



**HAL**  
open science

# Dirty Paper Coding for Consecutive Messages with Heterogeneous Decoding Deadlines in the Finite Blocklength Regime

Homa Nikbakht, Malcolm Egan, Jean-Marie Gorce

► **To cite this version:**

Homa Nikbakht, Malcolm Egan, Jean-Marie Gorce. Dirty Paper Coding for Consecutive Messages with Heterogeneous Decoding Deadlines in the Finite Blocklength Regime. [Research Report] RR-9462, Inria - Research Centre Grenoble – Rhône-Alpes. 2022, pp.1-21. hal-03556888

**HAL Id: hal-03556888**

**<https://inria.hal.science/hal-03556888v1>**

Submitted on 24 Feb 2022

**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.



# Dirty Paper Coding for Consecutive Messages with Heterogeneous Decoding Deadlines in the Finite Blocklength Regime

Homa Nikbakht, Malcolm Egan, Jean-Marie Gorce

**RESEARCH  
REPORT**

**N° 9462**

February 2022

Project-Teams MARACAS





# Dirty Paper Coding for Consecutive Messages with Heterogeneous Decoding Deadlines in the Finite Blocklength Regime

Homa Nikbakht, Malcolm Egan, Jean-Marie Gorce

Project-Teams MARACAS

Research Report n° 9462 — version 1 — initial version February 2022 —  
revised version Février 2022 — 21 pages

**Abstract:** To improve reliability in latency-critical applications, a point-to-point communication system with heterogeneous decoding deadlines is considered. Unlike existing work, this system allows for a message to arrive before the decoding deadline of a prior message. A new coding scheme with finite blocklength codewords is introduced exploiting the dirty paper coding principle. Rigorous bounds are derived for achievable error probabilities. Moreover, numerical results illustrate that the proposed scheme outperforms time sharing for a wide range of blocklengths.

**Key-words:** URLLC, Finite Blocklength Regime, Heterogeneous Decoding Deadline, Dirty Paper Coding

---

Homa Nikbakht, Malcolm Egan and Jean-Marie Gorce are with the Laboratoire CITI, a joint laboratory between the Institut National de Recherche en Informatique et en Automatique (INRIA), the Université de Lyon and the Institut National de Sciences Appliquées (INSA) de Lyon. 6 Av. des Arts 69621 Villeurbanne, France. ({homa.nikbakht, malcom.egan, jean-marie.gorce}@inria.fr)

This work was funded by INRIA Nokia Bell Labs ADR "Network Information Theory" and by the French National Agency for Research (ANR) under grant ANR-16-CE25-0001 - ARBURST.

**RESEARCH CENTRE  
GRENOBLE – RHÔNE-ALPES**

Inovallée  
655 avenue de l'Europe Montbonnot  
38334 Saint Ismier Cedex

**Résumé :** Pour améliorer la fiabilité des applications critiques en termes de latence, un système de communication point à point avec des délais de décodage hétérogènes est envisagé. Contrairement aux travaux existants, ce système permet à un message d'arriver avant la date limite de décodage d'un message précédent. Un nouveau schéma de codage avec des mots de code de longueur de bloc finie est introduit en exploitant le principe de codage du Dirty Paper Coding. Des limites rigoureuses sont dérivées pour les probabilités d'erreur atteignables. De plus, les résultats numériques montrent que le schéma proposé surpasse le partage de temps pour une large gamme de longueurs de bloc.

**Mots-clés :** URLLC, Régime de longueur de bloc fini, Date limite de décodage hétérogène

## Contents

<b>1</b>	<b>Introduction</b>	<b>4</b>
<b>2</b>	<b>Problem Setup</b>	<b>5</b>
<b>3</b>	<b>Random Coding Scheme</b>	<b>6</b>
3.1	Encoding . . . . .	6
3.1.1	Channel 1 . . . . .	6
3.1.2	Channel 2 – 1 . . . . .	6
3.1.3	Channel 2 – 2 . . . . .	7
3.1.4	Channel 3 . . . . .	8
3.2	Decoding . . . . .	8
3.2.1	Decoding the power type . . . . .	9
3.2.2	Decoding $m_1$ . . . . .	9
3.2.3	Decoding $m_2$ . . . . .	9
<b>4</b>	<b>Main Result</b>	<b>10</b>
4.1	Numerical Results . . . . .	10
<b>5</b>	<b>Conclusions</b>	<b>11</b>
<b>A</b>	<b>Proof of Theorem 1</b>	<b>12</b>
A.1	Analyzing $\epsilon_1$ . . . . .	12
A.1.1	Analyzing $\mathbb{P}[\mathcal{E}_{1,1}]$ . . . . .	12
A.1.2	Analyzing $\mathbb{P}[\mathcal{E}_{1,2} \mathcal{E}_{1,1}^c]$ . . . . .	12
A.1.3	Analyzing $\mathbb{P}[\mathcal{E}_{1,3} \mathcal{E}_{1,1}^c, \mathcal{E}_{1,2}^c]$ . . . . .	14
A.1.4	Analyzing $\mathbb{P}[\mathcal{E}_{1,4} \mathcal{E}_{1,1}^c, \mathcal{E}_{1,2}^c, \mathcal{E}_{1,3}^c]$ . . . . .	14
A.2	Analyzing $\epsilon_2$ . . . . .	18
A.2.1	Analyzing $\mathbb{P}[\mathcal{E}_2]$ . . . . .	18
A.2.2	Analyzing $\mathbb{P}[  \mathbf{X}_{2,2}  ^2 > n_{2,2}\beta_2 P_2]$ . . . . .	20

## 1 Introduction

Mobile wireless networks in 5G and proposals for 6G are increasingly intended for use in latency-critical and high-reliability systems; notably in industrial control applications as well as for autonomous vehicles and remote surgery [1–4]. Commonly known as ultra-reliable low-latency communications (URLLC), packets are typically short. As a consequence, data transmission is no longer arbitrarily reliable and a key challenge is to design coding schemes that support high reliability requirements [5].

A standard assumption in the design of coding schemes is that consecutive messages have separated arrival times and decoding deadlines. As such, there is no choice but to encode the messages independently. On the other hand, 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. In general, each of the messages will have a different decoding deadline.

In such a situation, code design must account for two issues: (i) messages with close arrival times; and (ii) messages with heterogeneous decoding deadlines. One approach is to begin encoding only after the second message has arrived. In this case, joint encoding is possible; however, messages must be decoded at different times.

The problem of code design with heterogeneous deadlines was first considered in the context of static broadcasting [6], where a single message is decoded at multiple receivers under different relative decoding delay constraints. The work in [6] was recently generalized to multi-source and multi-terminal networks by Langberg and Effros in [7]. 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 [6] and [7] focused on the asymptotic regime. In the finite blocklength regime, a coding scheme for the Gaussian broadcast channel with heterogeneous blocklength constraints depending on channel signal-to-noise ratio was introduced in [8]. By exploiting an early decoding scheme, the authors showed that significant improvements are possible over standard successive interference cancellation.

Nevertheless, the work in [6–8] all focused on the case where both messages are available at the time of encoding. In our previous work [9], we introduced a coding scheme for the Gaussian point-to-point channel where encoding begins for the first message before the second message arrives. This scheme 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 accounting for the message set size and finite decoding deadline constraints.

In this paper, we introduce a coding scheme for the Gaussian point-to-point channel model with heterogeneous decoding deadlines in [9] exploiting the dirty paper coding (DPC) principle. Accounting for finite decoding deadline constraints (corresponding to fixed blocklengths), we derive rigorous bounds on the achievable error probabilities for each of the messages. This is achieved by combining techniques to analyze the Gel'fand-Pinsker channel in the finite blocklength regime in [10] and multiple parallel channels in [11].

A natural question is under what conditions our joint encoding scheme outperforms time sharing with each message allocated a disjoint set of channel uses? We show via a numerical study that for sufficiently large power constraints and a wide range of finite blocklengths error probability of *both* messages can outperform time sharing.

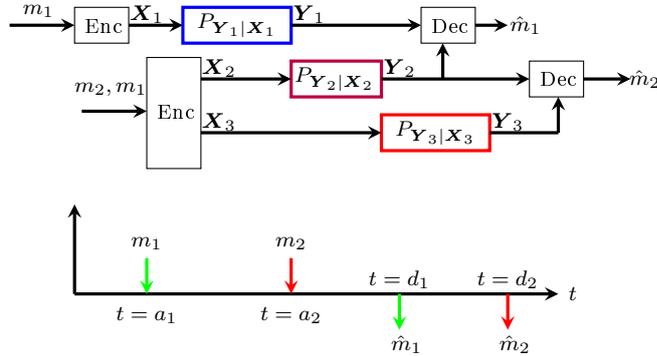


Figure 1: System model.

## 2 Problem Setup

Consider a sensor that seeks to send two messages, where each message lies in the set  $\{1, \dots, M\}$ . 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  $\{1, \dots, M\}$ . For the sake of consistency we only consider the symmetric case where both messages are of same size. The study can be easily extended to asymmetric cases.

Each message is subject to different decoding delay constraints. In particular, at time  $d_1$ , the receiver attempts to reconstruct the message  $m_1$ . Similarly, at time  $d_2 > d_1$ , the receiver attempts to reconstruct the message  $m_2$ .

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

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

where  $f, \psi, \phi$  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. We highlight that  $m_2$  is not known before time  $t = a_2$ ; i.e., encoding is causal.

The channel inputs can then be written as

$$\begin{aligned} \mathbf{X}_1 &= \{X_{a_1}, \dots, X_{a_2-1}\} \\ \mathbf{X}_2 &= \{X_{a_2}, \dots, X_{d_1}\} \\ \mathbf{X}_3 &= \{X_{d_1+1}, \dots, X_{d_2}\}. \end{aligned} \quad (2)$$

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

$$\mathbb{E}[\|\mathbf{X}_i\|^2] \leq n_i P_i, \quad i \in \{1, 2, 3\}, \quad (3)$$

where

$$n_1 = a_2 - a_1, \quad n_2 = d_1 - a_2 + 1, \quad n_3 = d_2 - d_1. \quad (4)$$

Denote the channel outputs by

$$\mathbf{Y}_1 = \{Y_{a_1}, \dots, Y_{a_2-1}\}$$

$$\begin{aligned}\mathbf{Y}_2 &= \{Y_{a_2}, \dots, Y_{d_1}\} \\ \mathbf{Y}_3 &= \{Y_{d_1+1}, \dots, Y_{d_2}\}.\end{aligned}\tag{5}$$

The conditional distributions governing the three channels are then denoted by  $P_{\mathbf{Y}_1|\mathbf{X}_1}$ ,  $P_{\mathbf{Y}_2|\mathbf{X}_2}$ , and  $P_{\mathbf{Y}_3|\mathbf{X}_3}$ . We assume that each channel is additive, memoryless, stationary, and Gaussian; that is,

$$\mathbf{Y}_i = \mathbf{X}_i + \mathbf{Z}_i,\tag{6}$$

where  $\mathbf{Z}_i \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_{n_i})$ . This setup is illustrated in Fig. 1.

At the receiver, the decoder attempts to reconstruct the two messages  $m_1, m_2$  based on the channel outputs via the decoding functions  $g_1$  and  $g_2$ ; i.e.,

$$\hat{m}_1 = g_1(\mathbf{Y}_1, \mathbf{Y}_2) \quad \text{and} \quad \hat{m}_2 = g_2(\mathbf{Y}_2, \mathbf{Y}_3).\tag{7}$$

Observe that both decoders are causal.

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

$$\epsilon_1 = \mathbb{P}[\hat{m}_1 \neq m_1], \quad \epsilon_2 = \mathbb{P}[\hat{m}_2 \neq m_2].\tag{8}$$

The focus of the remainder of this paper is to characterize the tradeoff between the size of the message set  $M$ , the error probabilities  $\epsilon_1, \epsilon_2$ , and the decoding deadlines  $d_1, d_2$ . Formally, we study the achievable region defined as follows.

**Definition 1** *Given the power constraints  $P_1, P_2, P_3$ , a tuple  $(a_1, a_2, d_1, d_2, M, \epsilon_1, \epsilon_2)$  is achievable if the messages  $m_1, m_2$  of cardinality  $M$  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  $\epsilon_1, \epsilon_2$ , respectively.*

### 3 Random Coding Scheme

In this section, we introduce a random coding scheme that is based on dirty paper coding [10], [13] and [15]. In Section 4, we provide upper bounds on  $\epsilon_1, \epsilon_2$  for this scheme.

#### 3.1 Encoding

Let  $\beta_1, \beta_2 \in [0, 1]$  such that  $\beta_1 + \beta_2 < 1$ . The encoding process consists of four phases.

##### 3.1.1 Channel 1

In the first channel, consisting of  $n_1$  channel uses, only  $m_1$  is known to the encoder. The channel input  $\mathbf{X}_1$  corresponding to message  $m_1$  is a codeword  $\mathbf{X}_1(m_1) \in \mathbb{R}^{n_1}$ , where  $\mathbf{X}_1(m_1) \sim \mathcal{N}(\mathbf{0}, P_1 \mathbf{I}_{n_1})$  with  $\mathbf{X}_1(m) \perp \mathbf{X}_1(m')$ ,  $m \neq m'$ .

##### 3.1.2 Channel 2 – 1

In the second channel, both  $m_1$  and  $m_2$  are known to the encoder. The channel input  $\mathbf{X}_2$  is based on the DPC scheme in [10], which makes use of the notion of a power type.

**Definition 2 (Power Types)** For a given blocklength  $n$ , fix  $\delta > 0$ . For each  $P_t = \frac{i\delta}{n}$  with  $i \in \{0, 1, 2, \dots\}$ , define the type class

$$T^n(P_t) = \{\mathbf{x} \in \mathbb{R}^n : nP_t \leq \|\mathbf{x}\|^2 < nP_t + \delta\}. \quad (9)$$

For each  $\mathbf{x} \in T^n(P_t)$ , we call  $P_t$  the type of  $\mathbf{x}$  and  $i$  the index of the type. The set of all types is given by

$$\mathcal{P}_n = \left\{ \frac{i\delta}{n} : i \in \{0, 1, 2, \dots\} \right\}. \quad (10)$$

In general, the type  $P_t$  of  $\mathbf{x}$  can be arbitrarily large with non-zero probability. It is therefore desirable to consider a *typical set of power types*, defined as

$$\tilde{\mathcal{P}}_n = \{P_t \in \mathcal{P}_n : P_t \leq \Pi\} \quad (11)$$

for some  $\Pi > 0$ . To produce the channel input  $\mathbf{X}_2$ , consider the first  $n_{2,1} < n_2$  symbols of  $n_2$  symbols allocated to the second channel. These  $n_{2,1}$  symbols are utilized to send the power type  $P_t$  to the receiver. Over the remaining  $n_{2,2} = n_2 - n_{2,1}$  symbols in the second channel, a codeword constructed from both  $m_1$  and  $m_2$  is sent to the receiver.

More precisely, in the second channel, the encoder first produces a codeword  $\mathbf{X}_{2,2} \in \mathbb{R}^{n_{2,2}}$  drawn from a Gaussian codebook of power  $\beta_2 P_2$ ,  $\beta_2 \in [0, 1]$  to encode  $m_2$ . The encoder then computes the power type of  $\mathbf{X}_{2,2}$ .

Indeed, for a fixed  $\delta_2$ , the power type of  $\mathbf{X}_{2,2}$  is of the form

$$P_t = \frac{k_2 \delta_2}{n_{2,2}}, \quad k_2 \in \{0, 1, 2, \dots\}. \quad (12)$$

The set of all power types  $P_t$  is then given by

$$\mathcal{P}_{n_{2,2}} = \left\{ \frac{k_2 \delta_2}{n_{2,2}} : k_2 \in \{0, 1, 2, \dots\} \right\}. \quad (13)$$

The corresponding typical set is given by

$$\tilde{\mathcal{P}}_{n_{2,2}} = \{P_t \in \mathcal{P}_{n_{2,2}} : P_t \leq P_{\max}\} \quad (14)$$

for some  $P_{\max} > 0$ . A power type  $P_t$  belongs to  $\tilde{\mathcal{P}}_{n_{2,2}}$  if its index belongs to the set  $\{0, \dots, \lfloor \frac{n_{2,2} P_{\max}}{\delta_2} \rfloor\}$ . Denoting by  $P_t$  as the power type of the signal  $\mathbf{X}_{2,2}$ , we have the following error event:

$$\mathcal{E}_{1,1} \triangleq \{P_t \notin \tilde{\mathcal{P}}_{n_{2,2}}\}. \quad (15)$$

To send the power type, the index  $k_2$  is encoded to produce a codeword  $\mathbf{X}_T \in \mathbb{R}^{n_{2,1}}$  drawn from a Gaussian point-to-point codebook of power  $(1 - \beta_1 - \beta_2)P_2$ .

### 3.1.3 Channel 2 – 2

Over the remaining  $n_{2,2}$  symbols allocated to the second channel, called channel 2–2, the encoder exploits DPC treating  $\mathbf{X}_{2,2}$  associated to  $m_2$  as the state. To employ DPC, we generate the following codebooks.

*Codebook Generation:* Denote by  $L^{(P_t)}$  the random coding parameter illustrating the number of auxiliary codewords for the type  $P_t \in \tilde{\mathcal{P}}_{n_{2,2}}$ . For each message  $m_1 \in \{1, \dots, M\}$  and each

power type  $P_t$  a random codebook  $C_U^{(P_t)}$  containing  $ML^{(P_t)}$  auxiliary codewords  $\{\mathbf{U}^{(P_t)}(m_1, \ell)\}$  with  $\ell \in \{1, \dots, L^{(P_t)}\}$ , where each codeword is independently distributed on the sphere  $\mathbb{S}^{n_{2,2}-1}$  with power  $n_{2,2}(\beta_1 P_2 + \alpha^2 P_t)$ ,  $\alpha \in (0, 1)$ . That is, the probability density function of  $\mathbf{U}$  is given by

$$f_{\mathbf{U}}^{(P_t)}(\mathbf{u}) = \frac{\delta(\|\mathbf{u}\|^2 - n_{2,2}(\beta_1 P_2 + \alpha^2 P_t))}{S_{n_{2,2}}(\sqrt{n_{2,2}(\beta_1 P_2 + \alpha^2 P_t)})}, \quad (16)$$

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.

Given  $\mathbf{X}_{2,2}$ , the second channel input over the  $n_{2,2}$  symbols allocated for transmission of  $m_1, m_2$  is given by

$$\mathbf{X}_2 = \mathbf{X}_{2,1} + \mathbf{X}_{2,2}, \quad (17)$$

where

$$\mathbf{X}_{2,1} = \mathbf{U} - \alpha \mathbf{X}_{2,2}. \quad (18)$$

Here,  $\mathbf{U} \in C_U^{(P_t)}$  with  $\ell$  chosen such that  $\mathbf{X}_{2,1} \in \mathcal{D}$ , where

$$\mathcal{D} = \{\mathbf{x}_{2,1} : n_{2,2}\beta_1 P_2 - \delta_1 \leq \|\mathbf{x}_{2,1}\|^2 \leq n_{2,2}\beta_1 P_2\} \quad (19)$$

for some  $\delta_1 > 0$ . If more than one index exists, then one is selected arbitrarily. We have the encoding error event:

$$\mathcal{E}_{1,2} \triangleq \{\text{no } \ell \text{ exists such that } \mathbf{U} - \alpha \mathbf{X}_{2,2} \in \mathcal{D}\}. \quad (20)$$

### 3.1.4 Channel 3

Over the last  $n_3$  channel uses, the sensor encodes only  $m_2$  with a codeword  $\mathbf{X}_3(m_2) \sim \mathcal{N}(0, P_3 \mathbf{I}_{n_3})$  with  $\mathbf{X}_3(m_2) \perp \mathbf{X}_3(m'_2)$ ,  $m_2 \neq m'_2$ .

## 3.2 Decoding

Given the structure of the encoding functions, receiver observations can be viewed as arising from four channels: over the first channel of  $n_1$  blocks only  $m_1$  is transmitted; over the second channel of  $n_{2,1}$  blocks the power type index  $k_2$  is transmitted; over the third channel of  $n_{2,2}$  blocks, both  $m_1$  and  $m_2$  are transmitted; and over the fourth channel of  $n_3$  blocks only  $m_2$  is transmitted. We define the outputs of these four channels by  $\mathbf{Y}_1, \mathbf{Y}_T, \tilde{\mathbf{Y}}_2$  and  $\mathbf{Y}_3$ , respectively which are of the following forms

$$\mathbf{Y}_1 = \mathbf{X}_1 + \mathbf{Z}_1, \quad (21a)$$

$$\mathbf{Y}_T = \mathbf{X}_T + \mathbf{Z}_T, \quad (21b)$$

$$\tilde{\mathbf{Y}}_2 = \mathbf{U} - \alpha \mathbf{X}_{2,2} + \mathbf{X}_{2,2} + \tilde{\mathbf{Z}}_2, \quad (21c)$$

$$\mathbf{Y}_3 = \mathbf{X}_3 + \mathbf{Z}_3, \quad (21d)$$

where  $\mathbf{Z}_1 \in \mathbb{R}^{n_1}$ ,  $\mathbf{Z}_T \in \mathbb{R}^{n_{2,1}}$ ,  $\tilde{\mathbf{Z}}_2 \in \mathbb{R}^{n_{2,2}}$  and  $\mathbf{Z}_3 \in \mathbb{R}^{n_3}$  are i.i.d standard Gaussian noise sequences.

### 3.2.1 Decoding the power type

The receiver first decodes the power type based on the outputs of the second channel (channel 2 – 1). Denoting by  $\hat{k}_2$  the decoded index of the power type, then we have the following error event while decoding the power type

$$\mathcal{E}_{1,3} \triangleq \{\text{Decoder chooses an index } \hat{k}_2 \neq k_2\}. \quad (22)$$

In our scheme, we fix the maximum error probability of decoding the power type at  $\epsilon_T \in [0, 1]$ . Then for a given  $P_{\max}$ ,  $P_2$ ,  $\beta_2$ ,  $n_2$  and  $\delta_2$ , we fix  $n_{1,2}$  at the smallest value such that the probability that event  $\mathcal{E}_{1,3}$  occurs is less than  $\epsilon_T$ .

### 3.2.2 Decoding $m_1$

Given the power type  $P_t$ , the receiver decodes  $m_1$  based on the outputs of the first and third channels (channel 1 and channel 2 – 2). More specifically, given observations  $\mathbf{y}_1, \tilde{\mathbf{y}}_2$ , the receiver estimates  $m_1$  according to the pair  $(\hat{m}_1, \hat{\ell})$ , such that the corresponding sequences  $\mathbf{U}^{P_t}(\hat{m}_1, \hat{\ell})$  and  $\mathbf{X}_1(\hat{m}_1)$  maximize

$$i_1(\mathbf{u}, \mathbf{x}_1; \mathbf{y}_1, \tilde{\mathbf{y}}_2) = \ln \left( \frac{f_{\mathbf{Y}_1|\mathbf{X}_1}(\mathbf{y}_1|\mathbf{x}_1)f_{\tilde{\mathbf{Y}}_2|\mathbf{U}}(\tilde{\mathbf{y}}_2|\mathbf{u})}{f_{\mathbf{Y}_1}(\mathbf{y}_1)f_{\tilde{\mathbf{Y}}_2}(\tilde{\mathbf{y}}_2)} \right) \quad (24)$$

over all pairs of  $\mathbf{x}_1$  and  $\mathbf{u} \in C_{\mathbf{U}}^{(P_t)}$ . We have the following error event while decoding  $m_1$ :

$$\mathcal{E}_{1,4} \triangleq \{\text{Decoder chooses a message } \hat{m}_1 \neq m_1\}. \quad (25)$$

### 3.2.3 Decoding $m_2$

The receiver decodes  $m_2$  based on the outputs of the third and fourth channels (channel 2 – 2 and channel 3). After observing  $\tilde{\mathbf{y}}_2$  and  $\mathbf{y}_3$ , the receiver estimates  $m_2$  such that the corresponding sequences  $\mathbf{x}_{2,2}, \mathbf{x}_3$  maximize

$$i_2(\mathbf{x}_{2,2}, \mathbf{x}_3; \tilde{\mathbf{y}}_2, \mathbf{y}_3) = \ln \left( \frac{f_{\tilde{\mathbf{Y}}_2|\mathbf{X}_{2,2}}(\tilde{\mathbf{y}}_2|\mathbf{x}_{2,2})f_{\mathbf{Y}_3|\mathbf{X}_3}(\mathbf{y}_3|\mathbf{x}_3)}{f_{\tilde{\mathbf{Y}}_2}(\tilde{\mathbf{y}}_2)f_{\mathbf{Y}_3}(\mathbf{y}_3)} \right). \quad (26)$$

over all pairs of  $\mathbf{x}_{2,2}$  and  $\mathbf{x}_3$ . We have the following error event while decoding  $m_2$ :

$$\mathcal{E}_2 \triangleq \{\text{Decoder chooses a message } \hat{m}_2 \neq m_2\}. \quad (27)$$

$$\begin{aligned} \tilde{\gamma}_1 \triangleq & -2\gamma_1 + 2\ln(J_1) + n_1\ln(1 + P_1) + n_{2,2} \left( 1 + \ln \left( \frac{\sigma_{y_2}^2}{\sigma_{z^*}^2} \right) - \frac{\delta_y}{\sigma_{y_2}^2} \right) \\ & - \frac{n_{2,2}}{\sigma_{z^*}^2} \left( \sqrt{\sigma_{y_2}^2 + \delta_y} + \sqrt{\sigma_{u^*}^2 - 1} \right)^2, \end{aligned} \quad (23a)$$

$$\begin{aligned} \tilde{\gamma}_2 \triangleq & -2\gamma_2 + 2\ln(J_2) + n_3\ln(1 + P_3) + n_{2,2} \left( 1 + \ln \left( \frac{\sigma_{y_2}^2}{\sigma_{u^*}^2} \right) - \frac{\delta_y}{\sigma_{y_2}^2} \right) \\ & - \frac{n_{2,2}}{\sigma_{u^*}^2} \left( \sqrt{\sigma_{y_2}^2 + \delta_y} + (1 - \alpha)\sqrt{P_t + \delta_2/n_{2,2}} \right)^2. \end{aligned} \quad (23b)$$

## 4 Main Result

Fix  $\beta_1 \in [0, 1]$  and  $\beta_2 \in [0, 1]$  such that  $\beta_1 + \beta_2 < 1$  and define

$$Q_1 \triangleq v_1 - v_2 \quad \text{and} \quad Q_2 \triangleq v'_1 - v'_2, \quad (28)$$

with  $v_1 \sim \mathcal{X}^2(n_1)$ ,  $v_2 \sim \mathcal{X}^2(n_1)$ ,  $v'_1 \sim \mathcal{X}^2(n_3)$  and  $v'_2 \sim \mathcal{X}^2(n_3)$  where  $\mathcal{X}^2(s)$  is central chi-squared distribution of degree  $s$ . Also define

$$\sigma_{y_2^*}^2 \triangleq 1 + (\beta_1 + \beta_2)P_2, \quad (29a)$$

$$\sigma_{z^*}^2 \triangleq 1 + (1 - \alpha)^2\beta_2P_2, \quad (29b)$$

$$\sigma_{u^*}^2 \triangleq 1 + \beta_1P_2 + \alpha^2P_t. \quad (29c)$$

By employing the scheme proposed in Section 3 we have the following theorem on the upper bounds on the error probabilities  $\epsilon_1$  and  $\epsilon_2$ .

**Theorem 1 (Achievability bounds)** *For a fixed message set size  $M$ , the error probabilities  $\epsilon_1$  and  $\epsilon_2$  are upper bonded by*

$$\begin{aligned} \epsilon_1 \leq & 1 - F_{Q_1}(\tilde{\gamma}_1) + ML^{(P_t)}e^{-\gamma_1}(1 - e^{-n_{2,2}^{1/3}\kappa}) + \epsilon_T \\ & + \Gamma\left(\frac{n_{2,2}}{2}, \frac{n_{2,2}P_{\max}}{2\beta_2P_2}\right)/\Gamma\left(\frac{n_{2,2}}{2}\right) + G + e^{-n_{2,2}^{1/3}\kappa}, \end{aligned} \quad (30)$$

$$\begin{aligned} \epsilon_2 \leq & 1 - F_{Q_2}(\tilde{\gamma}_2) + Me^{-\gamma_2}(1 - e^{-n_{2,2}^{1/3}\kappa}) \\ & + \Gamma\left(\frac{n_{2,2}}{2}, \frac{n_{2,2}}{2}\right)/\Gamma\left(\frac{n_{2,2}}{2}\right) + G + e^{-n_{2,2}^{1/3}\kappa}, \end{aligned} \quad (31)$$

for any  $\gamma_1$  and  $\gamma_2$ , where  $\tilde{\gamma}_1$  and  $\tilde{\gamma}_2$  are defined in (23), and

$$G \triangleq \left(1 - A(\beta_1P_2/(\beta_1P_2 + \alpha^2P_t))^{\frac{n_{2,2}}{2}}\right)^{L^{(P_t)}}, \quad (32a)$$

$$\begin{aligned} A \triangleq & \frac{\delta_1(\beta_1P_2 + \alpha^2P_t)}{2\alpha n_{2,2}\beta_1P_2\sqrt{\pi P_{\max}}\beta_1P_2} \cdot \frac{\Gamma\left(\frac{n_{2,2}}{2}\right)}{\Gamma\left(\frac{n_{2,2}-1}{2}\right)} \\ & \times \left(1 - (c + \alpha^2\delta_2 + \delta_1\sqrt{P_t/P_{\max}})/n_{2,2}\beta_1P_2\right)^{\frac{n_{2,2}-3}{2}} \end{aligned} \quad (32b)$$

$$c \triangleq \left(\alpha\delta_2/(2\sqrt{n_{2,2}P_t}) + \delta_1/(2\alpha\sqrt{n_{2,2}P_{\max}})\right)^2. \quad (32c)$$

Note that  $J_1$ ,  $J_2$ ,  $\kappa$ ,  $\delta_1$ ,  $\delta_2$  and  $\delta_y$  are constants and  $\Gamma(\cdot)$  is the gamma function,  $\Gamma(\cdot, \cdot)$  is the upper incomplete gamma function,  $F_{Q_1}(\cdot)$  and  $F_{Q_2}(\cdot)$  are the cumulative distribution functions of  $Q_1$  and  $Q_2$ , respectively.

*Proof:* See Appendix A. ■

### 4.1 Numerical Results

In Figure 2, we evaluate the bounds in Theorem 1 for the case where the consecutive messages  $m_1$  and  $m_2$  are each of a transmission window of 20 channel uses, i.e.,  $|d_1 - a_1 + 1| = |d_2 - a_2 + 1| = 20$ . The solid lines correspond to the error probabilities under our scheme and the dashed lines to time sharing scheme, where each message is allocated the same number of channel uses.

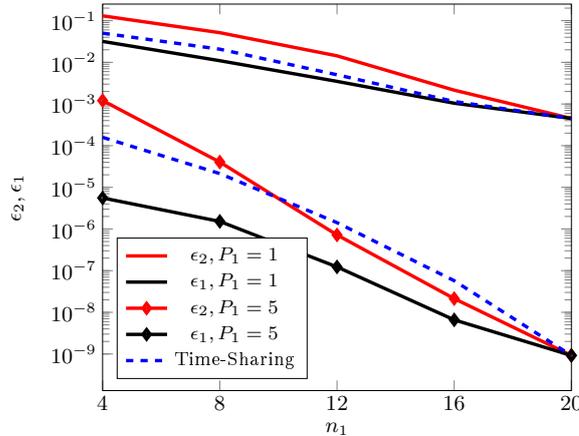


Figure 2: Upper bounds on  $\epsilon_1$  and  $\epsilon_2$  in terms of  $n_1$  for  $P_1 = P_2 = P_3$ ,  $P_{\max} = 2P_1$ ,  $\beta_1 = 0.4$ ,  $\beta_2 = 0.4$ ,  $R_1 = R_2 = 0.3466$  and  $\epsilon_T = 10^{-10}$ .

The bounds are evaluated for different values of  $n_1$ , different transmit power levels and for  $\alpha$  set at  $\frac{\beta_1 P_2}{1 + \beta_2 P_2}$ . Recall that  $n_1$  is equal to the number of channel uses between the arrivals of  $m_1$  and  $m_2$ . As  $n_1$  increases, the transmissions of  $m_1$  and  $m_2$  overlap over a fewer number of channel uses and thus their corresponding error probabilities decrease. The minimum error probabilities occur at  $n_1 = 20$ . In the time-sharing scheme, for example when  $n_1 = 4$ , we assign 12 channel uses to the transmission of each message. Note that in this figure we assign  $(1 - \beta_1 - \beta_2)P_2$  to the transmission of the power type. One can increase this power up to  $\frac{n_2}{n_{2,1}}(1 - \frac{n_{2,2}}{n_2}(\beta_1 + \beta_2))P_2$  which still satisfies the power constraint in (3).

Observe that when the transmit power is small, the error probabilities obtained under the time sharing scheme are slightly lower than the bound on  $\epsilon_2$ . For larger transmit powers, however, the error probabilities of *both* messages outperform time sharing over a wide range of  $n_1$ .

## 5 Conclusions

We considered a point-to-point communication where messages arrive at different times and are subject to heterogeneous decoding delay constraints. We proposed a coding scheme exploiting the dirty paper coding principle to jointly transmit consecutive messages with heterogeneous decoding deadlines. We derived rigorous bounds on the achievable error probabilities for each of the messages. We also numerically analyzed the obtained bounds in the finite block length regime and identified potential performance improvements over standard time sharing schemes.

## Acknowledgment

This work was funded by INRIA Nokia Bell Labs ADR “Network Information Theory” and by the French National Agency for Research (ANR) under grant ANR-16-CE25-0001 - ARBURST.

## A Proof of Theorem 1

### A.1 Analyzing $\epsilon_1$

The average error probability  $\epsilon_1$  is lower bounded based on the four error events  $\mathcal{E}_{1,1}$ ,  $\mathcal{E}_{1,2}$ ,  $\mathcal{E}_{1,3}$  and  $\mathcal{E}_{1,4}$  defined in (15), (20), (22) and (25), respectively and is as following:

$$\epsilon_1 \leq \mathbb{P}[\mathcal{E}_{1,1}] + \mathbb{P}[\mathcal{E}_{1,2}|\mathcal{E}_{1,1}^c] + \mathbb{P}[\mathcal{E}_{1,3}|\mathcal{E}_{1,1}^c, \mathcal{E}_{1,2}^c] + \mathbb{P}[\mathcal{E}_{1,4}|\mathcal{E}_{1,1}^c, \mathcal{E}_{1,2}^c, \mathcal{E}_{1,3}^c]. \quad (33)$$

#### A.1.1 Analyzing $\mathbb{P}[\mathcal{E}_{1,1}]$

We start by analyzing the probability that the power type of the signal  $\mathbf{X}_{2,2}$  denoted by  $P_t$  does not belong to the typical set  $\tilde{\mathcal{P}}_{n_{2,2}}$ .

$$\mathbb{P}[\mathcal{E}_{1,1}] = \mathbb{P}[P_t \notin \tilde{\mathcal{P}}_{n_{2,2}}] \quad (34)$$

$$= \mathbb{P}[\|\mathbf{X}_{2,2}\|^2 > n_{2,2}P_{\max}] \quad (35)$$

$$= \mathbb{P}\left[\left\|\frac{\mathbf{X}_{2,2}}{\sqrt{\beta_2 P_2}}\right\|^2 > \frac{n_{2,2}P_{\max}}{\beta_2 P_2}\right] \quad (36)$$

$$= 1 - F\left(\frac{n_{2,2}P_{\max}}{\beta_2 P_2}; n_{2,2}\right) \quad (37)$$

$$= 1 - \frac{\gamma\left(\frac{n_{2,2}}{2}, \frac{n_{2,2}P_{\max}}{2\beta_2 P_2}\right)}{\Gamma\left(\frac{n_{2,2}}{2}\right)} \quad (38)$$

$$= 1 - \frac{\Gamma\left(\frac{n_{2,2}}{2}\right) - \Gamma\left(\frac{n_{2,2}}{2}, \frac{n_{2,2}P_{\max}}{2\beta_2 P_2}\right)}{\Gamma\left(\frac{n_{2,2}}{2}\right)} \quad (39)$$

$$= \frac{\Gamma\left(\frac{n_{2,2}}{2}, \frac{n_{2,2}P_{\max}}{2\beta_2 P_2}\right)}{\Gamma\left(\frac{n_{2,2}}{2}\right)} \quad (40)$$

where in (36), the variable  $\frac{\mathbf{X}_{2,2}}{\sqrt{\beta_2 P_2}} \sim \mathcal{N}(0, I_{n_{2,2}})$  and therefore  $\left\|\frac{\mathbf{X}_{2,2}}{\sqrt{\beta_2 P_2}}\right\|^2$  has a central chi-squared distribution of degree  $n_{2,2}$ . In (37),  $F(\cdot; \cdot)$  is the cumulative distribution function (CDF) of the central chi-squared distribution and in (38) this CDF is replaced by its following form

$$F(x; k) = \frac{\gamma\left(\frac{k}{2}, \frac{x}{2}\right)}{\Gamma\left(\frac{k}{2}\right)}, \quad (41)$$

where  $\gamma(\cdot, \cdot)$  is the lower incomplete gamma function and  $\Gamma(\cdot)$  is the gamma function. In (39), the lower incomplete gamma function is replaced by

$$\gamma(s, x) = \Gamma(s) - \Gamma(s, x), \quad (42)$$

where  $\Gamma(\cdot, \cdot)$  is the upper incomplete gamma function.

#### A.1.2 Analyzing $\mathbb{P}[\mathcal{E}_{1,2}|\mathcal{E}_{1,1}^c]$

To calculate this probability we follow the same argument provided in [10, Appendix E]. Recall that for a fixed power type  $P_t$ , we have  $\|\mathbf{U}\|^2 = n_{2,2}(\beta_1 P_2 + \alpha^2 P_t)$  almost surely. From (19) we have that  $\mathbf{U} - \alpha \mathbf{X}_{2,2} \in \mathcal{D}$  if and only if

$$n_{2,2}\beta_1 P_2 - \delta_1 \leq \|\mathbf{U}\|^2 + \alpha^2 \|\mathbf{X}_{2,2}\|^2 - 2\alpha \langle \mathbf{X}_{2,2}, \mathbf{U} \rangle \leq n_{2,2}\beta_1 P_2 \quad (43)$$

or equivalently

$$\frac{n_{2,2}\alpha P_t}{2} \leq \langle \mathbf{X}_{2,2}, \mathbf{U} \rangle - \frac{\alpha}{2} \|\mathbf{X}_{2,2}\|^2 \leq \frac{n_{2,2}\alpha P_t}{2} + \frac{\delta_1}{2\alpha}. \quad (44)$$

Since  $\mathbf{U}$  is drawn uniformly from the sphere, the distribution of  $\langle \mathbf{X}_{2,2}, \mathbf{U} \rangle$  depends on  $\mathbf{X}_{2,2}$  only through its magnitude, this is seen by noting that the inner product of two vectors is unchanged when an orthogonal transformation is applied to both arguments, and the distribution of  $\mathbf{U}$  is unchanged under any orthogonal transformation. Thus assume that  $\mathbf{X}_{2,2} = (\|\mathbf{X}_{2,2}\|, 0, \dots, 0)$ . In this case,

$$\frac{n_{2,2}\alpha P_t}{2} \leq u_1 \|\mathbf{X}_{2,2}\| - \frac{\alpha}{2} \|\mathbf{X}_{2,2}\|^2 \leq \frac{n_{2,2}\alpha P_t}{2} + \frac{\delta_1}{2\alpha} \quad (45)$$

or equivalently

$$\frac{n_{2,2}\alpha P_t}{2\|\mathbf{X}_{2,2}\|} + \frac{\alpha}{2} \|\mathbf{X}_{2,2}\| \leq u_1 \leq \frac{n_{2,2}\alpha P_t}{2\|\mathbf{X}_{2,2}\|} + \frac{\delta_1}{2\alpha\|\mathbf{X}_{2,2}\|} + \frac{\alpha}{2} \|\mathbf{X}_{2,2}\|. \quad (46)$$

According to the definition of the typical power set in (14), for  $P_t \in \tilde{\mathcal{P}}_{n_{2,2}}$ , we have  $\|\mathbf{X}_{2,2}\| \leq \sqrt{n_{2,2}P_{\max}}$ , thus

$$\frac{n_{2,2}\alpha P_t}{2\|\mathbf{X}_{2,2}\|} + \frac{\alpha}{2} \|\mathbf{X}_{2,2}\| \leq u_1 \leq \frac{n_{2,2}\alpha P_t}{2\|\mathbf{X}_{2,2}\|} + \frac{\alpha}{2} \|\mathbf{X}_{2,2}\| + \frac{\delta_1}{2\alpha\sqrt{n_{2,2}P_{\max}}}. \quad (47)$$

We conclude that  $\mathbb{P}[\mathbf{U} - \alpha\mathbf{X}_{2,2} \in \mathcal{D}]$  is lower bounded by the probability of the first entry of  $\mathbf{U}$  falling within an interval of length  $\frac{\delta_1}{2\alpha\sqrt{n_{2,2}P_{\max}}}$  starting at  $\frac{n_{2,2}\alpha P_t}{2\|\mathbf{X}_{2,2}\|} + \frac{\alpha}{2} \|\mathbf{X}_{2,2}\|$ . The distribution of a given symbol in a length- $n_{2,2}$  random sequence distributed uniformly on the sphere is [16]

$$f_{U_1}(u_1) = \frac{1}{\sqrt{\pi n_{2,2}(\beta_1 P_2 + \alpha^2 P_t)}} \frac{\Gamma(\frac{n_{2,2}}{2})}{\Gamma(\frac{n_{2,2}-1}{2})} \left(1 - \frac{u_1^2}{n_{2,2}(\beta_1 P_2 + \alpha^2 P_t)}\right)^{\frac{n_{2,2}-3}{2}} \times \mathbb{1}\{u_1 \leq n_{2,2}(\beta_1 P_2 + \alpha^2 P_t)\}. \quad (48)$$

This density function is decreasing in  $u_1^2$ , which implies that

$$\mathbb{P}[\mathbf{U} - \alpha\mathbf{X}_{2,2} \in \mathcal{D}] \geq \frac{\delta_1}{2\alpha\sqrt{n_{2,2}P_t}} f_{U_1} \left( \frac{n_{2,2}\alpha P_t}{2\|\mathbf{X}_{2,2}\|} + \frac{\alpha}{2} \|\mathbf{X}_{2,2}\| + \frac{\delta_1}{2\alpha\sqrt{n_{2,2}P_{\max}}} \right) \quad (49)$$

Conditioning on  $P_t \in \tilde{\mathcal{P}}_{n_{2,2}}$ , then from (9) we have

$$\frac{n_{2,2}\alpha P_t}{2\|\mathbf{X}_{2,2}\|} + \frac{\alpha}{2} \|\mathbf{X}_{2,2}\| \leq \alpha \|\mathbf{X}_{2,2}\| \leq \alpha \sqrt{n_{2,2}P_t + \delta_2} \quad (50)$$

$$\leq \alpha \sqrt{n_{2,2}P_t} + \frac{\alpha\delta_2}{2\sqrt{n_{2,2}P_t}} \quad (51)$$

where the last inequality is because  $\sqrt{1+a} \leq 1 + \frac{a}{2}$ . Thus

$$\left( \frac{n_{2,2}\alpha P_t}{2\|\mathbf{X}_{2,2}\|} + \frac{\alpha}{2} \|\mathbf{X}_{2,2}\| + \frac{\delta_1}{2\alpha\sqrt{n_{2,2}P_{\max}}} \right)^2 \quad (52)$$

$$\leq \left( \alpha \sqrt{n_{2,2}P_t} + \frac{\alpha\delta_2}{2\sqrt{n_{2,2}P_t}} + \frac{\delta_1}{2\alpha\sqrt{n_{2,2}P_{\max}}} \right)^2 \quad (53)$$

$$\begin{aligned}
&= n_{2,2}\alpha^2 P_t + 2\alpha\sqrt{n_{2,2}P_t} \left( \frac{\alpha\delta_2}{2\sqrt{n_{2,2}P_t}} + \frac{\delta_1}{2\alpha\sqrt{n_{2,2}P_{\max}}} \right) \\
&\quad + \left( \frac{\alpha\delta_2}{2\sqrt{n_{2,2}P_t}} + \frac{\delta_1}{2\alpha\sqrt{n_{2,2}P_{\max}}} \right)^2
\end{aligned} \tag{54}$$

$$\leq n_{2,2}\alpha^2 P_t + \alpha^2\delta_2 + \sqrt{\frac{P_t}{P_{\max}}}\delta_1 + \left( \frac{\alpha\delta_2}{2\sqrt{n_{2,2}P_t}} + \frac{\delta_1}{2\alpha\sqrt{n_{2,2}P_{\max}}} \right)^2 \tag{55}$$

$$= n_{2,2}\alpha^2 P_t + \alpha^2\delta_2 + \sqrt{\frac{P_t}{P_{\max}}}\delta_1 + c, \tag{56}$$

where

$$c \triangleq \left( \frac{\alpha\delta_2}{2\sqrt{n_{2,2}P_t}} + \frac{\delta_1}{2\alpha\sqrt{n_{2,2}P_{\max}}} \right)^2. \tag{57}$$

By substituting (56) into (49), and again using the fact that  $f_{U_1}(u_1)$  is a decreasing function of  $u_1^2$ , we have

$$\mathbb{P}[U - \alpha\mathbf{X}_{2,2} \in \mathcal{D}] \geq A \cdot \left( \frac{\beta_1 P_2}{\beta_1 P_2 + \alpha^2 P_t} \right)^{\frac{n_{2,2}}{2}} \tag{58}$$

where

$$A \triangleq \frac{\delta_1(\beta_1 P_2 + \alpha^2 P_t)}{2\alpha n_{2,2} \beta_1 P_2 \sqrt{\pi P_{\max}} \beta_1 P_2} \cdot \left( 1 - \frac{c + \alpha^2\delta_2 + \sqrt{\frac{P_t}{P_{\max}}}\delta_1}{n_{2,2}\beta_1 P_2} \right)^{\frac{n_{2,2}-3}{2}} \cdot \frac{\Gamma(\frac{n_{2,2}}{2})}{\Gamma(\frac{n_{2,2}-1}{2})}. \tag{59}$$

Since the  $L^{(P_t)}$  codewords are generated independently, thus the probability of the error event  $\mathcal{E}_1$  conditioning on that  $\mathbf{X}_{2,2}$  has a given type  $P_t$  is equal to

$$\mathbb{P}[\mathcal{E}_{1,2} | \mathcal{E}_{1,1}^c] \leq \left( 1 - A \left( \frac{\beta_1 P_2}{\beta_1 P_2 + \alpha^2 P_t} \right)^{\frac{n_{2,2}}{2}} \right)^{L^{(P_t)}}. \tag{60}$$

### A.1.3 Analyzing $\mathbb{P}[\mathcal{E}_{1,3} | \mathcal{E}_{1,1}^c, \mathcal{E}_{1,2}^c]$

In our scheme, we fix the error probability of decoding the power type at  $\epsilon_T \in [0, 1]$ . Recall that the sensor transmits the power type over an AWGN channel during  $n_{2,1}$  channel uses. Then for a given  $P_{\max}$ ,  $P_2$ ,  $\beta_2$ ,  $n_2$  and  $\delta_2$ , we fix  $n_{1,2}$  at the smallest value such that this probability is less than  $\epsilon_T$ .

### A.1.4 Analyzing $\mathbb{P}[\mathcal{E}_{1,4} | \mathcal{E}_{1,1}^c, \mathcal{E}_{1,2}^c, \mathcal{E}_{1,3}^c]$

Note that the output sequence  $\tilde{\mathbf{Y}}_2$  does not follow a Gaussian distribution due to the fact that the vector  $\mathbf{U}$  is not Gaussian. Thus similar to [14, Section IV.B], for the outputs of the third channel, we define the following set of typical channel outputs

$$\mathcal{F} \triangleq \{ \tilde{\mathbf{y}}_2 \in \mathbb{R}^{n_{2,2}} : \frac{1}{n_{2,2}} \|\tilde{\mathbf{y}}_2\|^2 \in [\sigma_{y_2^*}^2 - \delta_y, \sigma_{y_2^*}^2 + \delta_y] \}. \tag{61}$$

for a fixed  $\delta_y > 0$  where  $\sigma_{y_2^*}^2$  is defined in (29a). By *Cramer's theorem* in [12], we have

$$\mathcal{E}_y \triangleq \mathbb{P}[\tilde{\mathbf{Y}}_2 \notin \mathcal{F}] < \exp(-n_{2,2}\kappa\delta_y^2) \quad (62)$$

for some constant  $\kappa > 0$ . By setting  $\delta_y = n_{2,2}^{-1/3}$ , then  $\exp(-n_{2,2}\kappa\delta_y^2)$  decays faster than any polynomial.

To evaluate this error event, we use the threshold bound for maximum-metric decoding. I.e.,

$$\begin{aligned} \mathbb{P}[\mathcal{E}_{1,4} | \mathcal{E}_{1,1}^c, \mathcal{E}_{1,2}^c, \mathcal{E}_{1,3}^c] &\leq \mathcal{E}_y + \mathbb{P}[i_1(\mathbf{U}, \mathbf{X}_1; \mathbf{Y}_1, \tilde{\mathbf{Y}}_2) \leq \gamma_1 | \tilde{\mathbf{Y}}_2 \in \mathcal{F}] \mathcal{E}_y^c \\ &\quad + ML^{(P_t)} \mathbb{P}[i_1(\bar{\mathbf{U}}, \bar{\mathbf{X}}_1; \mathbf{Y}_1, \tilde{\mathbf{Y}}_2) > \gamma_1 | \tilde{\mathbf{Y}}_2 \in \mathcal{F}] \mathcal{E}_y^c \end{aligned} \quad (63)$$

for any  $\gamma_1$ , where  $\bar{\mathbf{U}} \sim f_U$  and  $\bar{\mathbf{X}}_1 \sim f_{\mathbf{X}_1}$  are independent of  $(\mathbf{X}_{2,2}, \mathbf{U}, \mathbf{X}_1, \mathbf{Y}_1, \tilde{\mathbf{Y}}_2)$ .

We start by bounding  $\mathbb{P}[i_1(\bar{\mathbf{U}}, \bar{\mathbf{X}}_1; \mathbf{Y}_1, \tilde{\mathbf{Y}}_2) > \gamma_1]$ . By Bayes rule we have

$$f_{\mathbf{X}_1}(\bar{\mathbf{x}}_1) f_U(\bar{\mathbf{u}}) = \frac{f_{\mathbf{Y}_1}(\mathbf{y}_1) f_{\tilde{\mathbf{Y}}_2}(\tilde{\mathbf{y}}_2) f_{U|\tilde{\mathbf{Y}}_2}(\bar{\mathbf{u}}|\tilde{\mathbf{y}}_2) f_{\mathbf{X}_1|\mathbf{Y}_1}(\bar{\mathbf{x}}_1|\mathbf{y}_1)}{f_{\mathbf{Y}_1|\mathbf{X}_1}(\mathbf{y}_1|\bar{\mathbf{x}}_1) f_{\tilde{\mathbf{Y}}_2|U}(\tilde{\mathbf{y}}_2|\bar{\mathbf{u}})} \quad (64)$$

$$= f_{U|\tilde{\mathbf{Y}}_2}(\bar{\mathbf{u}}|\tilde{\mathbf{y}}_2) f_{\mathbf{X}_1|\mathbf{Y}_1}(\bar{\mathbf{x}}_1|\mathbf{y}_1) \exp(-i_1(\bar{\mathbf{u}}, \bar{\mathbf{x}}_1; \mathbf{y}_1, \tilde{\mathbf{y}}_2)) \quad (65)$$

For fixed sequences  $\mathbf{y}_1$  and  $\tilde{\mathbf{y}}_2 \in \mathcal{F}$ , by multiplying both side of the above equation by  $\mathbb{1}\{i_1(\bar{\mathbf{u}}, \bar{\mathbf{x}}_1; \mathbf{y}_1, \tilde{\mathbf{y}}_2) > \gamma_1\}$  and integrating over all  $\bar{\mathbf{x}}_1$  and  $\bar{\mathbf{u}}$ , we have

$$\int_{\bar{\mathbf{x}}_1} \int_{\bar{\mathbf{u}}} \mathbb{1}\{i_1(\bar{\mathbf{u}}, \bar{\mathbf{x}}_1; \mathbf{y}_1, \tilde{\mathbf{y}}_2) > \gamma_1\} f_{\mathbf{X}_1}(\bar{\mathbf{x}}_1) f_U(\bar{\mathbf{u}}) d\bar{\mathbf{u}} d\bar{\mathbf{x}}_1 \quad (66)$$

$$\begin{aligned} &= \int_{\bar{\mathbf{x}}_1} \int_{\bar{\mathbf{u}}} \exp(-i_1(\bar{\mathbf{u}}, \bar{\mathbf{x}}_1; \mathbf{y}_1, \tilde{\mathbf{y}}_2)) \\ &\quad \times \mathbb{1}\{i_1(\bar{\mathbf{u}}, \bar{\mathbf{x}}_1; \mathbf{y}_1, \tilde{\mathbf{y}}_2) > \gamma_1\} f_{U|\tilde{\mathbf{Y}}_2}(\bar{\mathbf{u}}|\tilde{\mathbf{y}}_2) f_{\mathbf{X}_1|\mathbf{Y}_1}(\bar{\mathbf{x}}_1|\mathbf{y}_1) d\bar{\mathbf{u}} d\bar{\mathbf{x}}_1. \end{aligned} \quad (67)$$

$$\begin{aligned} &= \int_{\bar{\mathbf{x}}_1} \int_{\bar{\mathbf{u}}} \exp(-i_1(\bar{\mathbf{u}}, \bar{\mathbf{x}}_1; \mathbf{y}_1, \tilde{\mathbf{y}}_2)) \\ &\quad \times \mathbb{1}\left\{ \frac{f_{\mathbf{Y}_1|\mathbf{X}_1}(\mathbf{y}_1|\bar{\mathbf{x}}_1) f_{\tilde{\mathbf{Y}}_2|U}(\tilde{\mathbf{y}}_2|\bar{\mathbf{u}})}{f_{\mathbf{Y}_1}(\mathbf{y}_1) f_{\tilde{\mathbf{Y}}_2}(\tilde{\mathbf{y}}_2)} > e^{\gamma_1} \right\} f_{U|\tilde{\mathbf{Y}}_2}(\bar{\mathbf{u}}|\tilde{\mathbf{y}}_2) f_{\mathbf{X}_1|\mathbf{Y}_1}(\bar{\mathbf{x}}_1|\mathbf{y}_1) d\bar{\mathbf{u}} d\bar{\mathbf{x}}_1. \end{aligned} \quad (68)$$

$$\begin{aligned} &\leq \int_{\bar{\mathbf{x}}_1} \int_{\bar{\mathbf{u}}} \exp(-i_1(\bar{\mathbf{u}}, \bar{\mathbf{x}}_1; \mathbf{y}_1, \tilde{\mathbf{y}}_2)) \\ &\quad \times \frac{f_{\mathbf{Y}_1|\mathbf{X}_1}(\mathbf{y}_1|\bar{\mathbf{x}}_1) f_{\tilde{\mathbf{Y}}_2|U}(\tilde{\mathbf{y}}_2|\bar{\mathbf{u}})}{f_{\mathbf{Y}_1}(\mathbf{y}_1) f_{\tilde{\mathbf{Y}}_2}(\tilde{\mathbf{y}}_2)} e^{-\gamma_1} f_{U|\tilde{\mathbf{Y}}_2}(\bar{\mathbf{u}}|\tilde{\mathbf{y}}_2) f_{\mathbf{X}_1|\mathbf{Y}_1}(\bar{\mathbf{x}}_1|\mathbf{y}_1) d\bar{\mathbf{u}} d\bar{\mathbf{x}}_1. \end{aligned} \quad (69)$$

$$\leq e^{-\gamma_1} \quad (70)$$

Equivalently

$$\mathbb{P}[i_1(\bar{\mathbf{U}}, \bar{\mathbf{X}}_1; \mathbf{Y}_1, \tilde{\mathbf{Y}}_2) > \gamma_1 | \tilde{\mathbf{Y}}_2 \in \mathcal{F}] \leq e^{-\gamma_1}. \quad (71)$$

Thus

$$\mathbb{P}[i_1(\bar{\mathbf{U}}, \bar{\mathbf{X}}_1; \mathbf{Y}_1, \tilde{\mathbf{Y}}_2) > \gamma_1 | \tilde{\mathbf{Y}}_2 \in \mathcal{F}] \cdot \mathcal{E}_y^c \leq (1 - e^{-n_{2,2}^{1/3}\kappa}) e^{-\gamma_1}. \quad (72)$$

To calculate  $\mathbb{P}[i_1(\mathbf{U}, \mathbf{X}_1; \mathbf{Y}_1, \tilde{\mathbf{Y}}_2) \leq \gamma_1 | \tilde{\mathbf{Y}}_2 \in \mathcal{F}]$ , for a fixed  $P_t$  we first define the output  $\mathbf{Y}_2^* \sim \mathcal{N}(0, I_{n_{2,2}} \sigma_{y_2^*}^2)$ . Then by [5, proof of Lemma 61], we have

$$\min_{\tilde{\mathbf{y}}_2 \in \mathcal{F}} \frac{f_{\tilde{\mathbf{Y}}_2}(\tilde{\mathbf{y}}_2)}{f_{\mathbf{Y}_2^*}(\tilde{\mathbf{y}}_2)} \leq j_1 \quad (73)$$

for a finite constant  $j_1 > 0$ .

To calculate  $f_{\tilde{\mathbf{Y}}_2|U}(\tilde{\mathbf{Y}}_2|U)$ , we define vector  $\mathbf{U}^*$  following the distribution  $\mathcal{N}(0, (\beta_1 P_2 + \alpha^2 P_t)I_{n_{2,2}})$ . Then by [5, proof of Lemma 61] and also by [17, Proposition 2]:

$$\min_{\mathbf{u}: \|\mathbf{u}\|^2 = n_{2,2}(\beta_1 P_2 + \alpha^2 P_t)} \frac{f_U(\mathbf{u})}{f_{U^*}(\mathbf{u})} \leq j_2 \quad (74)$$

where  $j_2 > 0$  is a constant. We also define  $f_{U^*, \mathbf{Y}_2^*}(\mathbf{u}, \tilde{\mathbf{y}}_2)$  as the joint distribution of  $\mathbf{Y}_2^*$  and  $U^*$ . We have

$$\frac{f_{U, \tilde{\mathbf{Y}}_2}(\mathbf{u}, \tilde{\mathbf{y}}_2)}{f_{U^*, \mathbf{Y}_2^*}(\mathbf{u}, \tilde{\mathbf{y}}_2)} = D_{f,Q}^* \quad (75)$$

where it can be shown that there is a constant  $j_3$  such that

$$D_{f,Q}^* \geq j_3. \quad (76)$$

By combining (73) and (74), we have

$$\frac{f_{\tilde{\mathbf{Y}}_2|U}(\tilde{\mathbf{y}}_2|\mathbf{u})}{f_{\tilde{\mathbf{Y}}_2}(\tilde{\mathbf{y}}_2)} \geq \frac{f_{U|\tilde{\mathbf{Y}}_2}(\mathbf{u}|\tilde{\mathbf{y}}_2)}{j_2 f_{U^*}(\mathbf{u})} \quad (77)$$

$$= \frac{f_{U, \tilde{\mathbf{Y}}_2}(\mathbf{u}, \tilde{\mathbf{y}}_2)}{j_2 f_{U^*}(\mathbf{u}) f_{\tilde{\mathbf{Y}}_2}(\tilde{\mathbf{y}}_2)} \quad (78)$$

$$\geq \frac{f_{U, \tilde{\mathbf{Y}}_2}(\mathbf{u}, \tilde{\mathbf{y}}_2)}{j_1 j_2 f_{U^*}(\mathbf{u}) f_{\mathbf{Y}_2^*}(\tilde{\mathbf{y}}_2)} \quad (79)$$

$$\geq \frac{j_3}{j_1 j_2} \cdot \frac{f_{U^*, \mathbf{Y}_2^*}(\mathbf{u}, \tilde{\mathbf{y}}_2)}{f_{U^*}(\mathbf{u}) f_{\mathbf{Y}_2^*}(\tilde{\mathbf{y}}_2)} \quad (80)$$

$$= J_1 \frac{f_{\mathbf{Y}_2^*|U^*}(\tilde{\mathbf{y}}_2, \mathbf{u})}{f_{\mathbf{Y}_2^*}(\tilde{\mathbf{y}}_2)} \quad (81)$$

where  $J_1 \triangleq \frac{j_3}{j_1 j_2}$ . We thus have

$$\mathcal{E}_y^c \cdot \mathbb{P}[i_1(U, \mathbf{X}_1; \mathbf{Y}_1, \tilde{\mathbf{Y}}_2) \leq \gamma_1 | \tilde{\mathbf{Y}}_2 \in \mathcal{F}] \quad (82)$$

$$\leq \mathbb{P} \left[ \ln \left( \frac{f_{\mathbf{Y}_1|\mathbf{X}_1}(\mathbf{Y}_1|\mathbf{X}_1) f_{\mathbf{Y}_2^*|U^*}(\tilde{\mathbf{Y}}_2|U)}{f_{\mathbf{Y}_1}(\mathbf{Y}_1) f_{\mathbf{Y}_2^*}(\tilde{\mathbf{Y}}_2)} \right) \leq \gamma_1 - \ln(J_1) | \tilde{\mathbf{Y}}_2 \in \mathcal{F} \right] \cdot \mathcal{E}_y^c \quad (83)$$

$$= \mathbb{P} \left[ \ln \left( \frac{\frac{1}{(\sqrt{2\pi})^{n_1}} \exp\left(-\frac{\|\mathbf{Y}_1 - \mathbf{X}_1\|^2}{2}\right)}{\frac{1}{(\sqrt{2\pi(1+P_1)})^{n_1}} \exp\left(-\frac{\|\mathbf{Y}_1\|^2}{2(1+P_1)}\right)} \right) \right. \\ \left. + \ln \left( \frac{\frac{1}{(\sqrt{2\pi\sigma_{z^*}^2})^{n_{2,2}}} \exp\left(-\frac{\|\tilde{\mathbf{Y}}_2 - U\|^2}{2\sigma_{z^*}^2}\right)}{\frac{1}{(\sqrt{2\pi\sigma_{y_2^*}^2})^{n_{2,2}}} \exp\left(-\frac{\|\tilde{\mathbf{Y}}_2\|^2}{2\sigma_{y_2^*}^2}\right)} \right) \leq \gamma_1 - \ln(J_1) | \tilde{\mathbf{Y}}_2 \in \mathcal{F} \right] \cdot \mathcal{E}_y^c \quad (84)$$

$$= \mathbb{P} \left[ \frac{n_1}{2} \ln(1+P_1) - \frac{\|\mathbf{Y}_1 - \mathbf{X}_1\|^2}{2} + \frac{\|\mathbf{Y}_1\|^2}{2(1+P_1)} + \frac{n_{2,2}}{2} \ln \left( \frac{\sigma_{y_2^*}^2}{\sigma_{z^*}^2} \right) \right. \\ \left. - \frac{\|\tilde{\mathbf{Y}}_2 - U\|^2}{2\sigma_{z^*}^2} + \frac{\|\tilde{\mathbf{Y}}_2\|^2}{2\sigma_{y_2^*}^2} \leq \gamma_1 - \ln(J_1) | \tilde{\mathbf{Y}}_2 \in \mathcal{F} \right] \cdot \mathcal{E}_y^c \quad (85)$$

$$= \mathbb{P} \left[ \|\mathbf{Y}_1 - \mathbf{X}_1\|^2 - \frac{\|\mathbf{Y}_1\|^2}{(1+P_1)} + \frac{\|\tilde{\mathbf{Y}}_2 - \mathbf{U}\|^2}{\sigma_{z^*}^2} - \frac{\|\tilde{\mathbf{Y}}_2\|^2}{\sigma_{y_2^*}^2} \geq \gamma_1'' | \tilde{\mathbf{Y}}_2 \in \mathcal{F} \right] \cdot \mathcal{E}_y^c \quad (86)$$

$$\leq \mathbb{P} \left[ \|\mathbf{Z}_1\|^2 - \frac{1}{1+P_1} \|\mathbf{Y}_1\|^2 + \frac{\|\tilde{\mathbf{Y}}_2 - \mathbf{U}\|^2}{\sigma_{z^*}^2} \geq \gamma_1'' + n_{2,2} - \frac{n_{2,2}\delta_y}{\sigma_{y_2^*}^2} | \tilde{\mathbf{Y}}_2 \in \mathcal{F} \right] \cdot \mathcal{E}_y^c \quad (87)$$

$$\leq \mathbb{P} \left[ \|\mathbf{Z}_1\|^2 - \frac{1}{1+P_1} \|\mathbf{Y}_1\|^2 + \frac{\|\tilde{\mathbf{Y}}_2\|^2 + \|\mathbf{U}\|^2 + 2\|\mathbf{U}\| \cdot \|\tilde{\mathbf{Y}}_2\|}{\sigma_{z^*}^2} \geq \gamma_1' | \tilde{\mathbf{Y}}_2 \in \mathcal{F} \right] \cdot \mathcal{E}_y^c \quad (88)$$

$$\leq \mathbb{P} \left[ \|\mathbf{Z}_1\|^2 - \frac{1}{1+P_1} \|\mathbf{Y}_1\|^2 \geq \gamma_1' - n_{2,2} \frac{\sigma_{y_2^*}^2}{\sigma_{z^*}^2} - \frac{n_{2,2}\delta_y}{\sigma_{z^*}^2} - \frac{\|\mathbf{U}\|^2}{\sigma_{z^*}^2} - \frac{2\|\mathbf{U}\| \sqrt{n_{2,2}(\sigma_{y_2^*}^2 + \delta_y)}}{\sigma_{z^*}^2} \right] \cdot \mathcal{E}_y^c \quad (89)$$

$$\leq \mathbb{P} \left[ \|\mathbf{Z}_1\|^2 - \left\| \frac{\mathbf{Y}_1}{\sqrt{1+P_1}} \right\|^2 \geq \tilde{\gamma}_1 \right], \quad (90)$$

where

$$\gamma_1'' \triangleq -2\gamma_1 + 2\ln(J_1) + n_1 \ln(1+P_1) + n_{2,2} \ln(\sigma_{y_2^*}^2) - n_{2,2} \ln(\sigma_{z^*}^2) \quad (91)$$

$$\gamma_1' \triangleq \gamma_1'' + n_{2,2} - \frac{n_{2,2}\delta_y}{\sigma_{y_2^*}^2} \quad (92)$$

$$\tilde{\gamma}_1 \triangleq \gamma_1' - n_{2,2} \frac{\sigma_{y_2^*}^2}{\sigma_{z^*}^2} - \frac{n_{2,2}\delta_y}{\sigma_{z^*}^2} - \frac{\|\mathbf{U}\|^2}{\sigma_{z^*}^2} - \frac{2\|\mathbf{U}\| \sqrt{n_{2,2}(\sigma_{y_2^*}^2 + \delta_y)}}{\sigma_{z^*}^2}. \quad (93)$$

Note that  $\frac{\mathbf{Y}_1}{\sqrt{1+P_1}} \sim \mathcal{N}(0, I_{n_1})$ . Define

$$v_1 \triangleq \|\mathbf{Z}_1\|^2 : v_1 \sim \mathcal{X}^2(n_1) \quad (94)$$

$$v_2 \triangleq \left\| \frac{\mathbf{Y}_1}{\sqrt{1+P_1}} \right\|^2 : v_2 \sim \mathcal{X}^2(n_1) \quad (95)$$

where  $\mathcal{X}^2(s)$  is central chi-squared distribution of degree  $s$ . We define

$$Q_1 \triangleq v_1 - v_2. \quad (96)$$

Thus

$$\mathcal{E}_y^c \cdot \mathbb{P}[i_1(\mathbf{U}, \mathbf{X}_1; \mathbf{Y}_1, \tilde{\mathbf{Y}}_2) \leq \gamma | \tilde{\mathbf{Y}}_2 \in \mathcal{F}] \quad (97)$$

$$\leq \mathbb{P}[Q_1 \geq \tilde{\gamma}_1] = 1 - F_{Q_1}(\tilde{\gamma}_1) \quad (98)$$

where  $F_{Q_1}$  is the CDF of  $Q_1$ . To calculate this CDF we follow the argument in [18] on the difference of correlated chi-squared random variables.

To sum up,

$$\begin{aligned} \epsilon_1 \leq & 1 - F_{Q_1}(\tilde{\gamma}_1) + \left( 1 - A \left( \frac{\beta_1 P_2}{\beta_1 P_2 + \alpha^2 P_t} \right)^{\frac{n_{2,2}}{2}} \right)^{L(P_t)} + \epsilon_T + e^{-n_{2,2}^{1/3} \kappa} \\ & + \frac{\Gamma(\frac{n_{2,2}}{2}, \frac{n_{2,2} P_{\max}}{2\beta_2 P_2})}{\Gamma(\frac{n_{2,2}}{2})} + ML(P_t) e^{-\gamma} (1 - e^{-n_{2,2}^{1/3} \kappa}). \end{aligned} \quad (99)$$

## A.2 Analyzing $\epsilon_2$

Recall the definition of  $\mathcal{E}_2$  from (27). We upper bound the average error probability  $\epsilon_2$  as

$$\epsilon_2 \leq \mathbb{P}[\mathcal{E}_2] + \mathbb{P}[\|\mathbf{X}_{2,2}\|^2 > n_{2,2}\beta_2 P_2] + \mathbb{P}[\mathcal{E}_{1,2}|\mathcal{E}_{1,1}^c]. \quad (100)$$

### A.2.1 Analyzing $\mathbb{P}[\mathcal{E}_2]$

Similar to Section A.1.4, to evaluate this error event, we use the threshold bound for maximum-metric decoding and we condition on  $\tilde{\mathbf{y}}_2 \in \mathcal{F}$ . I.e.,

$$\begin{aligned} \mathbb{P}[\mathcal{E}_2] &\leq \mathcal{E}_y^c \cdot \mathbb{P}[i_2(\mathbf{X}_{2,2}, \mathbf{X}_3; \tilde{\mathbf{Y}}_2, \mathbf{Y}_3) \leq \gamma_2 | \tilde{\mathbf{Y}}_2 \in \mathcal{F}] \\ &\quad + M \mathbb{P}[i_2(\bar{\mathbf{X}}_{2,2}, \bar{\mathbf{X}}_3; \tilde{\mathbf{Y}}_2, \mathbf{Y}_3) > \gamma_2 | \tilde{\mathbf{Y}}_2 \in \mathcal{F}] \cdot \mathcal{E}_y^c + \mathcal{E}_y \end{aligned} \quad (101)$$

for any  $\gamma_2$ , where  $\bar{\mathbf{X}}_{2,2} \sim f_{\mathbf{X}_{2,2}}$  and  $\bar{\mathbf{X}}_3 \sim f_{\mathbf{X}_3}$  are independent of  $(\mathbf{X}_{2,2}, \mathbf{X}_3, \mathbf{Y}_3, \tilde{\mathbf{Y}}_2)$ .

For given channel outputs  $\tilde{\mathbf{Y}}_2$  and  $\mathbf{Y}_3$ , we follow the same argument proposed in Section A.1.4 to bound  $P[i_2(\bar{\mathbf{X}}_{2,2}, \bar{\mathbf{X}}_3; \tilde{\mathbf{Y}}_2, \mathbf{Y}_3) > \gamma_2]$ . Thus

$$P[i_2(\bar{\mathbf{X}}_{2,2}, \bar{\mathbf{X}}_3; \tilde{\mathbf{Y}}_2, \mathbf{Y}_3) > \gamma_2 | \tilde{\mathbf{Y}}_2, \mathbf{Y}_3] \leq e^{-\gamma_2}. \quad (102)$$

We continue by calculating  $\mathbb{P}[i_2(\mathbf{X}_{2,2}, \mathbf{X}_3; \tilde{\mathbf{Y}}_2, \mathbf{Y}_3) \leq \gamma_2 | \tilde{\mathbf{Y}}_2 \in \mathcal{F}]$ . First note that similar to the argument provided for bonding  $\epsilon_1$ , here also  $\tilde{\mathbf{Y}}_2$  does not follow a Gaussian distribution. We again bound  $f_{\tilde{\mathbf{Y}}_2}$  as in (73). To calculate  $f_{\tilde{\mathbf{Y}}_2|\mathbf{X}_{2,2}}(\tilde{\mathbf{y}}_2|\mathbf{x}_{2,2})$ , note that  $\tilde{\mathbf{Y}}_2|\mathbf{X}_{2,2}$  is equivalent to the output of a Gaussian channel with a Gaussian noise  $\tilde{\mathbf{Z}} \sim \mathcal{N}((1-\alpha)\mathbf{X}_{2,2}, I_{n_{2,2}})$  and with the input vector  $\mathbf{U}$  that is on the power-shell. According to [17, equation 109],  $f_{\tilde{\mathbf{Y}}_2|\mathbf{X}_{2,2}}(\tilde{\mathbf{y}}_2|\mathbf{x}_{2,2})$  is of the following form:

$$f_{\tilde{\mathbf{Y}}_2|\mathbf{X}_{2,2}}(\tilde{\mathbf{y}}_2|\mathbf{x}_{2,2}) \quad (103)$$

$$\begin{aligned} &= \frac{1}{2(\sqrt{\pi})^{n_{2,2}}} \Gamma\left(\frac{n_{2,2}}{2}\right) \exp(-\|\mathbf{U}\|^2 - \|\tilde{\mathbf{y}}_2 - (1-\alpha)\mathbf{x}_{2,2}\|^2) \\ &\quad \times \frac{\mathcal{I}_{n_{2,2}/2-1}(\|\tilde{\mathbf{y}}_2 - (1-\alpha)\mathbf{x}_{2,2}\| \cdot \|\mathbf{u}\|)}{(\|\tilde{\mathbf{y}}_2 - (1-\alpha)\mathbf{x}_{2,2}\| \cdot \|\mathbf{U}\|)^{n_{2,2}/2-1}}, \end{aligned} \quad (104)$$

where  $\mathcal{I}_a(\cdot)$  is the modified Bessel function of the first kind and  $a$ -th order. Define  $Q_{\mathbf{Y}^*|\mathbf{X}_{2,2}}$  as the output distribution of the Gaussian channel with noise  $\tilde{\mathbf{Z}}$  but where the input sequence  $\mathbf{U}$  follows the Gaussian distribution  $\mathcal{N}(0, (\beta_1 P_2 + \alpha^2 P_t) I_{n_{2,2}})$ . Then according to [17, Appendix B]:

$$\frac{f_{\tilde{\mathbf{Y}}_2|\mathbf{X}_{2,2}}(\tilde{\mathbf{y}}_2|\mathbf{x}_{2,2})}{Q_{\mathbf{Y}^*|\mathbf{X}_{2,2}}(\tilde{\mathbf{y}}_2|\mathbf{x}_{2,2})} = D_{f,Q,2} \quad (105)$$

where one can prove that for all  $\tilde{\mathbf{y}}_2$  and  $\mathbf{x}_{2,2}$

$$D_{f,Q,2} \geq j_4 \quad (106)$$

where  $j_4 > 0$ . Define  $J_2 \triangleq \frac{j_4}{j_1}$ , then by employing above arguments, we have

$$\mathcal{E}_y^c \cdot \mathbb{P}[i_2(\mathbf{X}_{2,2}, \mathbf{X}_3; \tilde{\mathbf{Y}}_2, \mathbf{Y}_3) \leq \gamma_2 | \tilde{\mathbf{Y}}_2 \in \mathcal{F}] \quad (107)$$

$$\leq \mathbb{P}\left[\ln\left(\frac{Q_{\mathbf{Y}^*|\mathbf{X}_{2,2}}(\tilde{\mathbf{Y}}_2|\mathbf{X}_{2,2})f_{\mathbf{Y}_3|\mathbf{X}_3}(\mathbf{Y}_3|\mathbf{X}_3)}{f_{\tilde{\mathbf{Y}}_2}(\tilde{\mathbf{Y}}_2)f_{\mathbf{Y}_3}(\mathbf{Y}_3)}\right) \leq \gamma_2 - \ln(J_2) | \tilde{\mathbf{Y}}_2 \in \mathcal{F}\right] \cdot \mathcal{E}_y^c \quad (108)$$

$$\begin{aligned}
 &= \mathbb{P} \left[ \ln \left( \frac{\frac{1}{(\sqrt{2\pi\sigma_{u^*}^2})^{n_{2,2}}} \exp\left(-\frac{\|\tilde{\mathbf{Y}}_2 - (1-\alpha)\mathbf{X}_{2,2}\|^2}{2\sigma_{u^*}^2}\right)}{\frac{1}{(\sqrt{2\pi\sigma_{y_2^*}^2})^{n_{2,2}}} \exp\left(-\frac{\|\tilde{\mathbf{Y}}_2\|^2}{2\sigma_{y_2^*}^2}\right)} \right) \right. \\
 &\quad \left. + \ln \left( \frac{\frac{1}{(\sqrt{2\pi})^{n_3}} \exp\left(-\frac{\|\mathbf{Z}_3\|^2}{2}\right)}{\frac{1}{(\sqrt{2\pi(1+P_3)})^{n_3}} \exp\left(-\frac{\|\mathbf{Y}_3\|^2}{2(1+P_3)}\right)} \right) \leq \gamma_2 - \ln(J_2) \mid \tilde{\mathbf{Y}}_2 \in \mathcal{F} \right] \cdot \mathcal{E}_y^c \quad (109)
 \end{aligned}$$

$$\begin{aligned}
 &= \mathbb{P} \left[ \frac{n_{2,2}}{2} \ln(\sigma_{y_2^*}^2) - \frac{n_{2,2}}{2} \ln(\sigma_{u^*}^2) - \frac{\|\tilde{\mathbf{Y}}_2 - (1-\alpha)\mathbf{X}_{2,2}\|^2}{2\sigma_{u^*}^2} + \frac{\|\tilde{\mathbf{Y}}_2\|^2}{2\sigma_{y_2^*}^2} \right. \\
 &\quad \left. + \frac{n_3}{2} \ln(1+P_3) - \frac{\|\mathbf{Y}_3 - \mathbf{X}_3\|^2}{2} + \frac{\|\mathbf{Y}_3\|^2}{2(1+P_3)} \leq \gamma_2 - \ln(J_2) \mid \tilde{\mathbf{Y}}_2 \in \mathcal{F} \right] \cdot \mathcal{E}_y^c \quad (110)
 \end{aligned}$$

$$\begin{aligned}
 &= \mathbb{P} \left[ \frac{\|\tilde{\mathbf{Y}}_2 - (1-\alpha)\mathbf{X}_{2,2}\|^2}{\sigma_{u^*}^2} - \frac{\|\tilde{\mathbf{Y}}_2\|^2}{\sigma_{y_2^*}^2} + \|\mathbf{Y}_3 - \mathbf{X}_3\|^2 - \frac{\|\mathbf{Y}_3\|^2}{1+P_3} \geq \gamma_2'' \mid \tilde{\mathbf{Y}}_2 \in \mathcal{F} \right] \cdot \mathcal{E}_y^c \quad (111)
 \end{aligned}$$

$$\begin{aligned}
 &\leq \mathbb{P} \left[ \|\mathbf{Z}_3\|^2 - \frac{1}{1+P_3} \|\mathbf{Y}_3\|^2 + \frac{\|\tilde{\mathbf{Y}}_2 - (1-\alpha)\mathbf{X}_{2,2}\|^2}{\sigma_{u^*}^2} \geq \gamma_2'' + n_{2,2} - \frac{n_{2,2}\delta_y}{\sigma_{y_2^*}^2} \mid \tilde{\mathbf{Y}}_2 \in \mathcal{F} \right] \cdot \mathcal{E}_y^c \quad (112)
 \end{aligned}$$

$$\begin{aligned}
 &\leq \mathbb{P} \left[ \|\mathbf{Z}_3\|^2 - \frac{1}{1+P_3} \|\mathbf{Y}_3\|^2 \right. \\
 &\quad \left. + \frac{\|\tilde{\mathbf{Y}}_2\|^2 + (1-\alpha)^2 \|\mathbf{X}_{2,2}\|^2 + 2(1-\alpha) \|\mathbf{X}_{2,2}\| \cdot \|\tilde{\mathbf{Y}}_2\|}{\sigma_{u^*}^2} \geq \gamma_2' \mid \tilde{\mathbf{Y}}_2 \in \mathcal{F} \right] \cdot \mathcal{E}_y^c \quad (113)
 \end{aligned}$$

$$\begin{aligned}
 &\leq \mathbb{P} \left[ \|\mathbf{Z}_3\|^2 - \frac{1}{1+P_3} \|\mathbf{Y}_3\|^2 \geq \gamma_2' - n_{2,2} \frac{\sigma_{y_2^*}^2}{\sigma_{u^*}^2} - \frac{n_{2,2}\delta_y}{\sigma_{u^*}^2} \right. \\
 &\quad \left. - \frac{n_{2,2}P_t + \delta_2}{\sigma_{u^*}^2} - \frac{2\sqrt{n_{2,2}P_t + \delta_2} \sqrt{n_{2,2}(\sigma_{y_2^*}^2 + \delta_y)}}{\sigma_{u^*}^2} \right] \cdot \mathcal{E}_y^c \quad (114)
 \end{aligned}$$

$$\begin{aligned}
 &\leq \mathbb{P} \left[ \|\mathbf{Z}_3\|^2 - \left\| \frac{\mathbf{Y}_3}{\sqrt{1+P_3}} \right\|^2 \geq \tilde{\gamma}_2 \right], \quad (115)
 \end{aligned}$$

where

$$\gamma_2'' \triangleq -2\gamma_2 + 2\ln(J_2) + n_3 \ln(1+P_3) + n_{2,2} \ln(\sigma_{y_2^*}^2) - n_{2,2} \ln(\sigma_{u^*}^2) \quad (116)$$

$$\gamma_2' \triangleq \gamma_2'' + n_{2,2} - \frac{n_{2,2}\delta_y}{\sigma_{y_2^*}^2} \quad (117)$$

$$\tilde{\gamma}_2 \triangleq \gamma_2' - n_{2,2} \frac{\sigma_{y_2^*}^2}{\sigma_{u^*}^2} - \frac{n_{2,2}\delta_y}{\sigma_{u^*}^2} - \frac{n_{2,2}P_t + \delta_2}{\sigma_{u^*}^2} - \frac{2\sqrt{n_{2,2}P_t + \delta_2} \sqrt{n_{2,2}(\sigma_{y_2^*}^2 + \delta_y)}}{\sigma_{u^*}^2} \quad (118)$$

Define

$$v_1' \triangleq \|\mathbf{Z}_3\|^2 : v_1' \sim \mathcal{X}^2(n_3), \quad (119)$$

$$v_2' \triangleq \left\| \frac{\mathbf{Y}_3}{\sqrt{1+P_3}} \right\|^2 : v_2' \sim \mathcal{X}^2(n_3). \quad (120)$$

Also define

$$Q_2 \triangleq v_1' - v_2'. \quad (121)$$

Thus

$$\mathcal{E}_y^c \cdot \mathbb{P}[i_2(\mathbf{X}_{2,2}, \mathbf{X}_3; \tilde{\mathbf{Y}}_2, \mathbf{Y}_3) \leq \gamma_2] \leq \mathbb{P}[Q_2 > \tilde{\gamma}_2] = 1 - F_{Q_2}(\tilde{\gamma}_2). \quad (122)$$

where  $F_{Q_2}(\cdot)$  is the CDF of  $Q_2$ .

### A.2.2 Analyzing $\mathbb{P}[||\mathbf{X}_{2,2}||^2 > n_{2,2}\beta_2 P_2]$

We calculate  $\mathbb{P}[||\mathbf{X}_{2,2}||^2 > n_{2,2}\beta_2 P_2]$  by following the same argument as in that of Section A.1.1. We can prove that

$$\mathbb{P}[||\mathbf{X}_{2,2}||^2 > n_{2,2}\beta_2 P_2] = \frac{\Gamma(\frac{n_{2,2}}{2}, \frac{n_{2,2}}{2})}{\Gamma(\frac{n_{2,2}}{2})}. \quad (123)$$

To sum up

$$\begin{aligned} \epsilon_2 \leq & 1 - F_{Q_2}(\tilde{\gamma}_2) + M e^{-\gamma_2} (1 - e^{-n_{2,2}^{1/3} \kappa}) + \frac{\Gamma(\frac{n_{2,2}}{2}, \frac{n_{2,2}}{2})}{\Gamma(\frac{n_{2,2}}{2})} \\ & + e^{-n_{2,2}^{1/3} \kappa} + \left( 1 - A \left( \frac{\beta_1 P_2}{\beta_1 P_2 + \alpha^2 P_t} \right)^{\frac{n_{2,2}}{2}} \right)^{L^{(P_t)}}. \end{aligned} \quad (124)$$

## References

- [1] P. Popovski, Č. Stefanović, J. J. Nielsen, E. de Carvalho, M. Angjelichinoski, K. F. Trillingsgaard and A. Bana, "Wireless access in ultra-reliable low-latency communication (URLLC)," in *IEEE Trans. on Communications*, vol. 67, no. 8, pp. 5783–5801, Aug. 2019.
- [2] X. Ge, "Ultra-reliable low-latency communications in autonomous vehicular networks," in *IEEE Trans. on Vehicular Technology*, vol. 68, no. 5, pp. 5005–5016, May 2019.
- [3] M. Shirvanimoghaddam et al., "Short block-length codes for ultra-reliable low latency communications," in *IEEE Communications Magazine*, vol. 57, no. 2, pp. 130–137, Feb. 2019.
- [4] X. Zhang, J. Wang and H. V. Poor, "Statistical QoS-driven energy-efficiency optimization for URLLC over 5G mobile wireless networks in the finite blocklength regime," in *Info. Sciences and Systems*, pp. 1–6, 2019.
- [5] Y. Polyanskiy, H. V. Poor and S. Verdú, "Channel coding rate in the finite blocklength regime," in *IEEE Trans. on Info. Theory*, vol. 56, no. 5, pp. 2307–2359, May, 2010.
- [6] N. Shulman and M. Feder. "Static broadcasting", *IEEE International Symposium on Info. Theory (ISIT)*, 2000.
- [7] M. Langberg and M. Effros, "Beyond capacity: The joint time-rate region", *arXiv:2101.12236v1*, Jan. 2021.
- [8] P. H. Lin, S. C. Lin and E. A. Jorswieck, "Early decoding for Gaussian broadcast channels with heterogeneous blocklength constraints," in *IEEE International Symposium on Information Theory (ISIT)*, pp. 3243–3248, 2021.

- 
- [9] H. Nikbakht, M. Egan and J-M. Gorce, "Joint channel coding of consecutive packets with heterogeneous decoding deadlines in the finite blocklength regime", [Research Report] Inria - Research Centre Grenoble – Rhône-Alpes. 2021. Available: <https://hal.inria.fr/hal-03360567>
- [10] J. Scarlett, "On the dispersions of the Gel'fand - Pinsker channel and dirty paper coding," in *IEEE Trans. on Info Theory*, vol. 61, no. 9, pp. 4569-4586, Sept. 2015.
- [11] T. Erseghe, "Coding in the finite-blocklength regime: Bounds based on Laplace integrals and their asymptotic approximations," in *IEEE Trans. on Info. Theory*, vol. 62, no. 12, pp. 6854–6883, Dec. 2016.
- [12] I. Csiszàr and J. Körner, "Information theory", *New York: Academic Press*, 1981.
- [13] G. Caire and S. Shamai, "On the achievable throughput of a multiantenna Gaussian broadcast channel," in *IEEE Transactions on Information Theory*, vol. 49, no. 7, pp. 1691-1706, July 2003.
- [14] V. Y. F. Tan and M. Tomamichel, "The third-order term in the normal approximation for the AWGN channel," in *IEEE Trans. on Info. Theory*, vol. 61, no. 5, pp. 2430–2438, May 2015.
- [15] M. H. M. Costa, "Writing on dirty paper (Corresp.)," *IEEE Trans. Inf. Theory*, vol. 29, no. 3, pp. 439–441, May 1983.
- [16] A. J. Stam, "Limit theorems for uniform distributions on spheres in high-dimensional Euclidean spaces," *J. Appl. Probab.*, vol. 19, no. 1, pp. 221–228, Mar. 1982.
- [17] E. MolavianJazi and J. N. Laneman, "A second-order achievable rate region for Gaussian multi-access channels via a central limit theorem for functions," in *IEEE Transactions on Information Theory*, vol. 61, no. 12, pp. 6719–6733, Dec. 2015.
- [18] A. Ferrari, "A note on sum and difference of correlated chi-squared variables", [Online]. Available: [arXiv:1906.09982](https://arxiv.org/abs/1906.09982), 2019.



**RESEARCH CENTRE  
GRENOBLE – RHÔNE-ALPES**

Inovallée  
655 avenue de l'Europe Montbonnot  
38334 Saint Ismier Cedex

Publisher  
Inria  
Domaine de Voluceau - Rocquencourt  
BP 105 - 78153 Le Chesnay Cedex  
[inria.fr](http://inria.fr)

ISSN 0249-6399