2 > \frac{\tilde{\kappa} + \kappa + j}{2}\right]\right\} \quad (171)$$

$$\leq \max\left\{\mathbb{P}\left[k\left(\|\mathbf{Z}_{1,1}\| + \frac{b}{2}\right)^2 > \frac{\tilde{\kappa} + \kappa + j}{2}\right], \mathbb{P}\left[\tilde{k}\left(\|\mathbf{Z}_{1,2}\| + \frac{\tilde{b}}{2}\right)^2 > \frac{\tilde{\kappa} + \kappa + j}{2}\right]\right\} \quad (172)$$

$$= 1 - \min \left\{ \mathbb{P} \left[\|\mathbf{Z}_{1,1}\| \leq \sqrt{\frac{\tilde{\kappa} + \kappa + j}{2k}} - \frac{b}{2} \right], \mathbb{P} \left[\|\mathbf{Z}_{1,2}\| \leq \sqrt{\frac{\tilde{\kappa} + \kappa + j}{2\tilde{k}}} - \frac{\tilde{b}}{2} \right] \right\} \quad (173)$$

$$\stackrel{(i)}{=} 1 - \min \left\{ 1 - \frac{\Gamma \left(\frac{n_{1,1}}{2}, \frac{1}{2\sigma_{1,1}^4} \left(\sqrt{\frac{\tilde{\kappa} + \kappa + j}{2k}} - \frac{b}{2} \right)^2 \right)}{\Gamma \left(\frac{n_{1,1}}{2} \right)}, 1 - \frac{\Gamma \left(\frac{n_{1,2}}{2}, \frac{1}{2\sigma_{1,2}^4} \left(\sqrt{\frac{\tilde{\kappa} + \kappa + j}{2\tilde{k}}} - \frac{\tilde{b}}{2} \right)^2 \right)}{\Gamma \left(\frac{n_{1,2}}{2} \right)} \right\} \quad (174)$$

$$= \max \left\{ \frac{\Gamma \left(\frac{n_{1,1}}{2}, \frac{1}{2\sigma_{1,1}^4} \left(\sqrt{\frac{\tilde{\kappa} + \kappa + j}{2k}} - \frac{b}{2} \right)^2 \right)}{\Gamma \left(\frac{n_{1,1}}{2} \right)}, \frac{\Gamma \left(\frac{n_{1,2}}{2}, \frac{1}{2\sigma_{1,2}^4} \left(\sqrt{\frac{\tilde{\kappa} + \kappa + j}{2\tilde{k}}} - \frac{\tilde{b}}{2} \right)^2 \right)}{\Gamma \left(\frac{n_{1,2}}{2} \right)} \right\} \quad (175)$$

$$\leq \max \left\{ \frac{(\sigma_{1,1}^2 \sqrt{2})^{2-n_{1,1}} \zeta \left(\sqrt{\frac{\tilde{\kappa} + \kappa + j}{2k}} - \frac{b}{2} \right)^{n_{1,1}-2} e^{-\frac{1}{2\sigma_{1,1}^4} \left(\sqrt{\frac{\tilde{\kappa} + \kappa + j}{2k}} - \frac{b}{2} \right)^2}}{\Gamma \left(\frac{n_{1,1}}{2} \right)}, \frac{(\sigma_{1,2}^2 \sqrt{2})^{2-n_{1,2}} \tilde{\zeta} \left(\sqrt{\frac{\tilde{\kappa} + \kappa + j}{2\tilde{k}}} - \frac{\tilde{b}}{2} \right)^{n_{1,2}-2} e^{-\frac{1}{2\sigma_{1,2}^4} \left(\sqrt{\frac{\tilde{\kappa} + \kappa + j}{2\tilde{k}}} - \frac{\tilde{b}}{2} \right)^2}}{\Gamma \left(\frac{n_{1,2}}{2} \right)} \right\}, \quad (176)$$

where in (i) we use the fact that $\|\mathbf{Z}_{1,1}\|/\sigma_{1,1}^2$ and $\|\mathbf{Z}_{1,2}\|/\sigma_{1,2}^2$ follow a central chi-distribution of degree $n_{1,1}$ and $n_{1,2}$, respectively. In the last inequality, we use the following bound [36]:

$$\Gamma(a, x) < \zeta x^{a-1} e^{-x}, \quad (177)$$

for $a > 1, \zeta > 1, x > \zeta/(\zeta - 1)(a - 1)$. This concludes the proof.

REFERENCES

- [1] M. Bennis, M. Debbah, and H. V. Poor, "Ultrareliable and low-latency wireless communication: Tail, risk, and scale," *Proceedings of the IEEE*, vol. 106, no. 10, pp. 1834–1853, Oct. 2018.
- [2] Y. Wang, W. Chen, and H. V. Poor, "Ultra-reliable and low-latency wireless communications in the high SNR regime: A cross-layer tradeoff," *IEEE Transactions on Communications*, vol. 70, no. 1, pp. 149–162, Jan. 2022.
- [3] H. Nikbakht, M. Wigger, and S. Shamai (Shitz), "Random user activity with mixed delay traffic," in *Proceedings of the IEEE Information Theory Workshop*, April, 11–14, 2021.
- [4] J. Yao, Q. Zhang, and J. Qin, "Joint decoding in downlink NOMA systems with finite blocklength transmissions for ultrareliable low-latency tasks," *IEEE Internet of Things Journal*, vol. 9, no. 18, pp. 17705–17713, Sept. 2022.
- [5] H. Nikbakht, M. Wigger, M. Egan, S. Shamai (Shitz), J.-M. Gorce, and H. V. Poor, "An information-theoretic view of mixed-delay traffic in 5G and 6G," *Entropy*, vol. 24, no. 5, article 637, 2022.
- [6] N. Khan and S. Coleri, "Resource allocation for ultra-reliable low-latency vehicular networks in finite blocklength regime," in *Proceedings of the IEEE International Mediterranean Conference on Communications and Networking*, pp. 322–327, Athens, Greece, 2022.
- [7] Y. Liu, P. M. Olmos, and D. G. M. Mitchell, "Generalized LDPC codes for ultra reliable low latency communication in 5G and beyond," *IEEE Access*, vol. 6, pp. 72002–72014, 2018.
- [8] A. Sharma and M. Salim, "Polar code appropriateness for ultra-reliable and low-latency use cases of 5G systems," *International Journal of Networked and Distributed Computing*, vol. 7, no. 3, pp. 93–99, 2019.
- [9] Y. Polyanskiy, H. V. Poor and S. Verdú, "Channel coding rate in the finite blocklength regime," *IEEE Transactions on Information Theory*, vol. 56, no. 5, pp. 2307–2359, May, 2010.
- [10] M. Hayashi, "Information spectrum approach to second-order coding rate in channel coding," *IEEE Transactions on Information Theory*, vol. 55, no. 11, pp. 4947–4966, Nov. 2009.
- [11] V. Strassen, "Asymptotic estimates in Shannon's information theory [Asymptotische Abschätzungen in Shannon's Informationstheorie]," in *Transactions of the Third Prague Conference on Information Theory, Statistical Decision Functions, Random Processes*, pp. 689–723, Prague, Czechoslovak Academy of Sciences, June 1962, translated from German by Peter Luthy. [Online]. Available: <http://pi.math.cornell.edu/pmlut/strassen.pdf>.
- [12] M. C. Lucas-Estañ and J. Gozalvez, "Sensing-based grant-free scheduling for ultra reliable low latency and deterministic beyond 5G networks," *IEEE Transactions on Vehicular Technology*, vol. 71, no. 4, pp. 4171–4183, April 2022.
- [13] C. E. Shannon, R. G. Gallager, and E. R. Berlekamp, "Lower bounds to error probability for coding on discrete memoryless channels. I," *Information Control*, vol. 10, no. 1, pp. 65–103, Feb. 1967.
- [14] V. Y. F. Tan and M. Tomamichel, "The third-order term in the normal approximation for the AWGN channel," *IEEE Transactions on Information Theory*, vol. 61, no. 5, pp. 2430–2438, May 2015.
- [15] T. Erseghe, "Coding in the finite-blocklength regime: Bounds based on Laplace integrals and their asymptotic approximations," *IEEE Transactions on Information Theory*, vol. 62, no. 12, pp. 6854–6883, Dec. 2016.
- [16] N. Shulman and M. Feder. "Static broadcasting," in *Proceedings of the IEEE International Symposium on Information Theory (ISIT)*, Sorrento, Italy, 2000.
- [17] M. Langberg and M. Effros, "Beyond capacity: The joint time-rate region", arXiv:2101.12236v1, Jan. 2021.
- [18] P. H. Lin, S. C. Lin, P.-W. Chen, M. Mross, and E. A. Jorswieck, "Gaussian broadcast channels in heterogeneous blocklength constrained networks," arXiv:2109.07767v1, Sep. 2021.
- [19] M. Mross, P. H. Lin, and E. A. Jorswieck, "New inner and outer bounds for 2-user Gaussian broadcast channels with heterogeneous blocklength constraints", arXiv:2202.02110v1, Feb. 2022.
- [20] H. Nikbakht, M. Egan, and J.-M. Gorce, "Joint channel coding of consecutive messages with heterogeneous decoding deadlines in the finite blocklength regime," in *Proceedings of the IEEE Wireless Communications and Networking Conference*, pp. 2423–2428, Austin, TX, USA, 2022.

- [21] H. Nikbakht, M. Egan, and J. -M. Gorce, "Dirty paper coding for consecutive messages with heterogeneous decoding deadlines in the finite blocklength regime," in *Proceedings of the IEEE International Symposium on Information Theory*, pp. 2100–2105, Espoo, Finland, 2022.
- [22] M. H. M. Costa, "Writing on dirty paper (Corresp.)," *IEEE Transactions on Information Theory*, vol. 29, no. 3, pp. 439–441, May 1983.
- [23] J. Scarlett, "On the dispersions of the Gel'fand - Pinsker channel and dirty paper coding," *IEEE Transactions on Information Theory*, vol. 61, no. 9, pp. 4569–4586, Sept. 2015.
- [24] G. Caire and S. Shamai, "On the achievable throughput of a multiantenna Gaussian broadcast channel," *IEEE Transactions on Information Theory*, vol. 49, no. 7, pp. 1691–1706, July 2003.
- [25] A. Tajer, A. Steiner, and S. Shamai (Shitz), "The broadcast approach in communication networks," *Entropy*, vol. 23, no. 1, article 120, 2021.
- [26] P. Sheldon, D. Tuninetti, and B. Smida, "The Gaussian broadcast channels with a hard deadline and a global reliability constraint," in *Proceedings of the IEEE International Conference on Communications*, pp. 1–6, Montreal, QC, Canada, 2021.
- [27] P. Sheldon, B. Smida, N. Devroye, and D. Tuninetti, "Achievable rate regions for the Gaussian broadcast channel with fixed blocklength and per user reliability," in *Proceedings of the 2022 58th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, pp. 1-6, Monticello, IL, USA, 2022.
- [28] D. Tuninetti, P. Sheldon, B. Smida, and N. Devroye, "On second order rate regions for the static scalar Gaussian broadcast channel," *IEEE Journal on Selected Areas in Communications*, vol. 41, no. 7, pp. 1982–1999, July 2023.
- [29] D. Tuninetti, B. Smida, N. Devroye, and H. Seferoglu, "Scheduling on the Gaussian broadcast channel with hard deadlines," in *Proceedings of the IEEE International Conference on Communications (ICC)*, pp. 1–7, Kansas City, MO, USA, 2018.
- [30] K. Marton, "A coding theorem for the discrete memoryless broadcast channel," *IEEE Transactions on Information Theory*, vol. 25, no. 3, pp. 306–311, May 1979.
- [31] C. E. Shannon, "Probability of error for optimal codes in a Gaussian channel," *The Bell System Technical Journal*, vol. 38, no. 3, pp. 611–656, 1959.
- [32] E. MolavianJazi and J. N. Laneman, "A second-order achievable rate region for Gaussian multi-access channels via a central limit theorem for functions," *IEEE Transactions on Information Theory*, vol. 61, no. 12, pp. 6719–6733, Dec. 2015.
- [33] H. Nikbakht, M. Wigger, S. Shamai, J. M. Gorce, and H. V. Poor, "Joint coding of URLLC and eMBB in Wyner's soft-handoff network in the finite blocklength regime," in *Proceedings of the IEEE Global Communications Conference*, Rio de Janeiro, Brazil, pp. 1–6, 2022.
- [34] H. P. Mulholland, "On generalizations of Minkowski's inequality in the form of a triangle inequality," in *Proceedings of the London Mathematical Society*, s2-51 (1): 294–307, 1949.
- [35] H. M. Srivastava, R. K. Saxena, T. K. Pogány, and R. Saxena, "Integral and computational representations of the extended Hurwitz-Lerch Zeta function", *Integral Transforms and Special Functions*, vol. 22, no. 7, pp. 487–506, July 2011.
- [36] I. Pinelis, "Exact lower and upper bounds on the incomplete Gamma function," *Journal of Mathematical Inequalities and Applications*, vol. 23, no. 4, pp. 1261–1278, 2020.