On Asynchronous Interference Channels

by

Kamyar Moshksar

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Mathematics
in
Pure Mathematics

Waterloo, Ontario, Canada, 2016

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Kamyar Moshksar

Abstract

The first part of the thesis studies a decentralized network of separate transmitter-receiver (Tx-Rx) pairs. The users are asynchronous meaning there exists a mutual delay between their transmitted codewords. Each Tx stays silent permanently after it sends its codeword. The channel from each Tx to each Rx is modelled by a static and non-frequency selective coefficient followed by additive white Gaussian noise. Each Tx sends a preamble sequence before transmitting its codeword to ensure its affiliated Rx knows the exact arrival time for the codeword. The network being decentralized, different users are unaware of each other's preamble sequences. As such, the receivers can not determine the exact positions of interference bursts. We introduce a learning technique based on piecewise-linear regression where it is shown how each Rx successfully estimates the number of interferers, the coefficients of the channels carrying interference and the mutual delays. The estimates for the mutual delays are not perfect, however, they are reliable enough to guarantee successful decoding of the codewords.

The second part of the thesis addresses a centralized Gaussian interference channel of two Tx-Rx pairs under stochastic data arrival (GIC-SDA). The information bits arrive at the transmitters according to independent and asynchronous Bernoulli processes (Tx-Tx asynchrony). The transmissions are asynchronous (Tx-Rx asynchrony) in the sense that a Tx immediately sends a codeword to its Rx when there are enough information bits gathered in its buffer. Such immediate style of transmission is in contrast to the Tx-Rx synchronous style discussed in [20]. In a setting where the transmitters only know the statistics of Tx-Tx asynchrony, it is shown how each user designs its codebook rate in order to maximize the probability of successful decoding at the receivers. An achievable region is characterized for the codebook rates in a two-user GIC-SDA under the requirements that the transmissions be immediate and the receivers treat interference as noise. This region is described as the union of uncountably many polyhedrons and is in general disconnected and non-convex due to infeasibility of time sharing. Special attention is given to the symmetric case where closed-form expressions are developed for the achievable codebook rates.

Acknowledgements

I would like to thank my supervisors, Professor Brian Forrest and Professor Nico Spronk, for their kind guidance, constant support and encouragement during the completion of this thesis.

I am very grateful to have had Professors Alexandru Nica, Ken Davidson, Benoit Charbonneau and Spiro Karigiannis as my teachers from whom I learnt tremendously.

Dedication

This is dedicated to my mother Azadeh and my father Rahim.

Table of Contents

List of Figures

Chapter 1

Introduction

1.1 Motivation and Previous Work

Characterizing the capacity region for Gaussian interference channels has been an open problem for more than thirty years. Even the two-user case is only partially solved [1, 2, 3, 4]. Some pivotal assumptions made in [1, 2, 3, 4] and the references therein are:

1. The network is centralized, i.e., there is a central controller that assigns the resources (such as time or bandwidth) to the users. Moreover, users know each other's codebooks (through the central controller or direct communication among users) which enables the receivers to perform interference cancellation or multiuser decoding. In decentralized setups such as ad hoc networks [5], there is no central node to assign the resources to the users and issues such as fairness and rate assignment must be handled locally by each transmitter-receiver pair. The well-known ALOHA system [6] is a potential candidate for data communication in such scenarios. Modelled based on the so-called collision channel, each user randomly stays silent or transmits a packet independently from packet interval[1] to packet interval. If two users collide, the transmitters are required to retransmit the lost packets which in turn drops the overall spectral efficiency and leads to extra delay and stability[2] issues regarding the backlogged packets [7, 8, 9]. The seminal paper of Massey and Mathys [10] studies the collision channel without feedback in the context of random multiple access where the packets sent by different transmitters incur independent delays. The transmitters can

[1]A packet interval is the communication period it takes each transmitter to send a packet.

[2]An ALOHA system is called stable if the random process of backlogged packets is ergodic.

not be synchronized due to the lack of feedback from the common receiver. Instead, each transmitter is assigned a "protocol sequence" in order to enable the receiver to identify and decode the transmitted packets regardless of the delays. Design of protocol sequences (signature coding) is further explored in [11, 12].

2. The number of users is fixed. In a network with invariable underlying infrastructure such as frequency division (FD), if the number of active users is less than the design target, part of the spectrum remains unused. Despite its simplicity, FD is optimal in various setups that comply with (i) and (ii) above. For example, [13] demonstrates that in a Gaussian interference channel, every pareto-optimal rate assignment is realized by FD under the assumptions that users treat each other as noise and the crossover channel gains are sufficiently larger than the direct channel gains. In practice, the majority of users are most likely inactive in a typical snapshot of the network. This results in a low spectral efficiency for FD.

3. Users are block-synchronous. This assumption is not necessarily valid in practice, because different users do not become active simultaneously. Information theoretic studies on a network of block-asynchronous users is investigated in [14, 15] in the context of centralized multiple access channels. Reference [16] combines the scheme of [15] with the so-called generalized time sharing [17] to derive an achievable rate region for a centralized interference channel in the presence of block asynchrony. The authors in [18] study a centralized Gaussian interference channel with block-asynchronous users. Invoking the general formula for capacity in [19], a multi-letter expression is derived for the capacity region of such channels.

4. Each transmitter constantly communicates with its affiliated receiver. In some applications such as wireless sensor networks, transmitters stop sending data intermittently. Moreover, no user knows a priori the times that other users start or end their transmissions.

The goal of this thesis is to explore the possibility and limits of reliable communication in a network where the number of users is random, users are block-asynchronous with burst transmissions and the coordination among the users is minimal. In the next section, we present an overview of the main results.

1.2 Contributions

1. In the first part of the thesis, a decentralized wireless network of separate transmitter-receiver pairs is studied where there is no central controller to assign the resources to the users and users do not explicitly cooperate. For simplicity, we focus on a single-burst scenario where each transmitter sends a single codeword upon activation and remains silent afterwards. Users are block-asynchronous meaning there exists a mutual delay between their transmitted codewords.

   We show how the receivers learn about the number of active users, channel coefficients and activation times of the transmitters based on locally available measurements. It is essential that each receiver finds the exact arrival time of the codeword sent by its corresponding transmitter. To achieve this goal, preamble sequences are embedded at the beginning of a transmitted codeword. As different users do not necessarily know each other's preamble sequences, there is no guarantee that a receiver can estimate the arrival times of interference bursts along its desired data with vanishingly small probability of error. Nevertheless, the estimates are reliable enough to guarantee successful decoding at each receiver.

2. In the second part of the thesis, we study a two-user Gaussian interference channel (GIC) with stochastic data arrival (SDA). The input bit streams at the transmitters are independent and asynchronous Bernoulli processes. The information source at each transmitter turns off after randomly generating a given total number of bits. Let us consider two transmission schemes:

   1- Each transmitter begins to send a codeword only at a predetermined set of time slots $t_1 < t_2 < \cdots$ agreed upon between both Tx-Rx pairs. Transmission of a codeword at time instant t_1 is subject to availability of enough data bits to represent a codeword. If the number of available information bits is not enough, the transmitter waits until the earliest time slot t_m for $m \geq 2$ when enough data is gathered in its buffer. This scheme is introduced in [20] which we refer to as the Tx-Rx synchronous scheme.

   2- Each transmitter begins to send a codeword immediately when there are enough bits gathered in its buffer. In this case, each receiver does not know a priori the time slots when the codewords are dispatched by the transmitters. We refer to this style of transmission immediate or Tx-Rx asynchronous.

   Under both the Tx-Rx synchronous and the Tx-Rx asynchronous schemes, the data sent by both transmitters can potentially look like intermittent bursts along the time axis. In the synchronous scheme, all symbols in a transmitted codeword are received either in the absence of interference or in the presence of interference. In the asynchronous scheme, however, a number of symbols per transmitted codeword may be received interference-free while the rest are received in the presence of interference.

We compare both schemes in terms of the underlying *relative delay* induced in the transmission process. More precisely, if the first symbol of the j^{th} transmitted codeword is sent at time slots $T^{(j)}_{\text{synch}}$ and $T^{(j)}_{\text{asynch}}$ under the synchronous and asynchronous schemes, respectively, then there exists a $\delta > 0$ such that the probability of $\frac{T^{(j)}_{\text{synch}} - T^{(j)}_{\text{asynch}}}{T^{(j)}_{\text{asynch}}} > \delta$ occurring grows to 1 in the limit of large codeword length.

Applying sequential joint typicality decoding [21], the receivers estimate and learn about the positions of the transmitted bursts along the time axis. We study fundamental constraints on the codebook rates in order to guarantee immediate transmission at the transmitters and successful decoding at the receivers. For simplicity of presentation, we employ random Gaussian codebooks and assume all receivers treat interference as noise. The achievable region for codebook rates is characterized as the union of uncountably many polyhedrons which is in general non-convex and disconnected due to infeasibility of time sharing. In a setup where the exact asynchrony between the input bit streams is unknown to the transmitters, the number of transmitted codewords at each transmitter is optimized to achieve a target transmission rate and minimize the probability of unsuccessful decoding at the receivers.

Our analysis directly incorporates the burst-like nature of incoming data in the standard information-theoretic framework for reliable communications.

Chapter 2

Decentralized Networks with Asynchronous Users and Burst Transmissions

2.1 Notations and terminology

The set of real and complex numbers are shown by $\mathbb{R}$ and $\mathbb{C}$, respectively. Random quantities are shown in bold such as $\boldsymbol{x}$. A realization of $\boldsymbol{x}$ is denoted by x. Vectors are shown by an arrow on top, e.g., $\vec{x}$. A sequence $(a_1, \cdots, a_m)$ is denoted by $(a_l)_{l=1}^m$. A circularly symmetric complex Gaussian random sequence $(\boldsymbol{x}_i)_{i=1}^m$ of length m with zero mean and covariance matrix Σ is denoted by $\mathrm{CN}(\vec{0}, \Sigma)$. The probability density function (PDF) of a $\mathrm{CN}(0, \sigma^2)$ random variable is denoted by $\mathrm{g}(x; \sigma^2) = \frac{1}{\pi\sigma^2} e^{-\frac{|x|^2}{\sigma^2}}$. A Bernoulli random variable with parameter θ is denoted by $\mathrm{Ber}(\theta)$. We use $\mathbb{P}(\mathcal{E})$ for the probability of an event $\mathcal{E}$, $\mathbb{1}_{\mathcal{E}}$ for the indicator function of an event $\mathcal{E}$ and $p_{\boldsymbol{x}}(\cdot)$ for the PDF of a random variable $\boldsymbol{x}$. The mutual information between random variables $\boldsymbol{x}$ and $\boldsymbol{y}$ is denoted by $I(\boldsymbol{x}; \boldsymbol{y})$, $h(\boldsymbol{x})$ denotes the differential entropy of a continuous random variable $\boldsymbol{x}$ and $h_b(x) := -x \log x - (1-x) \log(1-x)$ denotes the binary entropy function. For $x \in \mathbb{R}$, $\lfloor x \rfloor$ is the floor of x and $x^+ = x\mathbb{1}_{x>0}$. We write $f(x) = \Theta(g(x))$ to mean there exist constants $c_1, c_2 > 0$ such that $c_1 < |\frac{f(x)}{g(x)}| < c_2$ for sufficiently large x. For any set $\mathcal{A}$, $|\mathcal{A}|$ and $\mathcal{A}^c$ show the cardinality and complement of $\mathcal{A}$, respectively. For any two sets $\mathcal{A}$ and $\mathcal{B}$, $\mathcal{A} \backslash \mathcal{B}$ denotes the difference between these sets. Throughout the chapter,

- i.i.d. stands for "independent and identically distributed".

- We refer to the inequality $a-1<\lfloor a\rfloor\leq a$ for any $a\in\mathbb{R}$ as "the floor inequality".
- Any equality, inequality or convergence involving random quantities is understood in the "almost sure" sense unless otherwise stated. We avoid using the term "almost surely" hereafter.

2.2 System Model

2.2.1 Channel model and the signalling scheme

We consider a wireless network of separate transmitter-receiver pairs. From the perspective of each active user, say user 0, there are a number K of other active users that are referred to as *potential interferers*. The static and non-frequency selective channel coefficient from transmitter i to receiver j is denoted by the complex number $h_{i,j}$. For notational simplicity, we denote $h_{i,0}$ by h_i and define $\vec{h}=(h_i)_{i=0}^{K}$. The channel from each transmitter to each receiver is slotted and the time slots on different channels coincide. Therefore, all active users are synchronous at the symbol level.[1] Transmitter i chooses its message from a set of size $2^{\lfloor nR_i\rfloor}$ where n and R_i are the codeword length and the code rate, respectively. The corresponding codeword $(x_{i,l})_{l=0}^{n-1}$ is transmitted during n consecutive time slots. The elements of $(x_{i,l})_{l=0}^{n-1}$ are realizations of independent $\mathrm{CN}(0,P_i)$ random variables. To address the fact that in practice users do not stay active forever, we assume each transmitter sends one codeword upon activation and remains silent afterwards. Due to fairness, the code rates and average transmission powers of all users are identical, i.e.,

$$R_i=R,\ P_i=P,\ 0\leq i\leq K. \tag{2.1}$$

2.2.2 Modelling the asynchrony

Fixing an origin $t=0$ on the discrete time axis as shown in Fig. 2.1 and recalling the codeword length is n for all users, transmitter i starts its activity at time slot $t_{i,n}$, referred to as the delay associated to user i. The delay $t_{i,n}$ takes its values in the set $\{0,1,\cdots,\lfloor n/\alpha\rfloor-1\}$ where $\lfloor n/\alpha\rfloor-1$ is the index of the latest time slot that any transmitter starts its activity

[1] See the discussion at the end of Section 2.2.2 on how synchronization at the symbol level is achieved in practice.

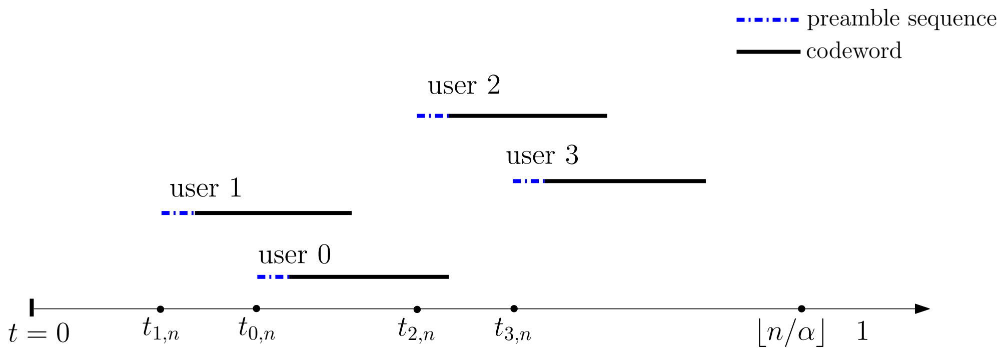


Figure 2.1: A decentralized network of four active users (user 0 together with $K = 3$ potential interferers). Different users become active at different time slots. The activation time $t_{i,n}$ for user i is a realization of a discrete uniform random variable over $\{0, 1, \cdots, \lfloor n/\alpha \rfloor - 1\}$.

and $0 < \alpha < 1$ is referred to as the *congestion factor*. A smaller α implies a less congested network. For any $0 \le i \le K$, one can write

$$t_{i,n} = \lfloor \lfloor n/\alpha \rfloor \nu_i \rfloor , \tag{2.2}$$

where $\nu_0, \cdots, \nu_K$ are distinct numbers in the interval $(0, 1)$. We denote the sequence $(\nu_i)_{i=0}^{K}$ by $\vec{\nu}$. No user is aware of $\vec{\nu}$ a priori, however, all users are aware of α.

The motivation for the model considered in this chapter is that in practical wireless systems, users become active after enough data (determined by the codeword length) is available in their buffers. Due to the random nature of the data arrival, buffers of different nodes will reach to such a state (from now on called the transmission state) at different times. The conclusion of this model is that, if each unit is allowed to decide for the start of its transmission based on the local content of its buffer, rather than waiting for all units to have enough data, the overall delay experienced by each unit will be significantly reduced. However, each unit, upon reaching its transmission state, can synchronize its transmission at a symbol level by waiting for the start of the next symbol interval (referred to as "time slot" in the article). This results in a minor additional delay (less than a single time slot). The question is how one can guarantee that such separate users can detect and synchronize their transmissions at a time slot level. This requires the following:

1. Access to a central clock (measure of time)
2. The ability of each unit to deal with small errors in synchronizing with such a central clock

- **Answer to (1)**: Access to a central clock is part of most modern wireless systems, e.g., through a GPS clock, through the timing information embedded in the TV signals [22], or through access to a backbone network using Network Time Protocol (NTP) [23]. Even in the absence of such systems, it is fairly easy to provide a coarse time synchronization mechanism by using a dedicated node to occasionally broadcast a training signal and having separate units to look for such a training signal in order to synchronize. The latter option, however, is against the definition of a decentralized network in this chapter where the presence of a central node is out of the question. This conflict can be settled if one restricts the role of a central node to only broadcasting training signals for time slot synchronization.

- **Answer to (2)**: In practice, many wireless systems rely on a mechanism that reduces their sensitivity to synchronization errors. In particular, most modern wireless systems rely on orthogonal frequency division multiple access (OFDM/OFDMA). In OFDM, the time slot can be defined as a single OFDM symbol, and it is well known that the use of cyclic prefix in OFDM provides total immunity to timing errors as long as such timing errors are less than the length of the cyclic prefix. The cyclic prefix is long enough (is always set to be longer than the length of the channel memory) to guarantee that the timing errors in synchronizing with the central clock are totally absorbed within the OFDM cyclic prefix.

2.3 Channel estimation and user identification

It is assumed that receiver 0 knows h_0, i.e., the channel coefficient from its affiliated transmitter. However, receiver 0 is required to estimate K, $(h_i)_{i=1}^{K}$ and $\vec{\nu}$. Towards this goal, each transmitter starts its activity by sending a preamble sequence, namely, $(\boldsymbol{x}'_{i,l})_{l=0}^{\lfloor n^{\delta}\rfloor}$ for transmitter i consisting of $\lfloor n^{\delta}\rfloor$ independent complex Gaussian random variables with average power P where $\frac{1}{2} < \delta < 1$.[2] This sequence is revealed to both ends of user i and is transmitted in $\lfloor n^{\delta}\rfloor$ consecutive time slots upon activation. Each user generates its preamble sequence independently of other users. As a shorthand notation, let $n_{\delta} = \lfloor n^{\delta}\rfloor$. The signal received by receiver 0 at time slot $0 \leq t \leq \lfloor n/\alpha\rfloor + n_{\delta} + n - 1$ is given by

$$\boldsymbol{y}_0[t] = h_0\boldsymbol{s}_0[t] + \sum_{i=1}^{K} h_i\,\boldsymbol{s}_i[t] + \boldsymbol{z}_0[t], \tag{2.3}$$

[2]The constraint $\frac{1}{2} < \delta < 1$ is required in the proof of Proposition 2.

where $\boldsymbol{z}_0[t] \sim \mathrm{CN}(0,1)$ is the ambient noise at receiver 0 at time slot t and $\boldsymbol{s}_i[t]$ is the signal transmitted by user i at time slot t given by

$$\boldsymbol{s}_i[t] = \begin{cases} \boldsymbol{x}'_{i,t-t_{i,n}} & t_{i,n} \le t \le t_{i,n} + n_\delta - 1 \\ \boldsymbol{x}_{i,t-t_{i,n}-n_\delta} & t_{i,n} + n_\delta \le t \le t_{i,n} + n_\delta + n - 1 \\ 0 & t < t_{i,n} \ \text{ or } \ t > t_{i,n} + n_\delta + n - 1 \end{cases} \tag{2.4}$$

Since the ambient noise has unit variance, P is a measure of SNR. Note that $|t_{i,n} - t_{0,n}| \ge n + n_\delta$ if and only if there is no interference between user 0 and user i.

2.3.1 Estimating the number of users, channel coefficients and the delays

Let T_n be the latest time slot that a transmitter may send the last symbol in its codeword, i.e., $T_n = \lfloor n/\alpha \rfloor + n_\delta + n - 1$. Define the random continuous function[3] $\boldsymbol{F}_n : [0,1] \to [0,\infty)$ by $\boldsymbol{F}_n(0) = 0$, $\boldsymbol{F}_n(\frac{l}{T_n}) = \frac{1}{PT_n} \sum_{t=1}^{l} |\boldsymbol{y}_0[t]|^2$ for $1 \le l \le T_n$ and let $\boldsymbol{F}_n$ be linear over each interval $[\frac{l-1}{T_n}, \frac{l}{T_n}]$. Also, define

$$\psi(\tau) := \tau \mathbb{1}_{\tau \le \frac{\alpha}{1+\alpha}} + \frac{\alpha}{1+\alpha} \mathbb{1}_{\tau > \frac{\alpha}{1+\alpha}}. \tag{2.5}$$

and

$$F(\tau) := \frac{\tau}{P} + \sum_{i=0}^{K} |h_i|^2 \psi\Big(\tau - \frac{\nu_i}{1+\alpha}\Big), \ \ 0 \le \tau \le 1. \tag{2.6}$$

The following proposition is essential towards estimating K, $(h_i)_{i=1}^K$ and $\vec{\nu}$:

Proposition 1. *For any $\epsilon > 0$, let n be large enough such that $n_\delta \ge 2$ and $\epsilon T_n > \max\{cn_\delta, c'\}$ where c and c' are constants that only depend on K, $\vec{h}$ and P. Then*

$$\mathbb{P}(\|\boldsymbol{F}_n - F\|_\infty > \epsilon) \le \Theta(1) e^{-\epsilon^2 \Theta(n)}. \tag{2.7}$$

In particular,

$$\lim_{n\to\infty} \|\boldsymbol{F}_n - F\|_\infty = 0, \tag{2.8}$$

where for any bounded real-valued function f on $[0,1]$, $\|f\|_\infty := \sup_{\tau\in[0,1]} |f(\tau)|$ is the so-called uniform norm of f.

[3] All realizations of $\boldsymbol{F}_n$ are continuous.

Proof. See Appendix A. □

Explicit expressions for the constants c and c' in Proposition 1 are

$$c = 45(K^2 + 5K + 6) \max\Big\{ \max_{0 \le i \le K} |h_i|, \max_{0 \le i,j \le K} |h_i||h_j| \Big\},$$
$$c' = 5\Big(\frac{1}{P} + \sum_{i=0}^{K} |h_i|^2\Big). \tag{2.9}$$

To explain Proposition 1, let k_i denote the index of the $(i+1)^{st}$ transmitter that becomes active, i.e., $\nu_{k_0} < \nu_{k_1} < \cdots < \nu_{k_K}$. The function F in (2.6) is a piecewise linear function. The slope of F is $\frac{1}{P}$ over the interval $[0, \frac{\nu_{k_0}}{1+\alpha}]$. As transmitter k_0 becomes active, the slope of F jumps to $\frac{1}{P} + |h_{k_0}|^2$. Let us distinguish two cases:

- If $\nu_{k_1} > \nu_{k_0} + \alpha$, the slope of F is $\frac{1}{P} + |h_{k_0}|^2$ over the interval $[\frac{\nu_{k_0}}{1+\alpha}, \frac{\nu_{k_0}+\alpha}{1+\alpha}]$, drops to $\frac{1}{P}$ over the interval $[\frac{\nu_{k_0}+\alpha}{1+\alpha}, \frac{\nu_{k_1}}{1+\alpha}]$ and jumps to $\frac{1}{P} + |h_{k_1}|^2$ as transmitter k_1 becomes active.
- If $\nu_{k_0} < \nu_{k_1} < \nu_{k_0} + \alpha$, the slope of F is $\frac{1}{P} + |h_{k_0}|^2$ over the interval $[\frac{\nu_{k_0}}{1+\alpha}, \frac{\nu_{k_1}}{1+\alpha}]$, jumps to $\frac{1}{P} + |h_{k_0}|^2 + |h_{k_1}|^2$ as transmitter k_1 becomes active and stays at this value over the interval $[\frac{\nu_{k_1}}{1+\alpha}, \frac{\nu_{k_0}+\alpha}{1+\alpha}]$.

The function F_n converges uniformly to F, i.e., for arbitrary $\epsilon > 0$, $\|F_n - F\|_\infty < \epsilon$ for sufficiently large n. As a specific example, let us consider a scenario where there are $K + 1 = 3$ active users in the network, $\nu_{k_0} = 0.3106, \nu_{k_1} = 0.4104, \nu_{k_2} = 0.6959$ and $h_{k_0} = 1.4878e^{0.2932\sqrt{-1}}, h_{k_1} = 0.9014e^{3.0291\sqrt{-1}}, h_{k_2} = 1.1134e^{-2.7195\sqrt{-1}}$. It is assumed that $\alpha = 0.5$, $\delta = 0.6$, $\theta = 0.6$ and $P = 20\,\text{dB}$. Fig. 2.2 presents plots of the sample paths $F_n(\tau)$ in terms of τ for different values of n. We observe that as n increases, F_n becomes closer to F in the sense that $\|F_n - F\|_\infty$ decreases. This property of F_n enables receiver 0 to obtain estimates $|\widehat{h}_{k_i}|$ and $\widehat{\nu}_{k_i}$ for $|h_{k_i}|$ and ν_{k_i}, respectively[4], by applying piecewise-linear regression to the graph of F_n. The slopes and breaking points of the resulting piecewise-linear approximation are used to obtain $|\widehat{h}_{k_i}|$ and $\widehat{\nu}_{k_i}$.

[4] It is more accurate to write $\widehat{|h_{k_i}|}$ instead of $|\widehat{h}_{k_i}|$ as receiver 0 only estimates the absolute value of h_{k_i}. However, we adopt the notation $|\widehat{h}_{k_i}|$ for simplicity. Moreover, $|\widehat{h}_{k_i}|$ and $\widehat{\nu}_{k_i}$ depend on n which is again dropped from notation for simplicity.

To describe the estimation procedure in more detail, Fig. 2.3 presents schematic plots of $F_n(\tau)$ and $F(\tau)\pm\epsilon$ in terms of τ near the point $\tau = \frac{\nu_{k_i}}{1+\alpha}$ where transmitter k_i becomes active. It is assumed that n is large enough so that $\|F_n - F\| \leq \epsilon$. Activation of transmitter k_i results in a jump in the slope of F. Let us denote the slope of F immediately before and after $\tau = \frac{\nu_{k_i}}{1+\alpha}$ by a_- and a_+, respectively. We assume ϵ is sufficiently small so that receiver 0 is able to find points (shown in filled squares) on the graph of F_n for $\tau < \frac{\nu_{k_i}}{1+\alpha}$, while the slope of F is a_- and for $\tau > \frac{\nu_{k_i}}{1+\alpha}$, while the slope of F is a_+. Receiver 0 constructs the regression lines (shown in dashed lines) for these two groups of points and announces the τ-coordinate of the point where the regression lines intersect as $\frac{\widehat{\nu}_{k_i}}{1+\alpha}$. Moreover, the jump in the slope of the regression line is the estimate $|\widehat{h}_{k_i}|^2$ for $|h_{k_i}|^2$. This estimation technique is analyzed in Appendix B where we prove the following corollary:

Corollary 1. *The estimators* $|\widehat{\boldsymbol{h}}_{k_i}|$ *and* $\widehat{\boldsymbol{\nu}}_{k_i}$ *satisfy*

$$\mathbb{P}\big(\big||\widehat{\boldsymbol{h}}_{k_i}| - |h_{k_i}|\big| > \epsilon\big) \leq \Theta(1)e^{-\epsilon^2\Theta(n)} \tag{2.10}$$

and

$$\mathbb{P}(|\widehat{\boldsymbol{\nu}}_{k_i} - \nu_{k_i}| > \epsilon) \leq \Theta(1)e^{-\epsilon^2\Theta(n)}. \tag{2.11}$$

In particular, $\lim_{n\to\infty} |\widehat{\boldsymbol{h}}_{k_i}| = |h_{k_i}|$ *and* $\lim_{n\to\infty} \widehat{\boldsymbol{\nu}}_{k_i} = \nu_{k_i}$.

Proof. See Appendix B. □

Motivated by Corollary 1, we assume receiver 0 has learnt the exact values of K and $|h_{k_i}|$. However, it is unrealistic to assume receiver 0 knows the exact values of ν_{k_i}. This is because a small difference $\widehat{\nu}_{k_i} - \nu_{k_i}$ can translate to a large difference $\widehat{t}_{k_i,n} - t_{k_i,n} = \lfloor\lfloor n/\alpha\rfloor\widehat{\nu}_{k_i}\rfloor - \lfloor\lfloor n/\alpha\rfloor\nu_{k_i}\rfloor$. Let us formalize our observations under Assumption 1:

Assumption 1. *Receiver 0 is able to find the exact values of* K *and* $|h_{k_i}|$ *for* $0 \leq i \leq K$. *Moreover, it is able to construct an estimator* $\widehat{\boldsymbol{\nu}}_{k_i}$ *for* ν_{k_i} *such that* $\mathbb{P}(|\widehat{\boldsymbol{\nu}}_{k_i} - \nu_{k_i}| > \epsilon) \leq \Theta(1)e^{-\epsilon^2\Theta(n)}$ *for any* $\epsilon > 0$.

2.3.2 Estimating the activation time $t_{0,n}$

It is essential for receiver 0 to identify the exact time slot $t_{0,n}$ when transmitter 0 starts its activity. According to our results thus far, receiver 0 can obtain the estimate

$$\widehat{\boldsymbol{t}}_{k_i,n} = \lfloor\lfloor n/\alpha\rfloor\widehat{\boldsymbol{\nu}}_{k_i}\rfloor, \tag{2.12}$$

for $t_{k_i,n}$. By Assumption 1, $\widehat{\boldsymbol{\nu}}_{k_i}$ can be made arbitrarily close to ν_{k_i}, however, there is no guarantee that $\mathbb{P}(\widehat{\boldsymbol{t}}_{k_i,n} \neq t_{k_i,n})$ vanishes as n increases. Sending a preamble sequence by transmitter 0 enables receiver 0 to obtain a finer estimate for $t_{0,n}$.

We start with the following lemma:

Proposition 2. *There exists a positive integer N depending on c, c', α and δ such that*

$$\mathbb{P}\big(|\widehat{\boldsymbol{t}}_{k_i,n} - t_{k_i,n}| > 2cn_{\delta}\big) \leq \Theta(1)e^{-\Theta(n^{2\delta-1})}, \tag{2.13}$$

for any i and $n \geq N$.

Proof. By the floor inequality, $|\widehat{\boldsymbol{t}}_{k_i,n} - t_{k_i,n}| > 2cn_{\delta}$ implies $|\widehat{\boldsymbol{\nu}}_{k_i} - \nu_{k_i}| > \frac{2\alpha(cn_{\delta}-1)}{n}$. Therefore,

$$\mathbb{P}(|\widehat{\boldsymbol{t}}_{k_i,n} - t_{k_i,n}| > 2cn_{\delta}) \leq \mathbb{P}\Big(|\widehat{\boldsymbol{\nu}}_{k_i} - \nu_{k_i}| > \frac{2\alpha(cn_{\delta}-1)}{n}\Big) \leq \Theta(1)e^{-\left(\frac{2\alpha(cn_{\delta}-1)}{n}\right)^2\Theta(n)}, \tag{2.14}$$

where the last step is due to (2.11) for the choice of $\epsilon = \epsilon_n := \frac{2\alpha(cn_{\delta}-1)}{n}$ and n is assumed to be large enough so that $cn_{\delta} > 1$. By Proposition 1, one needs to check if $\epsilon_n T_n > \max\{cn_{\delta}, c'\}$ in order to guarantee (2.14) holds. Since $\epsilon_n T_n = \frac{2\alpha(cn_{\delta}-1)T_n}{n} = \Theta(n_{\delta})$ and c' is a constant, there is N_1 such that $\epsilon_n T_n > c'$ for $n \geq N_1$. Moreover, $\frac{\alpha T_n}{n} > \frac{\alpha(\frac{n}{\alpha}-1+n+n_{\delta}-1)}{n} = 1+\frac{n+n_{\delta}-2}{n}\alpha \geq 1$. Therefore, $\epsilon_n T_n = \frac{\alpha T_n}{n} \times 2(cn_{\delta}-1) > 2(cn_{\delta}-1)$. As such, $\epsilon T_n > cn_{\delta}$ holds if $cn_{\delta} > 2$. Let N_2 be such that $cn_{\delta} > 2$ for $n \geq N_2$. Finally, $(\frac{2\alpha(cn_{\delta}-1)}{n})^2\Theta(n) = \Theta(\frac{n_{\delta}^2}{n}) = \Theta(n^{2\delta-1})$. This completes the proof of proposition by letting $N = \max\{N_1, N_2\}$. ☐

Motivated by Proposition 2, we make the following assumption:

Assumption 2. *Receiver 0 estimates the delay $t_{k_i,n}$ within an interval of length at most $4cn_{\delta}$ around its actual value, i.e., $|\widehat{t}_{k_i,n} - t_{k_i,n}| \leq 2cn_{\delta}$ for any i. Without loss of generality, $2c$ is an integer. Otherwise, one can replace $2c$ by $\max\{1, \lfloor 2c \rfloor\}$.*

In fact, for given $\epsilon > 0$, one can assume n is large enough such that $\mathbb{P}(\exists i : |\widehat{\boldsymbol{t}}_{k_i,n} - t_{k_i,n}| \leq 2cn_{\delta}) > 1-\epsilon$ and add ϵ to the probability of error in decoding the message sent by transmitter 0.

By Assumption 2, the preamble sequence for transmitter k_i starts no earlier than time slot $\widehat{t}_{k_i,n} - 2cn_{\delta}$ and ends no later than time slot $\widehat{t}_{k_i,n} + (2c+1)n_{\delta}$. Define

$$\mathsf{A}_{i,n} := \{\widehat{t}_{k_i,n} - 2cn_{\delta}, \cdots, \widehat{t}_{k_i,n} + (2c+1)n_{\delta}\}. \tag{2.15}$$

The burst for transmitter k_i starts no earlier than time slot $\widehat{t}_{k_i,n} - 2cn_\delta$ and no later than time slot $\widehat{t}_{k_i,n} + 2cn_\delta$. Define

$$\mathsf{B}_{i,n} := \{\widehat{t}_{k_i,n} - 2cn_\delta, \cdots, \widehat{t}_{k_i,n} + 2cn_\delta\}. \tag{2.16}$$

The burst for transmitter k_i ends no earlier than time slot $\widehat{t}_{k_i,n} + n + (1-2c)n_\delta - 1$ and no later than time slot $\widehat{t}_{k_i,n} + n + (1+2c)n_\delta - 1$. Define

$$\mathsf{C}_{i,n} := \{\widehat{t}_{k_i,n} + n + (1-2c)n_\delta - 1, \cdots, \widehat{t}_{k_i,n} + n + (1+2c)n_\delta - 1\}. \tag{2.17}$$

The next assumption simplifies the description and analysis for the sequential typicality decoder:

Assumption 3. *For any $i \neq j$, there exists a constant $N^* \geq 1$ such that $\mathsf{A}_{i,n} \bigcap \mathsf{B}_{j,n} = \emptyset$ and $\mathsf{A}_{i,n} \bigcap \mathsf{C}_{j,n} = \emptyset$ for any $n \geq N^*$. This implies that no transmitter starts or ends its transmission while another transmitter is sending its preamble sequence.*

The next proposition justifies Assumption 3:

Proposition 3. *For any $\epsilon > 0$, there is an integer N_ϵ such that the probability of Assumption 3 being false is less that ϵ if one selects $N^* = N_\epsilon$.*

Proof. See Appendix C. □

To find $t_{0,n}$, receiver 0 uses the so-called sequential typicality decoder [21]. Recall that for given $\epsilon > 0$, $m \geq 1$ and a PDF $p(\cdot,\cdot)$ on $\mathbb{C}^2$ with marginals p_1 and p_2, the typical set $A_\epsilon^{(m)}[p]$ is the set of all pairs $(\vec{x}, \vec{y})$ where $\vec{x}, \vec{y} \in \mathbb{C}^m$ and the following three inequalities hold:

$$\left| \frac{1}{m} \sum_{i=1}^{m} \log p_1(x_i) + h(p_1) \right| < \epsilon, \tag{2.18}$$

$$\left| \frac{1}{m} \sum_{i=1}^{m} \log p_2(y_i) + h(p_2) \right| < \epsilon \tag{2.19}$$

and

$$\left| \frac{1}{m} \sum_{i=1}^{m} \log p(x_i, y_i) + h(p) \right| < \epsilon. \tag{2.20}$$

We refer to any $(\vec{x}, \vec{y}) \in A_{\epsilon}^{(m)}[p]$ as an ϵ-jointly typical pair with respect to p [28].

In order to describe how the sequential typicality decoder operates, we find it best to look at the specific example in Fig. 2.2. By Assumption 3, the number of active users is fixed while transmitter 0 is sending its preamble sequence $(\boldsymbol{x}'_{l,0})_{l=0}^{n_\delta-1}$. Let us denote the PDF of noise plus interference in each time slot during the transmission of $(\boldsymbol{x}'_{l,0})_{l=0}^{n_\delta-1}$ by $p_{\text{NI}}(\cdot)$. According to Fig. 2.2, if transmitter 0 is the first active transmitter ($k_0 = 0$),

$$p_{\text{NI}}(w) = p_{\text{NI},k_0}(w) := \text{g}(w; 1), \tag{2.21}$$

if transmitter 0 is the second active transmitter ($k_1 = 0$),

$$p_{\text{NI}}(w) = p_{\text{NI},k_1}(w) := \text{g}\big(w; 1 + |h_{k_0}|^2 P\big) \tag{2.22}$$

and if transmitter 0 is the third active transmitter ($k_2 = 0$),

$$p_{\text{NI}}(w) = p_{\text{NI},k_2}(w) := \text{g}\big(w; 1 + (|h_{k_0}|^2 + |h_{k_1}|^2)P\big). \tag{2.23}$$

Let $p_{k_i}(x, y) = \text{g}(x; P)p_{\text{NI},k_i}(y - h_0 x)$ be the joint PDF of the input and output of user 0 while this user is transmitting its preamble sequence. Receiver 0 estimates $t_{0,n}$ as the unique integer $t \in \mathsf{B}_{i,n}$ such that

$$\big((\boldsymbol{x}'_{0,l})_{l=0}^{n_\delta-1}, (\boldsymbol{y}_0[l])_{l=t}^{t+n_\delta-1}\big) \in A_{\epsilon}^{(n_\delta)}[p_{k_i}], \tag{2.24}$$

for exactly one choice of $0 \leq i \leq 2$. We denote this estimate of $t_{0,n}$ by $\widehat{\widehat{t}}_{0,n}$. The next proposition shows that the probability of error in estimating $t_{0,n}$ vanishes as n grows to infinity.

Proposition 4. *Let $\widehat{\widehat{\boldsymbol{t}}}_{0,n}$ be the estimator of $t_{0,n}$ as described in above. Then*

$$\mathbb{P}\big(\widehat{\widehat{\boldsymbol{t}}}_{0,n} = t_{0,n}\big) \geq 1 - \Theta(n_\delta^2)e^{-\Theta(n_\delta)}. \tag{2.25}$$

In particular, $\lim_{n\to\infty} \mathbb{P}\big(\widehat{\widehat{\boldsymbol{t}}}_{0,n} = t_{0,n}\big) = 1$.

Proof. See Appendix D. ∎

2.4 Decoding strategy and achievable rates

In the previous section, we demonstrated how receiver 0 estimates the number K of potential interferers, the channel gains and delays of active transmitters. Moreover, it was established that the probability of error in estimating $t_{0,n}$ by receiver 0 vanishes as n grows. In this section, we explore reliable communication between transmitter 0 and receiver 0. For simplicity, it is assumed $K = 1$ throughout this section and the only potential interferer on user 0 is user 1. Extending the achievability results to general K is straightforward.

We begin by listing several assumptions that facilitate the description of the decoding strategy and error analysis:

Assumption 4. *Receiver 0 knows the exact value of* $t_{0,n}$.

Assumption 4 is motivated by Proposition 4. In fact, for given $\epsilon > 0$, one can assume n is large enough such that $\mathbb{P}\big(\widehat{\boldsymbol{t}}_{0,n} = t_{0,n}\big) > 1 - \epsilon$ and add ϵ to the probability of error in decoding the message sent by transmitter 0.

Having the assumptions in above, let us explain the decoding scheme. Description of the decoding scheme depends on the sign of $\nu_1 - \nu_0$ and whether $|\nu_1 - \nu_0|$ is smaller or larger than α. Here, we only consider the case $0 < \nu_1 - \nu_0 < \alpha$. Other cases can be explained similarly. By Assumption 2, transmitter 1 starts its activity no earlier than $\widehat{t}_{1,n} - 2cn_\delta$ and no later than $\widehat{t}_{1,n} + 2cn_\delta$. Recalling that $t_{0,n} + n_\delta$ and $t_{0,n} + n_\delta + n - 1$ are the time slots transmitter 0 sends the first and last symbols in its codeword, respectively, we have

$$\begin{aligned}
(\widehat{t}_{1,n} - 2cn_\delta) - (t_{0,n} + n_\delta) &\overset{(a)}{\geq} (t_{1,n} - 2cn_\delta - 2cn_\delta) - (t_{0,n} + n_\delta) \\
&= t_{1,n} - t_{0,n} - (4c+1)n_\delta \\
&\overset{(b)}{>} \big((n/\alpha - 1)\nu_1 - 1\big) - (n/\alpha)\nu_0 - (4c+1)n_\delta \\
&> \frac{\nu_1 - \nu_0}{\alpha} n - (4c+1)n_\delta - \nu_1 - 1, \qquad (2.26)
\end{aligned}$$

and

$$\begin{aligned}
(t_{0,n} + n_\delta + n - 1) - (\widehat{t}_{1,n} + 2cn_\delta) &\overset{(a)}{\geq} (t_{0,n} + n_\delta + n - 1) - (t_{1,n} + 2cn_\delta + 2cn_\delta) \\
&= n + t_{0,n} - t_{1,n} - (4c-1)n_\delta \\
&\overset{(b)}{>} n + \big((n/\alpha - 1)\nu_0 - 1\big) - (n/\alpha)\nu_1 - (4c-1)n_\delta \\
&> n\Big(1 - \frac{\nu_1 - \nu_0}{\alpha}\Big) - (4c+1)n_\delta - \nu_0 - 1, \qquad (2.27)
\end{aligned}$$

where in both (2.26) and (2.27), (a) is due to Assumption 2 and (b) is due to the floor inequality. Since $0 < \nu_1 - \nu_0 < \alpha$, the terms on the right sides of (2.26) and (2.27) are positive if n is sufficiently large. As such, user 0 experiences partial interference on its codeword in the sense that there exists $t_{0,n} + n_{\delta} < t < t_{0,n} + n_{\delta} + n - 1$ such that any symbol in the codeword of user 0 transmitted in time slot t or later is received in the presence of interference. By Assumption 4, receiver 0 knows the exact value of $t_{0,n}$ and therefore, it can divide the collection of time slots $t_{0,n} + n_{\delta}, \cdots, t_{0,n} + n_{\delta} + n - 1$ into three groups as shown in Fig. 2.4, namely,

- The initial interval consisting of time slots $t_{0,n} + n_{\delta}, \cdots, \widehat{t}_{1,n} - 2cn_{\delta} - 1$. The number of time slots in this interval is $\widehat{t}_{1,n} - t_{0,n} - (2c+1)n_{\delta}$.
- The ambiguity interval consisting of time slots $\widehat{t}_{1,n} - 2cn_{\delta}, \cdots, \widehat{t}_{1,n} + 2cn_{\delta}$. The number of time slots in this interval is $4cn_{\delta} + 1$.
- The final interval consisting of time slots $\widehat{t}_{1,n} + 2cn_{\delta} + 1, \cdots, t_{0,n} + n_{\delta} + n - 1$. The number of time slots in this interval is $t_{0,n} - \widehat{t}_{1,n} - (2c-1)n_{\delta} + n - 1$.

Receiver 0 knows for a fact that there is no interference during the initial interval, it is uncertain about the presence of interference in any time slot during the ambiguity interval and it knows for a fact that interference is present in each time slot during the final interval. Therefore, the PDF of noise plus interference per time slot at receiver 0 during the initial interval and the final interval is given by

$$p_{\text{NI,I}}(w) = \mathrm{g}(w; 1) \tag{2.28}$$

and

$$p_{\text{NI,F}}(w) = \mathrm{g}(w; 1 + |h_1|^2 P), \tag{2.29}$$

respectively. Let

$$p_{\text{I}}(x, y) = \mathrm{g}(x; P) p_{\text{NI,I}}(y - h_0 x) \tag{2.30}$$

and

$$p_{\text{F}}(x, y) = \mathrm{g}(x; P) p_{\text{NI,F}}(y - h_0 x). \tag{2.31}$$

Receiver 0 finds the unique codeword $(\boldsymbol{x}_{0,l})_{l=0}^{n-1}$ such that both

$$\left((\boldsymbol{x}_{0,l})_{l=0}^{\widehat{t}_{1,n} - t_{0,n} - (2c+1)n_{\delta} - 1}, (\boldsymbol{y}_0[l])_{l=t_{0,n}+n_{\delta}}^{\widehat{t}_{1,n} - 2cn_{\delta} - 1} \right) \in A_{\epsilon}^{(\widehat{t}_{1,n} - t_{0,n} - (2c+1)n_{\delta})}[p_{\text{I}}] \tag{2.32}$$

and

$$\Big((\boldsymbol{x}_{0,l})_{l=\widehat{t}_{1,n}-t_{0,n}+(2c-1)n_\delta+1}^{n-1}, (\boldsymbol{y}_0[l])_{l=\widehat{t}_{1,n}+2cn_\delta+1}^{t_{0,n}+n_\delta+n-1}\Big) \in A_\epsilon^{(\,t_{0,n}-\widehat{t}_{1,n}-(2c-1)n_\delta+n-1)}[\,p_{\mathrm{F}}] \tag{2.33}$$

are satisfied. The decoding rule in (2.32) requires the first $\widehat{t}_{1,n} - t_{0,n} - (2c+1)n_\delta$ symbols in a codeword be jointly typical with the received signals during the initial interval. Also, the decoding rule in (2.33) requires that the last $t_{0,n} - \widehat{t}_{1,n} - (2c-1)n_\delta + n - 1$ symbols in a codeword be jointly typical with the received signals during the final interval.

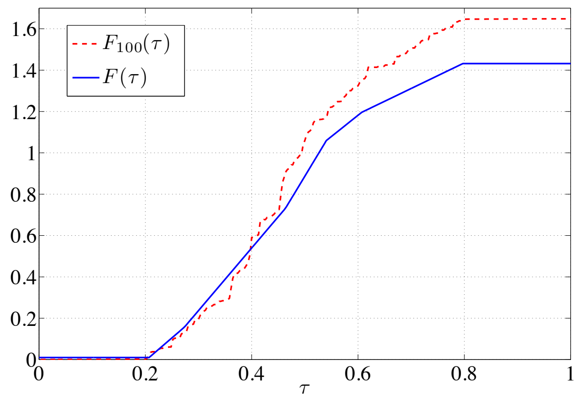
We are ready to present the main result of this section:

Proposition 5. *Define the achievable rate for user 0 by*

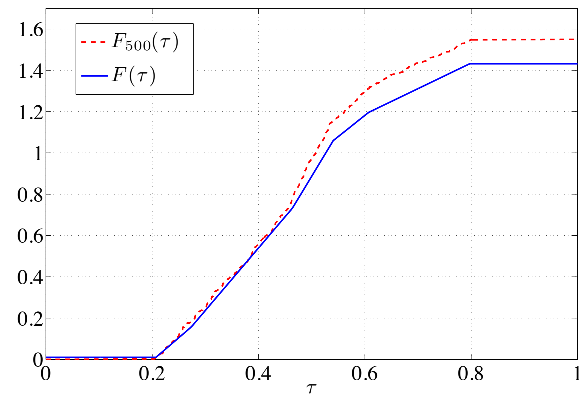
$$\underline{R} := \min\{1, |\nu_1 - \nu_0|/\alpha\} \log(1 + |h_0|^2 P) + (1 - |\nu_1 - \nu_0|/\alpha)^+ \log\Big(1 + \frac{|h_0|^2 P}{1 + |h_1|^2 P}\Big). \tag{2.34}$$

If $R < \underline{R}$, the probability of error in decoding the message of user 0 can be made arbitrarily small by letting n be sufficiently large.

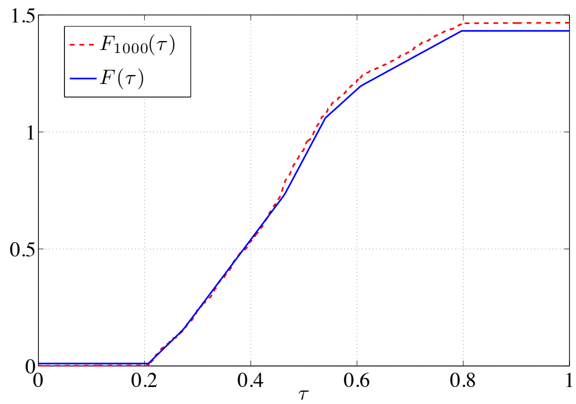
Proof. See Appendix E. □



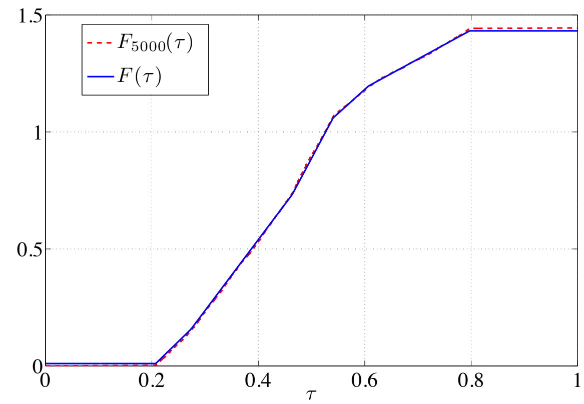
(a) $n = 100$



(b) $n = 500$



(c) $n = 1000$



(d) $n = 5000$

Figure 2.2: Plots of $F(\tau)$ and $F_n(\tau)$ in terms of τ for different values of n. It is assumed that there are $K+1=3$ active transmitters, $\nu_{k_0} = 0.3106, \nu_{k_1} = 0.4104, \nu_{k_2} = 0.6959$, $h_{k_0} = 1.4878e^{0.2932\sqrt{-1}}, h_{k_1} = 0.9014e^{3.0291\sqrt{-1}}, h_{k_2} = 1.1134e^{-2.7195\sqrt{-1}}$, $P = 20\,\text{dB}$, $\alpha = 0.5$, $\delta = 0.6$ and $\theta = 0.6$. Note that activation of user k_i corresponds to $\tau = \frac{\nu_{k_i}}{1+\alpha}$ on the τ-axis, i.e., $\frac{\nu_{k_0}}{1+\alpha} = 0.2071$, $\frac{\nu_{k_1}}{1+\alpha} = 0.2736$ and $\frac{\nu_{k_2}}{1+\alpha} = 0.464$. By increasing n, F_n becomes closer to F in the sense that $\|F_n - F\|_\infty$ decreases.

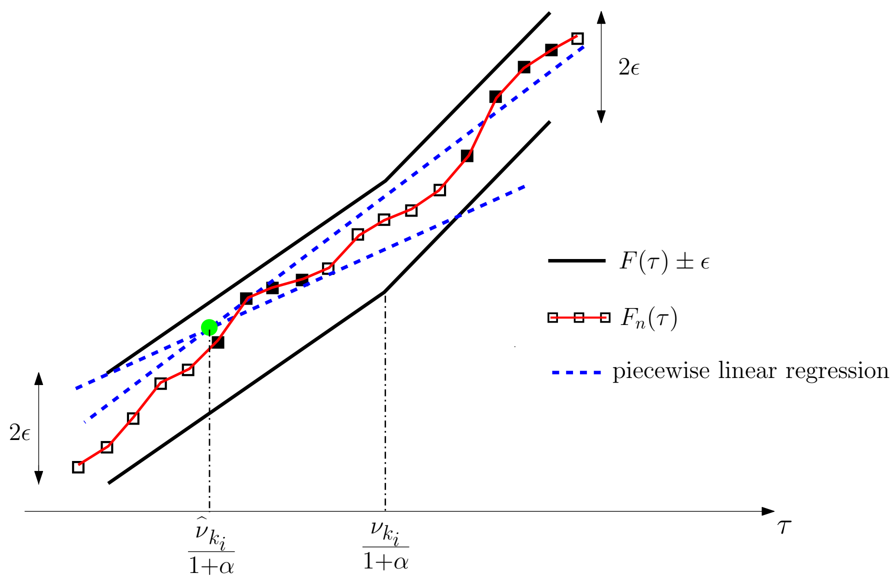


Figure 2.3: Plots of $F(\tau) \pm \epsilon$ and $F_n(\tau)$ in terms of τ. Transmitter k_i become active at $\tau = \frac{\nu_{k_i}}{1+\alpha}$ which results in an increase in the slope of F. Let us denote the slope of F immediately before and after $\tau = \frac{\nu_{k_i}}{1+\alpha}$ by $a_{\_}$ and a_{+}, respectively. We assume ϵ is sufficiently small so that receiver 0 is able to find points (shown in filled squares) over the graph of F_n for $\tau < \frac{\nu_{k_i}}{1+\alpha}$, while the slope of F is $a_{\_}$ and for $\frac{\nu_{k_i}}{1+\alpha} < \tau$, while the slope of F is a_{+}. Receiver 0 finds the regression lines (shown in dashed lines) for these two groups of points and announces the τ-coordinate of the point where these regression lines intersect as $\frac{\widehat{\nu}_{k_i}}{1+\alpha}$. Moreover, the increase in slope of the regression lines is $|\widehat{h}_{k_i}|^2$.

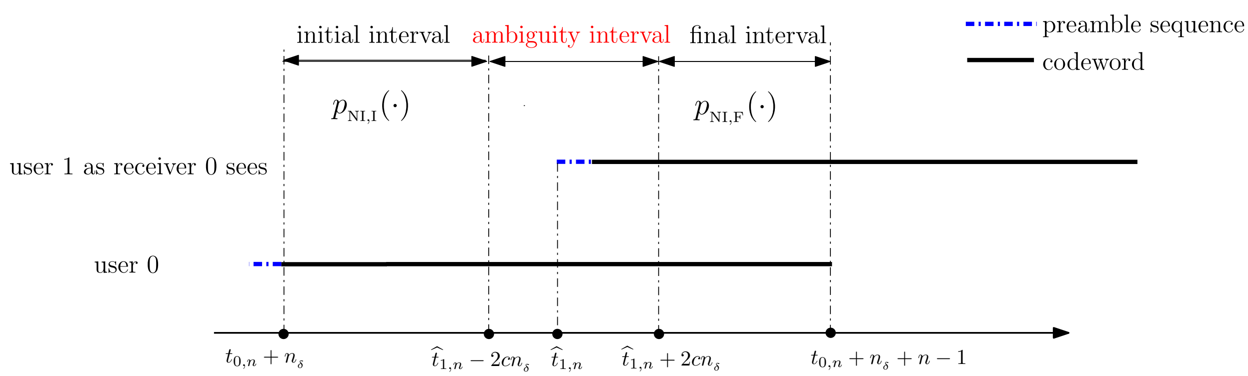


Figure 2.4: Relative positions of the codewords sent by transmitter 0 and transmitter 1 across the time axis from the viewpoint of receiver 0. It is assumed that $0 < \nu_1 - \nu_0 < \alpha$. By Assumption 4, receiver 0 knows the value of $t_{0,n}$, however, it does not know the value of $t_{1,n}$. Receiver 0 is able to find the estimate $\widehat{t}_{1,n}$ such that $|\widehat{t}_{1,n} - t_{1,n}| \leq 2cn_\delta$ by Assumption 2. The codeword of user 0 is received during the time slots $t_{0,n}+n_\delta, \cdots, t_{0,n}+n_\delta+n-1$. These time slots are divided into three intervals, namely, the initial interval where receiver 0 knows for a fact that interference is absent, the ambiguity interval where receiver 0 is unable to identify the presence of interference and the final interval where receiver 0 knows for a fact that interference is present.

Chapter 3

Asynchronous Communication over Gaussian Interference Channels with Stochastic Data Arrival

3.1 Notations and terminology

Random quantities are shown in bold such as $\boldsymbol{x}$ with realization x. Sets and in particular, events are shown using capital calligraphic or cursive letters such that $\mathcal{A}$ or $\mathscr{A}$. The set difference for two sets $\mathcal{A}$ and $\mathcal{B}$ is denoted by $\mathcal{A} \setminus \mathcal{B}$. The underlying probability measure and the expectation operator are denoted by $\mathbb{P}(\cdot)$ and $\mathbb{E}[\cdot]$, respectively. For a real number x, the floor of x is $\lfloor x \rfloor$ and the ceiling of x is $\lceil x \rceil$. A binomial random variable with parameters n (number of trials) and p (probability of success) is denoted by $\mathrm{Bin}(n, p)$. A negative binomial random variable with parameters k and p, denoted by $\mathrm{NB}(k, p)$, is defined to be the number of trials until k successes are observed where p is the probability of success. The probability density function (PDF) of a Gaussian random variable with zero mean and variance σ^2 is denoted by $\mathrm{g}(x;\sigma^2) := \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}}$. The differential entropy of a continuous random variable $\boldsymbol{x}$ with PDF $p(\cdot)$ is denoted by $h(\boldsymbol{x})$ or $h(p)$. For two functions f and g of a real variable x, we write $f = \Theta(g)$ if there is x_0 and constants c_1 and c_2 such that $c_1 g(x) \leq f(x) \leq c_2 g(x)$ for $x \geq x_0$. We define

$$\mathsf{C}(x) := \frac{1}{2} \log(1 + x). \tag{3.1}$$

All logarithms have base 2. Throughout the chapter,

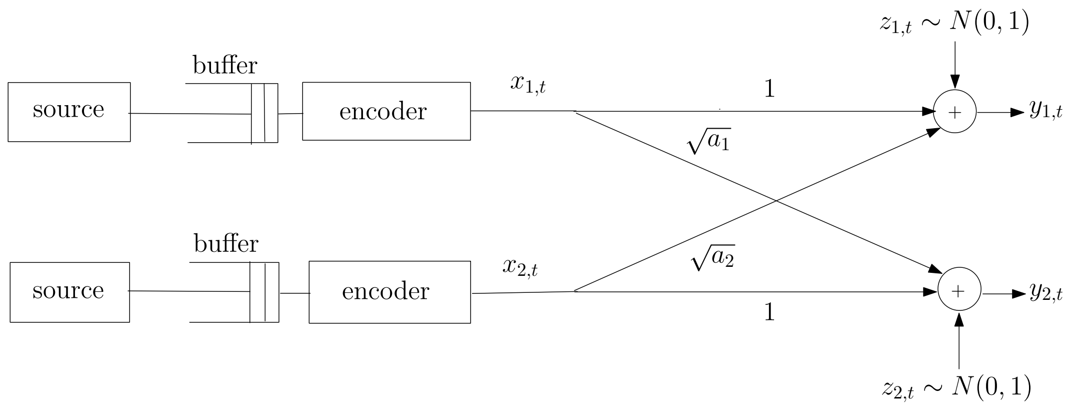


Figure 3.1: This figure shows a two-user GIC with stochastic data arrival (GIC-SDA). The source of Tx i generates k_i bits per time slot with a probability of q_i and turns off after a total of $k_i n$ bits are generated. The links from each transmitter to each receiver are modelled by static and non-frequency selective coefficients. The signals at the transmitters are subject to an average power constraint and the noise at each receiver is an AWGN process with unit variance.

- Any equality or inequality involving random variables is understood in the "almost sure" sense unless otherwise stated. We avoid repeating "almost surely" throughout the chapter.
- "SLLN" stands for "the strong law of large numbers".
- The symbol ":=" means "is defined by"

3.2 System Model

3.2.1 Signalling and channel model

We consider a GIC with two users of separate Tx-Rx pairs shown in Fig. 3.1. The channel from Tx i to Rx j is modelled by a static and non-frequency selective coefficient $h_{i,j}$ where $h_{1,1} = h_{2,2} = 1$, $h_{2,1} = \sqrt{a_2}$ and $h_{1,2} = \sqrt{a_1}$. The channel from each transmitter to each receiver is slotted in time and the time slots on any of the four channels from different transmitters to different receivers coincide. Therefore, the two users are symbol-synchronous. Throughout the chapter, we show the time slots using the index $t = 1, 2, \cdots$.

If we are describing a property for user i, the index i' refers to the other user, i.e., $i' = 3-i$ for $i = 1, 2$. Denoting the signal at Tx i in time slot t by $x_{i,t}$, we impose the average power constraint

$$Q_i := \frac{1}{|\mathcal{T}_i|} \sum_{t \in \mathcal{T}_i} x_{i,t}^2 \leq P_i, \tag{3.2}$$

where $\mathcal{T}_i$ is the communication period of interest for Tx i and $|\mathcal{T}_i|$ denotes the length of $\mathcal{T}_i$. The signal $y_{i,t}$ received at Rx i in time slot t is given by

$$y_{i,t} = x_{i,t} + \sqrt{a_{i'}}\, x_{i',t} + z_{i,t}, \quad i = 1, 2, \tag{3.3}$$

where $z_{i,t}$ is the additive noise at Rx i in time slot t. The noise at each receiver is an additive white Gaussian noise (AWGN) process with unit variance.

3.2.2 Data arrival

Each transmitter is connected to an information source through a buffer as shown in Fig. 3.1. At the "beginning" of time slot $t = 1$, the buffers are empty. At the "end" of each time slot, a number of k_i bits arrive at the buffer of Tx i with a probability of q_i or no bits arrive with a probability of $1 - q_i$.[1] The rate of data arrival at Tx i is denoted by

$$\lambda_i := k_i q_i. \tag{3.4}$$

The bit streams entering the buffers of the two users are independent processes. Source i is turned off permanently after it generates a total number of $k_i n$ bits where n runs in the set of positive integers. To transmit its data, Tx i employs a codebook consisting of $2^{\lfloor n\eta_i \rfloor}$ codewords of length

$$n_i := \lfloor n\theta_i \rfloor, \tag{3.5}$$

where $\eta_i, \theta_i > 0$. Note that the codebook rate for Tx i is $\frac{\lfloor n\eta_i \rfloor}{\lfloor n\theta_i \rfloor}$. Assuming η_i has the particular expression

$$\eta_i = \frac{k_i}{N_i}, \tag{3.6}$$

for integers N_1 and N_2, we see that Tx i sends a total number of N_i codewords where each codeword represents $\lfloor n\eta_i \rfloor = \lfloor \frac{k_i n}{N_i} \rfloor$ of the bits stored in its buffer. The number of bits that are not transmitted is equal to $k_i n - N_i \lfloor \frac{k_i n}{N_i} \rfloor \leq N_i$ which is negligible in the limit of large

[1] All results in the chapter are still valid as long as the incoming bit stream is a random process with independent and identically distributed inter-arrival periods with finite mean value.

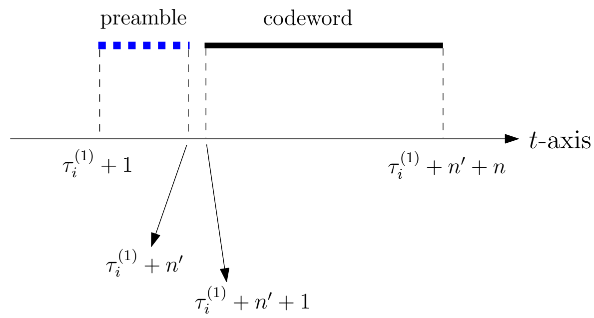


Figure 3.2: This figure shows the first transmission burst of Tx i along the t-axis. At the end of time slot $\tau_i^{(1)}$ the number of bits in the buffer of Tx i become larger than or equal to $\lfloor n\eta_i \rfloor$ for the first time. A number of $\lfloor n\eta_i \rfloor$ bits in the buffer of Tx i are represented by a codeword which together with the preamble sequence are sent during time slots $\tau_i^{(1)}+1, \cdots, \tau_i^{(1)}+n'+n_i$.

n. Before a codeword is transmitted over the channel, each transmitter sends a preamble sequence of length n' where[2]

$$n' = o(n). \tag{3.7}$$

Each preamble sequence enables the receivers to identify the arrival of a codeword. Details on the preamble sequences and how they are utilized are provided in Section 3.3.

Let $b_{i,t}$ be the number of bits in the buffer of user i at the "beginning" of time slot t, $b'_{i,t}$ be the number of bits entering the buffer of user i at the "end" of time slot t and $\tau_i^{(1)}$ be the smallest index $t \geq 1$ such that $b_{i,t} + b'_{i,t} \geq \lfloor n\eta_i \rfloor$. At time slot $t = \tau_i^{(1)}+1$, a number of $\lfloor n\eta_i \rfloor$ bits in the buffer of Tx i are represented by a codeword which together with the preamble sequence are sent during time slots $\tau_i^{(1)}+1, \cdots, \tau_i^{(1)}+n'+n_i$. This is referred to as a *transmission burst* or simply a burst as shown in Fig. 3.2. These $\lfloor n\eta_i \rfloor$ bits are erased from the buffer of Tx i, i.e.,

$$b_{i,\tau_i^{(1)}+1} = b_{i,\tau_i^{(1)}} + b'_{i,\tau_i^{(1)}} - \lfloor n\eta_i \rfloor. \tag{3.8}$$

Let $\tau_i^{(2)}$ be the smallest index $t \geq \tau_i^{(1)} + n' + n_i$ such that $b_{i,t} + b'_{i,t} \geq \lfloor n\eta_i \rfloor$. At time slot $t = \tau_i^{(2)}$, a second group of $\lfloor n\eta_i \rfloor$ bits in the buffer are scheduled for transmission. These bits are represented by a codeword which together with the preamble sequence are sent

[2]This means $\lim_{n\to\infty} \frac{n'}{n} = 0$.

during time slots $\tau_i^{(2)}+1,\cdots,\tau_i^{(2)}+n'+n_i$ and we have

$$b_{i,\tau_i^{(2)}+1}=b_{i,\tau_i^{(2)}}+b'_{i,\tau_i^{(2)}}-\lfloor n\eta_i\rfloor. \tag{3.9}$$

In general, $\tau_i^{(j)}$ is defined by

$$\tau_i^{(j)}=\min\left\{t\geq\tau_i^{(j-1)}+n'+n_i : b_{i,t}+b'_{i,t}\geq\lfloor n\eta_i\rfloor\right\}. \tag{3.10}$$

At time slot $\tau_i^{(j)}$, a number of $\lfloor n\eta_i\rfloor$ bits in the buffer of Tx i are represented by a codeword which together with the preamble sequence are sent during time slots $\tau_i^{(j)}+1,\cdots,\tau_i^{(j)}+n'+n_i$. Moreover,

$$b_{i,\tau_i^{(j)}+1}=b_{i,\tau_i^{(j)}}+b'_{i,\tau_i^{(j)}}-\lfloor n\eta_i\rfloor. \tag{3.11}$$

This style of transmission is *Tx-Rx asynchronous* in the sense that Rx i does not know a priori the time slots $\tau_i^{(1)}+1,\tau_i^{(2)}+1,\cdots$ when Tx i begins to send its bursts.

A few remarks are in order:

1. The Tx-Rx asynchronous transmission considered in this paper is in contrast to the Tx-Rx synchronous scheme[3] studied in [20] in the context of networking and information theory. In this scheme, Tx i sends its codewords only at time slots mn_i+1 where $m\geq 1$ is an integer.[4] The so-called augmented codebook of Tx i consists of $2^{\lfloor n\eta_i\rfloor}$ *data codewords* of length n_i and one additional codeword referred to as the *null codeword* with the same length n_i. At the "end" of time slot mn_i, if there are at least $\lfloor n\eta_i\rfloor$ bits in the buffer, a data codeword is transmitted over the channel during time slots $mn_i+1,\cdots,(m+1)n_i$. If the number of bits at the "end" of time slot mn_i is less than $\lfloor n\eta_i\rfloor$, the null codeword is transmitted over the channel during time slots $mn_i+1,\cdots,(m+1)n_i$ and Tx i repeats this process at time slot $(m+1)n_i$. Transmission of the null codeword facilitates the synchronization between a receiver and its corresponding transmitter. Lemma 24.1 in [20] guarantees that the buffer of Tx i is stable, i.e., $\sup_{t\geq 0}\mathbb{E}[\boldsymbol{b}_{i,t}]<\infty$, if and only if

$$\mu_i:=\frac{\eta_i}{\lambda_i}=\frac{1}{N_iq_i}>\theta_i. \tag{3.12}$$

[3] See chapter 24 on page 600.

[4] The description provided here for the scheme in [20] is given in terms of the notations introduced in this paper. Moreover, the communication scenarios studied in [20] are the point to point channel and the multiple access channel.

In the scheme considered in this paper, stability of the buffers is not an issue because Tx i only transmits a finite number $k_i n$ of bits and hence, the backlog (buffer content) is bounded from above by $k_i n$ at any time slot. However, we still impose the constraint in (3.12) for $N_i > 1$ because it guarantees immediate data transmission described in the next remark.

2. It is desirable that the transmissions be *immediate* in the following sense:

   We say the transmissions of Tx i are immediate if Tx i sends a codeword immediately after there are at least $\lfloor n\eta_i \rfloor$ bits stored in its buffer.

   Such immediate transmission is not possible if a previously scheduled codeword is not completely transmitted. More precisely, let

$$\widetilde{\tau}_i^{(1)} := \min\left\{t \geq \tau_i^{(1)} + 1 : b_{i,t} + b'_{i,t} \geq \lfloor n\eta_i \rfloor\right\}. \tag{3.13}$$

   Then $\widetilde{\tau}_i^{(1)}$ is the earliest time slot such that the buffer of Tx i contains at least $\lfloor n\eta_i \rfloor$ bits after the transmission of the first burst begun at time slot $\tau_i^{(1)} + 1$. If $\widetilde{\tau}_i^{(1)} \leq \tau_i^{(1)} + n' + n_i - 1$, these $\lfloor n\eta_i \rfloor$ bits must stay in the buffer until time slot $\tau_i^{(1)} + n' + n_i$ when the transmission of the first scheduled codeword is complete. In Appendix F it is shown that if (3.12) holds, then

$$\mathbb{P}\left(\widetilde{\boldsymbol{\tau}}_i^{(1)} \leq \boldsymbol{\tau}_i^{(1)} + n' + n_i - 1\right) \leq e^{-c_i n}, \tag{3.14}$$

   where $c_i > 0$ is a constant that does not depend on n. In virtue of (3.14) and for sufficiently large n, the second transmission is immediate with arbitrarily large probability. Next, define

$$\widetilde{\tau}_i^{(2)} := \min\left\{t \geq \tau_i^{(2)} + 1 : b_{i,t} + b'_{i,t} \geq \lfloor n\eta_i \rfloor\right\}. \tag{3.15}$$

   Then $\widetilde{\tau}_i^{(2)}$ is the earliest time slot such that the buffer of Tx i contains at least $\lfloor n\eta_i \rfloor$ bits after the transmission of the second burst begun at time slot $\tau_i^{(2)} + 1$. If $\widetilde{\tau}_i^{(2)} \leq \tau_i^{(2)} + n' + n_i - 1$, then these $\lfloor n\eta_i \rfloor$ bits must stay in the buffer until time slot $\tau_i^{(2)} + n' + n_i$ when the transmission of the second scheduled codeword is complete. Similar to (3.14),

$$\mathbb{P}\left(\widetilde{\boldsymbol{\tau}}_i^{(2)} \leq \boldsymbol{\tau}_i^{(2)} + n' + n_i - 1 \,\middle|\, \widetilde{\boldsymbol{\tau}}_i^{(1)} \geq \boldsymbol{\tau}_i^{(1)} + n' + n_i\right) \leq e^{-c_i n} \tag{3.16}$$

   holds under the condition $\mu_i > \theta_i$. By (3.14) and (3.16), the probability that both

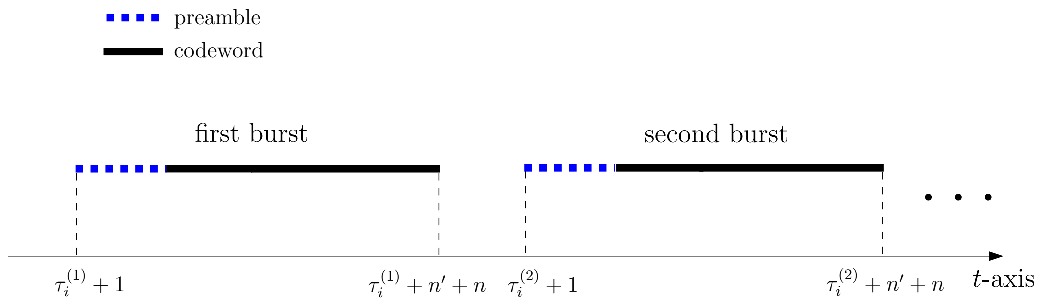


Figure 3.3: If $\mu_i > \theta_i$, the signals sent by Tx i look like intermittent bursts along the t-axis with high probability.

the second and third transmissions are immediate is bounded from above by $2e^{-c_i n}$. Simple induction shows that the probability of all N_i transmissions by Tx i being immediate is bounded from above by $N_i e^{-c_i n}$.

3. By the previous remark and under the constraint in (3.12), the signals sent by Tx i look like intermittent bursts along the t-axis with high probability as shown in Fig. 3.3. After sending a codeword, the transmitter must wait to receive enough bits in its buffer to transmit the next codeword. In contrast to [20], no "null codeword" is utilized in this paper, i.e., Tx i stays silent if it does not have enough bits in its buffer to represent a codeword.

4. Since transmissions are immediate, $\boldsymbol{\tau}_i^{(j)} = \widetilde{\boldsymbol{\tau}}_i^{(j)}$ with high probability and one may redefine $\tau_i^{(j)}$ in (3.10) by

$$\tau_i^{(0)} := 0, \quad \tau_i^{(j)} := \min\left\{t \geq \tau_i^{(j-1)} + 1 : b_{i,t} + b'_{i,t} \geq \lfloor n\eta_i \rfloor\right\}, \quad j \geq 1. \tag{3.17}$$

Without loss of generality, let n be a multiple of $N_1 N_2$. Then $\lfloor n\eta_i \rfloor$ becomes divisible by k_i for both $i = 1, 2$ and the inequality $b_{i,t} + b'_{i,t} \geq \lfloor n\eta_i \rfloor$ in (3.17) can be replaced by $b_{i,t} + b'_{i,t} = \lfloor n\eta_i \rfloor$, i.e.,

$$\tau_i^{(0)} := 0, \quad \tau_i^{(j)} := \min\left\{t \geq \tau_i^{(j-1)} + 1 : b_{i,t} + b'_{i,t} = \lfloor n\eta_i \rfloor\right\}, \quad j \geq 1. \tag{3.18}$$

Moreover, the buffer dynamics can be written as

$$b_{i,t+1} = \begin{cases} b_{i,t} + b'_{i,t} & b_{i,t} + b'_{i,t} < \lfloor n\eta_i \rfloor \\ 0 & b_{i,t} + b'_{i,t} = \lfloor n\eta_i \rfloor \end{cases}. \tag{3.19}$$

The following proposition compares the times when Tx i begins to send its j^{th} burst under the immediate Tx-Rx asynchronous scheme considered in this paper and the Tx-Rx synchronous scheme in [20]:

Proposition 6. *Assume μ_i is not an integer multiple of $\theta_i, \frac{\theta_i}{2}, \cdots, \frac{\theta_i}{N_i}$. Let $\boldsymbol{\varsigma}_i^{(j)} + 1$ for $1 \leq j \leq N_i$ be the time slot that Tx i begins to send its j^{th} codeword under the Tx-Rx synchronous scheme in [20]. There exists $\delta > 0$ such that*

$$\lim_{n\to\infty} \mathbb{P}\big(\boldsymbol{\varsigma}_i^{(j)} > (1+\delta)\boldsymbol{\tau}_i^{(j)}\big) = 1, \tag{3.20}$$

for any $1 \leq j \leq N_i$.

Proof. See Appendix G. □

Remark- Under the assumptions in Proposition 6, one can prove the existence of $\delta > 0$ such that $\lim_{n\to\infty} \frac{\boldsymbol{\varsigma}_i^{(j)} - \boldsymbol{\tau}_i^{(j)}}{\boldsymbol{\tau}_i^{(j)}} > \delta$ for any $1 \leq j \leq N_i$ which is stronger than the statement in (3.20).

3.2.3 Tx-Tx asynchrony

In the previous section the incoming bit streams at the transmitters where assumed to be synchronous in the sense that both start to run at time slot $t = 1$. In practice, the activation times for these processes are different. Let Tx 1 and Tx 2 start their activity at time slots $t = \lfloor n\nu_1 \rfloor$ and $t = \lfloor n\nu_2 \rfloor$, respectively, where $\nu_1, \nu_2 > 0$. Then (3.18) is rewritten as

$$\tau_i^{(0)} := \lfloor n\nu_i \rfloor, \quad \tau_i^{(j)} := \min\Big\{t \geq \tau_i^{(j-1)} + 1 : b_{i,t} + b'_{i,t} = \lfloor n\eta_i \rfloor\Big\}, \quad j \geq 1. \tag{3.21}$$

We see that $\tau_i^{(j)}$ is the smallest $t \geq \tau_i^{(j-1)} + 1$ such that Tx i receives packets of k_i bits in exactly $\frac{\lfloor n\eta_i \rfloor}{k_i}$ slots among the time slots with indices $\tau_i^{(j-1)}+1, \cdots, t$. Therefore, $\boldsymbol{\tau}_i^{(j)} - \boldsymbol{\tau}_i^{(j-1)}$ is a negative binomial random variable with parameters $\frac{\lfloor n\eta_i \rfloor}{k_i}$ and q_i, i.e.,

$$\boldsymbol{\tau}_i^{(j)} - \boldsymbol{\tau}_i^{(j-1)} \sim \mathrm{NB}\Big(\frac{\lfloor n\eta_i \rfloor}{k_i}, q_i\Big), \quad j \geq 1. \tag{3.22}$$

Alternatively, if $\boldsymbol{\xi}_{i,1}, \cdots, \boldsymbol{\xi}_{i,j}$ is a sequence of independent $\mathrm{NB}(\frac{\lfloor n\eta_i \rfloor}{k}, q)$ random variables, one can write

$$\boldsymbol{\tau}_i^{(j)} = \boldsymbol{\xi}_{i,1} + \cdots + \boldsymbol{\xi}_{i,j} + \lfloor n\nu_i \rfloor - 1, \quad j \geq 1. \tag{3.23}$$

Defining

$$\boldsymbol{\xi}_i^{(j)} := \boldsymbol{\xi}_{i,1} + \cdots + \boldsymbol{\xi}_{i,j}, \tag{3.24}$$

then $\boldsymbol{\xi}_i^{(j)} \sim \mathrm{NB}(\frac{j\lfloor n\eta_i \rfloor}{k_i}, q_i)$ and we get our final expression for $\boldsymbol{\tau}_i^{(j)}$, i.e.,

$$\boldsymbol{\tau}_i^{(j)} = \boldsymbol{\xi}_i^{(j)} + \lfloor n\nu_i \rfloor - 1. \tag{3.25}$$

We end this subsection with the following remarks:

Remark- Throughout the chapter, ν_1 and ν_2 are realizations of independent and continuous random variables $\boldsymbol{\nu}_1$ and $\boldsymbol{\nu}_2$.

Remark- In Remark (iv) in the previous subsection we assumed that n is a multiple of $N_1 N_2$. If we do not make such an assumption, each $\boldsymbol{\xi}_{i,l}$ for $1 \leq l \leq j$ turns out to be a $\mathrm{NB}\big(\lfloor \frac{\lfloor n\eta_i \rfloor}{k_i} \rfloor + m_l, q_i\big)$ random variable where m_l is an integer that depends on n, k_i and N_i and $0 \leq |m_l| < k_i$. This does not affect the results in the forthcoming sections. The assumption that n is a multiple of $N_1 N_2$ is made only for notational simplicity.

3.2.4 The Average transmission power and the average transmission rate

The incoming bit stream at the buffer of Tx i starts at time slot $\lfloor n\nu_i \rfloor$ and Tx i sends the last symbol in its N_i^{th} burst (last burst) at time slot $\tau_i^{(N_i)} + n' + n_i$. Therefore, the activity period $\mathcal{T}_i$ appearing in (3.2) is given by

$$\mathcal{T}_i = \Big\{\lfloor n\nu_i \rfloor, \lfloor n\nu_i \rfloor + 1, \cdots, \tau_i^{(N_i)} + n' + n_i\Big\}. \tag{3.26}$$

The elements of each codeword and the preamble sequence for Tx i are realizations of independent $\mathrm{N}(0, \gamma_i)$ random variables where $\gamma_i > 0$ is a constant designed to ensure that the average transmission power at Tx i does not exceed P_i. In the following, we compute the average transmission power Q_i and the average transmission rate R_i for Tx i:

Average transmission power

Tx i sends out N_i bursts where the j^{th} burst lasts from time slot $\boldsymbol{\tau}_i^{(j)}+1$ to time slot $\boldsymbol{\tau}_i^{(j)}+n'+n_i$. The average transmission power $\boldsymbol{Q}_i$ is a random variable given by

$$\boldsymbol{Q}_i=\frac{1}{|\boldsymbol{\mathcal{T}}_i|}\sum_{t=\lfloor n\nu_i\rfloor}^{\boldsymbol{\tau}_i^{(N_i)}+n'+n_i}\boldsymbol{x}_{i,t}^2=\frac{n'+n_i}{|\boldsymbol{\mathcal{T}}_i|}\sum_{j=1}^{N_i}\frac{1}{n'+n_i}\sum_{t=\boldsymbol{\tau}_i^{(j)}+1}^{\boldsymbol{\tau}_i^{(j)}+n'+n_i}\boldsymbol{x}_{i,t}^2. \tag{3.27}$$

By SLLN,

$$\lim_{n\to\infty}\frac{1}{n'+n_i}\sum_{l=\boldsymbol{\tau}_i^{(j)}+1}^{\boldsymbol{\tau}_i^{(j)}+n'+n_i}\boldsymbol{x}_{i,t}^2=\gamma_i, \tag{3.28}$$

for any $1\le j\le N_i$. Recalling the expression for $\boldsymbol{\tau}_i^{(j)}$ in (3.25),

$$\begin{aligned}\frac{n'+n_i}{|\boldsymbol{\mathcal{T}}_i|} &= \frac{n'+n_i}{\boldsymbol{\tau}_i^{(N_i)}+n'+n_i-\lfloor n\nu_i\rfloor+1}\\ &= \frac{n'+n_i}{\boldsymbol{\xi}_i^{(N_i)}+n'+n_i}\\ &= \frac{n'+n_i}{\frac{N_i\lfloor n\eta_i\rfloor}{k_i}\frac{\boldsymbol{\xi}_i^{(N_i)}}{\frac{N_i\lfloor n\eta_i\rfloor}{k_i}}+n'+n_i}.\end{aligned} \tag{3.29}$$

Since $\boldsymbol{\xi}_i^{(N_i)}\sim \mathrm{NB}(\frac{N_i\lfloor n\eta_i\rfloor}{k_i},q_i)$ is the sum of $\frac{N_i\lfloor n\eta_i\rfloor}{k_i}$ independent geometric random variables with parameter q_i, we invoke SLLN one more time to write

$$\lim_{n\to\infty}\frac{\boldsymbol{\xi}_i^{(N_i)}}{\frac{N_i\lfloor n\eta_i\rfloor}{k_i}}=\frac{1}{q_i}. \tag{3.30}$$

By (3.27), (3.28), (3.29) and (3.30), the average transmission power in the limit of large n is given by

$$\lim_{n\to\infty}\boldsymbol{Q}_i=\lim_{n\to\infty}\frac{(n'+n_i)N_i\gamma_i}{\frac{N_i\lfloor n\eta_i\rfloor}{k_i}\frac{1}{q_i}+n'+n_i}=\frac{\theta_iN_i\gamma_i}{\frac{N_i\eta_i}{k_iq_i}+\theta_i}=\frac{N_i\gamma_i}{1+\frac{1}{q_i\theta_i}}, \tag{3.31}$$

where we replaced $\eta_i = \frac{k_i}{N_i}$ in the last step.

Average transmission rate

Tx i sends a total number of $k_i n$ bits over its whole period of activity $\mathcal{T}_i$. Then the average transmission rate $\boldsymbol{R}_i$ is a random variable given by

$$\boldsymbol{R}_i = \frac{k_i n}{|\boldsymbol{\mathcal{T}}_i|} = \frac{k_i n}{\boldsymbol{\xi}_i^{(N_i)} + n' + n_i}. \tag{3.32}$$

Using (3.30), the average transmission rate in the limit of large n is

$$\lim_{n\to\infty} \boldsymbol{R}_i = \frac{k_i}{\frac{N_i \eta_i}{\lambda_i} + \theta_i} = \frac{\lambda_i}{1 + q_i \theta_i}. \tag{3.33}$$

We will use the expressions in (3.31) and (3.33) in Section 3.5.1 where we study system design.

3.3 Estimating the arrival times and transmitter identification at the receivers

Let $(s_{i,l}^{(j)})_{l=0}^{n_i-1}$ for $1 \leq j \leq N_i$ be the N_i codewords of length n_i sent by Tx i. Also, let $(s'_{i,l})_{l=0}^{n'-1}$ be the preamble sequence for user i. The signal $x_{i,t}$ in (3.3) can be written as

$$x_{i,t} = \begin{cases} s'_{i,t-\tau_i^{(j)}-1} & \tau_i^{(j)} + 1 \leq t \leq \tau_i^{(j)} + n' \\ s^{(j)}_{i,t-\tau_i^{(j)}-n'-1} & \tau_i^{(j)} + n' + 1 \leq t \leq \tau_i^{(j)} + n' + n_i \\ 0 & \text{otherwise} \end{cases}, \tag{3.34}$$

for $1 \leq i \leq 2$ and $1 \leq j \leq N_i$. The preambles $(s'_{1,l})_{l=0}^{n'-1}$ and $(s'_{2,l})_{l=0}^{n'-1}$ are revealed to both receivers. The following assumption considerably simplifies the analysis in this section.
Assumption- For any integers $1 \leq j_1 \leq N_1$ and $1 \leq j_2 \leq N_2$,

$$j_2\mu_2 - j_1\mu_1 + \nu_2 - \nu_1 \notin \{0, \theta_1, -\theta_2, \theta_1 - \theta_2\}. \tag{3.35}$$

Remark- Since we are assuming that ν_1 and ν_2 are realizations of independent and continuous random variables $\boldsymbol{\nu}_1$ and $\boldsymbol{\nu}_2$, the restrictions in (3.35) are considered "mild" in the sense that the probability of $j_2\mu_2 - j_1\mu_1 + \boldsymbol{\nu}_2 - \boldsymbol{\nu}_1$ lying in $\{0, \theta_1, -\theta_2, \theta_1 - \theta_2\}$ for some j_1 and j_2 is equal to zero.

We will use the assumption in (3.35) throughout the chapter. Its first application appears in the following proposition:

Proposition 7. *Assuming (3.35) holds, the probability of Tx i starting or ending a transmission burst while Tx i' is sending a preamble sequence tends to zero as n grows.*

Proof. See Appendix H. □

In view of Proposition 7 and for given $\epsilon > 0$, we assume n is large enough so that the probability of Tx i starting or ending a transmission burst while Tx i' is sending a preamble sequence is less than ϵ and add ϵ to the probability of error in decoding the codewords. In other words, we assume no transmitter starts or ends a transmission burst while the other transmitter is sending a preamble sequence.

Next, we study the detection/estimation procedure at Rx 1. A similar procedure is carried out at Rx 2. Define the PDFs $p^{(1)}(\cdot,\cdot), \cdots, p^{(4)}(\cdot,\cdot)$ on $\mathbb{R}^2$ as follows:

- For $1 \le j \le N_1$ and $\tau_1^{(j)} + 1 \le t \le \tau_1^{(j)} + n'$, Tx 1 is sending the preamble sequence in its j^{th} burst. If Tx 2 is not transmitting during this time interval, then $p_{\boldsymbol{x}_{1,t},\boldsymbol{y}_{1,t}}(x,y) = \mathrm{g}(x;\gamma_1)\mathrm{g}(y-x;1)$. We define

$$p^{(1)}(x,y) := \mathrm{g}(x;\gamma_1)\mathrm{g}(y-x;1). \tag{3.36}$$

- For $1 \le j \le N_1$ and $\tau_1^{(j)} + 1 \le t \le \tau_1^{(j)} + n'$, Tx 1 is sending the preamble sequence in its j^{th} burst. If Tx 2 is transmitting during this time interval, then $p_{\boldsymbol{x}_{1,t},\boldsymbol{y}_{1,t}}(x,y) = \mathrm{g}(x;\gamma_1)\mathrm{g}(y-x;1+a_2\gamma_2)$. We define

$$p^{(2)}(x,y) := \mathrm{g}(x;\gamma_1)\mathrm{g}(y-x;1+a_2\gamma_2). \tag{3.37}$$

- For $1 \le j \le N_2$ and $\tau_2^{(j)} + 1 \le t \le \tau_2^{(j)} + n'$, Tx 2 is sending the preamble sequence in its j^{th} burst. If Tx 1 is not transmitting during this time interval, then $p_{\boldsymbol{x}_{2,t},\boldsymbol{y}_{1,t}}(x,y) = \mathrm{g}(x;\gamma_2)\mathrm{g}(y-a_2x;1)$. We define

$$p^{(3)}(x,y) := \mathrm{g}(x;\gamma_2)\mathrm{g}(y-a_2x;1). \tag{3.38}$$

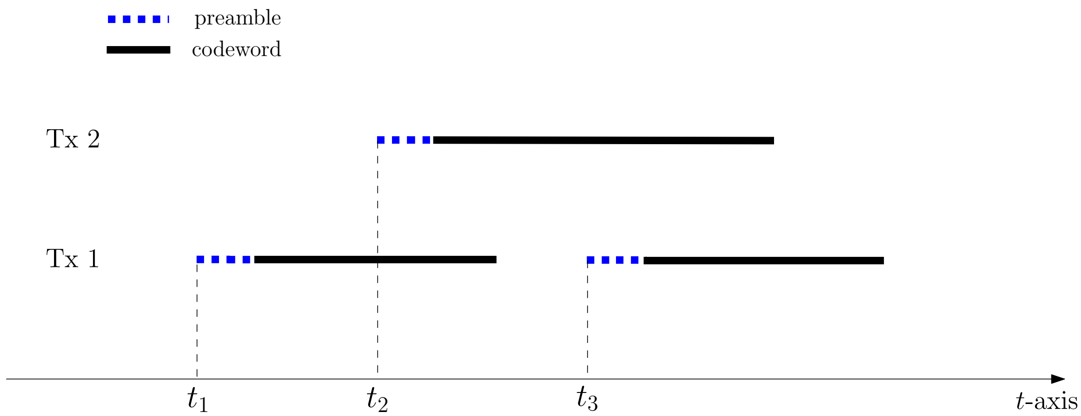


Figure 3.4: A scenario where where Tx 1 and Tx 2 send only two and one bursts, respectively, i.e., $N_1 = 2$ and $N_2 = 1$. It is assumed that $n_1 < n_2$. For simplicity of presentation, we call $t_1 := \tau_1^{(1)} + 1$, $t_2 := \tau_2^{(1)} + 1$ and $t_3 := \tau_1^{(2)} + 1$.

- For $1 \leq j \leq N_2$ and $\tau_2^{(j)} + 1 \leq t \leq \tau_2^{(j)} + n'$, Tx 2 is sending the preamble sequence in its j^{th} burst. If Tx 1 is transmitting during this time interval, then $p_{\boldsymbol{x}_{2,t},\boldsymbol{y}_{1,t}}(x, y) = \mathrm{g}(x; \gamma_2)\mathrm{g}(y - a_2x; 1 + \gamma_1)$. We define

$$p^{(4)}(x, y) := \mathrm{g}(x; \gamma_2)\mathrm{g}(y - a_2x; 1 + \gamma_1). \tag{3.39}$$

To identify the arrival time of a transmission burst, each receiver applies the so-called sequential joint typicality decoder [21]. To describe how Rx 1 estimates $\tau_i^{(j)}$ for different i and j and without loss of generality, we find it best to consider the particular situation shown in Fig. 3.4 where $N_1 = 2$, $N_2 = 1$ and $n_1 < n_2$. The arrival time estimation and user identification are performed in the following steps:

1. By Fig. 3.4, $t_1 := \tau_1^{(1)} + 1$ is the time slot that the first active transmitter sends the first symbol in its preamble sequence. Rx 1 estimates t_1 by

$$\hat{t}_1 = \min\left\{t \geq 0 : \left((s'_{1,l})_{l=0}^{n'-1}, (y_{1,l})_{l=t}^{t+n'-1}\right) \in A_\epsilon^{(n')}[p^{(1)}] \text{ or } \left((s'_{2,l})_{l=0}^{n'-1}, (y_{1,l})_{l=t}^{t+n'-1}\right) \in A_\epsilon^{(n')}[p^{(3)}]\right\}, \tag{3.40}$$

where the PDFs $p^{(1)}$ and $p^{(3)}$ are defined in (3.36) and (3.38), respectively. By Proposition 7, we can assume $p_{\boldsymbol{x}_{1,l},\boldsymbol{y}_{1,l}} = p^{(1)}$ for any $t_1 \leq l \leq t_1 + n' - 1$. Then the weak law of large numbers yields

$$\lim_{n\to\infty} \mathbb{P}\Big(\left((\boldsymbol{s}'_{1,l})_{l=0}^{n'-1}, (\boldsymbol{y}_{1,l})_{l=\boldsymbol{t}_1}^{\boldsymbol{t}_1+n'-1}\right) \in A_\epsilon^{(n')}[p^{(1)}]\Big) = 1. \tag{3.41}$$

By (3.41), $\lim_{n\to\infty} \mathbb{P}(\hat{\boldsymbol{t}}_1 \leq \boldsymbol{t}_1) = 1$. As such, to show that $\hat{\boldsymbol{t}}_1 = \boldsymbol{t}_1$ holds with high probability, it is enough to show that $\mathbb{P}(\hat{\boldsymbol{t}}_1 < \boldsymbol{t}_1)$ is negligible for sufficiently large n. This is the content of the following proposition:

Proposition 8. *We have*

$$\mathbb{P}(\hat{\boldsymbol{t}}_1 < \boldsymbol{t}_1) \leq \Theta(n)e^{-\Theta(n')}. \tag{3.42}$$

In particular, $\lim_{n\to\infty} \mathbb{P}(\hat{\boldsymbol{t}}_1 < \boldsymbol{t}_1) = 0$.

Proof. See Appendix I. □

Motivated by Proposition 8, we assume Rx 1 knows the exact value of t_1. In fact, for given $\epsilon > 0$, we assume n is large enough so that the probability of error in estimating the first arrival time is less than ϵ and add ϵ to the probability of error in decoding the codewords. Not only does Rx 1 know the exact value of t_1, but also it realizes that t_1 is $\tau_1^{(1)} + 1$ and not $\tau_2^{(1)} + 1$, i.e., it knows the first arriving burst belongs to Tx 1. This is described in the next step.

2. After finding t_1, Rx 1 decides whether the first burst belongs to Tx 1 or Tx 2. Towards this goal, Rx 1 verifies if

$$\left((s'_{1,l})_{l=0}^{n'-1}, (y_{1,l})_{l=t_1}^{t_1+n'-1}\right) \in A_\epsilon^{(n')}[p^{(1)}] \tag{3.43}$$

or

$$\left((s'_{2,l})_{l=0}^{n'-1}, (y_{1,l})_{l=t_1}^{t_1+n'-1}\right) \in A_\epsilon^{(n')}[p^{(3)}]. \tag{3.44}$$

If (3.43) holds, the first arriving burst is assumed to belong to Tx 1. If (3.44) holds, the first arriving burst is assumed to belong to Tx 2. As mentioned earlier in (3.41), (3.43) holds with high probability in the limit of large n. In Appendix J, it is shown that

$$\mathbb{P}\Big(\left((\boldsymbol{s}'_{2,l})_{l=0}^{n'-1}, (\boldsymbol{y}_{1,l})_{l=t_1}^{t_1+n'-1}\right) \in A_\epsilon^{(n')}[p^{(3)}]\Big) \leq e^{-\Theta(n')}. \tag{3.45}$$

Therefore, (3.44) holds with a probability that decays exponentially with n' and hence, Rx 1 can identify the sender of the first burst with high probability.

3. Up to this point, Rx 1 knows that the first burst belongs to Tx 1 and it lasts from time slot t_1 to time slot $t_1 + n' + n_1 - 1$. If another burst arrives during this period, it

must belong to Tx 2. As shown in Fig. 3.4, a burst belonging to Tx 2 indeed arrives at time slot $t_2 := \tau_2^{(1)} + 1$ when the first burst of Tx 1 is still arriving. The preamble sequence in the first burst by Tx 2 extends from time slot t_2 to time slot $t_2 + n' - 1$. By Proposition 7, $p_{\boldsymbol{x}_{2,l},\boldsymbol{y}_{1,l}} = p^{(4)}$ for any $t_2 \leq l \leq t_2 + n' - 1$ where $p^{(4)}$ is defined in (3.39). Based on these observations, Rx 1 estimates t_2 by

$$\hat{t}_2 = \min\left\{t_1 \leq t \leq t_1 + n' + n_1 - 1 : \left((s'_{2,l})_{l=0}^{n'-1}, (y_{1,l})_{l=t}^{t+n'-1}\right) \in A_\epsilon^{(n')}[p^{(4)}]\right\}. \quad (3.46)$$

Following similar lines of reasoning in the proof of Proposition 8, one can show that $\mathbb{P}(\hat{\boldsymbol{t}}_2 \neq \boldsymbol{t}_2) \leq \Theta(n)e^{-\Theta(n')}$. As such, we can assume that Rx 1 knows the exact value of t_2, i.e., Rx 1 knows $\tau_2^{(1)}$.

Remark- If (3.47) fails to return an estimate for t_2, Rx 1 concludes that no burst of Tx 2 is received by the time Tx 1 finishes its first burst. As such, starting at time slot $t_1 + n' + n_1$, Rx 1 looks for the arrival time t^* of a new transmission burst that might belong to Tx 1 or Tx 2. The time slot t^* is estimated similar to (3.40), i.e., $\hat{t}^*$ is the smallest value of $t \geq t_1 + n' + n_1$ such that $\left((s'_{1,l})_{l=0}^{n'-1}, (y_{1,l})_{l=t}^{t+n'-1}\right) \in A_\epsilon^{(n')}[p^{(1)}]$ or $\left((s'_{2,l})_{l=0}^{n'-1}, (y_{1,l})_{l=t}^{t+n'-1}\right) \in A_\epsilon^{(n')}[p^{(3)}]$.

4. After finding t_2 in step (iii), Rx 1 knows that the first burst of Tx 2 lasts from time slot t_2 to time slot $t_2 + n' + n_2 - 1$. Since the first burst of Tx 1 ends at time slot $t_1 + n + n_1 - 1$, Rx 1 looks for possible arrival of the second burst of Rx 1 during time slots $t_1 + n' + n_1$ to $t_2 + n' + n_2 - 1$. In fact, as shown in Fig. 3.4, the second burst of Tx 1 arrives at time slot $t_3 := \tau_1^{(2)} + 1$ when the first burst by Tx 2 is still arriving. The preamble sequence in the second burst of Tx 1 extends from time slot t_3 to $t_3 + n' - 1$. By Proposition 7, $p_{\boldsymbol{x}_{1,l},\boldsymbol{y}_{1,l}} = p^{(2)}$ for any $t_3 \leq l \leq t_3 + n' - 1$ where $p^{(2)}$ is defined in (3.37). Based on these observations, Rx 1 estimates t_3 by

$$\hat{t}_3 = \min\left\{t_1 + n' + n_1 \leq t \leq t_2 + n' + n_2 - 1 : \left((s'_{1,l})_{l=0}^{n'-1}, (y_{1,l})_{l=t}^{t+n'-1}\right) \in A_\epsilon^{(n')}[p^{(2)}]\right\}, \quad (3.47)$$

where following the proof of Proposition 8, it can be shown that $\mathbb{P}(\hat{\boldsymbol{t}}_3 \neq \boldsymbol{t}_3) \leq \Theta(n)e^{-\Theta(n')}$.

The detection/estimation procedure described here can be easily extended to scenarios other than the one depicted in Fig. 3.4. Throughout the rest of the chapter, we assume both receivers know the values of $\tau_i^{(j)}$ for any $i = 1, 2$ and $1 \leq j \leq N_i$.

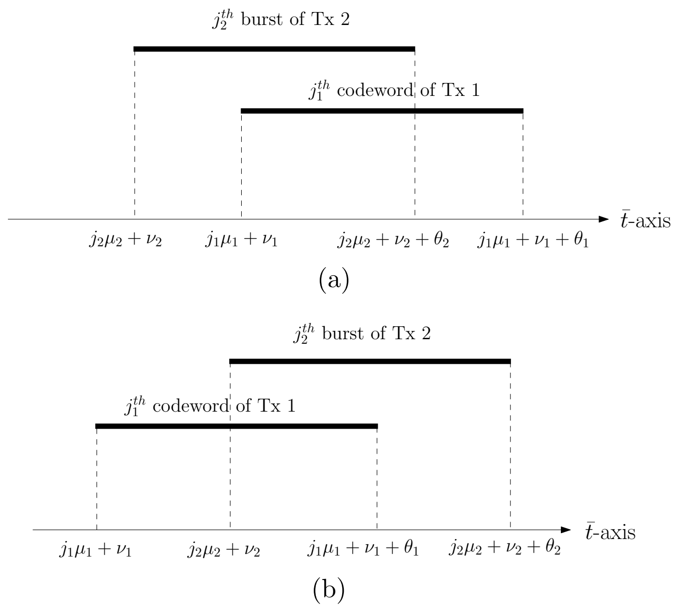


Figure 3.5: If (3.50) holds for $i = 1$, the j_2^{th} burst of Tx 2 overlaps (with high probability for large n) with the j_1^{th} codeword of Tx 1 over its "left end" as shown in panel (a), while if (3.51) holds for $i = 1$, the j_2^{th} burst of Tx 2 overlaps with the j_1^{th} codeword of Tx 1 over its "right end" as shown in panel (b).

3.4 Decoding strategy and achievability results

In the previous section, we described how each receiver is capable of estimating $\tau_i^{(j)}$ with vanishingly small probability of error. Our analysis heavily relied on the conditions in (3.35) which are also used in the following proposition:

Proposition 9. *Let $i = 1, 2$ and $1 \le j_i \le N_i$. Assuming (3.35) holds, the j_i^{th} codeword of Tx i and the $j_{i'}^{th}$ burst of Tx i' overlap with arbitrarily large probability in the limit of large n if and only if*

$$j_{i'}\mu_{i'} - j_i\mu_i + \nu_{i'} - \nu_i \in (0, \theta_i) \bigcup (-\theta_{i'}, \theta_i - \theta_{i'}). \tag{3.48}$$

Proof. See Appendix K. □

The result of Proposition 9 can be intuitively described as follows. Define the scaled time variable

$$\bar{t} := \frac{t}{n}. \tag{3.49}$$

On the $\bar{t}$-axis, Tx i sends its j_i^{th} codeword at times $\frac{1}{n}(\boldsymbol{\tau}_i^{(j_i)} + n' + 1)$ to $\frac{1}{n}(\boldsymbol{\tau}_i^{(j_i)} + n' + n_i)$. In virtue of SLLN, $\lim_{n\to\infty} \frac{1}{n}\boldsymbol{\tau}_i^{(j_i)} = j_i\mu_i + \nu_i$. As such, in the limit as n grows to infinity, the j_i^{th} codeword of Tx i lies on the interval $\big(j_i\mu_i + \nu_i, j_i\mu_i + \nu_i + \theta_i\big)$ along the $\bar{t}$-axis. Similarly, one sees that the $j_{i'}^{th}$ burst of Tx i' lies on the interval $\big(j_{i'}\mu_{i'} + \nu_{i'}, j_{i'}\mu_{i'} + \nu_{i'} + \theta_{i'}\big)$ along the $\bar{t}$-axis. Provided (3.35) holds, the condition in (3.48) is equivalent to saying that the two intervals $\big(j_i\mu_i + \nu_i, j_i\mu_i + \nu_i + \theta_i\big)$ and $\big(j_{i'}\mu_{i'} + \nu_{i'}, j_{i'}\mu_{i'} + \nu_{i'} + \theta_{i'}\big)$ overlap. More specifically, if

$$-\theta_{i'} < j_{i'}\mu_{i'} - j_i\mu_i + \nu_{i'} - \nu_i < \min\{0, \theta_i - \theta_{i'}\}, \tag{3.50}$$

the $j_{i'}^{th}$ burst of Tx i' overlaps (with high probability for large n) with the j_i^{th} codeword of Tx i at its "left end" as shown in Fig. O.1(a) for $i = 1$, while if

$$\max\{0, \theta_i - \theta_{i'}\} < j_{i'}\mu_{i'} - j_i\mu_i + \nu_{i'} - \nu_i < \theta_i, \tag{3.51}$$

the $j_{i'}^{th}$ burst of Tx i' overlaps with the j_i^{th} codeword of Tx i at its "right end" as shown in Fig. O.1(b) for $i = 1$. Finally,

$$\min\{0, \theta_i - \theta_{i'}\} < j_{i'}\mu_{i'} - j_i\mu_i + \nu_{i'} - \nu_i < \max\{0, \theta_i - \theta_{i'}\}, \tag{3.52}$$

implies that the j_i^{th} codeword of Tx i is contained in the $j_{i'}^{th}$ burst of Tx i' or the other way around depending on whether $\theta_i < \theta_{i'}$ or $\theta_{i'} < \theta_i$, respectively.

Remark- The geometric interpretation of the conditions in (3.35) is that the endpoints of the intervals $\big(j_i\mu_i + \nu_i, j_i\mu_i + \nu_i + \theta_i\big)$ and $\big(j_{i'}\mu_{i'} + \nu_{i'}, j_{i'}\mu_{i'} + \nu_{i'} + \theta_{i'}\big)$ do not coincide.

We make the following definitions:

- Fixing $j_i = j$, there exists at most one positive integer $j_{i'}$ that satisfies (3.50). We denote this value of $j_{i'}$ by $\omega_{i,j}^-$. In fact, $\omega_{i,j}^-$ is the index of the burst of Tx i' that overlaps with the left end of the j^{th} codeword of Tx i. In case $\omega_{i,j}^-$ does not exist, we write $\omega_{i,j}^- = 0$.

- Fixing $j_i = j$, there exists at most one positive integer $j_{i'}$ that satisfies (3.51). We denote this value of $j_{i'}$ by $\omega_{i,j}^+$. In fact, $\omega_{i,j}^+$ is the index of the burst of Tx i' that overlaps with the right end of the j^{th} codeword of Tx i. In case $\omega_{i,j}^+$ does not exist, we write $\omega_{i,j}^+ = 0$.

- We define $\omega_{i,j}$ as the number of bursts of Tx i' that are completely contained within the j^{th} codeword of Tx i.

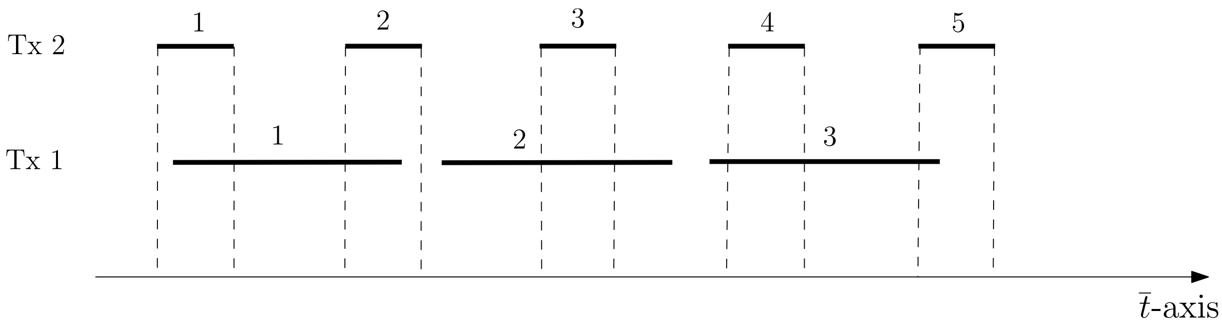


Figure 3.6: Positions of the bursts along the $\bar{t}$-axis in a scenario where $N_1 = 3$ and $N_2 = 5$.

For example, Fig. 3.6 presents a scenario where $N_1 = 3$ and $N_2 = 5$ and we have

$$(\omega_{1,1}^-, \omega_{1,1}^+, \omega_{1,1}) = (1, 2, 0), \ \ (\omega_{1,2}^-, \omega_{1,2}^+, \omega_{1,2}) = (0, 0, 1), \ \ (\omega_{1,3}^-, \omega_{1,3}^+, \omega_{1,3}) = (0, 5, 1). \tag{3.53}$$

Next, we study achievability results for the j^{th} transmitted codeword of Tx i, i.e., we look for sufficient conditions that guarantee the j^{th} transmitted codeword by Tx i is decoded successfully at Rx i. In order to describe the decoding strategy, we focus on Rx 1. For notational simplicity, in some equations we show $\omega_{1,j}^-$ and $\omega_{1,j}^+$ by ω^- and ω^+, respectively.

- Assume $\omega^- \neq 0$, $\omega^+ \neq 0$ and $\omega_{1,j} = 0$. Then

$$\tau_2^{(\omega^-)} + 1 < \tau_1^{(j)} + n' + 1 \leq \tau_2^{(\omega^-)} + n' + n_2 \leq \tau_2^{(\omega^+)} + 1 \leq \tau_1^{(j)} + n' + n_1 < \tau_2^{(\omega^+)} + n' + n_2. \tag{3.54}$$

  This situation is shown in Fig. 3.7. The j^{th} codeword of Tx 1 is transmitted during time slots $\tau_1^{(j)} + n' + 1$ to $\tau_1^{(j)} + n' + n_1$. The interference pattern over this codeword is described as follows:

  - Any symbol of the codeword transmitted during time slots $\tau_1^{(j)} + n' + 1$ to $\tau_2^{(\omega^-)} + n' + n_2$ is received in the presence of interference. For any $\tau_1^{(j)} + n' + 1 \leq l \leq \tau_2^{(\omega^-)} + n' + n_2$, we have $p_{\boldsymbol{x}_{1,l},\boldsymbol{y}_{1,l}} = p^{(2)}$ where the PDF $p^{(2)}$ is defined in (3.37).
  - Any symbol of the codeword transmitted during time slots $\tau_2^{(\omega^-)} + n' + n_2 + 1$ to $\tau_2^{(\omega^+)}$ does not experience interference. For any $\tau_2^{(\omega^-)} + n' + n_2 + 1 \leq l \leq \tau_2^{(\omega^+)}$, we have $p_{\boldsymbol{x}_{1,l},\boldsymbol{y}_{1,l}} = p^{(1)}$ where the PDF $p^{(1)}$ is defined in (3.36).
  - Any symbol of the codeword transmitted during time slots $\tau_2^{(\omega^+)} + 1$ to $\tau_1^{(j)} + n' +$

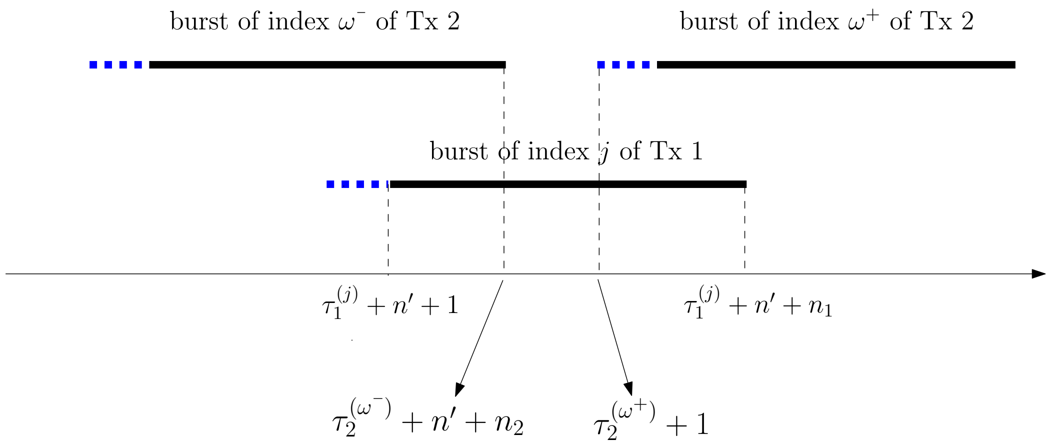


Figure 3.7: A scenario where $\omega_{1,j}^- \neq 0$, $\omega_{1,j}^+ \neq 0$ and $\omega_{1,j}=0$. For notational simplicity, we have shown $\omega_{1,j}^-$ and $\omega_{1,j}^+$ by ω^- and ω^+, respectively.

n_1 is received in the presence of interference. For any $\tau_2^{(\omega^+)}+1 \leq l \leq \tau_1^{(j)}+n'+n_1$, we have $p_{\boldsymbol{x}_{1,l},\boldsymbol{y}_{1,l}} = p^{(2)}$ where the PDF $p^{(2)}$ is defined in (3.37).

According to the interference pattern just described, Rx 1 finds the unique codeword $(s_{1,l})_{l=0}^{n_1-1}$ such that all three statements

$$\left((s_{1,l})_{l=0}^{\tau_2^{(\omega^-)}-\tau_1^{(j)}+n_2-1}, (y_{1,l})_{l=\tau_1^{(j)}+n'+1}^{\tau_2^{(\omega^-)}+n'+n_2}\right) \in A_\epsilon^{(\tau_2^{(\omega^-)}-\tau_1^{(j)}+n_2)}[p^{(2)}], \tag{3.55}$$

$$\left((s_{1,l})_{l=\tau_2^{(\omega^-)}-\tau_1^{(j)}+n_2}^{\tau_2^{(\omega^+)}-\tau_1^{(j)}-n'-1}, (y_{1,l})_{l=\tau_2^{(\omega^-)}+n'+n_2+1}^{\tau_2^{(\omega^+)}}\right) \in A_\epsilon^{(\tau_2^{(\omega^+)}-\tau_2^{(\omega^-)}-n'-n_2)}[p^{(1)}] \tag{3.56}$$

and

$$\left((s_{1,l})_{l=\tau_2^{(\omega^+)}-\tau_1^{(j)}-n'}^{n_1-1}, (y_{1,l})_{l=\tau_2^{(\omega^+)}+1}^{\tau_1^{(j)}+n'+n_1}\right) \in A_\epsilon^{(\tau_1^{(j)}-\tau_2^{(\omega^+)}+n'+n_1)}[p^{(2)}] \tag{3.57}$$

hold. We have the following proposition:

Proposition 10. *Given the index j of a transmitted codeword of Tx i, assume $\omega_{i,j}^- \neq 0$ and $\omega_{i,j}^+ \neq 0$. If $\omega_{i,j}^- \neq \omega_{i,j}^+$, then*

$$\eta_i - \frac{1}{\lambda_{i'}}(1+\omega_{i,j})(\kappa_i-\kappa_i')\eta_{i'} < \theta_i\kappa_i' - \theta_{i'}(1+\omega_{i,j})(\kappa_i-\kappa_i'), \tag{3.58}$$

is a sufficient condition for reliable decoding of the j^{th} message of Tx i where

$$\kappa_i := \mathsf{C}(\gamma_i), \;\; \kappa_i' := \mathsf{C}\Big(\frac{\gamma_i}{1+a_{i'}\gamma_{i'}}\Big) \tag{3.59}$$

and the function $\mathsf{C}(\cdot)$ is defined in (3.1). If $\omega_{i,j}^- = \omega_{i,j}^+$, then (3.58) is replaced by

$$\eta_i < \theta_i \kappa_i'. \tag{3.60}$$

Proof. See Appendix L. □

- Assume $\omega^- \neq 0$, $\omega^+ = 0$ and $\omega_{1,j} = 0$. Then

$$\tau_2^{(\omega^-)} + 1 < \tau_1^{(j)} + n' + 1 \leq \tau_2^{(\omega^-)} + n' + n_2 \leq \tau_1^{(j)} + n' + n_1 < \tau_2^{(\omega^-+1)} + 1. \tag{3.61}$$

This situation is shown in Fig. 3.7 after removing the bust with index ω^+ of Tx 2 from the picture. The interference pattern over the j^{th} codeword of Tx 1 is described as follows:

  - Any symbol of the codeword transmitted during time slots $\tau_1^{(j)} + n' + 1$ to $\tau_2^{(\omega^-)} + n' + n_2$ is received in the presence of interference. For any $\tau_1^{(j)} + n' + 1 \leq l \leq \tau_2^{(\omega^-)} + n' + n_2$, we have $p_{\boldsymbol{x}_{1,l},\boldsymbol{y}_{1,l}} = p^{(2)}$ where the PDF $p^{(2)}$ is defined in (3.37).
  - Any symbol of the codeword transmitted during time slots $\tau_2^{(\omega^-)} + n' + n_2 + 1$ to $\tau_1^{(j)} + n' + n_1$ does not experience interference. For any $\tau_2^{(\omega^-)} + n' + n_2 + 1 \leq l \leq \tau_1^{(j)} + n' + n_1$, we have $p_{\boldsymbol{x}_{1,l},\boldsymbol{y}_{1,l}} = p^{(1)}$ where the PDF $p^{(1)}$ is defined in (3.36).

According to the interference pattern just described, Rx 1 finds the unique codeword $(s_{1,l})_{l=0}^{n_1-1}$ such that the two constraints

$$\Big((s_{1,l})_{l=0}^{\tau_2^{(\omega^-)}-\tau_1^{(j)}+n_2-1}, (y_{1,l})_{l=\tau_1^{(j)}+n'+1}^{\tau_2^{(\omega^-)}+n'+n_2}\Big) \in A_\epsilon^{(\tau_2^{(\omega^-)}-\tau_1^{(j)}+n_2)}[p^{(2)}] \tag{3.62}$$

and

$$\Big((s_{1,l})_{l=\tau_2^{(\omega^-)}-\tau_1^{(j)}+n_2}^{n_1-1}, (y_{1,l})_{l=\tau_2^{(\omega^-)}+n'+n_2+1}^{\tau_1^{(j)}+n'+n_1}\Big) \in A_\epsilon^{(\tau_1^{(j)}-\tau_2^{(\omega^-)}+n_1-n_2)}[p^{(1)}] \tag{3.63}$$

hold.

Proposition 11. *Given the index j of a transmitted codeword of Tx i, assume $\omega^{-}_{i,j} \neq 0$ and $\omega^{+}_{i,j} = 0$. Then*

$$\big(1 - \frac{j}{\lambda_i}(\kappa_i - \kappa_i')\big)\eta_i + \frac{\omega^{-}_{i,j}}{\lambda_{i'}}(\kappa_i - \kappa_i')\eta_{i'} < \theta_i\kappa_i - \big(\nu_{i'} - \nu_i + \theta_{i'}(1 + \omega_{i,j})\big)(\kappa_i - \kappa_i') \tag{3.64}$$

is a sufficient condition for reliable decoding of the j^{th} message of Tx i where κ_i and κ_i' are defined in (3.59).

Proof. See Appendix M. □

- Assume $\omega^{+} \neq 0$, $\omega^{-} = 0$ and $\omega_{1,j} = 0$. We have

$$\tau_2^{(\omega^{+}-1)} + n' + n_2 < \tau_1^{(j)} + n' + 1 \leq \tau_2^{(\omega^{+})} + 1 \leq \tau_1^{(j)} + n' + n_1 < \tau_2^{(\omega^{+})} + n' + n_2. \tag{3.65}$$

This situation is shown in Fig. 3.7 after removing the burst with index ω^{-} of Tx 2 from the picture. The interference pattern over the j^{th} codeword of Tx 1 is described as follows:

  - Any symbol of the codeword transmitted during time slots $\tau_1^{(j)} + n' + 1$ to $\tau_2^{(j^{+})}$ does not experience interference. For any $\tau_1^{(j)} + n' + 1 \leq l \leq \tau_2^{(\omega^{+})}$, we have $p_{\boldsymbol{x}_{1,l},\boldsymbol{y}_{1,l}} = p^{(1)}$ where the PDF $p^{(1)}$ is defined in (3.36).
  - Any symbol of the codeword transmitted during time slots $\tau_2^{(\omega^{+})} + 1$ to $\tau_1^{(j)} + n' + n_1$ is received in the presence of interference. For any $\tau_2^{(\omega^{+})} + 1 \leq l \leq \tau_1^{(j)} + n' + n_1$, we have $p_{\boldsymbol{x}_{1,l},\boldsymbol{y}_{1,l}} = p^{(2)}$ where the PDF $p^{(2)}$ is defined in (3.37).

According the to the interference pattern just described, Rx 1 finds the unique codeword $(s_{1,l})_{l=0}^{n_1-1}$ such that the two constraints

$$\Big((s_{1,l})_{l=0}^{\tau_2^{(\omega^{+})} - \tau_1^{(j)} - n' - 1}, (y_{1,l})_{l=\tau_1^{(j)}+n'+1}^{\tau_2^{(\omega^{+})}}\Big) \in A_\epsilon^{(\tau_2^{(\omega^{+})} - \tau_1^{(j)} - n')}[p^{(1)}] \tag{3.66}$$

and

$$\Big((s_{1,l})_{l=\tau_2^{(\omega^{+})} - \tau_1^{(j)} - n'}^{n_1-1}, (y_{1,l})_{l=\tau_2^{(\omega^{+})}+1}^{\tau_1^{(j)}+n'+n_1}\Big) \in A_\epsilon^{(\tau_1^{(j)} - \tau_2^{(\omega^{+})} + n' + n_1)}[p^{(2)}] \tag{3.67}$$

hold.

Proposition 12. *Given the index j of a transmitted codeword of Tx i, assume $\omega_{i,j}^- = 0$ and $\omega_{i,j}^+ \neq 0$. Then*

$$\big(1 + \frac{j}{\lambda_i}(\kappa_i - \kappa_i')\big)\eta_i - \frac{\omega_{i,j}^+}{\lambda_{i'}}(\kappa_i - \kappa_i')\eta_{i'} < \theta_i\kappa_i' + (\nu_{i'} - \nu_i - \theta_{i'}\omega_{i,j})(\kappa_i - \kappa_i'). \quad (3.68)$$

is a sufficient condition for reliable decoding of the j^{th} message of Tx i where κ_i and κ_i' are defined in (3.59).

Proof. The proof is similar to the proof of Proposition 11 and is omitted. □

- If $\omega_{i,j}^- = \omega_{i,j}^+ = 0$,

$$\eta_i < \theta_i\kappa_i - \theta_{i'}\omega_{i,j}(\kappa_i - \kappa_i') \quad (3.69)$$

is a sufficient condition for reliable decoding of the j^{th} message of Tx i.

Corollary 1. *If $\eta_i < \theta_i\kappa_i'$, the probability of error in decoding the j^{th} message of Tx i vanishes by increasing n for any j regardless of the values of $\omega_{i,j}^-$, $\omega_{i,j}^+$, $\omega_{i,j}$, ν_i and $\nu_{i'}$. If $\eta_i \geq \theta_i\kappa_i$, the probability of error in decoding any message of Tx i approaches one by increasing n.*

Proof. Let $\eta_i < \theta_i\kappa_i'$. We consider the following cases:

- Assume $\omega_{i,j}^- \neq 0$, $\omega_{i,j}^+ \neq 0$ and $\omega_{i,j}^- \neq \omega_{i,j}^+$. Since $\eta_i < \theta_i\kappa_i'$ and $\theta_{i'} < \mu_{i'}$ by (3.12), we must have $\eta_i < \theta_i\kappa_i' + (\mu_{i'} - \theta_{i'})(1 + \omega_{i,j})(\kappa_i - \kappa_i')$ which is exactly (3.58) after rearranging terms.

- Assume $\omega_{i,j}^- \neq 0$ and $\omega_{i,j}^+ = 0$. On the $\bar{t}$-axis, the interval $\mathcal{J}_0 = (j\mu_i + \nu_i, \omega^-\mu_{i'} + \nu_{i'} + \theta_{i'})$ together with $\omega_{i,j}$ intervals $\mathcal{J}_1, \cdots, \mathcal{J}_{\omega_{i,j}}$ each of length $\theta_{i'}$ corresponding to the bursts of indices $\omega^- + 1, \cdots, \omega^- + \omega_{i,j}$ of Tx i' are disjoint intervals and all are included in the interval $\mathcal{J} = (j\mu_i + \nu_i, j\mu_i + \nu_i + \theta_i)$ corresponding to the j^{th} codeword of Tx i. Hence, the sum of the lengths of the intervals $\mathcal{J}_0, \mathcal{J}_1, \cdots, \mathcal{J}_{\omega_{i,j}}$ which is $\vartheta := \omega^-\mu_{i'} + \nu_{i'} + \theta_{i'} - (j\mu_i + \nu_i) + \theta_{i'}\omega_{i,j}$ must be less than or equal to the length θ_i of $\mathcal{J}$. Then one can write $\theta_i\kappa_i' \leq (\theta_i - \vartheta)\kappa_i + \vartheta\kappa_i'$ due to $\theta_i - \vartheta \geq 0$. Since $\eta_i < \theta_i\kappa_i'$ by assumption, we get $\eta_i < (\theta_i - \vartheta)\kappa_i + \vartheta\kappa_i'$ which is exactly (3.64) after rearranging terms.

- The cases $\omega_{i,j}^- = 0, \omega_{i,j}^+ \neq 0$ and $\omega_{i,j}^- = \omega_{i,j}^+ = 0$ are analyzed similarly. We omit the details.

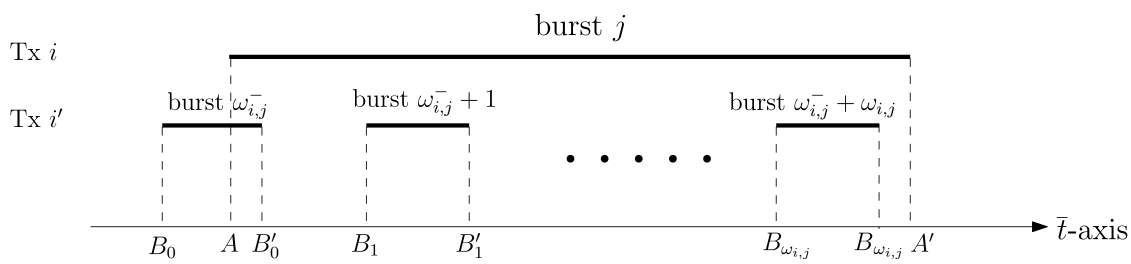


Figure 3.8: This picture shows the bursts of both users on the $\bar{t}$-axis in a situation where $\omega_{i,j}^- \neq 0$ and $\omega_{i,j}^+ = 0$

$A = j\mu_i + \nu_i$	$A' = j\mu_i + \nu_i + \theta_i$
$B_m = (\omega_{i,j}^- + m)\mu_{i'} + \nu_{i'}$ $m = 0, 1, \cdots, \omega_{i,j}$	$B'_m = (\omega_{i,j}^- + m)\mu_{i'} + \nu_{i'} + \theta_{i'}$ $m = 0, 1, \cdots, \omega_{i,j}$

(3.70)

If $\eta_i > \theta_i\kappa_i$, reliable communication is impossible due to the fact that the capacity of an AWGN channel with SNR γ_i is κ_i. The codebook rate for Tx i is $\lim_{n\to\infty} \frac{\lfloor n\eta_i \rfloor}{\lfloor n\theta_i \rfloor} = \frac{\eta_i}{\theta_i}$. The probability of error tends to one if $\frac{\eta_i}{\theta_i} \geq \kappa_i$. □

Remark- Let us describe a simple method to obtain the constraints (3.58), (3.64) and (3.68) in Propositions 10, 11 and 12, respectively. For example, let us discuss how to obtain (3.64) by looking at the positions of the bursts on the $\bar{t}$-axis. Fig. 3.8 depicts the j^{th} codeword of Tx i in a situation where $\omega_{i,j}^- \neq 0$ and $\omega_{i,j}^+ = 0$. The table in (3.70) shows the numbers on the $\bar{t}$-axis corresponding to different points in Fig 3.8. The interval during which the j^{th} burst of Tx i is sent can be divided into two subintervals, i.e., subinterval 1 where the two users interfere and subinterval 2 where there is no interference. Using (3.70) it is easy to see that

$$\begin{aligned}\text{length of subinterval 1} &= AB'_0 + B_1B'_1 + \cdots + B_{\omega_{i,j}}B'_{\omega_{i,j}} \\ &= \omega_{i,j}^-\mu_{i'} + \nu_{i'} + \theta_{i'} - (j\mu_i + \nu_i) + \omega_{i,j}\theta_{i'} \qquad (3.71)\end{aligned}$$

and

$$\begin{aligned}\text{length of subinterval 2} &= \theta_i - \text{length of subinterval 1} \\ &= \theta_i - \big(\omega_{i,j}^{-}\mu_{i'} + \nu_{i'} + \theta_{i'} - (j\mu_i + \nu_i) + \omega_{i,j}\theta_{i'}\big). \end{aligned} \quad (3.72)$$

Finally, the criterion for successful decoding of the j^{th} transmitted codeword of Tx i is given by

$$\eta_i < (\text{length of subinterval 1})\,\kappa_i' + (\text{length of subinterval 2})\,\kappa_i. \quad (3.73)$$

Substituting (3.71) and (3.72) in (3.73) and rearranging terms, we get (3.64).

3.5 The admissible set $\mathcal{A}$ for (ν_1, ν_2) and the probability of outage

3.5.1 System Design

In Section 3.2.4 we obtained the average transmission power and the average transmission rate for Tx i as $Q_i = \frac{N_i\gamma_i}{1+\frac{1}{q_i\theta_i}}$ and $R_i = \frac{\lambda_i}{1+q_i\theta_i}$, respectively, in the limit of large n. Throughout this section we assume none of the transmitters performs power control and both transmit at full power, i.e.,

$$Q_i = P_i, \;\; i = 1, 2. \quad (3.74)$$

We get

$$\theta_i = \hat{\theta}_i := \frac{1}{q_i}\Big(\frac{\lambda_i}{R_i} - 1\Big) \quad (3.75)$$

and

$$\gamma_i = \hat{\gamma}_i := \frac{P_i}{N_i}\Big(1 + \frac{1}{q_i\hat{\theta}_i}\Big) = \frac{P_i}{N_i}\frac{\lambda_i}{\lambda_i - R_i}. \quad (3.76)$$

If $N_i > 1$, then (3.12) together with (3.75) imply that

$$\frac{N_i}{N_i+1}\lambda_i < R_i < \lambda_i. \quad (3.77)$$

If $N_i = 1$, we simply have

$$0 < R_i < \lambda_i. \quad (3.78)$$

In Section 3.4, we derived sufficient conditions for successful decoding at the receivers. Letting $\theta_i = \hat{\theta}_i$ and $\gamma_i = \hat{\gamma}_i$, we aim to characterize an admissible region $\mathcal{A}$ for (ν_1, ν_2) such that reliable communication is guaranteed for all transmitted codewords. Define

$$\hat{\kappa}_i := \kappa_i\big|_{\gamma_i=\hat{\gamma}_i}, \quad \hat{\kappa}'_i := \kappa'_i\big|_{\gamma_1=\hat{\gamma}_1, \gamma_2=\hat{\gamma}_2} \tag{3.79}$$

If $\eta_i \geq \hat{\theta}_i \hat{\kappa}_i$, reliable communication is impossible for Tx i as stated in Corollary 1. For any value of N_i, define[5]

$$\overline{R}_i(N_i) := \sup\{R_i : \eta_i < \hat{\theta}_i \hat{\kappa}_i\}. \tag{3.80}$$

By (3.77), (3.78) and (3.80), we demand that R_i be in the interval

$$\frac{N_i}{N_i+1} \mathbb{1}_{N_i>1} \lambda_i < R_i < \min\{\lambda_i, \overline{R}_i(N_i)\} = \overline{R}_i(N_i), \tag{3.81}$$

where we have used the fact that[6] $\overline{R}_i(N_i) < \lambda_i$. For any R_1, R_2 define

$$\mathcal{N}_{R_1,R_2} := \big\{(N_1, N_2) : R_i \text{ satisfies (3.81) for } i = 1, 2\big\}. \tag{3.82}$$

We call $\mathcal{N}_{R_1,R_2}$ the *active set* for the pair (R_1, R_2). Note that $\mathcal{N}_{R_1,R_2}$ is finite due to the constraint $\frac{N_i}{N_i+1} \mathbb{1}_{N_i>1} < \frac{R_i}{\lambda_i} < 1$ in (3.81).

By Corollary 1, if $\eta_i < \hat{\theta}_i \hat{\kappa}'_i$, the codewords of user i are received reliably regardless of the values of ν_1 and ν_2. A more interesting situation occurs when R_1, R_2 satisfy

$$\min_{(N_1,N_2)\in\mathcal{N}_{R_1,R_2}} \max\left\{\frac{\eta_1}{\hat{\theta}_1 \hat{\kappa}'_1}, \frac{\eta_2}{\hat{\theta}_2 \hat{\kappa}'_2}\right\} \geq 1. \tag{3.83}$$

The inequality in (3.83) implies that for any $(N_1, N_2) \in \mathcal{N}_{R_1,R_2}$, there is $i \in \{1, 2\}$ such that $\eta_i \geq \hat{\theta}_i \hat{\kappa}'_i$ and hence, the values of ν_1 and ν_2 *may* potentially affect reliable communication for Tx i.

Remark- Throughout the rest of this section, we are only interested in rate pairs

[5] Using the change of variable $x := -\frac{1}{N_i} \frac{\lambda_i}{\lambda_i - R_i} - \frac{1}{P_i}$, the inequality $\eta_i < \hat{\theta}_i \hat{\kappa}_i$ can be written as $x2^x > -\frac{1}{P_i} 2^{-\frac{1}{N_i} - \frac{1}{P_i}}$ which is equivalent to $x > x_0$ for some x_0. This in turn results in the solution $0 < R_i < \overline{R}_i(N_i)$ for R_i.

[6] Recall from Footnote 5 that $\eta_i < \hat{\theta}_i \hat{\kappa}_i$ is equivalent to $0 < R_i < \overline{R}_i(N_i)$. It is easy to see that $\lim_{R_i \to \lambda_i^-} \hat{\theta}_i \hat{\kappa}_i = 0$. Hence, there exists a $0 < \delta < \lambda_i$ such that $\hat{\theta}_i \hat{\kappa}_i < \eta_i$ for $R_i = \lambda_i - \delta$. This gives $\overline{R}_i(N_i) \leq \lambda_i - \delta$ as desired.

(R_1, R_2) and system parameters $q_1, q_2, \lambda_1, \lambda_2, a_1, a_2, P_1, P_2$ such that (3.83) holds.

Towards characterizing an admissible region $\mathcal{A}$ for (ν_1, ν_2), we need to define the concept of the *state* in a GIC-SDA under immediate transmissions. Let $A_l = l\mu_1 + \nu_1$ and $A'_l = l\mu_1 + \nu_1 + \theta_1$ be the starting point and the ending point of the l^{th} codeword of Tx 1 on the $\bar{t}$-axis. The points A_l, A'_l for $1 \leq l \leq N_1$ partition the $\bar{t}$-axis into $2N_1 + 1$ disjoint intervals

$$\begin{array}{c} \mathcal{I}_1 := (-\infty, A_1) \\ \mathcal{I}_2 := (A_1, A'_1) \\ \mathcal{I}_3 := (A'_1, A_2) \\ \vdots \\ \mathcal{I}_{2N_1-1} := (A'_{N_1-1}, A_{N_1}) \\ \mathcal{I}_{2N_1} := (A_{N_1}, A'_{N_1}) \\ \mathcal{I}_{2N_1+1} := (A'_{N_1}, \infty) \end{array} . \tag{3.84}$$

For any $1 \leq j \leq N_2$, we assign a tuple (u_j, v_j) to the j^{th} transmitted codeword of Tx 2 where

- u_j is the unique index m such that the *starting* point of the j^{th} codeword of Tx 2 lies in interval $\mathcal{I}_m$.
- v_j is the unique index m such that the *ending* point of the j^{th} codeword of Tx 2 lies in interval $\mathcal{I}_m$.

Define the state of the asynchronous GIC-SDA by

$$S = \{(j; u_j, v_j) : 1 \leq j \leq N_2\}. \tag{3.85}$$

The set of all states is denoted by $\mathscr{S}$. For example, Fig. 3.9 depicts a situation where $N_1 = 3$ and $N_2 = 4$. The $\bar{t}$-axis is partitioned into seven intervals $\mathcal{I}_1, \cdots, \mathcal{I}_7$. We have $(u_1, v_1) = (1, 2), (u_2, v_2) = (3, 4), (u_3, v_3) = (5, 6), (u_4, v_4) = (7, 7)$ and the state of the channel is given by

$$S = \{(1; 1, 2), (2; 3, 4), (3; 5, 6), (4; 7, 7)\}. \tag{3.86}$$

In general, the number of states in a two-user asynchronous GIC-SDA is $|\mathscr{S}| = \binom{2N_1+2N_2}{2N_2}$. A proof of this fact is given in Appendix N. Note that any state uniquely determines the parameters $\omega^-_{i,j}, \omega^+_{i,j}$ and $\omega_{i,j}$ for any $1 \leq i \leq 2$ and $1 \leq j \leq N_i$.

For any state S, we impose two sets of constraints on (ν_1, ν_2) referred to as the geometric constraints $\mathcal{A}_S^{(\mathrm{geom})}$ and the reliability constraints $\mathcal{A}_S^{(\mathrm{rel})}$. The admissible region $\mathcal{A}$ is defined

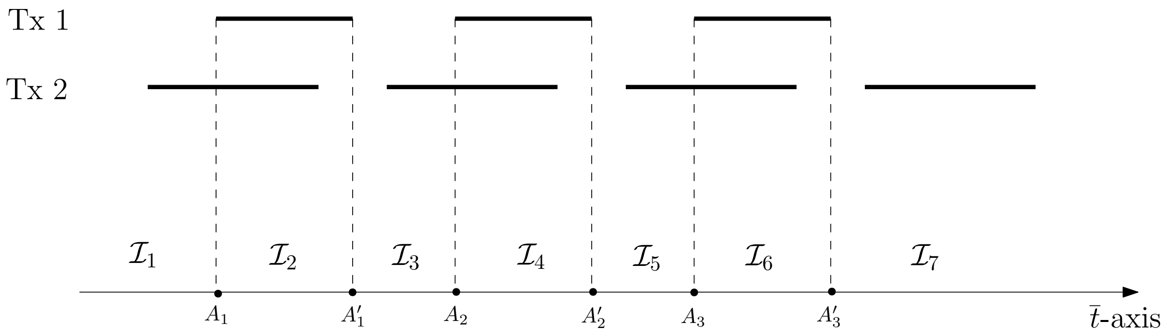


Figure 3.9: This picture shows the positions of the bursts of both users on the $\bar{t}$-axis in a situation where $N_1 = 3$ and $N_2 = 4$. For $1 \leq l \leq 3$, the points $A_l = l\mu_1 + \nu_1$ and $A'_l = l\mu_1 + \nu_1 + \theta_1$ are the starting point and the ending point of the l^{th} codeword of Tx 1. These points partition the $\bar{t}$-axis into seven disjoint intervals $\mathcal{I}_1, \cdots, \mathcal{I}_7$.

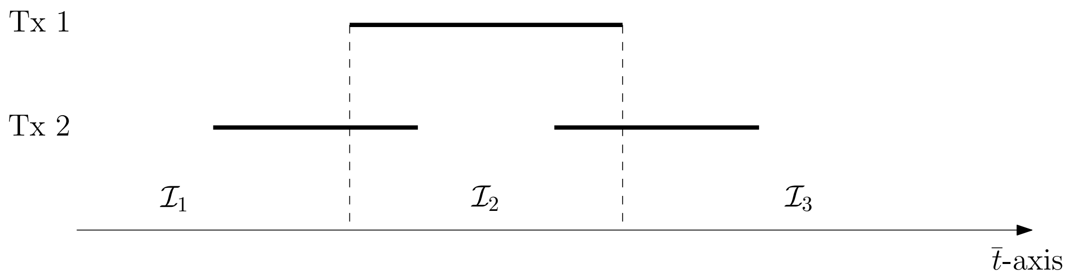


Figure 3.10: This picture shows the positions of the bursts of both users on the $\bar{t}$-axis in a situation where $N_1 = 1$ and $N_2 = 2$. The state of the channel is $\{(1;1,2),(2;2,3)\}$.

by

$$\mathcal{A} := \bigcup_{S \in \mathscr{S}} \left(\mathcal{A}_S^{(\text{geom})} \bigcap \mathcal{A}_S^{(\text{rel})}\right). \tag{3.87}$$

The geometric constraints are dictated by the positions of the bursts on the $\bar{t}$-axis. For example, let $N_1 = 1$, $N_2 = 2$ and the state of the channel be $S = \{(1;1,2),(2;2,3)\}$ as shown in Fig. 3.10. Then

$$\mathcal{A}_S^{(\text{geom})} = \left\{(\nu_1,\nu_2) : \mu_2+\nu_2 < \mu_1+\nu_1 < \mu_2+\nu_2+\theta_2 < 2\mu_2+\nu_2 < \mu_1+\nu_1+\theta_1 < 2\mu_2+\nu_2+\theta_2\right\}. \tag{3.88}$$

The reliability constraints guarantee reliable communications for all transmitted codewords. For example, for the situation in Fig. 3.10 there are three reliability constraints:

- For the first codeword of Tx 1, $\omega^-_{1,1} = 1, \omega^+_{1,1} = 2$ and $\omega_{1,1} = 0$. By Proposition 10,

$$\eta_1 - \frac{1}{\lambda_2}(\hat{\kappa}_1 - \hat{\kappa}'_1)\eta_2 < \hat{\theta}_1\hat{\kappa}'_1 - \hat{\theta}_2(\hat{\kappa}_1 - \hat{\kappa}'_1). \tag{3.89}$$

- For the first codeword of Tx 2, $\omega^-_{2,1} = 0, \omega^+_{2,1} = 1$ and $\omega_{2,1} = 0$. By Proposition 12,

$$\big(1 + \frac{1}{\lambda_2}(\hat{\kappa}_2 - \hat{\kappa}'_2)\big)\eta_2 - \frac{1}{\lambda_1}(\hat{\kappa}_2 - \hat{\kappa}'_2)\eta_1 < \hat{\theta}_2\hat{\kappa}'_2 + (\nu_1 - \nu_2)(\hat{\kappa}_2 - \hat{\kappa}'_2). \tag{3.90}$$

- For the second codeword of Tx 2, $\omega^-_{2,2} = 1, \omega^+_{2,2} = 0$ and $\omega_{2,2} = 0$. By Proposition 11,

$$\big(1 - \frac{2}{\lambda_2}(\hat{\kappa}_2 - \hat{\kappa}'_2)\big)\eta_2 + \frac{1}{\lambda_1}(\hat{\kappa}_2 - \hat{\kappa}'_2)\eta_1 < \hat{\theta}_2\hat{\kappa}_2 - (\nu_1 - \nu_2 + \hat{\theta}_1)(\hat{\kappa}_2 - \hat{\kappa}'_2). \tag{3.91}$$

Then $\mathcal{A}_S^{(\text{rel})}$ for $S = \{(1;1,2),(2;2,3)\}$ is the set of all (ν_1, ν_2) such that the three inequalities in (3.89), (3.90) and (3.91) hold.

We are ready to state our design problem. Let (R_1, R_2) be such that (3.83) holds and ν_1 and ν_2 be realizations of independent *uniform* random variables[7] $\boldsymbol{\nu}_1$ and $\boldsymbol{\nu}_2$, respectively, with support $[0, d]$ for some $d > 0$. We aim to find $(N_1, N_2) \in \mathcal{N}_{R_1,R_2}$ such that the probability of $(\boldsymbol{\nu}_1, \boldsymbol{\nu}_2)$ not being in the admissible region $\mathcal{A}$ is minimized, i.e.,

$$(\widehat{N}_1, \widehat{N}_2) = \arg \min_{(N_1,N_2)\in\mathcal{N}_{R_1,R_2}} \mathbb{P}\big((\boldsymbol{\nu}_1, \boldsymbol{\nu}_2) \notin \mathcal{A}\big). \tag{3.92}$$

In words, (3.92) answers the following question:

Given the value of d and assuming the rate pair (R_1, R_2) is such that (3.83) holds, What is the optimum number of transmission bursts N_i for Tx i in order to minimize the probability of the outage event, i.e., the event that $(\boldsymbol{\nu}_1, \boldsymbol{\nu}_2)$ does not lie in the admissible region $\mathcal{A}$?

Since the sets $\mathcal{A}_S^{(\text{geom})} \bigcap \mathcal{A}_S^{(\text{rel})}$ are disjoint[8] for different states S, we can write

$$\mathbb{P}\big((\boldsymbol{\nu}_1, \boldsymbol{\nu}_2) \notin \mathcal{A}\big) = 1 - \sum_{S\in\mathscr{S}} \mathbb{P}\big((\boldsymbol{\nu}_1, \boldsymbol{\nu}_2) \in \mathcal{A}_S^{(\text{geom})} \bigcap \mathcal{A}_S^{(\text{rel})}\big). \tag{3.93}$$

[7]One can consider any arbitrary continuous distribution for $\boldsymbol{\nu}_1$ and $\boldsymbol{\nu}_2$. We consider the uniform distribution due to its realistic nature.

[8]This is due to the fact that the geometric constraints $\mathcal{A}_S^{(\text{geom})}$ are disjoint for different states S.

Define

$$\boldsymbol{\alpha} := \boldsymbol{\nu}_2 - \boldsymbol{\nu}_1. \tag{3.94}$$

Each constraint $(\boldsymbol{\nu}_1, \boldsymbol{\nu}_2) \in \mathcal{A}_S^{(\text{geom})} \bigcap \mathcal{A}_S^{(\text{rel})}$ is in fact a constraint on $\boldsymbol{\alpha}$, i.e., for any state S there are real numbers $\alpha_S^{(l)}$ and $\alpha_S^{(u)}$ such that[9]

$$(\boldsymbol{\nu}_1, \boldsymbol{\nu}_2) \in \mathcal{A}_S^{(\text{geom})} \bigcap \mathcal{A}_S^{(\text{rel})} \iff \alpha_S^{(l)} < \boldsymbol{\alpha} < \alpha_S^{(u)}. \tag{3.95}$$

By (3.93) and (3.95),

$$\mathbb{P}\big((\boldsymbol{\nu}_1, \boldsymbol{\nu}_2) \notin \mathcal{A}\big) = 1 - \sum_{S \in \mathscr{S}} \big(F_{\boldsymbol{\alpha}}(\alpha_S^{(u)}) - F_{\boldsymbol{\alpha}}(\alpha_S^{(l)})\big), \tag{3.96}$$

where $F_{\boldsymbol{\alpha}}(\alpha) = \frac{1}{d}(1 - \frac{|\alpha|}{d})\mathbb{1}_{|\alpha| \leq d}$ is the cumulative distribution function of $\boldsymbol{\alpha}$.

The next proposition provides conditions on the parameter $d > 0$ such that reliable communication is guaranteed for both users regardless of the values of $\nu_1, \nu_2 \in [0, d]$:

Proposition 13. *Let $\mathcal{S}_d := (0, d) \times (0, d)$. Then $\mathbb{P}\big((\boldsymbol{\nu}_1, \boldsymbol{\nu}_2) \notin \mathcal{A}\big) = 0$ if and only if*

$$\mathcal{S}_d \bigcap (\mathcal{A}_S^{(\text{geom})} \setminus \mathcal{A}_S^{(\text{rel})}) = \emptyset, \tag{3.97}$$

for any $S \in \mathscr{S}$.

Proof. Assume $\mathbb{P}\big((\boldsymbol{\nu}_1, \boldsymbol{\nu}_2) \in \mathcal{A}\big) = \sum_{S \in \mathscr{S}} \mathbb{P}\big((\boldsymbol{\nu}_1, \boldsymbol{\nu}_2) \in \mathcal{A}_S^{(\text{geom})} \bigcap \mathcal{A}_S^{(\text{rel})}\big) = 1$. Since $\{\mathcal{A}_S^{(\text{geom})} : S \in \mathscr{S}\}$ is a partition of the sample space, $\sum_{S \in \mathscr{S}} \mathbb{P}\big((\boldsymbol{\nu}_1, \boldsymbol{\nu}_2) \in \mathcal{A}_S^{(\text{geom})}\big) = 1$. But, $\mathbb{P}\big((\boldsymbol{\nu}_1, \boldsymbol{\nu}_2) \in \mathcal{A}_S^{(\text{geom})} \bigcap \mathcal{A}_S^{(\text{rel})}\big) \leq \mathbb{P}\big((\boldsymbol{\nu}_1, \boldsymbol{\nu}_2) \in \mathcal{A}_S^{(\text{geom})}\big)$ for any $S \in \mathscr{S}$. Hence, $\mathbb{P}\big((\boldsymbol{\nu}_1, \boldsymbol{\nu}_2) \in \mathcal{A}_S^{(\text{geom})} \bigcap \mathcal{A}_S^{(\text{rel})}\big) = \mathbb{P}\big((\boldsymbol{\nu}_1, \boldsymbol{\nu}_2) \in \mathcal{A}_S^{(\text{geom})}\big)$ or equivalently, $\mathbb{P}\big((\boldsymbol{\nu}_1, \boldsymbol{\nu}_2) \in \mathcal{A}_S^{(\text{geom})} \setminus \mathcal{A}_S^{(\text{rel})}\big) = 0$ for any $S \in \mathscr{S}$. This in turn means $\mathcal{S}_d \bigcap (\mathcal{A}_S^{(\text{geom})} \setminus \mathcal{A}_S^{(\text{rel})})$ has Lebesgue measure zero. But, $\mathcal{A}_S^{(\text{geom})} \setminus \mathcal{A}_S^{(\text{rel})}$ is a union of a finite number of (disjoint) strips of the form $\{(\nu_1, \nu_2) : \beta^{(l)} < \nu_2 - \nu_1 < \beta^{(u)}\}$ where $\beta^{(l)}$ and $\beta^{(u)}$ are real numbers. As such, $\mathcal{S}_d \bigcap (\mathcal{A}_S^{(\text{geom})} \setminus \mathcal{A}_S^{(\text{rel})})$ is empty by virtue of being an *open* set with Lebesgue measure zero. □

[9] For example, look at the geometric and reliability constraints given for the state depicted in Fig. 3.10.

Motivated by Proposition 13, we define

$$d_{\max} := \sup\left\{d > 0 : \mathcal{S}_d \bigcap (\mathcal{A}_S^{(\text{geom})} \setminus \mathcal{A}_S^{(\text{rel})}) = \emptyset \text{ for any } S \in \mathscr{S}\right\}. \tag{3.98}$$

If $d_{\max} > 0$, then the probability of outage is zero for any $d < d_{\max}$. In the next subsection, we offer simulation results to study the effects of different system parameters on the optimal choices for N_1, N_2 in (3.92). In particular, we will see an example of a rate tuple (R_1, R_2) that satisfies (3.83) and still there exists $d > 0$ such that (3.97) holds for any $S \in \mathscr{S}$. Therefore, if (R_1, R_2) satisfies (3.83), it does not necessarily mean that $\mathbb{P}\big((\boldsymbol{\nu}_1, \boldsymbol{\nu}_2) \notin \mathcal{A}\big) > 0$.

3.5.2 Simulations

In this subsection, we study the optimum choices for N_1, N_2 in (3.92) in a few examples.

Example- Let $k_1 = 3, k_2 = 2, q_1 = 0.3, q_2 = 0.4, a_1 = 0.5, a_2 = 0.7$ and $P_1 = P_2 = 30\,\text{dB}$. We consider several possibilities for (R_1, R_2):

- Let $R_1 = 0.4\lambda_1 = 0.36$ and $R_2 = 0.4\lambda_2 = 0.32$. Then $\mathcal{N}_{R_1,R_2} = \{(1,1)\}$, however, condition (3.83) is not satisfied. This means that by setting $N_1 = N_2 = 1$, both receivers successfully decode the messages regardless of the values of ν_1 and ν_2.

- Let $R_1 = 0.5\lambda_1 = 0.45$ and $R_2 = 0.5\lambda_2 = 0.4$. Then $\mathcal{N}_{R_1,R_2} = \{(1,1)\}$ and the condition in (3.83) is satisfied. We fix $N_1 = N_2 = 1$. Fig. 3.11 in panel (a) shows the region $\bigcup_{S\in\mathscr{S}}(\mathcal{A}_S^{(\text{geom})} \setminus \mathcal{A}_S^{(\text{rel})})$. By Proposition 13, if $d < d_{\max} \approx 0.83$, the probability of outage is zero. Fig. 3.11 in panel (b) shows the probability of outage in terms of d.

- Let $R_1 = 0.7\lambda_1 = 0.63$ and $R_2 = 0.7\lambda_2 = 0.56$. Then $\mathcal{N}_{R_1,R_2} = \{(1,1),(1,2),(2,1),(2,2)\}$ and the condition in (3.83) is satisfied. Fig. 3.12 in panel (a) shows the probability of outage in terms of d for different values of $(N_1, N_2) \in \mathcal{N}_{R_1,R_2}$. If $d < 2$, $(N_1, N_2) = (1,2)$ and if $d > 2$, $(N_1, N_2) = (1,1)$ are the optimum choices.

- Let $R_1 = 0.8\lambda_1 = 0.72$ and $R_2 = 0.8\lambda_2 = 0.64$. Then $\mathcal{N}_{R_1,R_2} = \{(m_1, m_2) : 1 \leq m_1, m_2 \leq 3\}$ and the condition in (3.83) is satisfied. It turns out that depending on the value of d, the best choices are $(N_1, N_2) = (1,1), (1,2)$ or $(1,3)$. Fig. 3.12 in panel (b) shows the probability of outage in terms of d for these values of (N_1, N_2). If $d < 1.43$, $(N_1, N_2) = (1,3)$, if $1.43 < d < 2.51$, $(N_1, N_2) = (1,2)$ and if $d > 2.51$, $(N_1, N_2) = (1,1)$ are the best choices.

Example- It is possible that $d_{\max} = 0$. By Proposition 13, this happens when the line $\nu_1 = \nu_2$ lies in the interior of the region $\bigcup_{S\in\mathscr{S}}(\mathcal{A}_S^{(\text{geom})} \setminus \mathcal{A}_S^{(\text{rel})})$ in the ν_1-ν_2 plane. An example of this situation is a symmetric scenario where $k_1 = k_2 = 5$, $q_1 = q_2 = 0.2$, $a_1 = a_2 = 0.5$, $R_1 = 0.7\lambda_1 = 0.7$ and $R_2 = 0.7\lambda_2 = 0.7$. We consider two cases for the average transmission power, i.e., $P_1 = P_2 = 10\,\text{dB}$ and $P_1 = P_2 = 30\,\text{dB}$. It turns out that in both cases, $\mathcal{N}_{R_1,R_2} = \{(1,1),(1,2),(2,1),(2,2)\}$ and the condition in (3.83) is satisfied.

- Let $P_1 = P_2 = 10\,\text{dB}$. Fig. 3.13 in panel (a) presents the probability of outage in terms of d for different values of (N_1, N_2). Due to symmetry, the cases $N_1 = 1, N_2 = 2$ and $N_1 = 2, N_2 = 1$ offer the the same performance. We see that $N_1 = N_2 = 1$ is the optimum choice for any value of d.

- Let $P_1 = P_2 = 30\,\text{dB}$. Fig. 3.13 in panel (b) presents the probability of outage in terms of d for different values of (N_1, N_2). In contrast to the case $P_1 = P_2 = 10\,\text{dB}$ in panel (a), the situation is reversed. Here, $N_1 = N_2 = 2$ is the optimum choice for any value of d.

3.6 An achievable region for the asynchronous GIC-SDA

3.6.1 The General Model

In this section we consider a different setting where

- The source of Tx i no longer turns off after generating a number of $k_i n$ bits.

- The parameters ν_1 and ν_2 are known at both transmitters[10]. As before, we let $\alpha := \nu_2 - \nu_1$.

Accordingly, we adopt a slightly different notation where we assume Tx i has a codebook with rate $R_{c,i}$ consisting of $2^{\lfloor n_i R_{c,i} \rfloor}$ codewords of length $n_i = \lfloor n\theta_i \rfloor$ where $\theta_i > 0$ is a constant. We pose the following problem:

Given positive integers N_1 and N_2, determine the possible values for the codebook rates $(R_{c,1}, R_{c,2})$ such that

[10]This requires a certain level of coordination between the two transmitters in order to inform each other about their initial instants of activity.

1. *The first N_i codewords sent by Tx i are transmitted immediately in the sense defined in Section 3.2.2 and decoded successfully at Rx i.*

2. *The average transmission power for Tx i satisfies (3.2) where $\mathcal{T}_i$ is the period of activity for Tx i until the time slot it transmits the last symbol in its N_i^{th} burst.*

Since the codewords are transmitted immediately, the content of the buffer of Tx i never exceeds $\lfloor n_i R_{c,i} \rfloor + k_i$ and therefore, the buffers are stable. The first N_i bursts sent by Tx i represent a total number of $N_i \lfloor n_i R_{c,i} \rfloor$ bits. Therefore, the average transmission rate for Tx i is $R_i = \frac{N_i \lfloor n_i R_{c,i} \rfloor}{|\mathcal{T}_i|}$. Following similar steps in Section 3.2.4,

$$R_i = \frac{N_i R_{c,i}}{1 + \frac{N_i R_{c,i}}{\lambda_i}}, \tag{3.99}$$

and the average transmission power for Tx i is

$$Q_i = \frac{N_i \gamma_i}{1 + \frac{N_i R_{c,i}}{\lambda_i}}, \tag{3.100}$$

in the limit of large n. Then the average power constraint $Q_i \leq P_i$ in (3.2) becomes

$$0 \leq \gamma_i \leq \left(\frac{1}{N_i} + \frac{R_{c,i}}{\lambda_i} \right) P_i. \tag{3.101}$$

Before proceeding further, let us reiterate the major differences between the current setup and the setup in the previous section:

- We show the codebook rate of Tx i by $R_{c,i}$. No notation was selected for the codebook rate $\frac{\eta_i}{\theta_i}$ in the previous section. All the achievability results in Propositions 10, 11 and 12 remain valid after replacing η_i by $\theta_i R_{c,i}$.

- N_i and θ_i are constants, while they served as design parameters in the previous section.

- The parameters ν_1 and ν_2 are known at both transmitters, while ν_i was unknown to Tx i' in the previous section.

- The information source at Tx i turns on at time slot $\lfloor n\nu_i \rfloor$ and remains active indeterminately.

We aim to characterize a region $\mathcal{R}$ of all codebook rate tuples $(R_{c,1}, R_{c,2})$ such that both transmitters send their codewords immediately and reliably and such that the power constraints in (3.2) are not violated. We call $\mathcal{R}$ the achievable (codebook) rate region. One can also define an achievable rate region $\mathcal{R}'$ of all transmission rate tuples (R_1, R_2) such that the aforementioned properties hold. Since $\mathcal{R}'$ and $\mathcal{R}$ are related through the mappings in (3.99), we only focus on $\mathcal{R}$. Immediate transmission of a scheduled codeword is impossible if a previously scheduled codeword is not fully transmitted. By (3.12), immediate transmission of the codewords is guaranteed if $\frac{\theta_i R_{c,i}}{\lambda_i} > \theta_i$ for $N_i > 1$, i.e.,

$$R_{c,i} > \lambda_i \mathbb{1}_{N_i>1}, \;\; i = 1, 2. \tag{3.102}$$

An achievable $R_{c,i}$ must satisfy[11]

$$R_{c,i} < \frac{1}{2}\log(1 + \gamma_i). \tag{3.103}$$

Combining (3.101) and (3.103),

$$R_{c,i} < \frac{1}{2}\log\Big(1 + \Big(\frac{1}{N_i} + \frac{R_{c,i}}{\lambda_i}\Big)P_i\Big). \tag{3.104}$$

This is equivalent to

$$R_{c,i} < \overline{R}_{c,i}, \quad i = 1, 2, \tag{3.105}$$

where $\overline{R}_{c,i}$ is the unique positive solution[12] for $R_{c,i}$ in the equation $R_{c,i} = \frac{1}{2}\log(1 + (\frac{1}{N_i} + \frac{R_{c,i}}{\lambda_i})P_i)$. By (3.102) and (3.105), $\mathcal{R}$ lies inside the rectangle $[\lambda_1 \mathbb{1}_{N_1>1}, \overline{R}_{c,1}] \times [\lambda_2 \mathbb{1}_{N_2>1}, \overline{R}_{c,2}]$.

To describe $\mathcal{R}$, we need the concept of the state introduced in Section 3.5.1 for a GIC-SDA with immediate transmissions. For each state S, we impose two sets of constraints on $(R_{c,1}, R_{c,2})$, i.e., the geometric constraints and the reliability constraints shown by $\mathcal{R}_S^{(\text{geom})}$ and $\mathcal{R}_S^{(\text{rel})}$, respectively. To describe these constraints, let us consider the situation shown in Fig. 3.14 where $N_1 = N_2 = 2$ and the state of the channel is $S = \{(1; 1, 2), (2; 2, 4)\}$. This is only one of $\binom{8}{4} = 70$ possible states. The table in (3.106) shows the numbers on the $\bar{t}$-axis corresponding to different points in Fig 3.14.

[11] By Corollary 1, we must have $\theta_i R_{c,i} < \theta_i \kappa_i$ which simplifies to (3.103).

[12] Define $f(x) = x - \frac{1}{2}\log(1 + (\frac{1}{N_i} + \frac{x}{\lambda_i})P_i)$. Note that $f(0) = -\frac{1}{2}\log(1 + \frac{P_i}{N_i}) < 0$, $\lim_{x\to\infty} f(x) = \infty$ and $f'(x) = 1 - \frac{1}{2\ln 2}\frac{\frac{P_i}{\lambda_i}}{1+(\frac{1}{N_i}+\frac{x}{\lambda_i})P_i}$. If $f'(x) > 0$ for all $x > 0$, then $f(x) = 0$ has only one positive solution. If there is $x_0 > 0$ such that $f'(x_0) = 0$, then f is decreasing over $(0, x_0)$ and increasing over (x_0, ∞). This again implies that $f(x) = 0$ has only one positive solution.

- The geometric constraints are imposed by the positions of the bursts along the $\bar{t}$-axis. For example, point B_1 is on left of point A_1 which gives $\frac{\theta_2 R_{c,2}}{\lambda_2} + \nu_2 < \frac{\theta_1 R_{c,1}}{\lambda_1} + \nu_1$. A complete list of the geometric constraints is given by the polyhedron

$$\begin{bmatrix} -\frac{\theta_1}{\lambda_1} & \frac{\theta_2}{\lambda_2} \\ \frac{\theta_1}{\lambda_1} & -\frac{\theta_2}{\lambda_2} \\ -\frac{\theta_1}{\lambda_1} & \frac{2\theta_2}{\lambda_2} \\ \frac{2\theta_1}{\lambda_1} & -\frac{2\theta_2}{\lambda_2} \\ -\frac{2\theta_1}{\lambda_1} & \frac{2\theta_2}{\lambda_2} \\ -1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} R_{c,1} \\ R_{c,2} \end{bmatrix} < \begin{bmatrix} -\alpha \\ \theta_2 + \alpha \\ \theta_1 - \alpha \\ \theta_2 + \alpha \\ \theta_1 - \theta_2 - \alpha \\ -\lambda_1 \\ -\lambda_2 \end{bmatrix}, \tag{3.107}$$

i.e., $\mathcal{R}_S^{(\mathrm{geom})}$ for $S = \{(1;1,2),(2;2,4)\}$ in the set of all $(R_{c,1}, R_{c,2})$ such that (3.107) holds. The last two constraints in (3.107) are the inequalities in (3.102).

- The reliability constraints guarantee successful decoding for all $N_1 + N_2 = 2 + 2 = 4$ transmitted codewords subject to the power conditions in (3.101). For example, the second codeword of Tx 1 in Fig. 3.14 only interferes with the second codeword of Tx 2 at its "left end", i.e., $\omega_{2,2}^- = 2$, $\omega_{2,2}^+ = 0$, $\omega_{2,2} = 0$. We invoke Proposition 11 to write

$$\big(1 - \frac{2}{\lambda_1}(\kappa_1 - \kappa_1')\big)\theta_1 R_{c,1} + \frac{2}{\lambda_2}(\kappa_1 - \kappa_1')\theta_2 R_{c,2} < \theta_1\kappa_1 - (\alpha + \theta_2)(\kappa_1 - \kappa_1'). \tag{3.108}$$

A complete list of reliability constraints is given by the polyhedra

$$\begin{bmatrix} \theta_1 & -\frac{1}{\lambda_2}(\kappa_1 - \kappa_1')\theta_2 \\ (1 - \frac{2}{\lambda_1}(\kappa_1 - \kappa_1'))\theta_1 & \frac{2}{\lambda_2}(\kappa_1 - \kappa_1')\theta_2 \\ -\frac{1}{\lambda_1}(\kappa_2 - \kappa_2')\theta_1 & (1 + \frac{1}{\lambda_2}(\kappa_2 - \kappa_2'))\theta_2 \\ -\frac{1}{\lambda_1}(\kappa_2 - \kappa_2')\theta_1 & \theta_2 \\ -1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} R_{c,1} \\ R_{c,2} \end{bmatrix} < \begin{bmatrix} \theta_1\kappa_1' - \theta_2(\kappa_1 - \kappa_1') \\ \theta_1\kappa_1 - (\alpha + \theta_2)(\kappa_1 - \kappa_1') \\ \theta_2\kappa_2' - \alpha(\kappa_2 - \kappa_2') \\ \theta_2\kappa_2' - \theta_1(\kappa_2 - \kappa_2') \\ \lambda_1(\frac{1}{2} - \frac{\gamma_1}{P_1}) \\ \lambda_2(\frac{1}{2} - \frac{\gamma_2}{P_2}) \end{bmatrix}, \tag{3.109}$$

for some $\gamma_1, \gamma_2 \geq 0$, i.e., $\mathcal{R}_S^{(\mathrm{rel})}$ for $S = \{(1;1,2),(2;2,4)\}$ is the set of all $(R_{c,1}, R_{c,2})$ such that (3.109) holds for some $\gamma_1, \gamma_2 \geq 0$. The last two constraints in (3.109) are the inequalities in (3.101). Note that $\mathcal{R}_S^{(\mathrm{rel})}$ is the union of infinitely many polyhedra. More precisely, if we denote the polyhedron in (3.109) for fixed $\gamma_1, \gamma_2 \geq 0$ by

$\mathcal{P}_S(\gamma_1, \gamma_2)$, then

$$\mathcal{R}_S^{(\text{rel})} = \bigcup_{\gamma_1, \gamma_2 \geq 0} \mathcal{P}_S(\gamma_1, \gamma_2). \tag{3.110}$$

It is needless to mention that κ_i and κ_i' are functions of γ_1, γ_2.

Having $\mathcal{R}_S^{(\text{geom})}$ and $\mathcal{R}_S^{(\text{rel})}$ defined for any state S, the achievable rate region $\mathcal{R}$ is given by

$$\mathcal{R} = \bigcup_{S \in \mathscr{S}} \left(\mathcal{R}_S^{(\text{geom})} \bigcap \mathcal{R}_S^{(\text{rel})}\right). \tag{3.111}$$

A few remarks are in order:

- In general, none of $\mathcal{R}_S^{(\text{geom})}$ and $\mathcal{R}_S^{(\text{rel})}$ is a subset of the other.
- Depending on system parameters, there may exist a state S such that $\mathcal{R}_S^{(\text{geom})} \bigcap \mathcal{R}_S^{(\text{rel})} = \emptyset$.
- Full power transmission, i.e., $\gamma_i = (\frac{1}{N_i} + \frac{R_{c,i}}{\lambda_i})P_i$ is not in general optimum. For example, let $N_1 = 3$, $N_2 = 2$, $\theta_1 = \theta_2 = 1, k_1 = 2, k_2 = 3, q_1 = 0.2, q_2 = 0.1, a_1 = 1.5, a_2 = 0.5, P_1 = 20\,\text{dB}, P_2 = 30\,\text{dB}$ and $\alpha = 1$. Fig. 3.15 in panel (a) shows the regions $\mathcal{R}_S^{(\text{geom})}$ in grey and $\mathcal{R}_S^{(\text{rel})}$ in black for $S = \{(1; 2, 3), (2; 3, 4)\}$. Fig. 3.15 in panel (b) shows the same regions under full power transmission. It is seen that $\mathcal{R}_S^{(\text{rel})}$ under full power transmission is strictly smaller than $\mathcal{R}_S^{(\text{rel})} = \bigcup_{\gamma_1, \gamma_2 \geq 0} \mathcal{P}_S(\gamma_1, \gamma_2)$.
- In order to plot $\mathcal{R}_S^{(\text{rel})}$ for a given state S, we choose a finite set of values for γ_i, namely Γ_i, and approximate $\mathcal{R}_S^{(\text{rel})}$ by

$$\widetilde{\mathcal{R}}_S^{(\text{rel})} := \bigcup_{\gamma_1 \in \Gamma_1, \gamma_2 \in \Gamma_2} \mathcal{P}_S(\gamma_1, \gamma_2) \subseteq \mathcal{R}_S^{(\text{rel})}. \tag{3.112}$$

To choose Γ_i, we observe that

$$0 \leq \gamma_i < \overline{\gamma}_i := \left(\frac{1}{N_i} + \frac{\overline{R}_{c,i}}{\lambda_i}\right) P_i, \tag{3.113}$$

due to (3.101) and (3.105). Fix a natural number m and let

$$\Gamma_i := \left\{\frac{l}{m}\overline{\gamma}_i : 1 \leq l \leq m-1\right\}. \tag{3.114}$$

The set difference $\mathcal{R}_S^{(\mathrm{rel})} \backslash \widetilde{\mathcal{R}}_S^{(\mathrm{rel})}$ becomes smaller as m increases. For example, Fig. 3.16 shows the region $\widetilde{R}_S^{(\mathrm{rel})}$ for $S = \{(1;2,3),(2;3,4)\}$ in a setup where $N_1 = 3, N_2 = 2, \theta_1 = \theta_2 = 1, k_1 = 2, k_2 = 3, q_1 = 0.2, q_2 = 0.1, a_1 = 1.5, a_2 = 0.5, P_1 = 20\,\mathrm{dB}, P_2 = 30\,\mathrm{dB}$ and $\alpha = 1$. In panel (a), we have $\Gamma_i = \{\frac{l}{5}\overline{\gamma}_i : 1 \leq l \leq 4\}$ and in panel (b), $\Gamma_i = \{\frac{l}{10}\overline{\gamma}_i : 1 \leq l \leq 9\}$.

In the next two examples, we fix $\theta_1 = \theta_2 = 1, k_1 = k_2 = 2, q_1 = q_2 = 0.3, a_1 = a_2 = 0.5$, $P_1 = P_2 = 20\,\mathrm{dB}$ and study the effects of N_1, N_2 and α on $\mathcal{R}$ in (3.111). We also fix $\Gamma_i = \{\frac{l}{10}\overline{\gamma}_i : 1 \leq l \leq 9\}$.

Example- Let $\alpha = 0$, i.e., both users become active at the same time. Fig. 3.17 shows the region $\mathcal{R}$ for different values of $N_1 = N_2$. As the number of transmitted codewords increases, $\mathcal{R}$ becomes strictly smaller.

Example- Let $N_1 = N_2 = 2$. Fig. 3.18 shows the region $\mathcal{R}$ for different values of α. As α increases, the region $\mathcal{R}$ converges to the square $\{(R_{c,1}, R_{c,2}) : \lambda_i < R_{c,i} < \overline{R}_{c,i}, , i = 1, 2\}$ where $\overline{R}_{c,1} = \overline{R}_{c,2} \approx 4.8774$.

3.6.2 The Symmetric Model

In this section we study a symmetric setting where except for ν_1 and ν_2, other system parameters for the two users are identical. In this case, we drop the index $i = 1, 2$, i.e., $N_1 = N_2 = N$, $\lambda_1 = \lambda_2 = \lambda$, $R_{c,1} = R_{c,2} = R_c$, $\theta_1 = \theta_2 = \theta$, $a_1 = a_2 = a$, $\gamma_1 = \gamma_2 = \gamma$ and $P_1 = P_2 = P$. Without loss of generality,

$$\nu_2 \geq \nu_1. \tag{3.115}$$

Let[13] $\mathcal{R}_{\mathrm{sym}}$ be the set of all $R_c > \lambda \mathbb{1}_{N>1}$ such that

- All $2N$ transmitted codewords are decoded successfully.
- The average transmission power for Tx i satisfies $\frac{1}{|\mathcal{T}_i|}\sum_{t\in\mathcal{T}_i} x_{i,t}^2 \leq P$ where $\mathcal{T}_i$ is the period of activity for Tx i until the time slot it transmits the last symbol in its N^{th} burst.

[13]More precisely, $\mathcal{R}_{\mathrm{sym}}$ is the set of codebook rates $R_c > \lambda \mathbb{1}_{N>1}$ such that $\gamma < (\frac{1}{N} + \frac{R_c}{\lambda})P$ and all $2N$ transmitted codewords are decoded successfully according to the sufficient conditions put forth by Propositions 10, 11 and 12.

Towards characterizing $\mathcal{R}_{\text{sym}}$, we define the set $\mathcal{P}(x, y; \gamma)$ for real numbers x and y by

$$\mathcal{P}(x, y; \gamma) = \left\{ R_c > \lambda \mathbb{1}_{N>1} : \big(1 - \frac{x}{\lambda}(\kappa_\gamma - \kappa'_\gamma)\big) R_c < \kappa'_\gamma - (\kappa_\gamma - \kappa'_\gamma)\frac{y}{\theta}, \gamma < \big(\frac{1}{N} + \frac{R_c}{\lambda}\big) P \right\}, \tag{3.116}$$

where

$$\kappa_\gamma := \mathsf{C}(\gamma), \quad \kappa'_\gamma := \mathsf{C}\big(\frac{\gamma}{1 + a\gamma}\big). \tag{3.117}$$

One can rephrase the statements in Propositions 10, 11 and 12 in Proposition 14:

Proposition 14. *For $1 \leq i \leq 2$ and $1 \leq j \leq N$ assume the following conditions hold:*

- *If $\omega^-_{i,j}, \omega^+_{i,j} \neq 0$, then $R_c \in \mathcal{P}(1, \theta; \gamma)$.*
- *If $\omega^-_{i,j} \neq 0$ and $\omega^+_{i,j} = 0$, then $R_c \in \mathcal{P}(j - \omega^-_{i,j}, \nu_{i'} - \nu_i; \gamma)$.*
- *If $\omega^-_{i,j} = 0$ and $\omega^+_{i,j} \neq 0$, then $R_c \in \mathcal{P}(\omega^+_{i,j} - j, \nu_i - \nu_{i'}; \gamma)$.*
- *If $\omega^-_{i,j} = \omega^+_{i,j} = 0$, then $R_c \in \mathcal{P}(0, -\theta; \gamma)$.*

Then the probability of error in decoding the j^{th} message of Tx i can be made arbitrarily small by choosing n sufficiently large.

If $N = 1$, one can easily find $\mathcal{R}_{\text{sym}}$ by considering the cases $\alpha < \theta$ and $\alpha > \theta$, separately. If $\alpha < \theta$, the two transmitted codewords overlap and we have $(\omega^-_{1,1}, \omega^+_{1,1}, \omega_{1,1}) = (0, 1, 0)$ and $(\omega^-_{2,1}, \omega^+_{2,1}, \omega_{2,1}) = (1, 0, 0)$. Applying Proposition 14, $R_c \in \mathcal{P}(1 - 1, \nu_1 - \nu_2; \gamma) = \mathcal{P}(0, -\alpha; \gamma)$. If $\alpha > \theta$, none of the transmitted codewords experiences interference and hence, $R_c \in \mathcal{P}(0, -\theta; \gamma)$. Therefore,

$$\mathcal{R}_{\text{sym}} = \begin{cases} \bigcup_{\gamma \geq 0} \mathcal{P}(0, -\alpha; \gamma) & \alpha < \theta \\ \bigcup_{\gamma \geq 0} \mathcal{P}(0, -\theta; \gamma) & \alpha > \theta \end{cases}. \tag{3.118}$$

Define

$$\overline{R}_c := \overline{R}_{c,1} = \overline{R}_{c,2}, \tag{3.119}$$

where $\overline{R}_{c,i}$ is given in (3.105) and let γ^* be the unique positive solution for γ in the equation $\kappa'_\gamma + \frac{\alpha}{\theta}(\kappa_\gamma - \kappa'_\gamma) = \big(\frac{\gamma}{P} - 1\big)\lambda$. Then it is easy to see that $\mathcal{R}_{\text{sym}}\big|_{N=1} = \big(0, R_{c,\max}\big)$ where $R_{c,\max}$ is given by

$$R_{c,\max} = \begin{cases} \big(\frac{\gamma^*}{P} - \frac{1}{N}\big)\lambda & \alpha < \theta \\ \overline{R}_c & \alpha > \theta \end{cases}. \tag{3.120}$$

For $N \geq 2$, it is not necessarily the case that $\mathcal{R}_{\text{sym}} = \big(\lambda, R_{\max}\big)$. For example, consider the setup in panel (a) of Fig. 3.18 where the line $R_{c,1} = R_{c,2}$ is shown in red. We see that $\mathcal{R}_{\text{sym}}$ is the union of two disjoint intervals.

Throughout the rest of this section let $N \geq 2$. Our goal is to characterize $\mathcal{R}_{\text{sym}}$. Define[14]

$$\mu := \frac{\theta R_c}{\lambda}. \tag{3.121}$$

Recall from Section 3.4 that the burst with index j of Tx i extends from $j\mu+\nu_i$ to $j\mu+\nu_i+\theta$ on the $\bar{t}$-axis. Let $j^* \geq 1$ be such that

$$j^*\mu + \nu_1 < \mu + \nu_2 < (j^*+1)\mu + \nu_1, \tag{3.122}$$

or equivalently,[15]

$$\frac{\alpha}{j^*} < \mu < \frac{\alpha}{j^*-1}. \tag{3.123}$$

i.e., the starting point of the first burst of Tx 2 lies between the starting points of the bursts with indices j^* and j^*+1 of Tx 1 as shown in Fig. 3.19. The interference pattern on the transmitted codewords depends on how the numbers μ, $\frac{\alpha-\theta}{j^*-1}$ and $\frac{\alpha+\theta}{j^*}$ compare with each other. For example, Fig. 3.19 shows the case where $\frac{\alpha-\theta}{j^*-1} < \mu < \frac{\alpha+\theta}{j^*}$. As a result, each codeword of Tx 1 with index $j \geq j^*+1$ experiences interference at both ends, the codeword of Tx 1 with index j^* experiences interference only at its right end and any codeword of Tx 1 with index $j \leq j^*-1$ does not experience any interference. In general, it is easy to see that

$$\omega_{1,j}^{-} = \begin{cases} 0 & j \leq j^* \\ j - j^* & \mu < \frac{\alpha+\theta}{j^*},\ j \geq j^*+1 \\ 0 & \mu > \frac{\alpha+\theta}{j^*},\ j \geq j^*+1 \end{cases}, \tag{3.124}$$

$$\omega_{1,j}^{+} = \begin{cases} 0 & j \leq j^*-1 \\ j - j^* + 1 & \mu > \frac{\alpha-\theta}{j^*-1},\ j \geq j^* \\ 0 & \mu < \frac{\alpha-\theta}{j^*-1},\ j \geq j^* \end{cases}, \tag{3.125}$$

[14]In Section 3.2.2 we defined $\mu_i := \frac{\eta_i}{\lambda_i}$ in (3.12). Since η_i is replaced by $\theta_i R_i$ in our new system model in this section, the choice of the letter μ for the quotient $\frac{\theta R}{\lambda}$ in (3.121) is in accordance with the one in (3.12).

[15]If $j^* = 1$, we drop the upper bound in (3.123).

$$\omega_{2,j}^{-} = \begin{cases} j + j^* - 1 & \mu > \frac{\alpha-\theta}{j^*-1},\ j \leq N - j^* + 1 \\ 0 & \mu < \frac{\alpha-\theta}{j^*-1},\ j \leq N - j^* + 1 \\ 0 & j \geq N - j^* + 2 \end{cases} \tag{3.126}$$

and

$$\omega_{2,j}^{+} = \begin{cases} j + j^* & \mu < \frac{\alpha+\theta}{j^*},\ j \leq N - j^* \\ 0 & \mu > \frac{\alpha+\theta}{j^*},\ j \leq N - j^* \\ 0 & j \geq N - j^* + 1 \end{cases}. \tag{3.127}$$

In view of the interference pattern described in (3.124) to (3.127) and considering the constraints in (3.123), we define the four disjoint sets

$$\begin{aligned} \mathscr{A}_{j^*} &:= \left\{ R_c > \lambda : \frac{\alpha-\theta}{j^*-1} < \mu < \frac{\alpha+\theta}{j^*},\ \frac{\alpha}{j^*} < \mu < \frac{\alpha}{j^*-1} \right\} \\ &= \left\{ R_c > \lambda : \max\left\{ \frac{\alpha-\theta}{j^*-1}, \frac{\alpha}{j^*} \right\} < \mu < \min\left\{ \frac{\alpha+\theta}{j^*}, \frac{\alpha}{j^*-1} \right\} \right\}, \end{aligned} \tag{3.128}$$

$$\begin{aligned} \mathscr{B}_{j^*} &:= \left\{ R_c > \lambda : \mu < \min\left\{ \frac{\alpha-\theta}{j^*-1}, \frac{\alpha+\theta}{j^*} \right\},\ \frac{\alpha}{j^*} < \mu < \frac{\alpha}{j^*-1} \right\} \\ &= \left\{ R_c > \lambda : \frac{\alpha}{j^*} < \mu < \min\left\{ \frac{\alpha-\theta}{j^*-1}, \frac{\alpha+\theta}{j^*} \right\} \right\}, \end{aligned} \tag{3.129}$$

$$\begin{aligned} \mathscr{C}_{j^*} &:= \left\{ R_c > \lambda : \mu > \max\left\{ \frac{\alpha-\theta}{j^*-1}, \frac{\alpha+\theta}{j^*} \right\},\ \frac{\alpha}{j^*} < \mu < \frac{\alpha}{j^*-1} \right\} \\ &= \left\{ R_c > \lambda : \max\left\{ \frac{\alpha-\theta}{j^*-1}, \frac{\alpha+\theta}{j^*} \right\} < \mu < \frac{\alpha}{j^*-1} \right\} \end{aligned} \tag{3.130}$$

and

$$\begin{aligned} \mathscr{D}_{j^*} &:= \left\{ R_c > \lambda : \frac{\alpha+\theta}{j^*} < \mu < \frac{\alpha-\theta}{j^*-1},\ \frac{\alpha}{j^*} < \mu < \frac{\alpha}{j^*-1} \right\} \\ &= \left\{ R_c > \lambda : \frac{\alpha+\theta}{j^*} < \mu < \frac{\alpha-\theta}{j^*-1} \right\}. \end{aligned} \tag{3.131}$$

Next, we explicitly compute the sets $\mathcal{R}_{\text{sym}} \bigcap \mathscr{A}_{j^*}$, $\mathcal{R}_{\text{sym}} \bigcap \mathscr{B}_{j^*}$, $\mathcal{R}_{\text{sym}} \bigcap \mathscr{C}_{j^*}$ and $\mathcal{R}_{\text{sym}} \bigcap \mathscr{D}_{j^*}$. We will frequently invoke Proposition 14 without specific mention.

- **Computing** $\mathcal{R}_{\text{sym}} \bigcap \mathscr{A}_{j^*}$:

  1. Conditions for successful decoding at Rx 1
     - Any codeword of Tx 1 with index $j \leq j^* - 1$ does not experience any interference.
     - The codeword of Tx 1 with index j^* experiences interference only at its right end and $\omega^+_{1,j^*} = 1$. We require $R_c \in \mathcal{P}(1 - j^*, \nu_1 - \nu_2; \gamma) = \mathcal{P}(1 - j^*, -\alpha; \gamma)$.
     - Any codeword of Tx 1 with index $j^* + 1 \leq j \leq N$ experiences interference at both ends. We require $R_c \in \mathcal{P}(1, \theta; \gamma)$.
  2. Conditions for successful decoding at Rx 2
     - Any codeword of Tx 2 with index $1 \leq j \leq N - j^*$ experiences interference at both ends. We require $R_c \in \mathcal{P}(1, \theta; \gamma)$.
     - The codeword of Tx 2 with index $j = N - j^* + 1$ experiences interference only at its left end and $\omega^-_{2,N-j^*+1} = N$. We require $R_c \in \mathcal{P}\big((N - j^* + 1) - N, \nu_1 - \nu_2; \gamma\big) = \mathcal{P}(1 - j^*, -\alpha; \gamma)$.
     - Any codeword of Tx 2 with index $N - j^* + 2 \leq j \leq N$ does not experience any interference.

  It follows that

$$\mathcal{R}_{\text{sym}} \bigcap \mathscr{A}_{j^*} = \begin{cases} \bigcup_{\gamma \geq 0} \big(\mathcal{P}(1, \theta; \gamma) \bigcap \mathcal{P}(1 - j^*, -\alpha; \gamma)\big) \bigcap \mathscr{A}_{j^*} & 1 \leq j^* \leq N - 1 \\ \bigcup_{\gamma \geq 0} \mathcal{P}(1 - N, -\alpha; \gamma) \bigcap \mathscr{A}_N & j^* = N \\ \bigcup_{\gamma \geq 0} \mathcal{P}(0, -\theta; \gamma) \bigcap \mathscr{A}_{j^*} & j^* \geq N + 1 \end{cases}. \tag{3.132}$$

- **Computing** $\mathcal{R}_{\text{sym}} \bigcap \mathscr{B}_{j^*}$:

  1. Conditions for successful decoding at Rx 1
     - Any codeword of Tx 1 with index $j \leq j^*$ does not experience any interference.
     - Any codeword of Tx 1 with index $j^* + 1 \leq j \leq N$ experiences interference only at its left end and $\omega^-_{1,j} = j - j^*$. We require $R_c \in \mathcal{P}\big(j - (j - j^*), \nu_2 - \nu_1; \gamma\big) = \mathcal{P}(j^*, \alpha; \gamma)$.
  2. Conditions for successful decoding at Rx 2

- Any codeword of Tx 2 with index $1 \le j \le N - j^*$ experiences interference only at its right end and $\omega^+_{2,j} = j + j^*$. We require $R_c \in \mathcal{P}\big((j+j^*) - j, \nu_2 - \nu_1; \gamma\big) = \mathcal{P}(j^*, \alpha; \gamma)$.
- Any codeword of Tx 2 with index $N - j^* + 1 \le j \le N$ does not experience any interference.

It follows that

$$\mathcal{R}_{\text{sym}} \bigcap \mathscr{B}_{j^*} = \begin{cases} \bigcup_{\gamma \ge 0} \mathcal{P}(j^*, \alpha; \gamma) \bigcap \mathscr{B}_{j^*} & 1 \le j^* \le N-1 \\ \bigcup_{\gamma \ge 0} \mathcal{P}(0, -\theta; \gamma) \bigcap \mathscr{B}_{j^*} & j^* \ge N \end{cases}. \tag{3.133}$$

- **Computing** $\mathcal{R}_{\text{sym}} \bigcap \mathscr{C}_{j^*}$:

  1. Conditions for successful decoding at Rx 1
     - Any codeword of Tx 1 with index $j \le j^* - 1$ does not experience any interference.
     - Any codeword of Tx 1 with index $j^* \le j \le N$ experiences interference only at its right end and $\omega^+_{1,j} = j - j^* + 1$. We require $R_c \in \mathcal{P}\big((j - j^* + 1) - j, \nu_1 - \nu_2; \gamma\big) = \mathcal{P}(1 - j^*, -\alpha; \gamma)$.
  2. Conditions for successful decoding at Rx 2
     - Any codeword of Tx 2 with index $1 \le j \le N - j^* + 1$ experiences interference only at its left end and $\omega^-_{2,j} = j + j^* - 1$. We require $R_c \in \mathcal{P}\big(j - (j + j^* - 1), \nu_1 - \nu_2; \gamma\big) = \mathcal{P}(1 - j^*, -\alpha; \gamma)$.
     - Any codeword of Tx 2 with index $N - j^* + 2 \le j \le N$ does not experience any interference.

  It follows that

$$\mathcal{R}_{\text{sym}} \bigcap \mathscr{C}_{j^*} = \begin{cases} \bigcup_{\gamma \ge 0} \mathcal{P}(1 - j^*, -\alpha; \gamma) \bigcap \mathscr{C}_{j^*} & 1 \le j^* \le N \\ \bigcup_{\gamma \ge 0} \mathcal{P}(0, -\theta; \gamma) \bigcap \mathscr{C}_{j^*} & j^* \ge N+1 \end{cases}. \tag{3.134}$$

- **Computing** $\mathcal{R}_{\text{sym}} \bigcap \mathscr{D}_{j^*}$: In this case, any codeword sent by Tx 1 or Tx 2 is received in the absence of interference. Hence,

$$\mathcal{R}_{\text{sym}} \bigcap \mathscr{D}_{j^*} = \bigcup_{\gamma \ge 0} \mathcal{P}(0, -\theta; \gamma) \bigcap \mathscr{D}_{j^*}, \quad j^* \ge 1. \tag{3.135}$$

In general, one can characterize $\mathcal{R}_{\text{sym}}$ for $N \ge 2$ by taking the following steps:

1. Write

$$\mathcal{R}_{\text{sym}} = \bigcup_{j^*=1}^{\infty} \mathcal{R}_{j^*}, \tag{3.136}$$

where

$$\mathcal{R}_{j^*} = \mathcal{R}_{\text{sym}} \bigcap \big(\mathscr{A}_{j^*} \bigcup \mathscr{B}_{j^*} \bigcup \mathscr{C}_{j^*} \bigcup \mathscr{D}_{j^*}\big). \tag{3.137}$$

2. Use (3.132), (3.133), (3.134) and (3.135) to describe $\mathcal{R}_{j^*}$ for any $j^* \geq 1$.

If $\alpha < \theta$,

$$\mathscr{A}_{j^*} = \mathscr{B}_{j^*} = \mathscr{C}_{j^*} = \mathscr{D}_{j^*} = \emptyset, \tag{3.138}$$

for any $j^* \geq 2$, i.e., $\mathcal{R}_{\text{sym}} = \mathcal{R}_1$. The following proposition characterizes $\mathcal{R}_{\text{sym}}$ provided that $\alpha < \theta$.

Proposition 15. *Assume $\alpha < \theta$. Let γ_0, γ_1 and γ_2 be the solutions for γ in $2\kappa'_\gamma = \kappa_\gamma$, $\kappa'_\gamma = \lambda$ and $\kappa'_\gamma + \frac{\alpha}{\theta}(\kappa_\gamma - \kappa'_\gamma) = \lambda$, respectively. If γ_1 does not exist, let $\gamma_1 = \infty$.*

- *If $\lambda < \kappa'_{\gamma_0}$, define*

$$f(\gamma) := \begin{cases} 0 & \gamma \leq \gamma_1 \\ \lambda & \gamma > \gamma_1 \end{cases}, \qquad g(\gamma) = \begin{cases} 0 & \gamma \leq \gamma_1 \\ \frac{2\kappa'_\gamma - \kappa_\gamma}{1 - \frac{1}{\lambda}(\kappa_\gamma - \kappa'_\gamma)} & \gamma_1 < \gamma \leq \gamma_2 \\ \kappa'_\gamma + \frac{\alpha}{\theta}(\kappa_\gamma - \kappa'_\gamma) & \gamma > \gamma_2 \end{cases}. \tag{3.139}$$

- *If $\lambda \geq \kappa'_{\gamma_0}$, define*

$$f(\gamma) := \begin{cases} 0 & \gamma \leq \gamma_2 \\ \frac{2\kappa'_\gamma - \kappa_\gamma}{1 - \frac{1}{\lambda}(\kappa_\gamma - \kappa'_\gamma)} & \gamma_2 < \gamma \leq \gamma_1 \\ \lambda & \gamma > \gamma_1 \end{cases}, \qquad g(\gamma) = \begin{cases} 0 & \gamma \leq \gamma_2 \\ \kappa'_\gamma + \frac{\alpha}{\theta}(\kappa_\gamma - \kappa'_\gamma) & \gamma > \gamma_2 \end{cases} \tag{3.140}$$

Then

$$\mathcal{R}_{\text{sym}} = \bigcup_{\gamma \geq 0} \Big(\max\Big\{f(\gamma), (\frac{\gamma}{P} - \frac{1}{N})\lambda\Big\}, g(\gamma)\Big). \tag{3.141}$$

Proof. See Appendix J. □

A few remarks are in order:

- Explicit expressions for γ_0 and γ_1 are

$$\gamma_0 = \frac{1+\sqrt{1+4a^2}}{2a^2}, \quad \gamma_1 = \frac{2^{2\lambda}-1}{1-a(2^{2\lambda}-1)}. \tag{3.142}$$

  There is no closed-form expression for γ_2 and it must be computed numerically.

- If P is sufficiently large, e.g., $P > N\max\{\gamma_1, \gamma_2\}$, it is easy to see that $\mathcal{R}_{\text{sym}} = (\lambda, R_{c,\max})$ where $R_{c,\max}$ is given in (3.120). For "smaller" values of P, $\inf_{R_c \in \mathcal{R}_{\text{sym}}} R_c$ can be larger than λ as we will see in the example in below.

- In (3.141), $\mathcal{R}_{\text{sym}}$ is given as the union of uncountably many intervals. It is more convenient to represent $\mathcal{R}_{\text{sym}}$ as follows. Define the regions Ω', Ω'' and Ω by

$$\Omega' := \big\{(\gamma, R_c) : f(\gamma) < R_c < g(\gamma)\big\}, \tag{3.143}$$

$$\Omega'' := \Big\{(\gamma, R_c) : R_c \geq \big(\frac{\gamma}{P} - \frac{1}{N}\big)\lambda\Big\} \tag{3.144}$$

  and

$$\Omega := \Omega' \bigcap \Omega'', \tag{3.145}$$

  i.e., Ω' is the set of all (γ, R_c) such that the $2N$ transmitted codewords are sent immediately and decoded successfully at the receivers and Ω'' is the set of all (γ, R_c) such that the average power constraint in (3.101) holds. Then

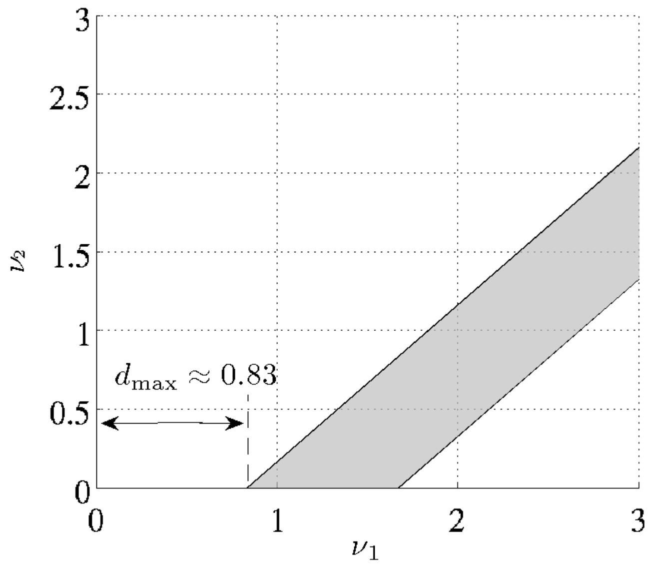
$$\mathcal{R}_{\text{sym}} = \Pi(\Omega) \tag{3.146}$$

  where the map $\Pi(\gamma, R_c) = R_c$ is the projection on the R_c-axis.

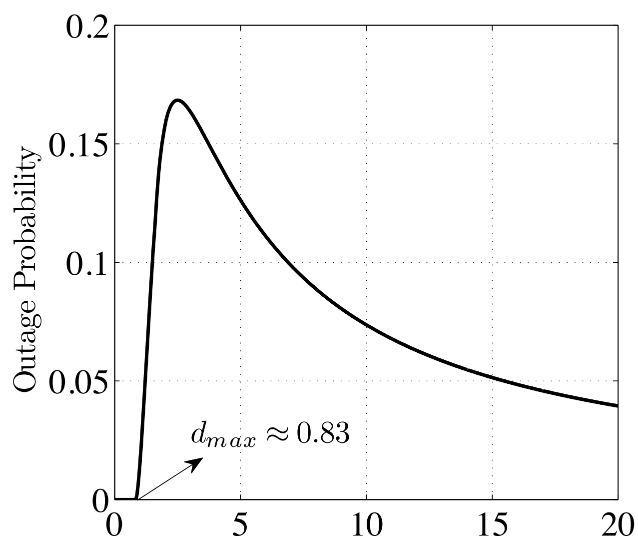
Example- Let $N = 4$, $\theta = 1$, $a = 0.5$ and $\alpha = 0.5$. Then $\gamma_0 \approx 6.84\,\text{dB}$ and $\kappa'_{\gamma_0} \approx 0.6358$. We consider two cases:

- If $\lambda < 0.9\kappa'_{\gamma_0} = 0.5722$, then $f(\gamma)$ and $g(\gamma)$ are given by (3.139). Assuming $P \approx 3.617\,\text{dB}$, Fig. 3.20 in panel (a) shows $f(\gamma)$, $g(\gamma)$ and $(\frac{\gamma}{P} - \frac{1}{N})\lambda$ as functions of γ. The set $\mathcal{R}_{\text{sym}}$ is shown by a red strip as the projection of the region Ω on the R_c-axis.

- If $\lambda \geq 0.9\kappa'_{\gamma_0} = 0.5722$, then $f(\gamma)$ and $g(\gamma)$ are given by (3.140). Assuming $P = 10\,\text{dB}$, Fig. 3.20 in panel (b) shows $f(\gamma)$, $g(\gamma)$ and $(\frac{\gamma}{P} - \frac{1}{N})\lambda$ as functions of γ. The set $\mathcal{R}_{\text{sym}}$ is shown by a red strip as the projection of the region Ω on the R_c-axis.

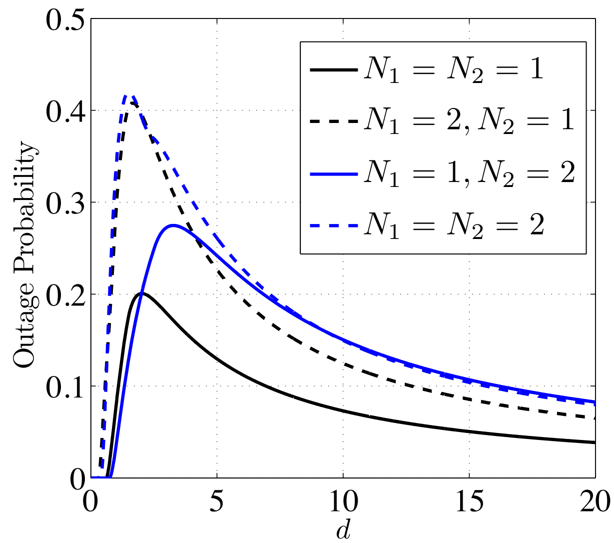


(a)

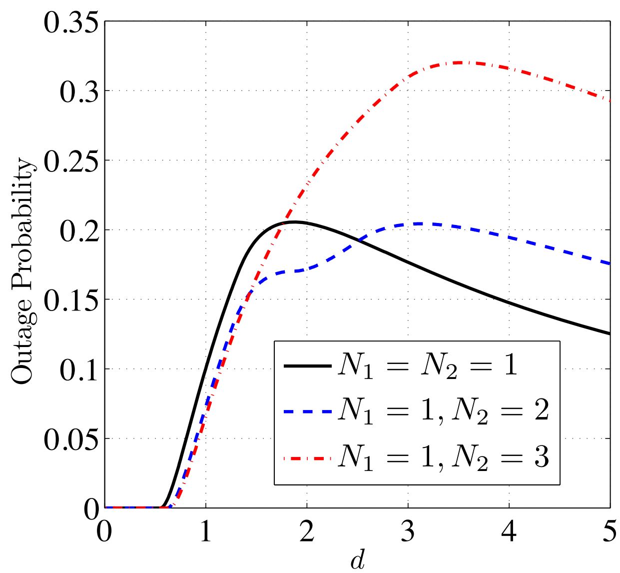


(b)

Figure 3.11: Let $k_1 = 3, k_2 = 2, q_1 = 0.3, q_2 = 0.4, a_1 = 0.5, a_2 = 0.7$, $P_1 = P_2 = 30\,\text{dB}$, $R_1 = 0.5\lambda_1 = 0.45$ and $R_2 = 0.5\lambda_2 = 0.4$. Then $\mathcal{N}_{R_1,R_2} = \{(1,1)\}$ and (3.83) is satisfied. Fixing $N_1 = N_2 = 1$, panel (a) shows the region $\bigcup_{S\in\mathscr{S}}(\mathcal{A}_S^{(\text{geom})} \setminus \mathcal{A}_S^{(\text{rel})})$ in grey shade. Panel (b) shows the probability of outage in terms of d.

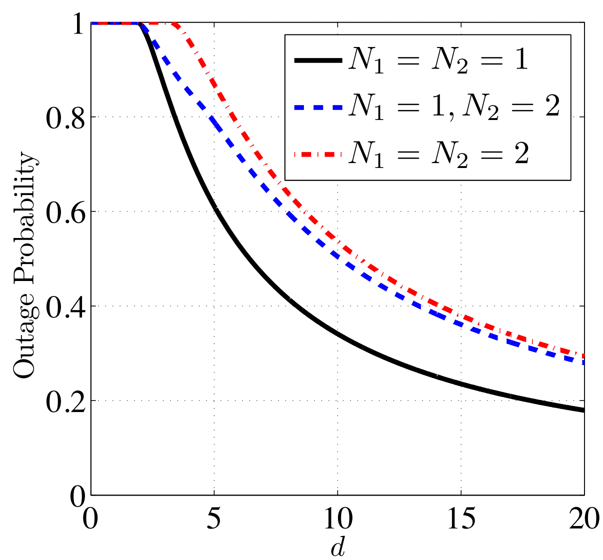


(a) $R_1 = 0.7\lambda_1, R_2 = 0.7\lambda_2$

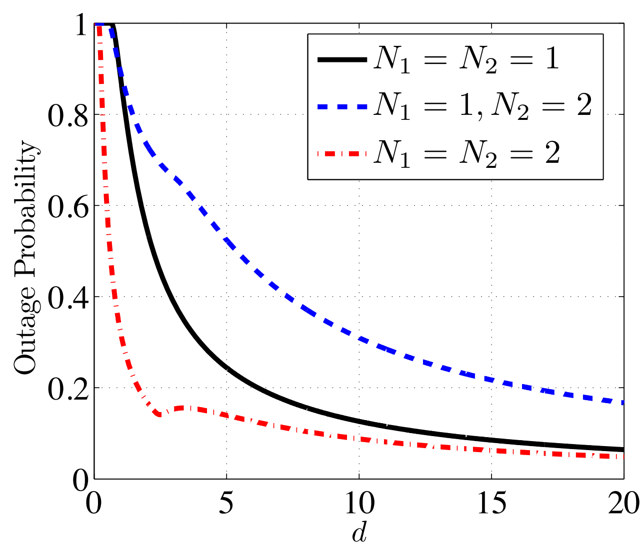


(b) $R_1 = 0.8\lambda_1, R_2 = 0.8\lambda_2$

Figure 3.12: Let $k_1 = 3, k_2 = 2, q_1 = 0.3, q_2 = 0.4, a_1 = 0.5, a_2 = 0.7$ and $P_1 = P_2 = 30\,\text{dB}$. If $R_1 = 0.7\lambda_1 = 0.63$ and $R_2 = 0.7\lambda_2 = 0.56$, then $\mathcal{N}_{R_1,R_2} = \{(m_1, m_2) : 1 \leq m_1, m_2 \leq 2\}$ and (3.83) is satisfied. Panel (a) shows the probability of outage in terms of d for different values of (N_1, N_2). If $R_1 = 0.8\lambda_1 = 0.72$ and $R_2 = 0.8\lambda_2 = 0.64$, then $\mathcal{N}_{R_1,R_2} = \{(m_1, m_2) : 1 \leq m_1, m_2 \leq 3\}$ and (3.83) is satisfied. It turns out that depending on the value of d, the best choices are $(N_1, N_2) = (1, 1), (1, 2)$ or $(1, 3)$ as shown in panel (b).



(a) $P_1 = P_2 = 10\,\mathrm{dB}$



(b) $P_1 = P_2 = 30\,\mathrm{dB}$

Figure 3.13: A symmetric scenario where $k_1 = k_2 = 5$, $q_1 = q_2 = 0.2$, $a_1 = a_2 = 0.5$, $R_1 = 0.7\lambda_1 = 0.7$ and $R_2 = 0.7\lambda_2 = 0.7$. Panel (a) presents the probability of outage in terms of d for different values of (N_1, N_2) for $P_1 = P_2 = 10\,\mathrm{dB}$. We see that $N_1 = N_2 = 1$ is the optimum choice for any value of d. Panel (b) presents the probability of outage in terms of d for different values of (N_1, N_2) for $P_1 = P_2 = 20\,\mathrm{dB}$. Here, $N_1 = N_2 = 2$ is the optimum choice for any value of d.

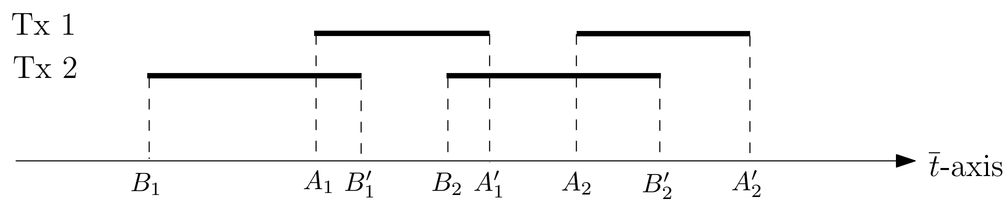


Figure 3.14: This picture shows the positions of different bursts on the $\bar{t}$-axis corresponding to the state $S = \{(1; 1, 2), (2; 2, 4)\}$ in a scenario where $N_1 = N_2 = 2$. The table in (3.106) shows the numbers on the $\bar{t}$-axis corresponding to different points on the $\bar{t}$-axis.

$A_m = \frac{m\theta_1 R_{c,1}}{\lambda_1} + \nu_1$	$B_m = \frac{m\theta_2 R_{c,2}}{\lambda_2} + \nu_2$
$A'_m = \frac{m\theta_1 R_{c,1}}{\lambda_1} + \nu_1 + \theta_1$	$B'_m = \frac{m\theta_2 R_{c,2}}{\lambda_2} + \nu_2 + \theta_2$

(3.106)

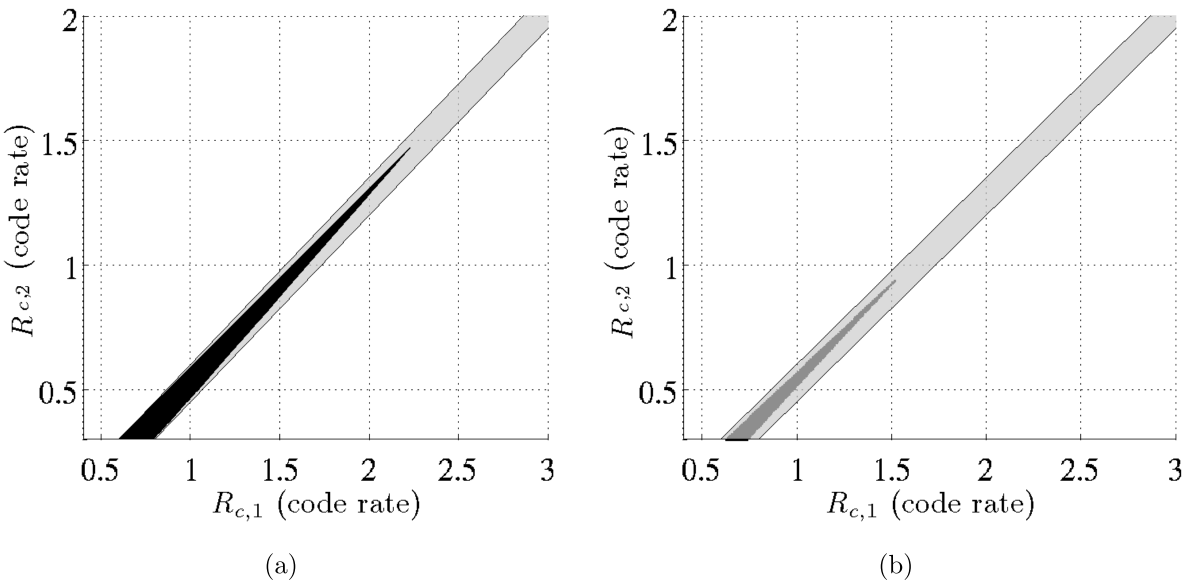


Figure 3.15: Consider a setup where $N_1 = 3, N_2 = 2, \theta_1 = \theta_2 = 1, k_1 = 2, k_2 = 3, q_1 = 0.2, q_2 = 0.1, a_1 = 1.5, a_2 = 0.5, P_1 = 20\,\mathrm{dB}, P_2 = 30\,\mathrm{dB}$ and $\alpha = 1$. Panel (a) shows the regions $\mathcal{R}_S^{(\mathrm{geom})}$ in grey and $\mathcal{R}_S^{(\mathrm{rel})}$ in black for $S = \{(1;2,3),(2;3,4)\}$. Panel (b) shows the same regions under full power transmission. It is seen that $\mathcal{R}_S^{(\mathrm{rel})}$ under full power transmission is strictly smaller than $\mathcal{R}_S^{(\mathrm{rel})}$ in its general form given in (3.110).

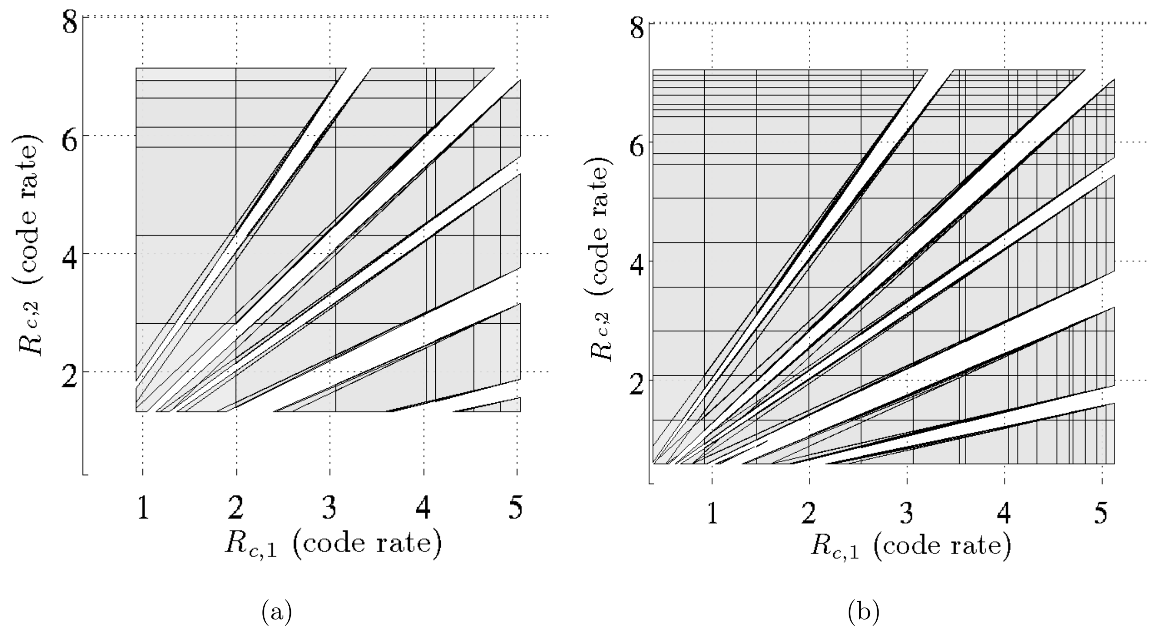


Figure 3.16: Consider a setup where $N_1 = 3, N_2 = 2, \theta_1 = \theta_2 = 1, k_1 = 2, k_2 = 3, q_1 = 0.2, q_2 = 0.1, a_1 = 1.5, a_2 = 0.5, P_1 = 20\,\text{dB}, P_2 = 30\,\text{dB}$ and $\alpha = 1$. Panel (a) shows the region $\widetilde{\mathcal{R}}_S^{(\text{rel})}$ in (3.112) where $S = \{(1; 2, 3), (2; 3, 4)\}$ and $\Gamma_i = \{\frac{l}{5}\overline{\gamma}_i : 1 \leq l \leq 4\}$. Panel (b) shows the same region for $\Gamma_i = \{\frac{l}{10}\overline{\gamma}_i : 1 \leq l \leq 9\}$.

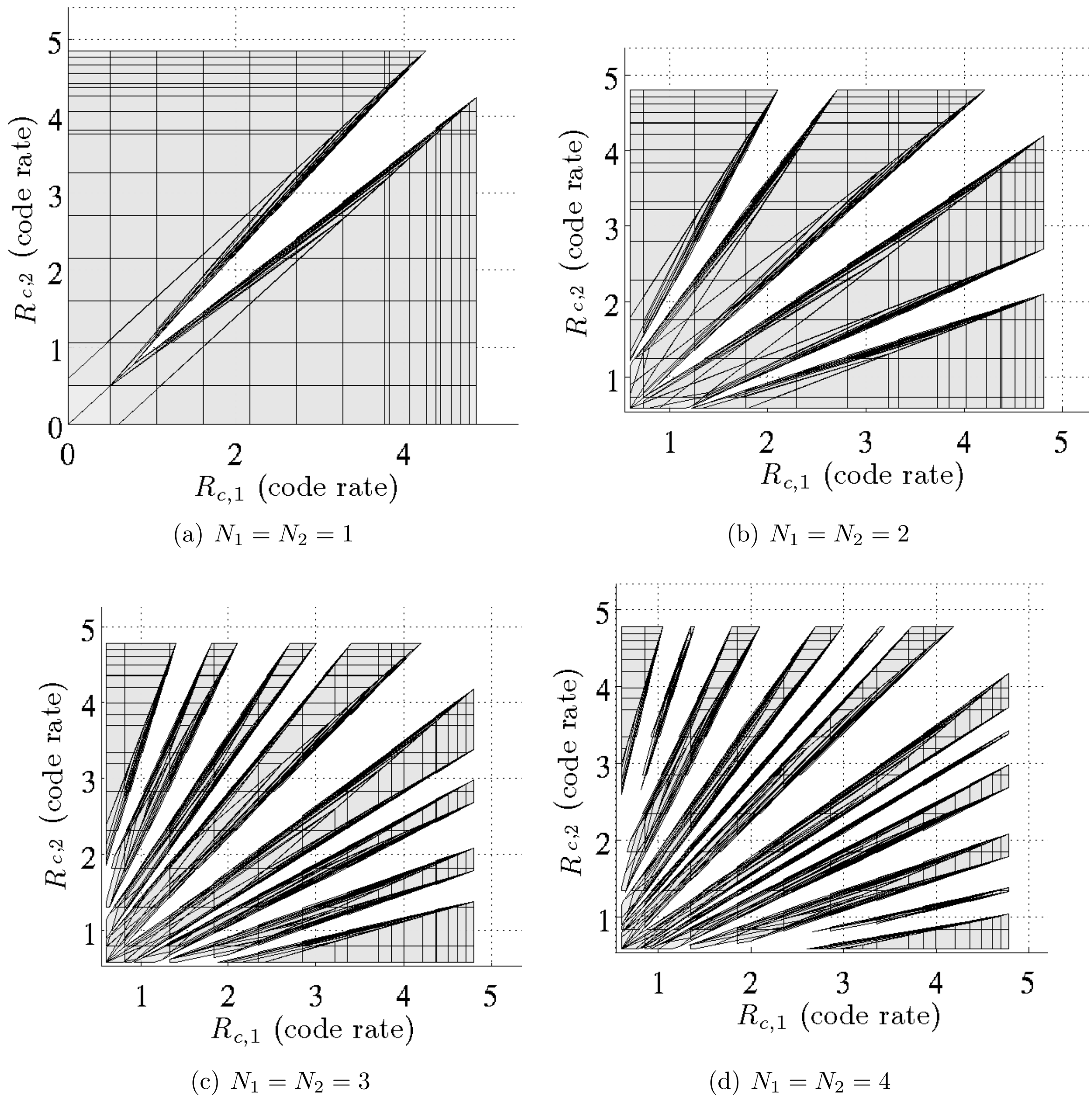
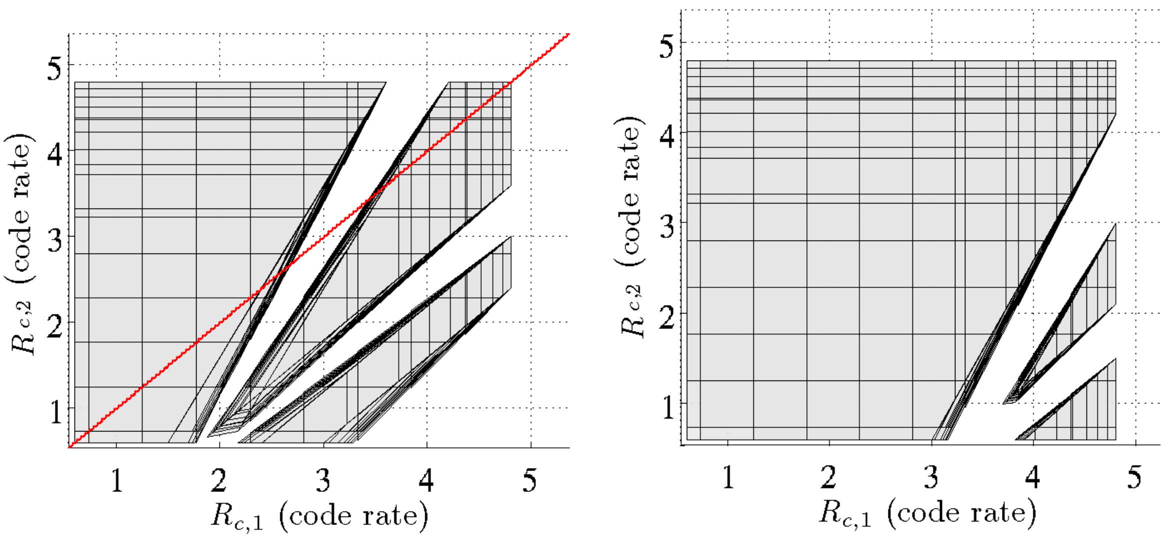


Figure 3.17: A setting where $\theta_1 = \theta_2 = 1, k_1 = k_2 = 2, q_1 = q_2 = 0.3, a_1 = a_2 = 0.5, P_1 = P_2 = 20\,\text{dB}$ and $\alpha = 0$. As the number of codewords $N_1 = N_2$ increases, $\mathcal{R}$ becomes strictly smaller.



(a) $\alpha = 5$ (b) $\alpha = 10$

Figure 3.18: A setting where $N_1 = N_2 = 2, \theta_1 = \theta_2 = 1, k_1 = k_2 = 2, q_1 = q_2 = 0.3, a_1 = a_2 = 0.5$ and $P_1 = P_2 = 20\,\text{dB}$. As α increases, the region $\mathcal{R}$ converges to the square $\{(R_{c,1}, R_{c,2}) : \lambda_i < R_{c,i} < \overline{R}_{c,i},\ i = 1,2\}$ where $\overline{R}_{c,1} = \overline{R}_{c,2} \approx 4.8774$.

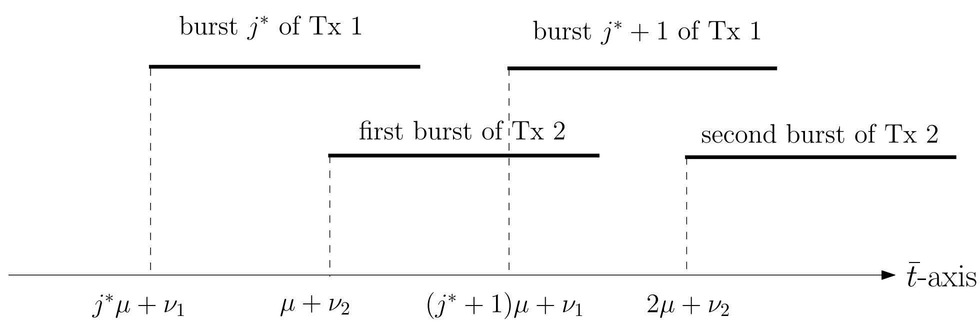
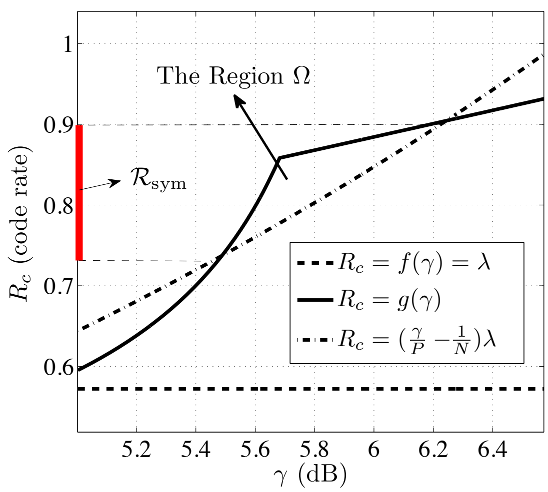
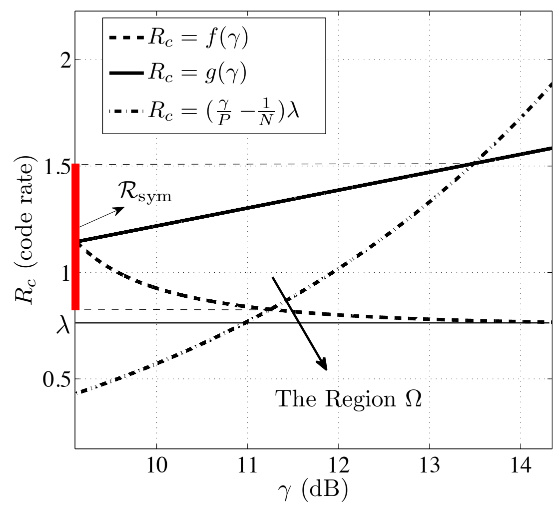


Figure 3.19: The integer $j^* \geq 1$ is such that the starting point of the first burst of Tx 2 lies between the bursts with indices j^* and $j^* + 1$ of Tx 1. The length of any burst is θ on the $\bar{t}$-axis. This picture shows the case where $(\mu + \nu_2) - (j^*\mu + \nu_1) < \theta$ and $((j^* + 1)\mu + \nu_1) - (\mu + \nu_2) < \theta$, or equivalently, $\frac{\alpha-\theta}{j^*-1} < \mu < \frac{\alpha+\theta}{j^*}$. This implies that each codeword of Tx 1 with index $j \geq j^*+1$ experiences interference at both ends, the codeword with index j^* experiences interference only at its right end and any codeword with index $j \leq j^* - 1$ does not experience any interference.



(a)



(b)

Figure 3.20: Panel (a) shows a setting where $N = 4$, $\theta = 1$, $\lambda = 0.5722$, $a = 0.5$, $P \approx 3.617\,\text{dB}$ and $\alpha = 0.5$. In this case, $\gamma_0 \approx 6.84\,\text{dB}$ and $\lambda < \kappa'_{\gamma_0} \approx 0.6358$, i.e., $f(\gamma)$ and $g(\gamma)$ are given by (3.139). Panel (b) shows a scenario where $N = 4$, $\theta = 1$, $\lambda = 0.7629$, $a = 0.5$, $P = 10\,\text{dB}$ and $\alpha = 0.5$. In this case, $\gamma_0 \approx 6.84\,\text{dB}$ and $\lambda > \kappa'_{\gamma_0} \approx 0.6358$, i.e., $f(\gamma)$ and $g(\gamma)$ are given by (3.140). The red strip on the R_c-axis is $\mathcal{R}_{\text{sym}}$ as the projection of the region Ω on the R_c-axis. In both panel (a) and panel (b), $\inf_{R_c \in \mathcal{R}_{\text{sym}}} R_c > \lambda$.

Chapter 4

Conclusion

In the first half of the thesis we proposed a distributed learning scheme in a decentralized wireless network with asynchronous users and burst transmission. It was shown how each user estimates the locations and intensities of interference bursts along its transmitted codeword. The main tool used in the learning process was piecewise linear regression.

In the second part of the thesis we have studied a two-user GIC-SDA with immediate transmissions under two different settings. In one scenario, the information source at each transmitter turned off after generating a given total number of bits and the transmitters only knew the statistics of the mutual delay between their bit streams. The codebook rate at each transmitter was optimized in order to achieve a target average transmission rate and transmission power and maximize the probability of successful decoding at the receivers. In another scenario, the information sources were active indeterminately and the transmitters were aware of the exact mutual delay between their bit streams. We characterized an achievable rate region for the codebook rates assuming the receivers treat interference as noise. This region was given as a union of uncountably many polyhedrons which is in general disconnected and non-convex due to infeasibility of time sharing.

References

[1] R. H. Etkin, D. N. C. Tse and H. Wang, "Gaussian interference channel capacity to within one bit", *IEEE Trans. Inf. Theory*, vol. 54, no. 12, pp. 5534-5562, Dec. 2008.

[2] X. Shang, G. Kramer and B. Chen, "A new outer bound and the noisy-interference sum-rate capacity for Gaussian interference channels", *IEEE Trans. Inf. Theory*, vol. 55, no. 2, pp. 689-699, Feb. 2009.

[3] A. Motahari and A.K. Khandani, "Capacity bounds for the Gaussian interference channel", *IEEE Trans. Inf. Theory*, vol. 55, no. 2, pp. 620-643, Feb. 2009.

[4] V. S. Annapureddy and V. V. Veeravalli, "Gaussian Interference Networks: Sum capacity in the low interference regime and new outer bounds on the capacity region", *IEEE Trans. Inf. Theory*, vol. 55, no. 7, pp. 3032-3050, July 2009.

[5] C. M. Cordeiro and D. P. Agrawal, "Ad hoc and sensor networks- theory and applications", *World Sci. Pub. Company*, 2006.

[6] N. Abramson, "The throughput of packet broadcasting channels", *IEEE Trans. on Communicatons*, vol. 25, no. 1, pp. 117-128, Jan. 1977.

[7] S. Ghez, S. Verdú and S. C. Schwartz, "Stability properties of slotted Aloha with multipacket reception capability", *IEEE Trans. on Automatic Control*, vol. 33, no. 7, pp. 640-649, July 1988.

[8] S. Ghez, S. Verdú and S. C. Schwartz, "Optimal decentralized control in the random access multipacket channel", *IEEE Trans. on Automatic Control*, vol. 34, no. 11, pp. 1153-1163, Nov. 1989.

[9] B. Hajek and T. Loon, "Decentralized dynamic control of a multiaccess broadcast channel", *IEEE Trans. on Automatic Control*, vol. 27, no. 3, pp. 559-569, June 1982.

[10] J. Massey and P. Mathys, "The collision channel without feedback", *IEEE Trans. Inf. Theory*, vol. 31, no. 2, pp. 192-204, Feb. 1985.

[11] W. S. Wong, "New protocol sequences for random access channels without feedback", *IEEE Trans. Inf. Theory*, vol. 53, no. 6, pp. 2060-2071, June 2007.

[12] K. W. Shum, W. S. Wong, C. W. Sung and C. S. Chen, "Shift-invariant protocol sequences for the collision channel without feedback", *IEEE Trans. Inf. Theory*, vol. 55, no. 7, pp. 3312-3322, July 2009.

[13] R. Etkin, A. P. Parekh and D. Tse, "Spectrum sharing for unlicensed bands", *IEEE Journal on Selected Areas in Communications*, vol. 25, no. 3, pp. 517-528, April 2007.

[14] T. M. Cover, R. J. Mceliece and E.C. Posner, "Asynchronous multiple-access channel capacity", *IEEE Trans. Inf. Theory*, vol. 27, no. 4, pp. 409-413, July 1981.

[15] J. Y. N. Hui and P. A. Humblet, "The capacity region of the totally asynchronous multiple access channel", *IEEE Trans. Inf. Theory*, vol 31, no. 2, pp. 207-216, March 1985.

[16] H. Yagi and H. V. Poor, "Multi-level rate splitting for synchronous and asynchronous interference channels", *International Symp. Inf. Theory*, Saint Petersburg, Russia, July 2011.

[17] B. Rimoldi, "Generalized time sharing: a low-complexity capacity achieving multiple access technique", *IEEE Trans. Inf. Theory*, vol. 47, no. 6, pp. 2432-2442, June 2001.

[18] E. Calvo, J. R. Fonollosa and J. Vidal, "On the totally asynchronous interference channel with single-user receivers", *International Symp. Inf. Theory*, Seoul, Korea, June 2009.

[19] S. Verdú and T.S. Han, "A general formula for channel capacity", *IEEE Trans. Inf. Theory*, vol. 40, no. 7, pp. 1147-1157, July 1994.

[20] A. El Gamal and Y. H. Kim, "Network information theory", *Cambridge University Press*, 2011.

[21] V. Chandar, A. Tchamkerten and G. Wornell, "Optimal sequential frame synchronization", *IEEE Trans. on Inf. Theory*, vol. 54, no. 8, pp. 3725-3728, August 2008.

[22] B. Hellstrom, "GPS-free synchronization of Digital Terrestrial TV and Mobile TV distribution networks", *Net Insight AB*, Sweden, August 2007. (info@netinsight.net)

[23] D. L. Mills, "Network time protocol version 4 reference and implementation guide", *NTP Working Group*, Technical Report 06-6-1, University of Delaware, June 2006.

[24] K. Moshksar, A. Bayesteh and A. K. Khandani, "Randomized resource allocation in decentralized wireless networks", *IEEE Trans. Inf. Theory*, vol. 57, no. 4, pp. 2115-2142, April 2011.

[25] K. Moshksar and A. K. Khandani, "Decentralized wireless networks with asynchronous users and burst transmissions", *IEEE Trans. on Inf. Theory*, vol. 61, no. 7, pp. 3851-3881, July 2015.

[26] K. Moshksar and A. K. Khandani, "Resource management in interference channels with asynchronous users", *International Symp. Inf. Theory*, ISIT 2009, Seoul, Korea, June 2009.

[27] K. Moshksar and A. K. Khandani, "Randomized on-off signalling for asynchronous interference channels", *Asilomar Conf. Signals Systems and Computers*, Monterey, USA, Nov. 2010.

[28] T. M. Cover and J. A. Thomas, "Elements of information theory", *John Wiley and Sons, Inc.*, 1991.

[29] R. M. Dudley, "Real analysis and probability", *Cambridge University Press*, 2002.

[30] A. Gutt, "Probability: A Graduate Course", *Springer-Verlag*, 2013.

[31] S. Boucheron, G. Lugosi and P. Massart, "Concentration inequalities: A nonasymptotic theory of independence", *Oxford University Press*, 2013.

[32] M. Mitzenmacher and E. Upfal, "Probability and computing: Randomized algorithms and probabilistic analysis", *Cambridge University Press*, 2005.

APPENDICES

Appendix A

Proof of Proposition 1

Let $\tau_n = \frac{\lceil \tau T_n \rceil}{T_n}$ and $\tau_n' = \frac{\lceil \tau T_n \rceil - 1}{T_n}$. We have

$$\begin{aligned} |\boldsymbol{F}_n(\tau) - F(\tau)| &\leq |\boldsymbol{F}_n(\tau_n) - \boldsymbol{F}_n(\tau)| + |\boldsymbol{F}_n(\tau_n) - F(\tau_n)| + |F(\tau_n) - F(\tau)| \\ &\leq |\boldsymbol{F}_n(\tau_n) - \boldsymbol{F}_n(\tau_n')| + |\boldsymbol{F}_n(\tau_n) - F(\tau_n)| + |F(\tau_n) - F(\tau)|, \quad (A.1) \end{aligned}$$

where in the last step we use the fact that $\boldsymbol{F}_n$ is nondecreasing and write $|\boldsymbol{F}_n(\tau_n) - \boldsymbol{F}_n(\tau)| \leq |\boldsymbol{F}_n(\tau_n) - \boldsymbol{F}_n(\tau_n')|$. Moreover,

$$|\boldsymbol{F}_n(\tau_n) - \boldsymbol{F}_n(\tau_n')| \leq |\boldsymbol{F}_n(\tau_n) - F(\tau_n)| + |\boldsymbol{F}_n(\tau_n') - F(\tau_n')| + |F(\tau_n) - F(\tau_n')|. \quad (A.2)$$

By (A.1) and (A.2),

$$\begin{aligned} |\boldsymbol{F}_n(\tau) - F(\tau)| &\leq 2|\boldsymbol{F}_n(\tau_n) - F(\tau_n)| + |\boldsymbol{F}_n(\tau_n') - F(\tau_n')| + |F(\tau_n) - F(\tau_n')| \\ &\quad + |F(\tau_n) - F(\tau)| \\ &\leq 2|\boldsymbol{F}_n(\tau_n) - F(\tau_n)| + |\boldsymbol{F}_n(\tau_n') - F(\tau_n')| + 2|F(\tau_n) - F(\tau_n')|, \end{aligned} \quad (A.3)$$

where the last step is due to the fact that F is increasing. Hence,

$$\|\boldsymbol{F}_n(\tau) - F(\tau)\|_\infty \leq 3 \sup_{\tau \in [0,1]} |\boldsymbol{F}_n(\tau_n) - F(\tau_n)| + 2 \sup_{\tau \in [0,1]} |F(\tau_n) - F(\tau_n')|. \quad (A.4)$$

In order to show $\lim_{n \to \infty} \|\boldsymbol{F}_n(\tau) - F(\tau)\|_\infty = 0$, we use the fact that a sequence of real-valued random variables $(\boldsymbol{a}_m)_{m \geq 1}$ tends to zero if $\sum_{m=1}^{\infty} \mathbb{P}(|\boldsymbol{a}_m| \geq \epsilon) < \infty$ for any $\epsilon > 0$.

This is a direct consequence of the Borel-Cantelli lemma [29]. By (A.4),

$$\begin{aligned}\mathbb{P}(\|\boldsymbol{F}_n(\tau)-F(\tau)\|_\infty \geq \epsilon) &\leq \mathbb{P}\Big(3\sup_{\tau\in[0,1]}|\boldsymbol{F}_n(\tau_n)-F(\tau_n)|+2\sup_{\tau\in[0,1]}|F(\tau_n)-F(\tau_n')|\geq\epsilon\Big)\\ &\leq \mathbb{P}\Big(\sup_{\tau\in[0,1]}|\boldsymbol{F}_n(\tau_n)-F(\tau_n)|\geq\frac{\epsilon}{5}\Big)+\mathbb{1}_{\sup_{\tau\in[0,1]}|F(\tau_n)-F(\tau_n')|\geq\frac{\epsilon}{5}},\end{aligned} \tag{A.5}$$

where the last step is due to the fact that if $3a+2b\geq\epsilon$ for two real numbers a and b, then at least one of a and b must be greater than or equal to $\frac{\epsilon}{5}$. The function F is piecewise linear and its slope never exceeds $\frac{1}{P}+\sum_{i=0}^K|h_i|^2$. Since $|\tau_n-\tau_n'|=\frac{1}{T_n}$, one can guarantee $\sup_{\tau\in[0,1]}|F(\tau_n)-F(\tau_n')|<\frac{\epsilon}{5}$ if $\frac{\epsilon T_n}{5}>\frac{1}{P}+\sum_{i=0}^K|h_i|^2$. Under this constraint and by (A.5),

$$\mathbb{P}(\|\boldsymbol{F}_n(\tau)-F(\tau)\|_\infty\geq\epsilon)\leq\mathbb{P}\Big(\sup_{\tau\in[0,1]}|\boldsymbol{F}_n(\tau_n)-F(\tau_n)|\geq\frac{\epsilon}{5}\Big). \tag{A.6}$$

We have

$$\boldsymbol{F}_n(\tau_n)=\frac{1}{PT_n}\sum_{t=1}^{\lceil\tau T_n\rceil}|\boldsymbol{y}_0[t]|^2 \tag{A.7}$$

$$\begin{aligned}&=\frac{1}{PT_n}\sum_{t=1}^{\lceil\tau T_n\rceil}|\boldsymbol{z}_0[t]|^2+\sum_{i=0}^{K}\frac{|h_i|^2}{PT_n}\sum_{t=1}^{\lceil\tau T_n\rceil}|\boldsymbol{s}_i[t]|^2\\ &\quad+2\sum_{i=0}^{K}\frac{1}{PT_n}\sum_{t=1}^{\lceil\tau T_n\rceil}\mathrm{Re}\left(h_i\boldsymbol{s}_i[t]\boldsymbol{z}_0^*[t]\right)\\ &\quad+2\sum_{0\leq i<j\leq K}\frac{1}{PT_n}\sum_{t=1}^{\lceil\tau T_n\rceil}\mathrm{Re}\left(h_ih_j^*\boldsymbol{s}_i[t]\boldsymbol{s}_j^*[t]\right).\end{aligned} \tag{A.8}$$

Recalling the definition of F in (2.6), we get

$$\begin{aligned}\sup_{\tau\in[0,1]}|\boldsymbol{F}_n(\tau_n)-F(\tau_n)| \;\leq\; & \frac{1}{P}\sup_{\tau\in[0,1]}\Big|\frac{1}{T_n}\sum_{t=1}^{\lceil\tau T_n\rceil}|\boldsymbol{z}_0[t]|^2-\tau_n\Big| \\ &+\sum_{i=0}^{K}|h_i|^2\sup_{\tau\in[0,1]}\Big|\frac{1}{PT_n}\sum_{t=1}^{\lceil\tau T_n\rceil}|\boldsymbol{s}_i[t]|^2-\psi\Big(\tau_n-\frac{\nu_i}{1+\alpha}\Big)\Big| \\ &+2\sum_{i=0}^{K}\sup_{\tau\in[0,1]}\Big|\frac{1}{PT_n}\sum_{t=1}^{\lceil\tau T_n\rceil}\mathrm{Re}\left(h_i\boldsymbol{s}_i[t]\boldsymbol{z}_0^*[t]\right)\Big| \\ &+2\sum_{0\leq i<j\leq K}\sup_{\tau\in[0,1]}\Big|\frac{1}{PT_n}\sum_{t=1}^{\lceil\tau T_n\rceil}\mathrm{Re}\left(h_ih_j^*\boldsymbol{s}_i[t]\boldsymbol{s}_j^*[t]\right)\Big|. \qquad \text{(A.9)}\end{aligned}$$

Writing $h_i=|h_i|e^{\sqrt{-1}\measuredangle h_i}$, the term $e^{\sqrt{-1}\measuredangle h_i}$ is absorbed into $\boldsymbol{s}_i[t]$ and we can write $\mathrm{Re}\left(h_i\boldsymbol{s}_i[t]\boldsymbol{z}_0^*[t]\right)=|h_i|\mathrm{Re}\left(\boldsymbol{s}_i[t]\boldsymbol{z}_0^*[t]\right)$ with a slight abuse of notation. Similarly, $\mathrm{Re}\left(h_ih_j^*\boldsymbol{s}_i[t]\boldsymbol{s}_j^*[t]\right)=|h_i||h_j|\mathrm{Re}\left(\boldsymbol{s}_i[t]\boldsymbol{s}_j^*[t]\right)$. This leads to

$$\begin{aligned}\sup_{\tau\in[0,1]}|\boldsymbol{F}_n(\tau_n)-F(\tau_n)| \;\leq\; & \frac{1}{P}\sup_{\tau\in[0,1]}\Big|\frac{1}{T_n}\sum_{t=1}^{\lceil\tau T_n\rceil}|\boldsymbol{z}_0[t]|^2-\tau_n\Big| \\ &+d\sum_{i=0}^{K}\sup_{\tau\in[0,1]}\Big|\frac{1}{PT_n}\sum_{t=1}^{\lceil\tau T_n\rceil}|\boldsymbol{s}_i[t]|^2-\psi\Big(\tau_n-\frac{\nu_i}{1+\alpha}\Big)\Big| \\ &+2d\sum_{i=0}^{K}\sup_{\tau\in[0,1]}\Big|\frac{1}{PT_n}\sum_{t=1}^{\lceil\tau T_n\rceil}\mathrm{Re}\left(\boldsymbol{s}_i[t]\boldsymbol{z}_0^*[t]\right)\Big| \\ &+2d\sum_{0\leq i<j\leq K}\sup_{\tau\in[0,1]}\Big|\frac{1}{PT_n}\sum_{t=1}^{\lceil\tau T_n\rceil}\mathrm{Re}\left(\boldsymbol{s}_i[t]\boldsymbol{s}_j^*[t]\right)\Big|, \qquad \text{(A.10)}\end{aligned}$$

where we define

$$d:=\max\Big\{\max_{0\leq i\leq K}|h_i|,\max_{0\leq i,j\leq K}|h_i||h_j|\Big\}. \qquad \text{(A.11)}$$

By (A.6) and (A.10), $\mathbb{P}(\|\boldsymbol{F}_n(\tau)-F(\tau)\|_\infty\geq\epsilon)$ is less than or equal to the probability that the term on the right side of (A.10) is greater than or equal to $\frac{\epsilon}{5}$. Thinking of the right side of (A.10) as a sum of $1+\sum_{i=0}^K 1+\sum_{i=0}^K 1+\sum_{0\leq i<j\leq K}1=1+2(K+1)+\frac{K(K+1)}{2}=\frac{K^2+5K+6}{2}$ terms, we conclude that $\mathbb{P}(\|\boldsymbol{F}_n(\tau)-F(\tau)\|_\infty\geq\epsilon)$ is less than or equal to the sum of the

probabilities that each of these terms is at least $\epsilon' := \frac{2\epsilon}{5(K^2+5K+6)}$. Hence,

$$\begin{aligned}\mathbb{P}(\|\boldsymbol{F}_n(\tau) - F(\tau)\|_\infty \geq \epsilon) \;\leq\; & \mathbb{P}\Big(\frac{1}{P} \sup_{\tau\in[0,1]} \Big|\frac{1}{T_n} \sum_{t=1}^{\lceil \tau T_n \rceil} |\boldsymbol{z}_0[t]|^2 - \tau_n\Big| \geq \epsilon'\Big) \\ & + \sum_{i=0}^{K} \mathbb{P}\Big(d \sup_{\tau\in[0,1]} \Big|\frac{1}{PT_n} \sum_{t=1}^{\lceil \tau T_n \rceil} |\boldsymbol{s}_i[t]|^2 - \psi\Big(\tau_n - \frac{\nu_i}{1+\alpha}\Big)\Big| \geq \epsilon'\Big) \\ & + \sum_{i=0}^{K} \mathbb{P}\Big(2d \sup_{\tau\in[0,1]} \Big|\frac{1}{PT_n} \sum_{t=1}^{\lceil \tau T_n \rceil} \mathrm{Re}\left(\boldsymbol{s}_i[t]\boldsymbol{z}_0^*[t]\right)\Big| \geq \epsilon'\Big) \\ & + \sum_{0\leq i<j\leq K} \mathbb{P}\Big(2d \sup_{\tau\in[0,1]} \Big|\frac{1}{PT_n} \sum_{t=1}^{\lceil \tau T_n \rceil} \mathrm{Re}\left(\boldsymbol{s}_i[t]\boldsymbol{s}_j^*[t]\right)\Big| \geq \epsilon'\Big) \end{aligned} \quad (A.12)$$

In the following, we find an upper bound of the form $\Theta(1)e^{-\epsilon^2\Theta(n)}$ on each term on the right side of (A.9). Throughout the proof, we invoke the following lemmas in several occasions:

Lemma 1 (Etemadi's inequality [30]). *Let $\boldsymbol{a}_1, \cdots, \boldsymbol{a}_N$ be independent real-valued random variables and $\epsilon > 0$. Then*

$$\mathbb{P}\Big(\max_{m=1,\cdots,N} \Big|\sum_{i=1}^{m} \boldsymbol{a}_i\Big| \geq \epsilon\Big) \leq 3 \max_{m=1,\cdots,N} \mathbb{P}\Big(\Big|\sum_{i=1}^{m} \boldsymbol{a}_i\Big| \geq \frac{\epsilon}{3}\Big). \quad (A.13)$$

Lemma 2 (Bernstein's inequality [31]). *Let $\boldsymbol{a}_1, \cdots, \boldsymbol{a}_N$ be independent zero mean real-valued random variables, $\boldsymbol{a}_i \leq 1$ for any $1 \leq i \leq N$ and $\epsilon > 0$. Then*

$$\mathbb{P}\Big(\frac{1}{N} \sum_{i=1}^{N} \boldsymbol{a}_i \geq \epsilon\Big) \leq e^{-\frac{N\epsilon^2}{2(\sigma^2+\frac{\epsilon}{3})}}, \quad (A.14)$$

where σ^2 is the arithmetic average of the variances of $\boldsymbol{a}_1, \cdots, \boldsymbol{a}_N$.

A.0.1 The term $\mathbb{P}\big(\sup_{\tau\in[0,1]}\big|\frac{1}{T_n}\sum_{t=1}^{\lceil \tau T_n\rceil}|\boldsymbol{z}_0[t]|^2-\tau_n\big|\geq \epsilon' P\big)$

We have

$$\begin{aligned}\mathbb{P}\Big(\sup_{\tau\in[0,1]}\Big|\frac{1}{T_n}\sum_{t=1}^{\lceil \tau T_n\rceil}|\boldsymbol{z}_0[t]|^2-\tau_n\Big|\geq \epsilon' P\Big) &= \mathbb{P}\Big(\sup_{\tau\in[0,1]}\Big|\sum_{t=1}^{\lceil \tau T_n\rceil}(|\boldsymbol{z}_0[t]|^2-1)\Big|\geq \epsilon' P T_n\Big)\\ &= \mathbb{P}\Big(\max_{m=1,\cdots,T_n}\Big|\sum_{t=1}^{m}(|\boldsymbol{z}_0[t]|^2-1)\Big|\geq \epsilon' P T_n\Big)\\ &\leq 3\max_{m=1,\cdots,T_n}\mathbb{P}\Big(\Big|\sum_{t=1}^{m}(|\boldsymbol{z}_0[t]|^2-1)\Big|\geq \frac{\epsilon' P T_n}{3}\Big),\end{aligned} \tag{A.15}$$

where the last step is due to Etemadi's inequality. But,

$$\begin{aligned}\mathbb{P}\Big(\Big|\sum_{t=1}^{m}(|\boldsymbol{z}_0[t]|^2-1)\Big|\geq \frac{\epsilon' P T_n}{3}\Big) &= \mathbb{P}\Big(\sum_{t=1}^{m}|\boldsymbol{z}_0[t]|^2\geq m+\frac{\epsilon' P T_n}{3}\Big)\\ &\quad+\mathbb{P}\Big(\sum_{t=1}^{m}(1-|\boldsymbol{z}_0[t]|^2)\geq \frac{\epsilon' P T_n}{3}\Big).\end{aligned} \tag{A.16}$$

Let us find upper bounds on the two terms on the right side of (A.16). For $0<a<1$ and by the Chernoff bound [28],

$$\mathbb{P}\Big(\sum_{t=1}^{m}|\boldsymbol{z}_0[t]|^2\geq m+\frac{\epsilon' P T_n}{3}\Big)\leq e^{-(m+\frac{\epsilon' P T_n}{3})a}\left(\mathbb{E}\left[e^{a|\boldsymbol{z}_0[0]|^2}\right]\right)^m=\frac{e^{-(m+\frac{\epsilon' P T_n}{3})a}}{(1-a)^m}, \tag{A.17}$$

where the last step is due to $|\boldsymbol{z}_0[0]|^2$ being an exponential random variable with parameter 1. Minimizing the right side of (A.17) over $0<a<1$, we get $a=1-\frac{m}{m+\frac{\epsilon' P T_n}{3}}$. Subbing this value of a in (A.17),

$$\mathbb{P}\Big(\sum_{t=1}^{m}|\boldsymbol{z}_0[t]|^2\geq m+\frac{\epsilon' P T_n}{3}\Big)\leq\Big(1+\frac{\epsilon' P T_n}{3m}\Big)^m e^{-\frac{\epsilon' P T_n}{3}}. \tag{A.18}$$

In order to find an upper bound on $\mathbb{P}\big(\sum_{t=1}^{m}(1-|\boldsymbol{z}_0[t]|^2)\geq\frac{\epsilon' P T_n}{3}\big)$, we use Bernstein's inequality. Since $\boldsymbol{a}_t=1-|\boldsymbol{z}_0[t]|^2$ are independent random variables with zero mean and unit variance and $\boldsymbol{a}_t \leq 1$ for any $1 \leq t \leq m$, we get

$$\mathbb{P}\Big(\sum_{t=1}^{m}(1-|\boldsymbol{z}_0[t]|^2) \geq \frac{\epsilon' PT_n}{3}\Big) \leq e^{-\frac{\epsilon'^2 P^2 T_n^2}{2(9m+\epsilon' PT_n)}}. \tag{A.19}$$

By (A.16), (A.18) and (A.19),

$$\begin{aligned}\max_{m=1,\cdots,T_n} \mathbb{P}\Big(\Big|\sum_{t=1}^{m}(|\boldsymbol{z}_0[t]|^2-1)\Big| \geq \frac{\epsilon' PT_n}{3}\Big) &\leq \max_{m=1,\cdots,T_n}\Big(1+\frac{\epsilon' PT_n}{3m}\Big)^m e^{-\frac{\epsilon' PT_n}{3}} \\ &\quad + \max_{m=1,\cdots,T_n} e^{-\frac{\epsilon'^2 P^2 T_n^2}{2(9m+\epsilon' PT_n)}} \\ &= e^{-(\frac{\epsilon' P}{3}-\ln(1+\frac{\epsilon' P}{3}))T_n} + e^{-\frac{\epsilon'^2 P^2 T_n}{2(9+\epsilon' P)}} \\ &\leq e^{-\frac{\epsilon'^2 P^2 T_n}{54}} + e^{-\frac{\epsilon'^2 P^2 T_n}{2(9+\epsilon' P)}}, \end{aligned} \tag{A.20}$$

where the penultimate step is due to the fact that both maximizations are achieved for $m = T_n$ and the last step is due to the inequality $a - \ln(1+a) \geq \frac{a^2}{6}$ for any $0 < a < 1$.[1]

A.0.2 The term $\mathbb{P}\Big(d \sup_{\tau\in[0,1]} \Big|\frac{1}{PT_n}\sum_{t=1}^{\lceil \tau T_n\rceil} |\boldsymbol{s}_0[t]|^2 - \psi\left(\tau_n - \frac{\nu_0}{1+\alpha}\right)\Big| \geq \epsilon'\Big)$

This is the most demanding part of the proof. Following similar lines in (A.15), we have

$$\begin{aligned}&\mathbb{P}\Big(d \sup_{\tau\in[0,1]}\Big|\frac{1}{PT_n}\sum_{t=1}^{\lceil \tau T_n\rceil}|\boldsymbol{s}_0[t]|^2 - \psi\Big(\tau_n - \frac{\nu_0}{1+\alpha}\Big)\Big| \geq \epsilon'\Big) \\ &\quad\leq 3\max_{m=1,\cdots,T_n}\mathbb{P}\Big(\Big|\frac{1}{PT_n}\sum_{t=1}^{m}|\boldsymbol{s}_0[t]|^2 - \psi\Big(\frac{m}{T_n} - \frac{\nu_0}{1+\alpha}\Big)\Big| \geq \frac{\epsilon'}{3d}\Big) \\ &\quad= 3\max_{\substack{1\leq i\leq 4\\ 1\leq j\leq 3}}\max_{m\in\mathcal{E}_i\cap\mathcal{F}_j}\mathbb{P}\Big(\Big|\frac{1}{PT_n}\sum_{t=1}^{m}|\boldsymbol{s}_0[t]|^2 - \psi\Big(\frac{m}{T_n} - \frac{\nu_0}{1+\alpha}\Big)\Big| > \frac{\epsilon'}{3d}\Big)\end{aligned} \tag{A.21}$$

where the last step is due to the fact that the maximum value of a sum of functions is less than the sum of the maximum values of those functions and the sets $\mathcal{E}_i$ and $\mathcal{F}_j$ are defined

[1] Having the Taylor series $\ln(1+a) = a - \frac{a^2}{2} + \frac{a^3}{3} - \frac{a^4}{4} + \cdots$ and invoking Leibniz lemma for alternating series, we get $\ln(1+a) \leq a - \frac{a^2}{2} + \frac{a^3}{3}$. Since $0 < a < 1$, we have $a^3 < a^2$ and $a - \ln(1+a) \geq \frac{a^2}{6}$ follows.

as

$$\begin{array}{c}\mathcal{E}_1 = \{m : m < t_{0,n}\} \\ \mathcal{E}_2 = \{m : t_{0,n} \leq m \leq t_{0,n} + n_\delta - 1\} \\ \mathcal{E}_3 = \{m : t_{0,n} + n_\delta \leq m \leq t_{0,n} + n_\delta + n - 1\} \\ \mathcal{E}_4 = \{m : m > t_{0,n} + n_\delta + n - 1\}\end{array} \tag{A.22}$$

and

$$\begin{array}{c}\mathcal{F}_1 = \left\{m : \frac{m}{T_n} - \frac{\nu_0}{1+\alpha} < 0\right\} \\ \mathcal{F}_2 = \left\{m : 0 \leq \frac{m}{T_n} - \frac{\nu_0}{1+\alpha} < \frac{\alpha}{1+\alpha}\right\} \\ \mathcal{F}_3 = \left\{m : \frac{m}{T_n} - \frac{\nu_0}{1+\alpha} \geq \frac{\alpha}{1+\alpha}\right\}\end{array}. \tag{A.23}$$

In (A.22) and (A.23), it is implicit that $m \in \{1, \cdots, T_n\}$. It is straightforward to see that $\mathcal{E}_i \bigcap \mathcal{F}_j = \emptyset$ if $(i,j) \in \{(1,2),(1,3),(2,3),(3,1),(4,1),(4,2)\}$ and n is sufficiently large such that $n_\delta \geq 2$. For example, let us verify $\mathcal{E}_3 \bigcap \mathcal{F}_1 = \emptyset$. Verification of the other cases can be carried out similarly and is omitted. We have $\mathcal{E}_3 \bigcap \mathcal{F}_1 \subseteq \{m : t_{0,n} + n_\delta \leq m \leq \lfloor \frac{\nu_0 T_n}{1+\alpha} \rfloor\}$. This requires

$$t_{0,n} + n_\delta \leq \lfloor \nu_0 T_n/(1+\alpha) \rfloor. \tag{A.24}$$

By the floor inequality, $t_{0,n} > \lfloor \frac{n}{\alpha} \rfloor \nu_0 - 1$ and $\lfloor \frac{\nu_0 T_n}{1+\alpha} \rfloor \leq \frac{\nu_0 T_n}{1+\alpha}$. Using these inequalities in (A.24),

$$(T_n - (1+\alpha)\lfloor n/\alpha \rfloor)\nu_0 > (1+\alpha)(n_\delta - 1). \tag{A.25}$$

But,

$$\begin{aligned} T_n - (1+\alpha)\lfloor n/\alpha \rfloor &= n_\delta + n - 1 - \alpha \lfloor n/\alpha \rfloor \\ &\geq n_\delta + n - 1 - \alpha \times n/\alpha \\ &= n_\delta - 1. \end{aligned} \tag{A.26}$$

If $n_\delta \geq 2$, then $T_n - (1+\alpha)\lfloor n/\alpha \rfloor > 0$ and (A.25) yields

$$\begin{aligned} \nu_0 &> \frac{(1+\alpha)(n_\delta - 1)}{T_n - (1+\alpha)\lfloor n/\alpha \rfloor} \\ &= 1 + \frac{n_\delta - 1 - (\frac{n}{\alpha} - \lfloor \frac{n}{\alpha} \rfloor)}{T_n - (1+\alpha)\lfloor n/\alpha \rfloor} \\ &\geq 1 + \frac{n_\delta - 2}{T_n - (1+\alpha)\lfloor n/\alpha \rfloor}, \end{aligned} \tag{A.27}$$

where the last step is due to $0 \leq \frac{n}{\alpha} - \lfloor \frac{n}{\alpha} \rfloor < 1$. Having $n_\delta \geq 2$, (A.27) results in $\nu_0 > 1$ which is a contradiction.

We proceed by investigating the cases $m \in \mathcal{E}_i \bigcap \mathcal{F}_j$ for (i,j) taking on $(1,1)$, $(2,1)$, $(2,2)$, $(3,2)$, $(3,3)$ and $(4,3)$:

- Let $m \in \mathcal{E}_1 \bigcap \mathcal{F}_1$. Since $\boldsymbol{s}_0[t] = 0$ for $n \in \mathcal{E}_1$ and $\psi\left(\frac{m}{T_n} - \frac{\nu_0}{1+\alpha}\right) = 0$ for $m \in \mathcal{F}_1$, we get $\mathbb{P}\big(\big|\frac{1}{PT_n}\sum_{t=1}^{m}|\boldsymbol{s}_0[t]|^2 - \psi\left(\frac{m}{T_n} - \frac{\nu_0}{1+\alpha}\right)\big| \geq \frac{\epsilon'}{3d}\big) = 0$.

- Let $m \in \mathcal{E}_2 \bigcap \mathcal{F}_1$. Then

$$\Big|\frac{1}{PT_n}\sum_{t=1}^{m}|\boldsymbol{s}_0[t]|^2 - \psi\Big(\frac{m}{T_n} - \frac{\nu_0}{1+\alpha}\Big)\Big| \overset{(a)}{=} \frac{1}{PT_n}\Big(\sum_{t=0}^{m-t_{0,n}}|\boldsymbol{x}'_{0,t}|^2 - |\boldsymbol{x}'_{0,0}|^2\mathbb{1}_{t_{0,n}=0}\Big)$$
$$\overset{(b)}{\leq} \frac{1}{PT_n}\sum_{t=0}^{n_\delta-1}|\boldsymbol{x}'_{0,t}|^2, \tag{A.28}$$

where (a) is due to the fact that if $\boldsymbol{t}_{0,n} = 0$, then the first symbol in the preamble sequence of user 0 is not included in the sum $\sum_{t=1}^{m}|\boldsymbol{s}_0[t]|^2$ and (b) is due to $m - t_{0,n} \leq n_\delta - 1$ and $|\boldsymbol{x}'_{0,0}|^2\mathbb{1}_{t_{0,n}=0} \geq 0$. Therefore,

$$\mathbb{P}\Big(\Big|\frac{1}{PT_n}\sum_{t=1}^{m}|\boldsymbol{s}_0[t]|^2 - \psi\Big(\frac{m}{T_n} - \frac{\nu_0}{1+\alpha}\Big)\Big| \geq \frac{\epsilon'}{3d}\Big) \leq \mathbb{P}\Big(\frac{1}{PT_n}\sum_{t=0}^{n_\delta-1}|\boldsymbol{x}'_{0,t}|^2 \geq \frac{\epsilon'}{3d}\Big)$$
$$\leq \Big(\frac{\epsilon' T_n}{3dn_\delta}\Big)^{n_\delta} e^{-\left(\frac{\epsilon' T_n}{3d} - n_\delta\right)}$$
$$= e^{-\frac{\epsilon' T_n}{3d}\left(1 - \frac{3dn_\delta}{\epsilon' T_n}\left(1 + \ln\frac{\epsilon' T_n}{3dn_\delta}\right)\right)}, \tag{A.29}$$

where the penultimate step follows by the Chernoff bounding technique as in (A.18). The function $g(a) = 1 - \frac{1+\ln a}{a}$ is increasing for $a \geq 1$. If n is large enough such that

$$\frac{\epsilon' T_n}{3dn_\delta} > 2, \tag{A.30}$$

then $1 - \frac{3dn_\delta}{\epsilon' T_n}\big(1 + \ln\frac{\epsilon' T_n}{3dn_\delta}\big) > \frac{1}{2}(1 - \ln 2)$. Using this in (A.29),

$$\mathbb{P}\Big(\Big|\frac{1}{PT_n}\sum_{t=1}^{m}|\boldsymbol{s}_0[t]|^2 - \psi\Big(\frac{m}{T_n} - \frac{\nu_0}{1+\alpha}\Big)\Big| \geq \frac{\epsilon'}{3d}\Big) \leq e^{-\frac{1-\ln 2}{6d}\epsilon' T_n} < e^{-\frac{1-\ln 2}{6d}\epsilon'^2 T_n}, \tag{A.31}$$

$$
\begin{aligned}
\Big|\frac{1}{PT_n}\sum_{t=1}^{m}|\boldsymbol{s}_0[t]|^2-\psi\Big(\frac{m}{T_n}-\frac{\nu_0}{1+\alpha}\Big)\Big| &= \Big|\frac{1}{PT_n}\Big(\sum_{t=0}^{m-t_{0,n}}|\boldsymbol{x}'_{0,t}|^2-|\boldsymbol{x}'_{0,0}|^2\mathbb{1}_{t_{0,n}=0}\Big)-\Big(\frac{m}{T_n}-\frac{\nu_0}{1+\alpha}\Big)\Big| \\
&\le \frac{1}{PT_n}\Big(\sum_{t=0}^{m-t_{0,n}}|\boldsymbol{x}'_{0,t}|^2-|\boldsymbol{x}'_{0,0}|^2\mathbb{1}_{t_{0,n}=0}\Big)+\Big|\frac{m}{T_n}-\frac{\nu_0}{1+\alpha}\Big| \\
&\le \frac{1}{PT_n}\sum_{t=0}^{n_\delta-1}|\boldsymbol{x}'_{0,t}|^2+\Big|\frac{m}{T_n}-\frac{\nu_0}{1+\alpha}\Big|. \qquad \text{(A.33)}
\end{aligned}
$$

where the last step is due to $\epsilon'^2 < \epsilon'$.

- Let $m \in \mathcal{E}_2 \bigcap \mathcal{F}_2$. Then we have (A.33) and hence,

$$
\begin{aligned}
&\mathbb{P}\Big(\Big|\frac{1}{PT_n}\sum_{t=1}^{m}|\boldsymbol{s}_0[t]|^2-\psi\Big(\frac{m}{T_n}-\frac{\nu_0}{1+\alpha}\Big)\Big|\ge\frac{\epsilon'}{3d}\Big) \\
&\le \mathbb{P}\Big(\frac{1}{PT_n}\sum_{t=0}^{n_\delta-1}|\boldsymbol{x}'_{0,t}|^2\ge\frac{\epsilon'}{6d}\Big)+\mathbb{1}_{|\frac{m}{T_n}-\frac{\nu_0}{1+\alpha}|\ge\frac{\epsilon'}{6d}} \\
&\le e^{-\frac{1-\ln 2}{12d}\epsilon'^2 T_n}+\mathbb{1}_{|\frac{m}{T_n}-\frac{\nu_0}{1+\alpha}|\ge\frac{\epsilon'}{6d}}, \qquad \text{(A.32)}
\end{aligned}
$$

where the last step follows by the same bounding technique in (A.29) and (A.31). Moreover,

$$
\begin{aligned}
\Big|\frac{m}{T_n}-\frac{\nu_0}{1+\alpha}\Big| &\overset{(a)}{=} \frac{m}{T_n}-\frac{\nu_0}{1+\alpha}\\
&\overset{(b)}{\leq} \frac{t_{0,n}+n_\delta-1}{T_n}-\frac{\nu_0}{1+\alpha}\\
&\leq \frac{\lfloor\frac{n}{\alpha}\rfloor\nu_0+n_\delta-1}{T_n}-\frac{\nu_0}{1+\alpha}\\
&= \frac{\nu_0((1+\alpha)\lfloor\frac{n}{\alpha}\rfloor-T_n)+(1+\alpha)(n_\delta-1)}{(1+\alpha)T_n}\\
&\overset{(c)}{\leq} \frac{-\nu_0(n_\delta-1)+(1+\alpha)(n_\delta-1)}{(1+\alpha)T_n}\\
&= \frac{(1+\alpha-\nu_0)(n_\delta-1)}{(1+\alpha)T_n}\\
&< \frac{n_\delta}{T_n},
\end{aligned}
\tag{A.34}
$$

where (a) is due to the fact that if $m\in\mathcal{F}_2$, then $\frac{m}{T_n}-\frac{\nu_0}{1+\alpha}\geq 0$, (b) is due to the fact that if $m\in\mathcal{E}_2$, then $m\leq t_{0,n}+n_\delta-1$ and (c) is due to (A.26). By (A.34), $|\frac{m}{T_n}-\frac{\nu_0}{1+\alpha}|\geq\frac{\epsilon'}{6d}$ implies $\frac{n_\delta}{T_n}>\frac{\epsilon'}{6d}$ which contradicts (A.30). Therefore, the second term on the right side of (A.32) is zero.

- Let $m\in\mathcal{E}_3\bigcap\mathcal{F}_2$. Following similar steps in (A.33), we arrive at

$$
\begin{aligned}
&\Big|\frac{1}{PT_n}\sum_{t=1}^{m}|\boldsymbol{s}_0[t]|^2-\psi\Big(\frac{m}{T_n}-\frac{\nu_0}{1+\alpha}\Big)\Big|\\
&\leq \frac{1}{PT_n}\sum_{t=0}^{n_\delta-1}|\boldsymbol{x}'_{0,t}|^2+\Big|\frac{1}{PT_n}\sum_{t=0}^{m-t_{0,n}-n_\delta}|\boldsymbol{x}_{0,t}|^2-\Big(\frac{m}{T_n}-\frac{\nu_0}{1+\alpha}\Big)\Big|\\
&\leq \frac{1}{PT_n}\sum_{t=0}^{n_\delta-1}|\boldsymbol{x}'_{0,t}|^2+\Big|\frac{1}{PT_n}\sum_{t=0}^{m-t_{0,n}-n_\delta}(|\boldsymbol{x}_{0,t}|^2-P)\Big|+\Big|\frac{\nu_0}{1+\alpha}-\frac{t_{0,n}+n_\delta-1}{T_n}\Big|.
\end{aligned}
\tag{A.35}
$$

where in the last step we have added and subtracted $\frac{m-t_{0,n}-n_\delta+1}{T_n}$ and applied triangle inequality. By (A.35),

$$\mathbb{P}\Big(\Big|\frac{1}{PT_n}\sum_{t=1}^{m}|\boldsymbol{s}_0[t]|^2-\psi\Big(\frac{m}{T_n}-\frac{\nu_0}{1+\alpha}\Big)\Big|\geq\frac{\epsilon'}{3d}\Big)$$
$$\leq\mathbb{P}\Big(\frac{1}{PT_n}\sum_{t=0}^{n_\delta-1}|\boldsymbol{x}'_{0,t}|^2\geq\frac{\epsilon'}{9d}\Big)+\mathbb{P}\Big(\Big|\frac{1}{PT_n}\sum_{t=0}^{m-t_{0,n}-n_\delta}(|\boldsymbol{x}_{0,t}|^2-P)\Big|\geq\frac{\epsilon'}{9d}\Big)$$
$$+\mathbb{1}_{\left|\frac{\nu_0}{1+\alpha}-\frac{t_{0,n}+n_\delta-1}{T_n}\right|\geq\frac{\epsilon'}{9d}}. \tag{A.36}$$

Let us investigate the three terms on the right side of (A.36) separately:

- Following the same bounding technique in (A.29) and (A.31),

$$\mathbb{P}\Big(\frac{1}{PT_n}\sum_{t=0}^{n_\delta-1}|\boldsymbol{x}'_{0,t}|^2\geq\frac{\epsilon'}{9d}\Big)\leq e^{-\frac{1-\ln 2}{18d}\epsilon'^2T_n}, \tag{A.37}$$

  where in contrast to (A.30), we require that

$$\frac{\epsilon' T_n}{9dn_\delta}>2. \tag{A.38}$$

- Following similar lines of reasoning that led to $\mathbb{1}_{|\frac{m}{T_n}-\frac{\nu_0}{1+\alpha}|\geq\frac{\epsilon'}{6d}}=0$ for $m\in\mathcal{E}_2\bigcap\mathcal{F}_2$, one can show that $\mathbb{1}_{\left|\frac{\nu_0}{1+\alpha}-\frac{t_{0,n}+n_\delta-1}{T_n}\right|\geq\frac{\epsilon'}{9d}}=0$ under the constraint in (A.38).
- Following similar steps as in (A.16), (A.18) and (A.19),

$$\mathbb{P}\Big(\Big|\frac{1}{PT_n}\sum_{t=0}^{m-t_{0,n}-n_\delta}(|\boldsymbol{x}_{0,t}|^2-P)\Big|\geq\frac{\epsilon'}{9d}\Big)\ \leq\ \Big(1+\frac{\epsilon' T_n}{9dm'}\Big)^{m'}e^{-\frac{\epsilon' T_n}{9d}}+e^{-\frac{\epsilon'^2T_n^2}{2d(81dm'+3\epsilon' T_n)}}, \tag{A.39}$$

  where $m'=m-t_{0,n}-n_\delta+1$. Since $1\leq m\leq T_n$, then $m'\leq T_n$. Moreover, the two terms on the right side of (A.39) are increasing in terms of m'. Hence,

$$\begin{aligned}\mathbb{P}\Big(\Big|\frac{1}{PT_n}\sum_{t=0}^{m-t_{0,n}-n_\delta}(|\boldsymbol{x}_{0,t}|^2-P)\Big|>\frac{\epsilon'}{9d}\Big)\ &\leq\ e^{-(\frac{\epsilon'}{9d}-\ln(1+\frac{\epsilon'}{9d}))T_n}+e^{-\frac{\epsilon'^2T_n}{2d(81d+3\epsilon')}}\\ &\leq\ e^{-\frac{\epsilon'^2T_n}{486d^2}}+e^{-\frac{\epsilon'^2T_n}{2d(81d+3\epsilon')}},\end{aligned} \tag{A.40}$$

where the last step is due to the inequality $a - \ln(1+a) \geq \frac{a^2}{6}$ for any $0 < a < 1$ verified in Footnote 1.

- Let $m \in \mathcal{E}_3 \bigcap \mathcal{F}_3$. We have (A.41).

$$\begin{aligned} \Big|\frac{1}{PT_n}\sum_{t=1}^{m}|\boldsymbol{s}_0[t]|^2 - \psi\Big(\frac{m}{T_n} - \frac{\nu_0}{1+\alpha}\Big)\Big| \;\leq\; & \frac{1}{PT_n}\sum_{t=0}^{n_\delta - 1}|\boldsymbol{x}'_{0,t}|^2 \\ & + \Big|\frac{1}{PT_n}\sum_{t=0}^{m-t_{0,n}-n_\delta}(|\boldsymbol{x}_{0,t}|^2 - P)\Big| \\ & + \Big|\frac{m - t_{0,n} - n_\delta + 1}{T_n} - \frac{\alpha}{1+\alpha}\Big| \end{aligned} \tag{A.41}$$

By (A.41),

$$\begin{aligned} & \mathbb{P}\Big(\Big|\frac{1}{PT_n}\sum_{t=1}^{m}|\boldsymbol{s}_0[t]|^2 - \psi\Big(\frac{m}{T_n} - \frac{\nu_0}{1+\alpha}\Big)\Big| \geq \frac{\epsilon'}{3d}\Big) \\ & \leq \mathbb{P}\Big(\frac{1}{PT_n}\sum_{t=0}^{n_\delta - 1}|\boldsymbol{x}'_{0,t}|^2 \geq \frac{\epsilon'}{9d}\Big) \\ & \quad + \mathbb{P}\Big(\Big|\frac{1}{PT_n}\sum_{t=0}^{m-t_{0,n}-n_\delta}(|\boldsymbol{x}_{0,t}|^2 - P)\Big| \geq \frac{\epsilon'}{9d}\Big) \\ & \quad + \mathbb{1}_{\big|\frac{m-t_{0,n}-n_\delta+1}{T_n} - \frac{\alpha}{1+\alpha}\big| \geq \frac{\epsilon'}{9d}}. \end{aligned} \tag{A.42}$$

We have obtained upper bounds on the first and second terms on the right side of (A.42) in (A.37) and (A.40), respectively. As for the term $\mathbb{1}_{|\frac{m-t_{0,n}-n_\delta+1}{T_n} - \frac{\alpha}{1+\alpha}| \geq \frac{\epsilon'}{9d}}$, note that if $m \in \mathcal{E}_3$, then $m \leq t_{0,n} + n_\delta + n - 1$. Therefore,

$$\begin{aligned} & \frac{m - t_{0,n} - n_\delta + 1}{T_n} - \frac{\alpha}{1+\alpha} \\ & \quad \leq \frac{n}{T_n} - \frac{\alpha}{1+\alpha} = -\frac{n_\delta - 1 - (\frac{n}{\alpha} - \lfloor\frac{n}{\alpha}\rfloor)}{(1+\alpha)T_n}\alpha \\ & \quad \leq -\frac{n_\delta - 2}{T_n}\alpha, \end{aligned} \tag{A.43}$$

where the last step uses the floor inequality. Therefore, as long as $n_\delta \geq 2$,

$$\begin{aligned}\Big|\frac{m-t_{0,n}-n_\delta+1}{T_n}-\frac{\alpha}{1+\alpha}\Big| &= \frac{\alpha}{1+\alpha}-\frac{m-t_{0,n}-n_\delta+1}{T_n}\\ &\overset{(a)}{\leq} \frac{\alpha}{1+\alpha}-\frac{\frac{\nu_0+\alpha}{1+\alpha}T_n-t_{0,n}-n_\delta+1}{T_n}\\ &= \frac{(1+\alpha)(t_{0,n}+n_\delta-1)-\nu_0T_n}{(1+\alpha)T_n}\\ &\leq \frac{(1+\alpha)(\lfloor\frac{n}{\alpha}\rfloor\nu_0+n_\delta-1)-\nu_0T_n}{(1+\alpha)T_n}\\ &= \frac{(1+\alpha)(n_\delta-1)-\nu_0(T_n-(1+\alpha)\lfloor\frac{n}{\alpha}\rfloor)}{(1+\alpha)T_n}\\ &\overset{(b)}{\leq} \frac{(1+\alpha)(n_\delta-1)-\nu_0(n_\delta-1)}{(1+\alpha)T_n}\\ &= \frac{(1+\alpha-\nu_0)(n_\delta-1)}{(1+\alpha)T_n}<\frac{n_\delta}{T_n},\end{aligned} \tag{A.44}$$

where (a) is due to the fact that if $m \in \mathcal{F}_3$, then $m \geq \frac{\nu_0+\alpha}{1+\alpha}T_n$ and (b) is due to (A.26). By (A.44) and under the constraint in (A.38), we conclude that the third term on the right side of (A.42) is zero.

- The case $m \in \mathcal{E}_4 \bigcap \mathcal{F}_3$ can be handled similar to the case of $m \in \mathcal{E}_3 \bigcap \mathcal{F}_2$. In fact, one obtains the same upper bound.

Looking back at (A.21), we have shown that $\mathbb{P}\big(d\sup_{\tau\in[0,1]}\big|\frac{1}{PT_n}\sum_{t=1}^{\lceil\tau T_n\rceil}|\boldsymbol{s}_0[t]|^2-\psi\big(\tau_n-\frac{\nu_0}{1+\alpha}\big)\big|\geq\epsilon'\big)$ is bounded from above by $\Theta(1)e^{-\epsilon^2\Theta(n)}$.

A.0.3 The term $\mathbb{P}\Big(2d \sup_{\tau\in[0,1]} \Big|\frac{1}{PT_n}\sum_{t=1}^{\lceil \tau T_n\rceil} \mathrm{Re}\left(\boldsymbol{s}_0[t]\boldsymbol{z}_0^*[t]\right)\Big| \geq \epsilon'\Big)$

We have

$$\begin{aligned}
&\mathbb{P}\Big(2d \sup_{\tau\in[0,1]} \Big|\frac{1}{PT_n}\sum_{t=1}^{\lceil \tau T_n\rceil} \mathrm{Re}\left(\boldsymbol{s}_0[t]\boldsymbol{z}_0^*[t]\right)\Big| \geq \epsilon'\Big)\\
&\leq \mathbb{P}\Big(2d \sup_{\tau\in[0,1]} \Big|\frac{1}{PT_n}\sum_{t=1}^{\lceil \tau T_n\rceil} \mathrm{Re}(\boldsymbol{s}_0[t])\mathrm{Re}(\boldsymbol{z}_0^*[t])\Big| \geq \frac{\epsilon'}{2}\Big)\\
&\quad +\mathbb{P}\Big(2d \sup_{\tau\in[0,1]} \Big|\frac{1}{PT_n}\sum_{t=1}^{\lceil \tau T_n\rceil} \mathrm{Im}(\boldsymbol{s}_0[t])\mathrm{Im}(\boldsymbol{z}_0^*[t])\Big| \geq \frac{\epsilon'}{2}\Big)\\
&= 2\mathbb{P}\Big(\sup_{\tau\in[0,1]} \Big|\frac{1}{PT_n}\sum_{t=1}^{\lceil \tau T_n\rceil} \mathrm{Re}(\boldsymbol{s}_0[t])\mathrm{Re}(\boldsymbol{z}_0^*[t])\Big| \geq \frac{\epsilon'}{4d}\Big),
\end{aligned} \tag{A.45}$$

where in the first step we write $\mathrm{Re}\left(\boldsymbol{s}_0[t]\boldsymbol{z}_0^*[t]\right) = \mathrm{Re}\left(\boldsymbol{s}_0[t]\right)\mathrm{Re}\left(\boldsymbol{z}_0^*[t]\right) - \mathrm{Im}\left(\boldsymbol{s}_0[t]\right)\mathrm{Im}\left(\boldsymbol{z}_0^*[t]\right)$ and the second step is due to $\mathrm{Re}\left(\boldsymbol{s}_0[t]\right)\mathrm{Re}\left(\boldsymbol{z}_0^*[t]\right)$ and $\mathrm{Im}\left(\boldsymbol{s}_0[t]\right)\mathrm{Im}\left(\boldsymbol{z}_0^*[t]\right)$ being identically distributed. Moreover, we have the thread of inequalities

$$\begin{aligned}
&\mathbb{P}\Big(\sup_{\tau\in[0,1]} \Big|\frac{1}{PT_n}\sum_{t=1}^{\lceil \tau T_n\rceil} \mathrm{Re}(\boldsymbol{s}_0[t])\mathrm{Re}(\boldsymbol{z}_0^*[t])\Big| \geq \frac{\epsilon'}{4d}\Big)\\
&= \mathbb{P}\Big(\max_{m=1,\cdots,T_n} \Big|\frac{1}{PT_n}\sum_{t=1}^{m} \mathrm{Re}(\boldsymbol{s}_0[t])\mathrm{Re}(\boldsymbol{z}_0^*[t])\Big| \geq \frac{\epsilon'}{4d}\Big)\\
&\leq 3\max_{m=1,\cdots,T_n} \mathbb{P}\Big(\Big|\frac{1}{PT_n}\sum_{t=1}^{m} \mathrm{Re}(\boldsymbol{s}_0[t])\mathrm{Re}(\boldsymbol{z}_0^*[t])\Big| \geq \frac{\epsilon'}{12d}\Big)\\
&= 3\max_{m=1,\cdots,T_n} \sum_{1\leq i\leq 4} \mathbb{P}\Big(\Big|\frac{1}{PT_n}\sum_{t=1}^{m} \mathrm{Re}(\boldsymbol{s}_0[t])\mathrm{Re}(\boldsymbol{z}_0^*[t])\Big| \geq \frac{\epsilon'}{12d}\Big)\mathbb{1}_{m\in\mathcal{E}_i}\\
&\leq 3\sum_{1\leq i\leq 4} \max_{\substack{m=1,\cdots,T_n\\ m\in\mathcal{E}_i}} \mathbb{P}\Big(\Big|\frac{1}{PT_n}\sum_{t=1}^{m} \mathrm{Re}(\boldsymbol{s}_0[t])\mathrm{Re}(\boldsymbol{z}_0^*[t])\Big| \geq \frac{\epsilon'}{12c}\Big).
\end{aligned} \tag{A.46}$$

As for the last term in (A.46), let us study the (more interesting) case of $i = 3$. For any $m \in \mathcal{E}_3$,

$$\Big|\sum_{t=1}^{m} \mathrm{Re}(\boldsymbol{s}_0[t])\mathrm{Re}(\boldsymbol{z}_0^*[t])\Big| \leq \Big|\sum_{t=0}^{m-t_{0,n}-n_\delta} \mathrm{Re}(\boldsymbol{x}_{0,t})\boldsymbol{z}'_t\Big| + \Big|\sum_{t=0}^{n_\delta-1} \mathrm{Re}(\boldsymbol{x}'_{0,t})\boldsymbol{z}''_t\Big|, \tag{A.47}$$

where $\boldsymbol{z}'_t = \mathrm{Re}(\boldsymbol{z}_0^*[t+t_{0,n}+n_\delta])$ and $\boldsymbol{z}''_t = \mathrm{Re}(\boldsymbol{z}_0^*[t+t_{0,n}])$ for $0 \leq t \leq m - t_{0,n} - n_\delta$ and $0 \leq t \leq n_\delta - 1$, respectively. Therefore,

$$\begin{aligned}
\mathbb{P}\Big(\Big|\frac{1}{PT_n}\sum_{t=1}^{m} \mathrm{Re}(\boldsymbol{s}_0[t])\mathrm{Re}(\boldsymbol{z}_0^*[t])\Big| \geq \frac{\epsilon'}{12d}\Big) &\leq \mathbb{P}\Big(\Big|\frac{1}{PT_n}\sum_{t=0}^{m-t_{0,n}-n_\delta} \mathrm{Re}(\boldsymbol{x}_{0,t})\boldsymbol{z}'_t\Big| \geq \frac{\epsilon'}{24d}\Big) \\
&\quad +\mathbb{P}\Big(\Big|\frac{1}{PT_n}\sum_{t=0}^{n_\delta-1} \mathrm{Re}(\boldsymbol{x}'_{0,t})\boldsymbol{z}''_t\Big| \geq \frac{\epsilon'}{24d}\Big) \\
&= 2\,\mathbb{P}\Big(\frac{1}{PT_n}\sum_{t=0}^{m-t_{0,n}-n_\delta} \mathrm{Re}(\boldsymbol{x}_{0,t})\boldsymbol{z}'_t \geq \frac{\epsilon'}{24d}\Big) \\
&\quad +2\,\mathbb{P}\Big(\frac{1}{PT_n}\sum_{t=0}^{n_\delta-1} \mathrm{Re}(\boldsymbol{x}'_{0,t})\boldsymbol{z}''_t \geq \frac{\epsilon'}{24d}\Big)
\end{aligned} \tag{A.48}$$

where the last step is due to the fact that the PDFs of $\sum_{t=0}^{m-t_{0,n}-n_\delta} \mathrm{Re}(\boldsymbol{x}_{0,t})\boldsymbol{z}'_t$ and $\sum_{t=0}^{n_\delta-1} \mathrm{Re}(\boldsymbol{x}'_{0,t})\boldsymbol{z}''_t$ are even functions. In order to develop Chernoff-type upper bounds on the terms on the right side of (A.48), one needs to know the moment generating function of the product of two Gaussian random variables. It is straightforward to see that if $\boldsymbol{b}_1$ and $\boldsymbol{b}_2$ are independent standard Gaussian random variables, then $\mathbb{E}[e^{a\boldsymbol{b}_1\boldsymbol{b}_2}] = \frac{1}{\sqrt{1-a^2}}$ for $|a| < 1$.[2] Then the Chernoff bound on $\mathbb{P}\big(\frac{1}{PT_n}\sum_{t=0}^{m-t_{0,n}-n_\delta} \mathrm{Re}(\boldsymbol{x}_{0,t})\boldsymbol{z}'_t \geq \frac{\epsilon'}{24d}\big)$ can be written as

$$\mathbb{P}\Big(\frac{1}{PT_n}\sum_{t=0}^{m-t_{0,n}-n_\delta} \mathrm{Re}(\boldsymbol{x}_{0,t})\boldsymbol{z}'_t \geq \frac{\epsilon'}{24d}\Big) \leq \Big(1 - \frac{a^2P}{4}\Big)^{-\frac{m'}{2}} e^{-\frac{a\epsilon' PT_n}{24d}}, \tag{A.49}$$

[2]We have $\mathbb{E}[e^{a\boldsymbol{b}_1\boldsymbol{b}_2}] = \frac{1}{2\pi}\iint_{\mathbb{R}^2} e^{-\frac{b_1^2+b_2^2-2ab_1b_2}{2}}\mathrm{d}b_1\mathrm{d}b_2$. Writing $b_1^2 + b_2^2 - 2ab_1b_2 = \begin{pmatrix} b_1 & b_2\end{pmatrix}^t \begin{pmatrix} 1 & -a \\ -a & 1\end{pmatrix}\begin{pmatrix} b_1 \\ b_2\end{pmatrix}$ and interpreting $\begin{pmatrix} 1 & -a \\ -a & 1\end{pmatrix}$ as the inverse covariance matrix of a bivariate Gaussian distribution yields the result.

where $m' = m - t_{0,n} - n_{\delta}$ and $0 < a < \frac{2}{\sqrt{P}}$. One can minimize the right side of (A.49) with respect to a to find the tightest upper bound. Instead, let us use the inequality $1 - b < e^{-b}$ for any real number b to get

$$\mathbb{P}\Big(\frac{1}{PT_n}\sum_{t=0}^{m-t_{0,n}-n_\delta}\mathrm{Re}(\boldsymbol{x}_{0,t})\boldsymbol{z}'_t \geq \frac{\epsilon'}{24d}\Big) \leq \big(e^{-\frac{a^2P}{4}}\big)^{-\frac{m'}{2}} e^{-\frac{a\epsilon' PT_n}{24d}} \leq e^{\frac{a^2PT_n}{8}-\frac{a\epsilon' PT_n}{24d}}, \quad \text{(A.50)}$$

where the last step is due to $m' \leq T_n$ for any choice of m. Let us assume $\epsilon' < \frac{8d}{\sqrt{P}}$ and select $a = \frac{\epsilon'}{4d}$. Then $\mathbb{P}\big(\frac{1}{PT_n}\sum_{t=0}^{m-t_{0,n}-n_\delta}\mathrm{Re}(\boldsymbol{x}_{0,t})\boldsymbol{z}'_t > \frac{\epsilon'}{24d}\big) \leq e^{-\frac{\epsilon'^2 PT_n}{384d^2}}$. Similarly, we obtain $\mathbb{P}\big(\frac{1}{PT_n}\sum_{t=0}^{n_\delta-1}\mathrm{Re}(\boldsymbol{x}'_{0,t})\boldsymbol{z}''_t \geq \frac{\epsilon'}{24d}\big) \leq e^{-\frac{\epsilon'^2 PT_n}{384d^2}}$. Using these upper bounds in (A.48),

$$\max_{\substack{m=1,\cdots,T_n\\ m\in\mathcal{E}_3}} \mathbb{P}\Big(\Big|\frac{1}{PT_n}\sum_{t=1}^{m}\mathrm{Re}(\boldsymbol{s}_0[t])\mathrm{Re}(\boldsymbol{z}_0^*[t])\Big| \geq \frac{\epsilon'}{12d}\Big) \leq 4e^{-\frac{\epsilon'^2 PT_n}{384d^2}}. \quad \text{(A.51)}$$

It is easy to see that $4e^{-\frac{\epsilon'^2 PT_n}{384c^2}}$ is also an upper bound on other terms in (A.46) for $i = 1, 2, 4$ under the constraint $\epsilon' < \frac{8d}{\sqrt{P}}$.

A.0.4 The term $\mathbb{P}\big(2d\sup_{\tau\in[0,1]}\big|\frac{1}{PT_n}\sum_{t=1}^{\lceil \tau T_n\rceil}\mathrm{Re}\left(\boldsymbol{s}_i[t]\boldsymbol{s}_j^*[t]\right)\big| \geq \epsilon'\big)$

The analysis for this term is quite similar to the analysis presented in the previous case and is omitted for brevity. The only difference is that we require $\epsilon' < 8d\sqrt{P}$.

We have shown that if $\epsilon' < 8d\min\big\{\sqrt{P}, \frac{1}{\sqrt{P}}\big\}$ and $\epsilon' T_n > \max\Big\{18dn_\delta, \frac{2(\frac{1}{P}+\sum_{i=0}^{K}|h_i|^2)}{K^2+5K+6}\Big\}$, then $\mathbb{P}(\|\boldsymbol{F}_n(\tau) - F(\tau)\|_\infty \geq \epsilon) \leq \Theta(1)e^{-\epsilon^2\Theta(n)}$. This completes the proof of Proposition 1.

Appendix B

Proof of Corollary 1

Appendix B; Proof of Corollary 1

Before proving Corollary 1, we observe a few facts about linear regression. Let (x_i, y_i) for $i = 1, \cdots, 2m$ be $2m$ points in the x-y plane that satisfy

$$ax_i + b_1 \leq y_i \leq ax_i + b_2, \quad 1 \leq i \leq 2m, \tag{B.1}$$

where a, b_1 and b_2 are constants. This implies that the points (x_i, y_i) are inside a strip whose boundaries are given by the lines $y = ax + b_1$ and $y = ax + b_2$. Let $\overline{x}$ and $\overline{y}$ be the arithmetic averages of x_i and y_i, respectively. We have the following:

- Fact 1: The point $(\overline{x}, \overline{y})$ lies on the regression line.
- Fact 2: Assume $x_i = ix_0$ where $x_0 > 0$ is a constant. By (B.1), $a\overline{x} + b_1 \leq \overline{y} \leq a\overline{x} + b_2$. Using these inequalities and the fact that the slope of the regression line is $\frac{\sum_{i=1}^{2m}(x_i-\overline{x})(y_i-\overline{y})}{\sum_{i=1}^{2m}(x_i-\overline{x})^2}$, it is easy to see that

$$\left|\frac{\sum_{i=1}^{2m}(x_i - \overline{x})(y_i - \overline{y})}{\sum_{i=1}^{2m}(x_i - \overline{x})^2} - a\right| \leq \frac{2(b_2 - b_1)\sum_{i=1}^{m}(\overline{x} - x_i)}{\sum_{i=1}^{2m}(x_i - \overline{x})^2} = \frac{6m(b_2 - b_1)}{(4m^2 - 1)x_0}, \tag{B.2}$$

  where the last step is due to $\sum_{i=1}^{m}(\overline{x} - x_i) = \frac{m^2}{2}x_0$ and $\sum_{i=1}^{2m}(x_i - \overline{x})^2 = \frac{m(4m^2-1)}{6}x_0^2$.

Next, let us describe the plot in Fig. B.1 and make some observations:

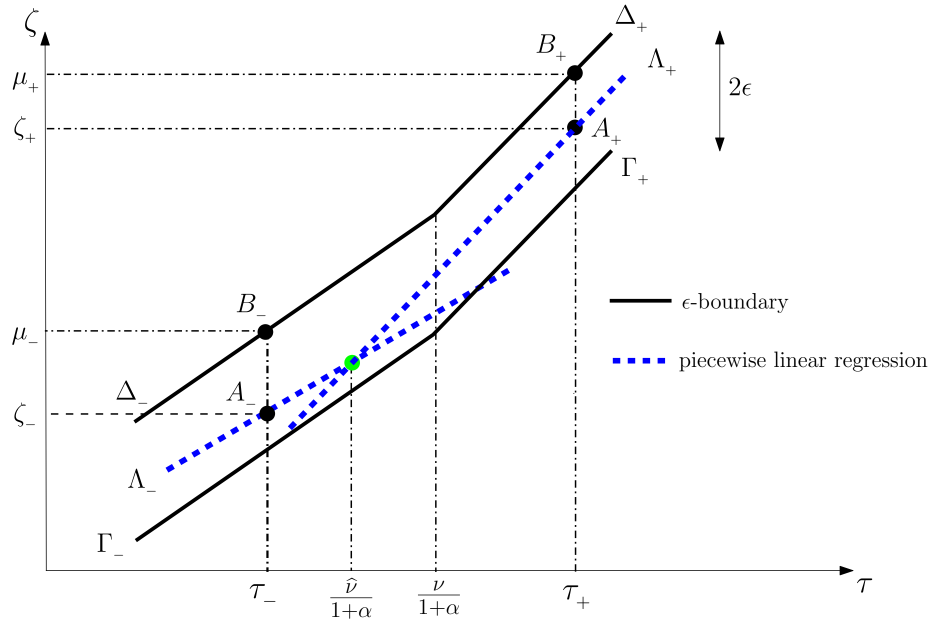


Figure B.1: The lines $\Delta_\pm$ and $\Gamma_\pm$ determine the ϵ-boundary for the function F in (2.6). In fact, $\Delta_- \bigcup \Delta_+ = \{(\tau, \zeta) : \zeta = F(\tau) + \epsilon\}$ and $\Gamma_- \bigcup \Gamma_+ = \{(\tau, \zeta) : \zeta = F(\tau) - \epsilon\}$. A new user has arrived at $\tau = \frac{\nu}{1+\alpha}$ which results in an increase in the slope of F. For given codeword length n, we assume receiver 0 is able to determine a number $N_n = \Theta(n)$ of points $(\tau_i^-, F_n(\tau_i^-))$ lying between Δ_- and Γ_-. The line Λ_- is the regression line for these points. Λ_- passes through the point $A_- = (\tau_-, \zeta_-)$ where τ_- and ζ_- are the arithmetic averages for τ_i^- and $F_n(\tau_i^-)$, respectively. The line Λ_+ and the point $A_+ = (\tau_+, \zeta_+)$ are defined similarly.

- The lines $\Delta_\pm$ and $\Gamma_\pm$ determine the ϵ-boundary for the function F in (2.6). In fact, $\Delta_- \bigcup \Delta_+ = \{(\tau, \zeta) : \zeta = F(\tau) + \epsilon\}$ and $\Gamma_- \bigcup \Gamma_+ = \{(\tau, \zeta) : \zeta = F(\tau) - \epsilon\}$. A new user has arrived at $\tau = \frac{\nu}{1+\alpha}$ which results in an increase in the slope of F.

- For given codeword length n, we assume receiver 0 is able to determine a number N_n of points $(\tau_i^-, F_n(\tau_i^-))$ for $1 \leq i \leq N_n$ lying between Δ_- and Γ_-. The line Λ_- is the regression line for these points. By Fact 1, Λ_- passes through the point $A_- = (\tau_-, \zeta_-)$ where τ_- and ζ_- are the arithmetic averages for τ_i^- and $F_n(\tau_i^-)$, respectively. The line Λ_+ and the point $A_+ = (\tau_+, \zeta_+)$ are defined similarly.

- $B_- = (\tau_-, \mu_-)$ and $B_+ = (\tau_+, \mu_+)$ are points on Δ_- and Δ_+ that have the same τ-coordinates as the points A_- and A_+, respectively.

- Let a_- and a_+ be the slopes of regression lines Δ_- and Δ_+, respectively. Then $\frac{\nu}{1+\alpha}$ is the τ-coordinate for the point of intersection of Δ_- and Δ_+, i.e., ν satisfies $a_-(\nu/(1+\alpha)-\tau_-)+\mu_- = a_+(\nu/(1+\alpha)-\tau_+)+\mu_+$. This yields

$$\nu = \frac{1+\alpha}{a_+ - a_-}(a_+\tau_+ - a_-\tau_- + \mu_- - \mu_+). \tag{B.3}$$

- Let $\widehat{a}_-$ and $\widehat{a}_+$ be the slopes of regression lines Λ_- and Λ_+, respectively. By definition, $\frac{\widehat{\nu}}{1+\alpha}$ is the τ-coordinate of the point of intersection of Λ_- and Λ_+, i.e., $\widehat{\nu}$ satisfies $\widehat{a}_-(\widehat{\nu}/(1+\alpha)-\tau_-)+\zeta_- = \widehat{a}_+(\widehat{\nu}/(1+\alpha)-\tau_+)+\zeta_+$. This yields

$$\widehat{\nu} = \frac{1+\alpha}{\widehat{a}_+ - \widehat{a}_-}(\widehat{a}_+\tau_+ - \widehat{a}_-\tau_- + \zeta_- - \zeta_+). \tag{B.4}$$

By Fact 2 and selecting $m = N_n$ and $x_0 = \frac{1}{T_n}$ in (B.2), we have $|\widehat{a}^\pm - a^\pm| \le \frac{12N_nT_n}{4N_n^2-1}\epsilon$. Let n be sufficiently large so that $N_n = \Theta(n)$, i.e., receiver 0 can identify a fraction, say one third, of the graph of F_n that lies within each arm of the ϵ-strip in Fig. B.1. For example, Fig. 2.2 shows that for large n, F_n is highly concentrated around F and the arms of the ϵ-strip can be identified. Then $d := \sup_n \frac{12N_nT_n}{4N_n^2-1}$ is finite and we get

$$|\widehat{a}^\pm - a^\pm| \le d\epsilon. \tag{B.5}$$

Moreover,

$$|\zeta_- - \zeta_+ - (\mu_- - \mu_+)| \le |\zeta_- - \mu_-| + |\zeta_+ - \mu_+| \le 2\epsilon + 2\epsilon = 4\epsilon. \tag{B.6}$$

By (B.4),

$$\begin{aligned}
\widehat{\nu} &= \frac{1+\alpha}{\widehat{a}_+ - \widehat{a}_-}\Big((\widehat{a}_+ - a_+)\tau_+ - (\widehat{a}_- - a_-)\tau_- + \zeta_- - \zeta_+ - (\mu_- - \mu_+)\Big) \\
&\quad + \frac{1+\alpha}{\widehat{a}_+ - \widehat{a}_-}(a_+\tau_+ - a_-\tau_- + \mu_- - \mu_+) \\
&= \frac{1+\alpha}{\widehat{a}_+ - \widehat{a}_-}\Big((\widehat{a}_+ - a_+)\tau_+ - (\widehat{a}_- - a_-)\tau_- + \zeta_- - \zeta_+ - (\mu_- - \mu_+)\Big) + \frac{a_+ - a_-}{\widehat{a}_+ - \widehat{a}_-}\nu,
\end{aligned} \tag{B.7}$$

where the last step is due to (B.3). Without loss of generality, assume $\epsilon < \frac{1}{4d} \min_{0\le i\le K} |h_i|^2$. By (B.7) and using triangle inequality,

$$\begin{aligned}
|\widehat{\nu} - \nu| &\le \frac{1+\alpha}{|\widehat{a}_+ - \widehat{a}_-|}\Big(|\widehat{a}_+ - a_+|\tau_+ + |\widehat{a}_- - a_-|\tau_- + |\zeta_- - \zeta_+ - (\mu_- - \mu_+)|\Big) \\
&\quad + \frac{\nu}{|\widehat{a}_+ - \widehat{a}_-|}\big(|\widehat{a}_+ - a_+| + |\widehat{a}_- - a_-|\big) \\
&\overset{(a)}{\le} \frac{(1+\alpha)(d\epsilon + d\epsilon + 4\epsilon)}{|\widehat{a}_+ - \widehat{a}_-|} + \frac{\nu(d\epsilon + d\epsilon)}{|\widehat{a}_+ - \widehat{a}_-|} \\
&\overset{(b)}{\le} \frac{(1+\alpha)(2d+4) + 2d\nu}{a_+ - a_- - 2d\epsilon}\epsilon \\
&\overset{(c)}{\le} \frac{2(3d+4)}{\min_{0\le i\le K} |h_i|^2 - 2d\epsilon}\epsilon \\
&\overset{(d)}{\le} \frac{4(3d+4)}{\min_{0\le i\le K} |h_i|^2}\epsilon, \qquad \text{(B.8)}
\end{aligned}$$

where (a) is due to (B.5), (B.6) and the fact that $0 < \tau_-, \tau_+ < 1$, (b) is due to $|\widehat{a}_+ - \widehat{a}_-| \ge |a_+ - a_-| - |\widehat{a}_+ - a_+| - |\widehat{a}_- - a_-| \ge |a_+ - a_-| - 2d\epsilon$, (c) is due to the fact that $a_+ - a_- = |h_i|^2$ for some $0 \le i \le K$ and (d) is due to $\epsilon < \frac{1}{4d}\min_{0\le i\le K} |h_i|^2$.

We conclude that if $\|F_n - F\|_\infty \le \epsilon$, there are constants $d_1 = 2$ and $d_2 = \frac{4(3d+4)}{\min_{0\le i\le K} |h_i|^2}$ such that $\big||\widehat{h}_{k_i}|^2 - |h_{k_i}|^2\big| \le d_1\epsilon$ and $|\widehat{\nu}_{k_i} - \nu_{k_i}| \le d_2\epsilon$. Therefore,

$$\begin{aligned}
\mathbb{P}\big(\big||\widehat{\boldsymbol{h}}_{k_i}|^2 - |h_{k_i}|^2\big| > \epsilon\big) &\le \mathbb{P}\Big(\|\boldsymbol{F}_n - F\|_\infty > \frac{\epsilon}{d_1}\Big) \\
&\le \Theta(1)e^{-\frac{\epsilon^2}{d_1^2}\Theta(n)} \\
&= \Theta(1)e^{-\epsilon^2\Theta(n)}, \qquad \text{(B.9)}
\end{aligned}$$

where the penultimate step follows by Proposition 1. Similarly, $\mathbb{P}\big(\big|\widehat{\boldsymbol{\nu}}_{k_i} - \nu_{k_i}\big| > \epsilon\big) \le \Theta(1)e^{-\epsilon^2\Theta(n)}$. This completes the proof of Corollary 1.

Appendix C

Proof of Proposition 3

Let $i \neq j$. We only show that $\mathbf{A}_{i,n} \bigcap \mathbf{B}_{j,n} = \emptyset$ with high probability if n is sufficiently large. A similar reasoning establishes the same fact if one replaces $\mathbf{A}_{i,n} \bigcap \mathbf{B}_{j,n} = \emptyset$ by $\mathbf{A}_{i,n} \bigcap \mathbf{C}_{j,n} = \emptyset$. Throughout this appendix, we let $\Delta\widehat{\boldsymbol{t}}_n = \widehat{\boldsymbol{t}}_{k_j,n} - \widehat{\boldsymbol{t}}_{k_i,n}$, $\Delta t_n = t_{k_j,n} - t_{k_i,n}$, $\Delta\nu = \nu_{k_j} - \nu_{k_i}$ and $\boldsymbol{d}_{i,n} = |\widehat{\boldsymbol{t}}_{k_i,n} - t_{k_i,n}|$. Define

$$\mathcal{T}_{N} := \bigcap_{n \geq N} \{\mathbf{A}_{i,n} \bigcap \mathbf{B}_{j,n} = \emptyset\}, \tag{C.1}$$

for any $N \geq 1$. Note that $\mathbf{A}_{i,n} \bigcap \mathbf{B}_{j,n} = \emptyset$ if and only if $\Delta\widehat{\boldsymbol{t}}_n < -4cn_\delta$ or $\Delta\widehat{\boldsymbol{t}}_n > (4c+1)n_\delta$.

$$\mathcal{T}_{N}^c = \bigcup_{n \geq N} \{\mathbf{A}_{i,n} \bigcap \mathbf{B}_{j,n} \neq \emptyset\} \subseteq \bigcup_{n \geq N} \{|\Delta\widehat{\boldsymbol{t}}_n| \leq (4c+1)n_\delta\} \tag{C.2}$$

Let us investigate the events $\{|\Delta\widehat{\boldsymbol{t}}_n| \leq (4c+1)n_\delta\}$ that appears in (C.2). We can write[1]

$$\begin{aligned} \{\, |\Delta\widehat{\boldsymbol{t}}_n| \leq (4c+1)n_\delta\} \;\subseteq\; & \{\, |\Delta\widehat{\boldsymbol{t}}_n| \leq (4c+1)n_\delta, \boldsymbol{d}_{i,n} \leq 2cn_\delta, \boldsymbol{d}_{j,n} \leq 2cn_\delta\} \\ & \bigcup \{\boldsymbol{d}_{i,n} > 2cn_\delta\} \bigcup \{\boldsymbol{d}_{j,n} > 2cn_\delta\}. \end{aligned} \tag{C.3}$$

Using triangle inequality, $|\Delta t_n| \leq \Delta\widehat{\boldsymbol{t}}_n + \boldsymbol{d}_{i,n} + \boldsymbol{d}_{j,n}$. This together with (C.3) yields

$$\{\, |\Delta\widehat{\boldsymbol{t}}_n| \leq (4c+1)n_\delta\} \subseteq \{\, |\Delta t_n| \leq (8c+1)n_\delta\} \bigcup \{\boldsymbol{d}_{i,n} > 2cn_\delta\} \bigcup \{\boldsymbol{d}_{j,n} > 2cn_\delta\}. \tag{C.4}$$

[1] We are using the fact that $\mathcal{A} \subseteq (\mathcal{A} \bigcap \mathcal{B}) \bigcup \mathcal{B}^c$ for events $\mathcal{A}, \mathcal{B}$.

We have[2] $|\Delta t_n - \frac{n}{\alpha}\Delta\nu| < 2$. Hence,

$$\{\, |\Delta t_n| \leq (8c+1)n_{\delta} \} \subseteq \{\, |\Delta\nu| < \alpha((8c+1)n_{\delta} + 2)/n \} \tag{C.5}$$

By assumption $|\Delta\nu| > 0$. Noting that $\lim_{n\to\infty} \frac{(8c+1)n_{\delta}+2}{n} = 0$, there exists an integer N', depending on α, δ, c and $\Delta\nu$, such that for $n \geq N'$, the set on the right side of (C.5) becomes empty. Then by (C.2) and (C.4),

$$\mathbb{P}(\mathcal{T}_{N}^{c}) \leq \sum_{n\geq N} \big(\mathbb{P}(\boldsymbol{d}_{i,n} > 2cn_{\delta}) + \mathbb{P}(\boldsymbol{d}_{j,n} > 2cn_{\delta})\big), \tag{C.6}$$

for any $N \geq N'$. By Proposition 2, one can further bound the right side of (C.6) to get

$$\mathbb{P}(\mathcal{T}_{N}^{c}) \leq \Theta(1) \sum_{n\geq N} e^{-\Theta(n^{2\delta-1})}. \tag{C.7}$$

Since the right side of (C.7) is the tail of a convergent series, we conclude that for arbitrary $\epsilon > 0$, there is $N_{\epsilon} \geq 1$ such that $\mathbb{P}(\mathcal{T}_{N}^{c}) < \epsilon$ if $N \geq N_{\epsilon}$.

[2] By the floor inequality, $\Delta t_m = t_{k_j,m} - t_{k_i,m} = \lfloor \lfloor \frac{m}{\alpha} \rfloor \nu_{k_j} \rfloor - \lfloor \lfloor \frac{m}{\alpha} \rfloor \nu_{k_i} \rfloor > (\lfloor \frac{m}{\alpha} \rfloor \nu_{k_j} - 1) - \lfloor \frac{m}{\alpha} \rfloor \nu_{k_i} > ((\frac{m}{\alpha} - 1)\nu_{k_j} - 1) - \frac{m}{\alpha}\nu_{k_i} = \frac{m}{\alpha}\Delta\nu - \nu_{k_j} - 1 > \frac{m}{\alpha}\Delta\nu - 2$. Similarly, $\Delta t_m < \frac{m}{\alpha}\Delta\nu + 2$.

Appendix D

Proof of Proposition 4

We prove the proposition for the situation in Fig. 2.2. We only consider the case that $k_0 = 0$, i.e., transmitter 0 is the first active transmitter. The cases $k_1 = 0$ and $k_2 = 0$ can be treated similarly. If p_{NI} is selected as in (2.21), then $p(x, y) = \mathrm{g}(x; P)p_{\text{NI}}(y - h_0 x)$ is the actual PDF of noise plus interference during transmission of $(\boldsymbol{x}'_{0,l})_{l=0}^{n_\delta - 1}$. Setting $t = t_{0,n}$ in the sequential typicality decoding rule in (2.24), we have

$$\mathbb{P}(\widehat{\hat{\boldsymbol{t}}}_{0,n} \leq t_{0,n}) \geq \mathbb{P}\left(\left((\boldsymbol{x}'_{0,l})_{l=0}^{n_\delta - 1}, (\boldsymbol{y}_0[l])_{l=t_{0,n}}^{t_{0,n}+n_\delta - 1}\right) \in A_\epsilon^{(n_\delta)}[p]\right). \tag{D.1}$$

Then

$$\lim_{n\to\infty} \mathbb{P}(\widehat{\hat{\boldsymbol{t}}}_{0,n} \leq t_{0,n}) \geq \lim_{n\to\infty} \mathbb{P}\left(\left((\boldsymbol{x}'_{0,l})_{l=0}^{n_\delta - 1}, (\boldsymbol{y}_0[l])_{l=t_{0,n}}^{t_{0,n}+n_\delta - 1}\right) \in A_\epsilon^{(n_\delta)}[p]\right) = 1, \tag{D.2}$$

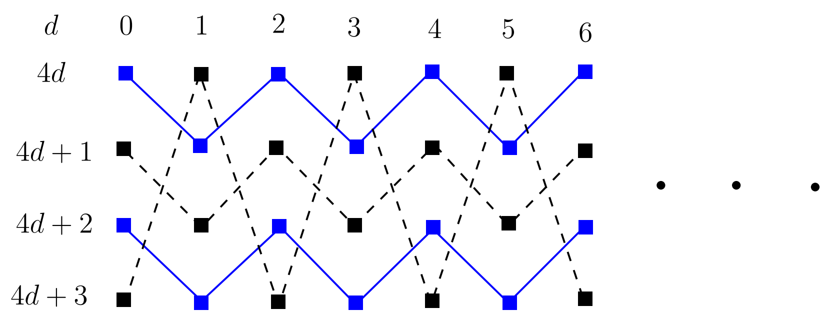


Figure D.1: Schematic diagram for the partitions $\mathcal{S}_e$ (all points crossed by solid lines) and $\mathcal{S}_o$ (all points crossed by dashed lines) in (D.9) and (D.10) for $m = 2$. The first row represents the numbers $4d$ for $d = 0, 1, \cdots$, the second row represents the numbers $4d + 1$ for $d = 0, 1, \cdots$ and so on.

where the last step is due to the weak law of large numbers. As such, in order to show $\lim_{n\to\infty} \mathbb{P}(\widehat{\widehat{\boldsymbol{t}}}_{0,n} = t_{0,n}) = 1$, it is enough to verify $\lim_{n\to\infty} \mathbb{P}(\widehat{\widehat{\boldsymbol{t}}}_{0,n} < t_{0,n}) = 0$. Let us write

$$\begin{aligned}
\mathbb{P}(\widehat{\widehat{\boldsymbol{t}}}_{0,n} < t_{0,n}) &= \mathbb{P}\big(\exists\, t \in \mathbf{B}_{0,n} : t < t_{0,n} \text{ and } \big((\boldsymbol{x}'_{0,l})_{l=0}^{n_\delta - 1}, (\boldsymbol{y}_0[l])_{l=t}^{t+n_\delta-1}\big) \in A_\epsilon^{(n_\delta)}[p]\big) \\
&= \sum_{s=t_{0,n}-2cn_\delta}^{t_{0,n}+2cn_\delta} \mathbb{P}\big(\exists\, t \in \mathbf{B}_{0,n} : t < t_{0,n} \text{ and } \big((\boldsymbol{x}'_{0,l})_{l=0}^{n_\delta - 1}, (\boldsymbol{y}_0[l])_{l=t}^{t+n_\delta-1}\big) \in A_\epsilon^{(n_\delta)}[p], \widehat{\boldsymbol{t}}_{0,n} = s\big) \\
&\le \sum_{s=t_{0,n}-2cn_\delta}^{t_{0,n}+2cn_\delta} \mathbb{P}\big(\exists\, t \in \mathsf{B}_{0,n}^{(s)} : t < t_{0,n} \text{ and } \big((\boldsymbol{x}'_{0,l})_{l=0}^{n_\delta - 1}, (\boldsymbol{y}_0[l])_{l=t}^{t+n_\delta-1}\big) \in A_\epsilon^{(n_\delta)}[p]\big) \\
&\le \sum_{s=t_{0,n}-2cn_\delta}^{t_{0,n}+2cn_\delta} \sum_{t \in \mathsf{B}_{0,n}^{(s)} : t < t_{0,n}} \mathbb{P}\big(\big((\boldsymbol{x}'_{0,l})_{l=0}^{n_\delta - 1}, (\boldsymbol{y}_0[l])_{l=t}^{t+n_\delta-1}\big) \in A_\epsilon^{(n_\delta)}[p]\big),
\end{aligned} \tag{D.3}$$

where for any $t_{0,n} - -2nc_\delta \le s \le t_{0,n} + -2nc_\delta$, we define $\mathsf{B}_{0,n}^{(s)} := \mathsf{B}_{0,n}\big|_{\widehat{t}_{0,n}=s}$. To compute an upper bound on each term on the right side of (D.3), we study the cases $t \le t_{0,n} - n_\delta$ and $t_{0,n} - n_\delta + 1 \le t \le t_{0,n} - 1$, separately:

- Let $t \le t_{0,n} - n_\delta$. Then $(\boldsymbol{x}'_{0,l})_{l=0}^{n_\delta - 1}$ and $(\boldsymbol{y}_0[l])_{l=t}^{t+n_\delta-1}$ are independent as $t + n_\delta - 1 \le t_{0,n} - 1$ and transmitter 0 is not active before time slot $t_{0,n}$. We explicitly write the constraint defining $A_\epsilon^{(n_\delta)}[p]$ in (2.20) as the set of all $(\vec{x}, \vec{y})$ such that

$\left|\frac{1}{n_\delta}\sum_{l=0}^{n_\delta-1}\left(\frac{1}{P}|x_l|^2+\left|y_l-h_0x_l\right|^2-2\right)\right|<\frac{\epsilon}{\log e}$. Then we have the thread of inequalities

$$
\begin{aligned}
&\mathbb{P}\Big(\,\big((\boldsymbol{x}'_{0,l})_{l=0}^{n_\delta-1},(\boldsymbol{y}_0[l])_{l=t}^{t+n_\delta-1}\big)\in A_\epsilon^{(n_\delta)}[p]\Big)\\
&\le \mathbb{P}\Big(\Big|\frac{1}{n_\delta}\sum_{l=0}^{n_\delta-1}\Big(\frac{1}{P}|\boldsymbol{x}'_{0,l}|^2+|\boldsymbol{y}_0[l+t]-h_0\boldsymbol{x}'_{0,l}|^2-2\Big)\Big|<\frac{\epsilon}{\log e}\Big)\\
&\le \mathbb{P}\Big(\frac{1}{n_\delta}\sum_{l=0}^{n_\delta-1}\Big(\frac{1}{P}|\boldsymbol{x}'_{0,l}|^2+\big|\boldsymbol{y}_0[l+t]-h_0\boldsymbol{x}'_{0,l}\big|^2-2\Big)<\frac{\epsilon}{\log e}\Big)\\
&\overset{(a)}{=} \mathbb{P}\Big(\frac{1}{n_\delta}\sum_{l=0}^{n_\delta-1}\Big(\frac{1}{P}|\boldsymbol{x}'_{0,l}|^2+\big|\boldsymbol{z}_0[l+t]-h_0\boldsymbol{x}'_{0,l}\big|^2-2\Big)<\frac{\epsilon}{\log e}\Big)\\
&\overset{(b)}{\le} \mathbb{P}\Big(\frac{1}{n_\delta}\sum_{l=0}^{n_\delta-1}\Big(2+|h_0|^2P-\frac{1}{P}|\boldsymbol{x}'_{0,l}|^2-\big|\boldsymbol{z}_0[l+t]-h_0\boldsymbol{x}'_{0,l}\big|^2\Big)>|h_0|^2P-\frac{\epsilon}{\log e}\Big)\\
&\overset{(c)}{=} \mathbb{P}\Big(\frac{1}{n_\delta}\sum_{l=0}^{n_\delta-1}\Big(1-\frac{\frac{1}{P}|\boldsymbol{x}'_{0,l}|^2+\big|\boldsymbol{z}_0[l+t]-h_0\boldsymbol{x}'_{0,l}\big|^2}{2+|h_0|^2P}\Big)>\frac{|h_0|^2P-\frac{\epsilon}{\log e}}{2+|h_0|^2P}\Big)\\
&\overset{(d)}{\le} e^{-\Theta(n_\delta)},
\end{aligned}
\tag{D.4}
$$

where (a) is due to the fact that the signal at receiver 0 during time slots t to $t+n_\delta-1$ consists of ambient noise only, in (b) we have added $|h_0|^2P$ to both sides of the inequality, in (c) both sides are divided by $2+|h_0|^2P$ and (d) is due to Lemma 2 (Bernstein's inequality)[1] and it is assumed that $\epsilon<|h_0|^2P\log e$.

[1]The random variables $\boldsymbol{a}_l=1-\frac{\frac{1}{P}|\boldsymbol{x}'_{0,l}|^2+\left|\boldsymbol{z}_0[l+t]-h_0\boldsymbol{x}'_{0,l}\right|^2}{2+|h_0|^2P}$ are zero mean with finite variance and $\boldsymbol{a}_l\le 1$. Therefore, one can apply Bernstein's inequality.

- Let $t_{0,n} - n_\delta + 1 \leq t \leq t_{0,n} - 1$. Then

$$\begin{aligned}
&\mathbb{P}\Big(\big((\boldsymbol{x}'_{0,l})_{l=0}^{n_\delta -1}, (\boldsymbol{y}_0[l])_{l=t}^{t+n_\delta-1}\big) \in A_\epsilon^{(n_\delta)}[p]\Big)\\
&\leq \mathbb{P}\Big(\frac{1}{n_\delta}\sum_{l=0}^{n_\delta-1}\Big(\frac{1}{P}|\boldsymbol{x}'_{0,l}|^2 + \big|\boldsymbol{y}_0[l+t] - h_0\boldsymbol{x}'_{0,l}\big|^2 - 2\Big) < \frac{\epsilon}{\log e}\Big)\\
&\overset{(a)}{=} \mathbb{P}\Big(\frac{1}{n_\delta}\sum_{l=0}^{t_{0,n}-t-1}\Big(\frac{1}{P}|\boldsymbol{x}'_{0,l}|^2 + \big|\boldsymbol{z}_0[l+t] - h_0\boldsymbol{x}'_{0,l}\big|^2 - 2\Big)\\
&\qquad + \frac{1}{n_\delta}\sum_{l=t_{0,n}-t}^{n_\delta-1}\Big(\frac{1}{P}|\boldsymbol{x}'_{0,l}|^2 + \big|h_0\boldsymbol{x}'_{0,l-t_{0,n}+t} + \boldsymbol{z}_0[l+t] - h_0\boldsymbol{x}'_{0,l}\big|^2 - 2\Big) < \frac{\epsilon}{\log e}\Big)\\
&\overset{(b)}{=} \mathbb{P}\Big(\frac{1}{n_\delta}\sum_{l=0}^{t_{0,n}-t-1}\Big(\frac{1}{P}|\boldsymbol{x}'_{0,l}|^2 + \big|\boldsymbol{z}_0[l+t] - h_0\boldsymbol{x}'_{0,l}\big|^2 - 2\Big) < \frac{\epsilon}{2\log e}\Big)\\
&\quad + \mathbb{P}\Big(\frac{1}{n_\delta}\sum_{l=t_{0,n}-t}^{n_\delta-1}\Big(\frac{1}{P}|\boldsymbol{x}'_{0,l}|^2 + \big|h_0\boldsymbol{x}'_{0,l-t_{0,n}+t} + \boldsymbol{z}_0[l+t] - h_0\boldsymbol{x}'_{0,l}\big|^2 - 2\Big) < \frac{\epsilon}{2\log e}\Big),
\end{aligned} \tag{D.5}$$

where (a) follows by noting that $\boldsymbol{y}_0[l+t]$ is the ambient noise for $l+t < t_{0,n}$ and (b) is due to the fact that if the sum of two numbers is less than $\frac{\epsilon}{\log e}$, then at least one of them is less than $\frac{\epsilon}{2\log e}$. Following a similar reasoning for the last step in (I.3), the first term on the right side of (I.4) is bounded from above by $e^{-\frac{\Theta(n_\delta^2)}{t_{0,n}-t}} \leq e^{-\Theta(n_\delta)}$ due to $t_{0,n} - t \leq n_\delta$. As for the second term on the right side of (I.4), we note that the summands are no longer independent. For simplicity, let us define

$$\boldsymbol{w}_l := \frac{1}{P}|\boldsymbol{x}'_{0,l}|^2 + \big|h_0\boldsymbol{x}'_{l-t_{0,n}+t} + \boldsymbol{z}_0[l+t] - h_0\boldsymbol{x}'_{0,l}\big|^2 - 2. \tag{D.6}$$

Any two $\boldsymbol{w}_l$ and $\boldsymbol{w}_{l'}$ are dependent if and only if $|l - l'| = t_{0,n} - t$. As such, one can not directly apply the concentration inequality of Lemma 2. To circumvent this difficulty, we use a trick in Appendix 24B in [20]. We consider two cases:

- If $t_{0,n} - t$ is odd, the terms with odd indices are independent. Similarly, the terms with even indices are independent. Then

$$\mathbb{P}\Big(\frac{1}{n_\delta}\sum_{l=t_{0,n}-t}^{n_\delta-1}\boldsymbol{w}_l < \frac{\epsilon}{2\log e}\Big) \;\leq\; \mathbb{P}\Big(\frac{1}{n_\delta}\sum_{\substack{l=t_{0,n}-t\\ l\text{ even}}}^{n_\delta-1}\boldsymbol{w}_l < \frac{\epsilon}{4\log e}\Big) + \mathbb{P}\Big(\frac{1}{n_\delta}\sum_{\substack{l=t_{0,n}-t\\ l\text{ odd}}}^{n_\delta-1}\boldsymbol{w}_l < \frac{\epsilon}{4\log e}\Big). \quad \text{(D.7)}$$

At this point, similar to (I.3), one can apply Bernstein's inequality to conclude that each term on the right side of (I.10) is bounded from above by $e^{-\Theta(n_\delta)}$.

– If $t_{0,n} - t$ is even, we need to partition the set of integers to a finite number of disjoint sets such that the difference of any two element in each set is not equal to $t_{0,n} - t$. Let $t_{0,n} - t = 2m$. It is easy to see that the required partition is given by

$$\{0, 1, 2, 3, \cdots\} = \mathcal{S}_e \bigcup \mathcal{S}_o, \quad \text{(D.8)}$$

where

$$\mathcal{S}_e = \bigcup_{\substack{b=0\\ b\text{ even}}}^{2m-1}\Big(\Big\{2ma + b : a = 0, 2, 4, \cdots\Big\}\bigcup\Big\{2ma + b + 1 : a = 1, 3, 5, \cdots\Big\}\Big) \quad \text{(D.9)}$$

and

$$\mathcal{S}_o = \bigcup_{\substack{b=0\\ b\text{ odd}}}^{2m-1}\Big(\Big\{2ma + b : a = 0, 2, 4, \cdots\Big\}\bigcup\Big\{2ma + b + 1 : a = 1, 3, 5, \cdots\Big\}\Big). \quad \text{(D.10)}$$

For example, Fig. D.1 presents the schematic diagram of $\mathcal{S}_e$ and $\mathcal{S}_o$ for $m = 2$.

We can write

$$\mathbb{P}\Big(\frac{1}{n_\delta}\sum_{l=t_{0,n}-t}^{n_\delta-1}\boldsymbol{w}_l<\frac{\epsilon}{2\log e}\Big) \;\le\; \mathbb{P}\Big(\frac{1}{n_\delta}\sum_{\substack{l=t_{0,n}-t\\ l\,\in\,\mathcal{S}_e}}^{n_\delta-1}\boldsymbol{w}_l<\frac{\epsilon}{4\log e}\Big) + \mathbb{P}\Big(\frac{1}{n_\delta}\sum_{\substack{l=t_{0,n}-t\\ l\,\in\,\mathcal{S}_o}}^{n_\delta-1}\boldsymbol{w}_l<\frac{\epsilon}{4\log e}\Big). \quad \text{(D.11)}$$

As each of $\sum_{l\,\in\,\mathcal{S}_e}\boldsymbol{w}_l$ and $\sum_{l\,\in\,\mathcal{S}_o}\boldsymbol{w}_l$ are sums of independent random variables, it follows that each term on the right side of (I.11) is bounded from above by $e^{-\Theta(n_\delta)}$.

We conclude that whether $t\le t_{0,n}-n_\delta$ or $t_{0,n}-n_\delta+1\le t\le t_{0,n}-1$, each term on the right side of (D.3) is bounded from above by $e^{-\Theta(n_\delta)}$. Since there are at most $\Theta(n_\delta^2)$ terms on the right side of (D.3), we get $\mathbb{P}(\widehat{\widehat{\boldsymbol{t}}}_{0,n}<t_{0,n})\le\Theta(n_\delta^2)e^{-\Theta(n_\delta)}$ which vanishes as n grows. This completes the proof of Proposition 4.

Appendix E

Proof of Proposition 5

We only consider the case where $0 < \nu_1 - \nu_0 < \alpha$. Other cases can be handled similarly. The proof is carried out in two steps:

- Let us show that (2.32) and (2.33) hold for the transmitted codeword with a probability that approaches 1 as n grows to infinity. We only verify (2.32). Verification of (2.33) is quite similar and is omitted for brevity. As mentioned earlier, let n be sufficiently large so that the right side of (2.26) is positive and hence, the initial interval is nonempty regardless of the value of $t_{1,n} - 2cn_{\delta} \le \widehat{t}_{1,n} \le t_{1,n} + 2cn_{\delta}$. Let

$(\boldsymbol{x}_{0,l})_{l=0}^{n-1}$ be the codeword sent by transmitter 0. We have

$$\begin{aligned}
&\mathbb{P}\Big(\big((\boldsymbol{x}_{0,l})_{l=0}^{\widehat{\boldsymbol{t}}_{1,n}-t_{0,n}-(2c+1)n_\delta-1},(\boldsymbol{y}_0[l])_{l=t_{0,n}+n_\delta}^{\widehat{\boldsymbol{t}}_{1,n}-2cn_\delta-1}\big)\notin A_\epsilon^{(\widehat{\boldsymbol{t}}_{1,n}-t_{0,n}-(2c+1)n_\delta)}[p_{\mathrm{I}}]\Big)\\
&=\sum_{t=t_{1,n}-2cn_\delta}^{t_{1,n}+2cn_\delta}\mathbb{P}\Big(\big((\boldsymbol{x}_{0,l})_{l=0}^{t-t_{0,n}-(2c+1)n_\delta-1},(\boldsymbol{y}_0[l])_{l=t_{0,n}+n_\delta}^{t-2cn_\delta-1}\big)\notin A_\epsilon^{(t-t_{0,n}-(2c+1)n_\delta)}[p_{\mathrm{I}}],\widehat{\boldsymbol{t}}_{1,n}=t\Big)\\
&\leq\sum_{t=t_{1,n}-2cn_\delta}^{t_{1,n}+2cn_\delta}\mathbb{P}\Big(\big((\boldsymbol{x}_{0,l})_{l=0}^{t-t_{0,n}-(2c+1)n_\delta-1},(\boldsymbol{y}_0[l])_{l=t_{0,n}+n_\delta}^{t-2cn_\delta-1}\big)\notin A_\epsilon^{(t-t_{0,n}-(2c+1)n_\delta)}[p_{\mathrm{I}}]\Big)\\
&=\sum_{t=t_{1,n}-2cn_\delta}^{t_{1,n}+2cn_\delta}\mathbb{P}\Bigg(\Big|\frac{1}{t-t_{0,n}-(2c+1)n_\delta}\\
&\qquad\times\sum_{l=0}^{t-t_{0,n}-(2c+1)n_\delta-1}\log p_{\mathrm{I}}(\boldsymbol{x}_{0,l},\boldsymbol{y}_0[l+t_{0,n}+n_\delta])-h(p_{\mathrm{I}})\Big|\geq\epsilon\Bigg).
\end{aligned}\tag{E.1}$$

For each time slot during the initial interval, p_{I} in (2.30) is the actual joint PDF between the transmitted code symbol by transmitter 0 and the received signal by receiver 0. In fact, for any $t_{1,n}-2cn_\delta\leq t\leq t_{1,n}+2cn_\delta$, the random variables $\log p_{\mathrm{I}}(\boldsymbol{x}_{0,l},\boldsymbol{y}_0[l+t_{0,n}+n_\delta])$ for $0\leq l\leq t-t_{0,n}-(2c+1)n_\delta-1$ are i.i.d. with common expectation $h(p_{\mathrm{I}})$. Fixing t and assuming the variance of $\log p_{\mathrm{I}}(\boldsymbol{x}_{0,0},\boldsymbol{y}_0[t_{0,n}+n_\delta])$ is finite, one can apply the Chebyshev's inequality [29] to get

$$\begin{aligned}
&\mathbb{P}\Bigg(\Big|\frac{1}{t-t_{0,n}-(2c+1)n_\delta}\sum_{l=0}^{t-t_{0,n}-(2c+1)n_\delta-1}\log p_{\mathrm{I}}(\boldsymbol{x}_{0,l},\boldsymbol{y}_0[l+t_{0,n}+n_\delta])-h(p_{\mathrm{I}})\Big|\geq\epsilon\Bigg)\\
&\qquad\leq\frac{\mathrm{var}(\log p_{\mathrm{I}}(\boldsymbol{x}_{0,0},\boldsymbol{y}_0[t_{0,n}+n_\delta]))}{(t-t_{0,n}-(2c+1)n_\delta)\epsilon},\quad t_{1,n}-2cn_\delta\leq t\leq t_{1,n}+2cn_\delta.
\end{aligned}\tag{E.2}$$

Denoting $\mathrm{var}(\log p_{\mathrm{I}}(\boldsymbol{x}_{0,0}, \boldsymbol{y}_0[t_{0,n}+n_\delta]))$ by $\Theta(1)$ and by (E.1) and (E.2), we get

$$
\begin{aligned}
&\mathbb{P}\Big(\big((\boldsymbol{x}_{0,l})_{l=0}^{\widehat{\boldsymbol{t}}_{1,n}-t_{0,n}-(2c+1)n_\delta-1}, (\boldsymbol{y}_0[l])_{l=t_{0,n}+n_\delta}^{\widehat{\boldsymbol{t}}_{1,n}-2cn_\delta-1}\big) \notin A_\epsilon^{(\widehat{\boldsymbol{t}}_{1,n}-t_{0,n}-(2c+1)n_\delta)}[p_{\mathrm{I}}]\Big) \\
&\le \sum_{t=t_{1,n}-2cn_\delta}^{t_{1,n}+2cn_\delta} \frac{\Theta(1)}{(t-t_{0,n}-(2c+1)n_\delta)\epsilon} \\
&\le \frac{\Theta(1)}{((t_{1,n}-2cn_\delta)-t_{0,n}-(2c+1)n_\delta)\epsilon} \sum_{t=t_{1,n}-2cn_\delta}^{t_{1,n}+2cn_\delta} 1 \\
&= \frac{(4cn_\delta+1)\Theta(1)}{(t_{1,n}-t_{0,n}-(4c+1)n_\delta)\epsilon}.
\end{aligned}
\tag{E.3}
$$

Since $t_{1,n} - t_{0,n} = \Theta(n)$, the right side of (E.3) tends to zero as n grows. This completes the proof of (2.32).

- Denoting the $2^{\lfloor nR \rfloor}$ messages of user 0 by message 1 to message $2^{\lfloor nR \rfloor}$, let message 1 be the transmitted message by transmitter 0 and $(\widetilde{\boldsymbol{x}}_{0,l})_{l=0}^{n-1}$ be the codeword that is assigned by user 0 to message 2. The probability p_{error} that a codeword different from the transmitted codeword satisfies both (2.32) and (2.33) is bounded from above as

$$p_{\mathrm{error}} \le 2^{nR}\,\mathbb{P}(\mathcal{E} \bigcap \mathcal{F}), \tag{E.4}$$

where $\mathcal{E}$ and $\mathcal{F}$ are given by

$$\mathcal{E} = \Big\{\Big((\widetilde{\boldsymbol{x}}_{0,l})_{l=0}^{\widehat{\boldsymbol{t}}_{1,n}-t_{0,n}-(2c+1)n_\delta-1}, (\boldsymbol{y}_0[l])_{l=t_{0,n}+n_\delta}^{\widehat{\boldsymbol{t}}_{1,n}-2cn_\delta-1}\Big) \in A_\epsilon^{(\widehat{\boldsymbol{t}}_{1,n}-t_{0,n}-(2c+1)n_\delta)}[p_{\mathrm{I}}]\Big\} \tag{E.5}$$

and

$$\mathcal{F} = \Big\{\Big((\widetilde{\boldsymbol{x}}_{0,l})_{l=\widehat{\boldsymbol{t}}_{1,n}-t_{0,n}+(2c-1)n_\delta+1}^{n-1}, (\boldsymbol{y}_0[l])_{\widehat{\boldsymbol{t}}_{1,n}+2cn_\delta+1}^{t_{0,n}+n_\delta+n-1}\Big) \in A_\epsilon^{(t_{0,n}-\widehat{\boldsymbol{t}}_{1,n}-(2c-1)n_\delta+n-1)}[p_{\mathrm{F}}]\Big\} \tag{E.6}$$

Fixing $\widehat{\boldsymbol{t}}_{1,n} = t$, we denote the events $\mathcal{E}$ and $\mathcal{F}$ by $\mathcal{E}_t$, $\mathcal{F}_t$, respectively. Then

$$\begin{aligned}\mathbb{P}(\mathcal{E}\bigcap\mathcal{F}) &= \sum_{t=t_{1,n}-2cn_\delta}^{t_{1,n}+2cn_\delta} \mathbb{P}(\mathcal{E}\bigcap\mathcal{F}, \widehat{\boldsymbol{t}}_{1,n} = t) \\ &= \sum_{t=t_{1,n}-2cn_\delta}^{t_{1,n}+2cn_\delta} \mathbb{P}(\mathcal{E}_t\bigcap\mathcal{F}_t, \widehat{\boldsymbol{t}}_{1,n} = t) \\ &\leq \sum_{t=t_{1,n}-2cn_\delta}^{t_{1,n}+2cn_\delta} \mathbb{P}(\mathcal{E}_t\bigcap\mathcal{F}_t) \\ &= \sum_{t=t_{1,n}-2cn_\delta}^{t_{1,n}+2cn_\delta} \mathbb{P}(\mathcal{E}_t)\mathbb{P}(\mathcal{F}_t), \end{aligned} \tag{E.7}$$

where the last step is due to independence of $\mathcal{E}_t$ and $\mathcal{F}_t$ as they only depend on the initial interval and the final interval, respectively. For any $t_{1,n}-2cn_\delta \leq t \leq t_{1,n}+2cn_\delta$, one can use the standard properties of typical sequences [28] to get

$$\mathbb{P}(\mathcal{E}_t) \leq 2^{-(t-t_{0,n}-(2c+1)n_\delta)(I(p_{\mathrm{I}})-3\epsilon)} \tag{E.8}$$

and

$$\mathbb{P}(\mathcal{F}_t) \leq 2^{-(t_{0,n}-t-(2c-1)n_\delta+n-1)(I(p_{\mathrm{F}})-3\epsilon)}, \tag{E.9}$$

where for any PDF $p(\cdot,\cdot)$ on $\mathbb{C}^2$ with marginals p_1 and p_2,

$$I(p) = \int p(x,y)\log\frac{p(x,y)}{p_1(x)p_2(y)}\mathrm{d}x\mathrm{d}y$$

is the mutual information between the marginals p_1 and p_2. By (E.7), (E.8) and

(E.9), we obtain the thread of inequalities

$$\mathbb{P}(\mathcal{E}\bigcap\mathcal{F}) \le \sum_{t=t_{1,n}-2cn_\delta}^{t_{1,n}+2cn_\delta} 2^{-(t-t_{0,n}-(2c+1)n_\delta)(I(p_{\mathrm{I}})-3\epsilon)}2^{-(t_{0,n}-t-(2c-1)n_\delta+n-1)(I(p_{\mathrm{F}})-3\epsilon)}$$

$$\begin{aligned}
&\overset{(a)}{\le} (4cn_\delta+1)2^{-((t_{1,n}-2cn_\delta)-t_{0,n}-(2c+1)n_\delta)(I(p_{\mathrm{I}})-3\epsilon)}2^{-(t_{0,n}-(t_{1,n}+2cn_\delta)-(2c-1)n_\delta+n-1)(I(p_{\mathrm{F}})-3\epsilon)}\\
&= (4cn_\delta+1)2^{-(t_{1,n}-t_{0,n}-(4c+1)n_\delta)(I(p_{\mathrm{I}})-3\epsilon)}2^{-(n-(t_{1,n}-t_{0,n})-(4c-1)n_\delta-1)(I(p_{\mathrm{F}})-3\epsilon)}\\
&\overset{(b)}{\le} (4cn_\delta+1)\,2^{-n\left(\frac{\nu_1-\nu_0}{\alpha}-\frac{2+(4c+1)n_\delta}{n}\right)(I(p_{\mathrm{I}})-3\epsilon)}\,2^{-n\left(1-\frac{\nu_1-\nu_0}{\alpha}-\frac{3+(4c-1)n_\delta}{n}\right)(I(p_{\mathrm{F}})-3\epsilon))}\\
&\overset{(c)}{\le} (4cn_\delta+1)\,2^{-n\left(\frac{\nu_1-\nu_0}{\alpha}-\frac{2+(4c+1)n_\delta}{n}\right)(I(p_{\mathrm{I}})-3\epsilon)}\,2^{-n\left(1-\frac{\nu_1-\nu_0}{\alpha}-\frac{2+(4c+1)n_\delta}{n}\right)(I(p_{\mathrm{F}})-3\epsilon))}\\
&\le (4cn_\delta+1)2^{-n\left(\frac{\nu_1-\nu_0}{\alpha}(I(p_{\mathrm{I}})-3\epsilon)+(1-\frac{\nu_1-\nu_0}{\alpha})(I(p_{\mathrm{F}})-3\epsilon)-\frac{2+(4c+1)n_\delta}{n}(I(p_{\mathrm{I}})+I(p_{\mathrm{F}})-6\epsilon)\right)}, \qquad \text{(E.10)}
\end{aligned}$$

where (a) is due to the fact that $2^{-(t-t_{0,n}-(2c+1)n_\delta)(I(p_{\mathrm{I}})-3\epsilon)}$ and $2^{-(t_{0,n}-t-(2c-1)n_\delta+n-1)(I(p_{\mathrm{F}})-3\epsilon)}$ assume their largest values for $t=t_{1,n}-2cn_\delta$ and $t_{1,n}+2cn_\delta$, respectively, (b) is due to $|t_{1,n}-t_{0,n}-\frac{n}{\alpha}(\nu_1-\nu_0)|<2$ which can be verified as in Footnote 2 and (c) is due to $3+(4c-1)n_\delta<2+(4c+1)n_\delta$. By (E.4) and (E.10), we have the bound

$$p_{\text{error}} \le (4n_\delta+1)2^{n\left(R-\frac{\nu_1-\nu_0}{\alpha}(I(p_{\mathrm{I}})-3\epsilon)-(1-\frac{\nu_1-\nu_0}{\alpha})(I(p_{\mathrm{F}})-3\epsilon)+\frac{2+(4c+1)n_\delta}{n}(I(p_{\mathrm{I}})+I(p_{\mathrm{F}})-6\epsilon)\right)} \qquad \text{(E.11)}$$

Assuming $R<R_{\mathrm{rm}}=\frac{\nu_1-\nu_0}{\alpha}I(p_{\mathrm{I}})+(1-\frac{\nu_1-\nu_0}{\alpha})I(p_{\mathrm{F}})$, one can choose ϵ sufficiently small and n sufficiently large so that $R-\frac{\nu_1-\nu_0}{\alpha}(I(p_{\mathrm{I}})-3\epsilon)-(1-\frac{\nu_1-\nu_0}{\alpha})(I(p_{\mathrm{F}})-3\epsilon)+\frac{2+(4c+1)n_\delta}{n}(I(p_{\mathrm{I}})+I(p_{\mathrm{F}})-6\epsilon)<0$ and hence, the right side of (E.11) tends to zero as n grows to infinity.

Appendix F

Proof of (3.14)

We need the following Lemma which is a slightly weaker version of Theorem 4.4 in [32]:

Lemma 3. *Let $\boldsymbol{x}$ be a* $\mathrm{Bin}(N, p)$ *random variable. Then*

$$\mathbb{P}\big(\boldsymbol{x} \geq (1+\epsilon)Np\big) \leq e^{-((1+\epsilon)\ln(1+\epsilon)-\epsilon)Np}, \tag{F.1}$$

for any $\epsilon > 0$.

At the "beginning" of time slot $t = 0$ the buffer is empty. Recall that $t_0 := \tau_i^{(1)}$ is the smallest t such that $b_{i,t} + b'_{i,t} \geq \lfloor n\eta_i \rfloor$. This implies that $b_{i,t_0} + b'_{i,t_0} = \lceil \frac{\lfloor n\eta_i \rfloor}{k_i} \rceil k_i$ is the smallest multiple of k_i which is larger than or equal to $\lfloor n\eta_i \rfloor$. The first codeword together with the preamble sequence are transmitted during the time slots $t_0 + 1, \cdots, t_0 + n' + n_i$ and the content of the buffer at the beginning of time slot $t_0 + 1$ becomes

$$b_{i,t_0+1} = \left\lceil \frac{\lfloor n\eta_i \rfloor}{k_i} \right\rceil k_i - \lfloor n\eta_i \rfloor. \tag{F.2}$$

We are interested in computing the probability of the event that a new codeword is scheduled for transmission before or at the time slot $t_0+n'+n_i-1$, i.e., $\boldsymbol{b}_{i,\boldsymbol{t}_0+1} + \sum_{t=\boldsymbol{t}_0+1}^{\boldsymbol{t}_0+n'+n_i-1} \boldsymbol{b}'_{i,t} \geq$

$\lfloor n\eta_i \rfloor$. We have

$$\begin{aligned}\mathbb{P}\Big(\boldsymbol{b}_{i,\boldsymbol{t}_0+1} + \sum_{t=\boldsymbol{t}_0+1}^{\boldsymbol{t}_0+n'+n_i-1} \boldsymbol{b}'_{i,t} \geq \lfloor n\eta_i \rfloor\Big) &\leq \mathbb{P}\Big(\sum_{t=\boldsymbol{t}_0+1}^{\boldsymbol{t}_0+n'+n_i-1} \boldsymbol{b}'_{i,t} \geq \lfloor n\eta_i \rfloor - k_i\Big) \\ &= \mathbb{P}\Big(\sum_{t=\boldsymbol{t}_0+1}^{\boldsymbol{t}_0+n'+n_i-1} \frac{\boldsymbol{b}'_{i,t}}{k_i} \geq \frac{\lfloor n\eta_i \rfloor}{k_i} - 1\Big), \end{aligned} \tag{F.3}$$

where the first step is due to the fact that $0 \leq \boldsymbol{b}_{i,\boldsymbol{t}_0+1} < k_i$ due to (F.2). By assumption,

$$\lim_{n\to\infty} \frac{1}{(n'+n_i-1)q_i}\left(\frac{\lfloor n\eta_i \rfloor}{k_i} - 1\right) = \frac{\eta_i}{\lambda_i\theta_i} = \frac{\mu_i}{\theta_i} > 1. \tag{F.4}$$

Let

$$\epsilon_i := \frac{1}{2}\Big(\frac{\mu_i}{\theta_i} - 1\Big). \tag{F.5}$$

In view of (F.4), assume n is large enough such that

$$\frac{1}{(n'+n_i-1)q_i}\left(\frac{\lfloor n\eta_i \rfloor}{k_i} - 1\right) > 1 + \epsilon_i. \tag{F.6}$$

Since $\sum_{t=\boldsymbol{t}_0+1}^{\boldsymbol{t}_0+n'+n_i-1} \frac{\boldsymbol{b}'_{i,l}}{k_i}$ is a $\text{Bin}(n'+n_i-1, q_i)$ random variable, the right side of (F.3) is bounded from above by $\mathbb{P}(\text{Bin}(n'+n_i-1, q_i) \geq (1+\epsilon_i)(n'+n_i-1)q_i)$ due to (F.6). Then Lemma 3 applies, i.e.,

$$\mathbb{P}(\text{Bin}(n'+n_i-1, q_i) \geq (1+\epsilon_i)(n'+n_i-1)q) \leq e^{-c_i n}, \tag{F.7}$$

where $c_i = q_i\theta_i((1+\epsilon_i)\ln(1+\epsilon_i) - \epsilon_i)$ and we have assumed n is large enough such that $n'+n_i-1 = \ln n + \lfloor n\theta_i \rfloor - 1 > n\theta_i$.

Appendix G

Proof of Proposition 6

We will use the following simple fact:

Lemma 4. *Let $\boldsymbol{x}_n$ for $n \geq 1$ be a sequence of real-valued random variables and $\lim_{n\to\infty} \boldsymbol{x}_n = a$ where $a > 0$ is a real number. Then $\lim_{n\to\infty} \mathbb{P}(\boldsymbol{x}_n > 0) = 1$.*

Proof. Since $\lim_{n\to\infty} \boldsymbol{x}_n = a$, then $\boldsymbol{x}_n$ also converges to a in probability. Fix $0 < \epsilon < a$. We have $\lim_{n\to\infty} \mathbb{P}(\boldsymbol{x}_n > 0) \geq \lim_{n\to\infty} \mathbb{P}(|\boldsymbol{x}_n - a| < \epsilon) = 1$. □

We only study the cases $j = 1$ and $j = 2$. The proof can be extended to any $j \geq 3$. Let $\beta_m := \sum_{t=1}^{mn_i} b'_{i,t}$ for $m \geq 1$ be the total number of bits arriving at the buffer of Tx i until the time slot of index mn_i. Under the Tx-Rx synchronous scheme, Tx i checks its buffer at time slots mn_i for $m \geq 1$ and if its buffer content is more than $\lfloor n\eta_i \rfloor$, a codeword is sent over the channel during time slots $mn_i + 1, \cdots, (m+1)n_i$. By SLLN,

$$\lim_{n\to\infty} \frac{\boldsymbol{\beta}_m}{n} = m\theta_i\lambda_i, \tag{G.1}$$

for any $m \geq 1$.

- Let $j = 1$. We have

$$\varsigma_i^{(1)} = mn_i \iff \beta_1, \cdots, \beta_{m-1} < \lfloor n\eta_i \rfloor, \;\; \beta_m \geq \lfloor n\eta_i \rfloor \tag{G.2}$$

  for any $m \geq 1$. Recall $\mu_i := \frac{\eta_i}{\lambda_i} > \theta_i$. By assumption, μ_i is not an integer multiple of θ_i. Let $m^* \geq 1$ be such that

$$m^*\theta_i < \mu_i < (m^*+1)\theta_i. \tag{G.3}$$

Putting (G.1), (G.2) and (G.3) together and invoking Lemma 4,

$$\lim_{n\to\infty} \mathbb{P}\big(\boldsymbol{\varsigma}_i^{(1)} = (m^*+1)n_i\big) = 1. \tag{G.4}$$

By (G.3), fix $\delta_1 > 0$ such that $(1+\delta_1)\mu_i < (m^*+1)\theta_i$. Then

$$\begin{aligned} \mathbb{P}\Big(\boldsymbol{\varsigma}_i^{(1)} > (1+\delta_1)\boldsymbol{\tau}_i^{(1)}\Big) &\geq \mathbb{P}\Big(\boldsymbol{\varsigma}_i^{(1)} > (1+\delta_1)\boldsymbol{\tau}_i^{(1)}, \boldsymbol{\varsigma}_i^{(1)} = (m^*+1)n_i\Big) \\ &= \mathbb{P}\Big((m^*+1)n_i > (1+\delta_1)\boldsymbol{\tau}_i^{(1)}, \boldsymbol{\varsigma}_i^{(1)} = (m^*+1)n_i\Big). \end{aligned} \tag{G.5}$$

Since $\boldsymbol{\tau}_i^{(1)} \sim \mathrm{NB}(\frac{\lfloor n\eta_i \rfloor}{k_i}, q_i)$, we can apply SLLN to get $\lim_{n\to\infty} \frac{1}{n}\big((m^*+1)n_i - (1+\delta_1)\boldsymbol{\tau}_i^{(1)}\big) = (m^*+1)\theta_i - (1+\delta_1)\mu_i > 0$. Using this together with Lemma 4,

$$\lim_{n\to\infty} \mathbb{P}\big((m^*+1)n_i > (1+\delta_1)\boldsymbol{\tau}_i^{(1)}\big) = 1. \tag{G.6}$$

By (G.4) and (G.11), the expression on the right side of (G.10) goes to 1 as n grows[1] and we get the desired result.

- Let $j = 2$. We have $\varsigma_i^{(2)} = mn_i$ for some $m \geq 2$ if and only if there is $1 \leq m' \leq m-1$ such that

$$\beta_1, \cdots, \beta_{m'-1} < \lfloor n\eta_i \rfloor,\ \beta_{m'} \geq \lfloor n\eta_i \rfloor,\ \beta_{m'+1}, \cdots, \beta_{m-1} < 2\lfloor n\eta_i \rfloor,\ \beta_m \geq 2\lfloor n\eta_i \rfloor \tag{G.7}$$

Recall that μ_i is not a multiple of $\frac{\theta_i}{2}$. Let $m^* \geq 2$ be such that

$$\frac{m^*\theta_i}{2} < \mu_i < \frac{(m^*+1)\theta_i}{2}. \tag{G.8}$$

Then we can invoke Lemma 4 together with (G.1), (G.7) and (G.8) to write

$$\lim_{n\to\infty} \mathbb{P}\big(\boldsymbol{\varsigma}_i^{(2)} = (m^*+1)n_i\big) = 1, \tag{G.9}$$

where m' in (G.7) is selected as $m' = \frac{m^*}{2} + 1$ for even m^* and $m' = \frac{m^*+1}{2}$ for odd m^*.

[1]If $\mathcal{A}_n$ and $\mathcal{B}_n$ are events such that $\lim_{n\to\infty} \mathbb{P}(\mathcal{A}_n) = \lim_{n\to\infty} \mathbb{P}(\mathcal{B}_n) = 1$, then $\lim_{n\to\infty} \mathbb{P}(\mathcal{A}_n \bigcap \mathcal{B}_n) = 1$.

By (G.8), fix $\delta_2 > 0$ such that $(1+\delta_2)\mu_i < \frac{(m^*+1)\theta_i}{2}$. Then

$$\begin{aligned}\mathbb{P}\Big(\boldsymbol{\varsigma}_i^{(2)} > (1+\delta_2)\boldsymbol{\tau}_i^{(2)}\Big) &\geq \mathbb{P}\Big(\boldsymbol{\varsigma}_i^{(2)} > (1+\delta_2)\boldsymbol{\tau}_i^{(1)}, \boldsymbol{\varsigma}_i^{(2)} = (m^*+1)n_i\Big) \\ &= \mathbb{P}\Big((m^*+1)n_i > (1+\delta_2)\boldsymbol{\tau}_i^{(2)}, \boldsymbol{\varsigma}_i^{(2)} = (m^*+1)n_i\Big) \end{aligned} \tag{G.10}$$

Since $\boldsymbol{\tau}_i^{(2)} \sim \mathrm{NB}(\frac{2\lfloor n\eta_i \rfloor}{k_i}, q_i)$, we can apply SLLN to get $\lim_{n\to\infty} \frac{1}{n}\big((m^*+1)n_i - (1+\delta_2)\boldsymbol{\tau}_i^{(2)}\big) = (m^*+1)\theta_i - 2(1+\delta_2)\mu_i > 0$. Using this together with Lemma 4,

$$\lim_{n\to\infty} \mathbb{P}\big((m^*+1)n_i > (1+\delta_2)\boldsymbol{\tau}_i^{(2)}\big) = 1. \tag{G.11}$$

By (G.4) and (G.11), the expression on the right side of (G.10) goes to 1 as n grows and we get the desired result.

Finally, we select $\delta := \min_{1\leq j\leq N_i} \delta_j$ to make sure (3.20) holds for any $1 \leq j \leq N_i$.

Appendix H

Proof of Proposition 7

Let $1 \leq j_1 \leq N_1$ and $1 \leq j_2 \leq N_2$. If $\tau_1^{(j_1)} + 1 \leq \tau_2^{(j_2)} + 1 \leq \tau_1^{(j_1)} + n'$, then the j_2^{th} burst of Tx 2 starts while Tx 1 is sending the preamble sequence in its j_1^{th} burst. If $\tau_1^{(j_1)} + 1 \leq \tau_2^{(j_2)} + n' + n_2 \leq \tau_1^{(j_1)} + n'$, then the j_2^{th} burst of Tx 2 ends while Tx 1 is sending the preamble sequence in its j_1^{th} burst. Let $\mathcal{E}_{j_1,j_2}$ be the union of these two events, i.e.,

$$\mathcal{E}_{j_1,j_2} := \big\{0 \leq \boldsymbol{\tau}_2^{(j_2)} - \boldsymbol{\tau}_1^{(j_1)} \leq n' - 1\big\} \bigcup \big\{n_2 \leq \boldsymbol{\tau}_1^{(j_1)} - \boldsymbol{\tau}_2^{(j_2)} \leq n' + n_2 - 1\big\}. \tag{H.1}$$

We have

$$\mathbb{P}\Big(\bigcup_{j_1,j_2} \mathcal{E}_{j_1,j_2}\Big) \leq \sum_{\substack{1 \leq j_1 \leq N_1 \\ 1 \leq j_2 \leq N_2}} \mathbb{P}(\mathcal{E}_{j_1,j_2}). \tag{H.2}$$

Moreover,

$$\mathbb{P}(\mathcal{E}_{j_1,j_2}) \leq \mathbb{P}\big(0 \leq \boldsymbol{\tau}_2^{(j_2)} - \boldsymbol{\tau}_1^{(j_1)} \leq n' - 1\big) + \mathbb{P}\big(n_2 \leq \boldsymbol{\tau}_1^{(j_1)} - \boldsymbol{\tau}_2^{(j_2)} \leq n' + n_2 - 1\big). \tag{H.3}$$

In view of (H.2) and (H.3), it is enough to show that

$$\lim_{n \to \infty} \mathbb{P}\big(0 \leq \boldsymbol{\tau}_2^{(j_2)} - \boldsymbol{\tau}_1^{(j_1)} \leq n' - 1\big) = 0 \tag{H.4}$$

and

$$\lim_{n \to \infty} \mathbb{P}\big(n_2 \leq \boldsymbol{\tau}_1^{(j_1)} - \boldsymbol{\tau}_2^{(j_2)} \leq n' + n_2 - 1\big) = 0 \tag{H.5}$$

for arbitrary choices of j_1 and j_2. To verify (H.4), define

$$\boldsymbol{\rho}_n := \frac{1}{n}\big(\boldsymbol{\tau}_2^{(j_2)} - \boldsymbol{\tau}_1^{(j_1)}\big), \;\; \boldsymbol{\rho}'_n := \frac{1}{n}\big(\boldsymbol{\tau}_1^{(j_1)} - \boldsymbol{\tau}_2^{(j_2)} + n' - 1\big). \tag{H.6}$$

Then

$$\mathbb{P}\big(0 \leq \boldsymbol{\tau}_2^{(j_2)} - \boldsymbol{\tau}_1^{(j_1)} \leq n' - 1\big) = \mathbb{P}\big(\boldsymbol{\rho}_n \geq 0, \boldsymbol{\rho}'_n \geq 0\big) = \mathbb{P}\big(\min\{\boldsymbol{\rho}_n, \boldsymbol{\rho}'_n\} \geq 0\big). \tag{H.7}$$

By (3.25),

$$\lim_{n\to\infty} \frac{\boldsymbol{\tau}_1^{(j_1)}}{n} \;=\; \lim_{n\to\infty} \frac{\boldsymbol{\xi}_1^{(j_1)}}{n} + \lim_{n\to\infty} \frac{\lfloor n\nu_1 \rfloor}{n} = \lim_{n\to\infty} \frac{\boldsymbol{\xi}_1^{(j_1)}}{n} + \nu_1. \tag{H.8}$$

Using SLLN, $\lim_{n\to\infty} \frac{\boldsymbol{\xi}_1^{(j_1)}}{n} = j_1\mu_1$ and we get

$$\lim_{n\to\infty} \frac{\boldsymbol{\tau}_1^{(j_1)}}{n} = j_1\mu_1 + \nu_1. \tag{H.9}$$

Similarly,

$$\lim_{n\to\infty} \frac{\boldsymbol{\tau}_2^{(j_2)}}{n} = j_2\mu_2 + \nu_2. \tag{H.10}$$

Define

$$\rho := j_2\mu_2 - j_1\mu_1 + \nu_2 - \nu_1. \tag{H.11}$$

By (H.6), (H.9) and (H.10) and noting that $\lim_{n\to\infty} \frac{n'-1}{n} = 0$, we get $\lim_{n\to\infty} \boldsymbol{\rho}_n = \rho$, $\lim_{n\to\infty} \boldsymbol{\rho}'_n = -\rho$ and hence,

$$\lim_{n\to\infty} \min\{\boldsymbol{\rho}_n, \boldsymbol{\rho}'_n\} = \min\{\rho, -\rho\} < 0, \tag{H.12}$$

where the last step is due to (3.35). By (H.12) and Lemma 4, the proof of (H.4) is complete. The proof of (H.5) is quite similar and is omitted for brevity.

Appendix I

Proof of Proposition 8

We have

$$\mathbb{P}(\hat{\boldsymbol{t}}_1 < \boldsymbol{t}_1) = \sum_{t_1 \geq 0} \mathbb{P}(\hat{\boldsymbol{t}}_1 < \boldsymbol{t}_1 | \boldsymbol{t}_1 = t_1)\mathbb{P}(\boldsymbol{t}_1 = t_1). \tag{I.1}$$

Moreover,

$$\begin{aligned}
&\mathbb{P}(\hat{\boldsymbol{t}}_1 < \boldsymbol{t}_1 | \boldsymbol{t}_1 = t_1) \\
&= \mathbb{P}\left(\exists\, t < t_1 : \left((\boldsymbol{s}'_{1,l})_{l=0}^{n'-1}, (\boldsymbol{y}_{1,l})_{l=t}^{t+n'-1}\right) \in A_\epsilon^{(n')}[p^{(1)}] \text{ or } \left((\boldsymbol{s}'_{2,l})_{l=0}^{n'-1}, (\boldsymbol{y}_{1,l})_{l=t}^{t+n'-1}\right) \in A_\epsilon^{(n')}[p^{(3)}]\right) \\
&\leq \sum_{t=0}^{t_1-1} \mathbb{P}\left(\left((\boldsymbol{s}'_{1,l})_{l=0}^{n'-1}, (\boldsymbol{y}_{1,l})_{l=t}^{t+n'-1}\right) \in A_\epsilon^{(n')}[p^{(1)}]\right) \\
&\quad + \sum_{t=0}^{t_1-1} \mathbb{P}\left(\left((\boldsymbol{s}'_{2,l})_{l=0}^{n'-1}, (\boldsymbol{y}_{1,l})_{l=t}^{t+n'-1}\right) \in A_\epsilon^{(n')}[p^{(3)}]\right).
\end{aligned} \tag{I.2}$$

In the following, we find an upper bound on each term on the right side of (I.2).

- The term $\sum_{t=0}^{t_1-1} \mathbb{P}\left(\left((\boldsymbol{s}'_{1,l})_{l=0}^{n'-1}, (\boldsymbol{y}_{1,l})_{l=t}^{t+n'-1}\right) \in A_\epsilon^{(n')}[p^{(1)}]\right)$: We study the cases $t \leq t_1 - n'$ and $t_1 - n' + 1 \leq t \leq t_1 - 1$, separately:
  - Let $t \leq t_1 - n'$. We explicitly write the constraint defining $A_\epsilon^{(m)}[p^{(1)}]$ in (2.20) as the set of all $(\vec{x}, \vec{y})$ such that $\left|\frac{1}{m}\sum_{l=1}^{m}\left(\frac{1}{2\gamma_1}x_l^2 + \frac{1}{2}\left(y_l - x_l\right)^2 - 1\right)\right| < \frac{\epsilon}{\log e}$. Then

we have the thread of inequalities

$$
\begin{aligned}
&\mathbb{P}\Big(\big((\boldsymbol{s}'_{1,l})_{l=0}^{n'-1}, (\boldsymbol{y}_{1,l})_{l=t}^{t+n'-1}\big) \in A_{\epsilon}^{(n')}[p^{(1)}]\Big) \\
&\leq \mathbb{P}\Big(\Big|\frac{1}{n'}\sum_{l=0}^{n'-1}\Big(\frac{1}{2\gamma_1}|\boldsymbol{s}'_{1,l}|^2 + \frac{1}{2}(\boldsymbol{y}_{1,l+t} - \boldsymbol{s}'_{1,l})^2 - 1\Big)\Big| < \frac{\epsilon}{\log e}\Big) \\
&\leq \mathbb{P}\Big(\frac{1}{n'}\sum_{l=0}^{n'-1}\Big(\frac{1}{2\gamma_1}|\boldsymbol{s}'_{1,l}|^2 + \frac{1}{2}\big(\boldsymbol{y}_{1,l+t} - \boldsymbol{s}'_{1,l}\big)^2 - 1\Big) < \frac{\epsilon}{\log e}\Big) \\
&\overset{(a)}{=} \mathbb{P}\Big(\frac{1}{n'}\sum_{l=0}^{n'-1}\Big(\frac{1}{2\gamma_1}|\boldsymbol{s}'_{1,l}|^2 + \frac{1}{2}\big(\boldsymbol{z}_{1,l+t} - \boldsymbol{s}'_{1,l}\big)^2 - 1\Big) < \frac{\epsilon}{\log e}\Big) \\
&\overset{(b)}{\leq} \mathbb{P}\Big(\frac{1}{n'}\sum_{l=0}^{n'-1}\Big(1 + \frac{\gamma_1}{2} - \frac{1}{2\gamma_1}|\boldsymbol{s}'_{1,l}|^2 - \frac{1}{2}\big(\boldsymbol{z}_{1,l+t} - \boldsymbol{s}'_{1,l}\big)^2\Big) > \frac{\gamma_1}{2} - \frac{\epsilon}{\log e}\Big) \\
&\overset{(c)}{=} \mathbb{P}\Big(\frac{1}{n'}\sum_{l=0}^{n'-1}\Big(1 - \frac{\frac{1}{2\gamma_1}|\boldsymbol{s}'_{1,l}|^2 + \frac{1}{2}\big(\boldsymbol{z}_{1,l+t} - \boldsymbol{s}'_{1,l}\big)^2}{1+\frac{\gamma_1}{2}}\Big) > \frac{\frac{\gamma_1}{2} - \frac{\epsilon}{\log e}}{1+\frac{\gamma_1}{2}}\Big) \\
&\overset{(d)}{\leq} e^{-\Theta(n')},
\end{aligned}
\tag{I.3}
$$

where (a) is due to the fact that the signal at Rx 1 during time slots t to $t+n'-1$ only consists of the ambient noise, in (b) we have added $\frac{\gamma_1}{2}$ to both sides of the inequality in (a) after multiplying both sides of the inequality by -1, in (c) both sides are divided by $1+\frac{\gamma_1}{2}$ and (d) is due to Lemma 2 (Bernstein's inequality) where it is assumed that $\epsilon < \frac{\gamma_1}{2}\log e$. In fact, the random variables $\boldsymbol{w}_l = 1 - \frac{\frac{1}{2\gamma_1}|\boldsymbol{s}'_{1,l}|^2 + \frac{1}{2}(\boldsymbol{z}_{1,l+t} - \boldsymbol{s}'_{1,l})^2}{1+\frac{\gamma_1}{2}}$ are independent with zero mean and finite variance and $\boldsymbol{w}_l \leq 1$. Therefore, one can apply Bernstein's inequality.

– Let $t_1 - n' + 1 \leq t \leq t_1 - 1$. Then we get

$$\begin{aligned}
&\mathbb{P}\Big(\big((\boldsymbol{s}'_{1,l})_{l=0}^{n'-1}, (\boldsymbol{y}_{1,l})_{l=t}^{t+n'-1}\big) \in A_\epsilon^{(n')}[p^{(1)}]\Big) \\
&\leq \mathbb{P}\Big(\frac{1}{n'}\sum_{l=0}^{n'-1}\Big(\frac{1}{2\gamma_1}|\boldsymbol{s}'_{1,l}|^2 + \frac{1}{2}\big(\boldsymbol{y}_{1,l+t} - \boldsymbol{s}'_{1,l}\big)^2 - 1\Big) < \frac{\epsilon}{\log e}\Big) \\
&\overset{(a)}{=} \mathbb{P}\Big(\frac{1}{n'}\sum_{l=0}^{t_1-t-1}\Big(\frac{1}{2\gamma_1}|\boldsymbol{s}'_{1,l}|^2 + \frac{1}{2}\big(\boldsymbol{z}_{1,l+t} - \boldsymbol{s}'_{1,l}\big)^2 - 1\Big) \\
&\quad + \frac{1}{n'}\sum_{l=t_1-t}^{n'-1}\Big(\frac{1}{2\gamma_1}|\boldsymbol{s}'_{1,l}|^2 + \frac{1}{2}\big(\boldsymbol{s}'_{1,l+t-t_1} + \boldsymbol{z}_{1,l+t} - \boldsymbol{s}'_{1,l}\big)^2 - 1\Big) < \frac{\epsilon}{\log e}\Big) \\
&\overset{(b)}{=} \mathbb{P}\Big(\frac{1}{n'}\sum_{l=0}^{t_1-t-1}\Big(1 + \frac{\gamma_1}{2} - \frac{1}{2\gamma_1}|\boldsymbol{s}'_{1,l}|^2 - \frac{1}{2}\big(\boldsymbol{z}_{1,l+t} - \boldsymbol{s}'_{1,l}\big)^2\Big) \\
&\quad + \frac{1}{n'}\sum_{l=t_1-t}^{n'-1}\Big(1 + \gamma_1 - \frac{1}{2\gamma_1}|\boldsymbol{s}'_{1,l}|^2 - \frac{1}{2}\big(\boldsymbol{s}'_{1,l+t-t_1} + \boldsymbol{z}_{1,l+t} - \boldsymbol{s}'_{1,l}\big)^2\Big) > \epsilon_n\Big),
\end{aligned} \tag{I.4}$$

where in (a) we have used the fact that $\boldsymbol{y}_{1,l+t}$ is the ambient noise for $l < t_1 - t$ and $\boldsymbol{y}_{1,l+t} = \boldsymbol{s}'_{1,l+t-t_1} + \boldsymbol{z}_{1,l+t}$ for $t_1 - t \leq l \leq n' - 1$ and in (b) we have added $\frac{\gamma_1}{2}$ to each term in the first sum and γ_1 to each term in the second sum after multiplying both sides of the inequality in (a) by -1. Moreover, ϵ_n is given by

$$\epsilon_n := \frac{t_1 - t}{2n'}\gamma_1 + \frac{n' + t - t_1}{n'}\gamma_1 - \frac{\epsilon}{\log e}. \tag{I.5}$$

We can write

$$\epsilon_n = \big(1 - \frac{t_1 - t}{2n'}\big)\gamma_1 - \frac{\epsilon}{\log e} \geq \big(1 - \frac{n' - 1}{2n'}\big)\gamma_1 - \frac{\epsilon}{\log e} \geq \frac{\gamma_1}{3} - \frac{\epsilon}{\log e}, \tag{I.6}$$

where the penultimate step is due to $t_1 - t \leq n' - 1$ and in the last step we are assuming n is sufficiently large[1] such that $1-\frac{n'-1}{2n'}\geq\frac{1}{3}$. By (I.4) and (I.6),

$$\begin{aligned}&\mathbb{P}\Big(\big((\boldsymbol{s}'_{1,l})_{l=0}^{n'-1},(\boldsymbol{y}_{1,l})_{l=t}^{t+n'-1}\big)\in A_\epsilon^{(n')}[p^{(1)}]\Big)\\&\leq\mathbb{P}\Big(\frac{1}{n'}\sum_{l=0}^{t_1-t-1}\Big(1+\frac{\gamma_1}{2}-\frac{1}{2\gamma_1}|\boldsymbol{s}'_{1,l}|^2-\frac{1}{2}\big(\boldsymbol{z}_{1,l+t}-\boldsymbol{s}'_{1,l}\big)^2\Big)\\&\quad+\frac{1}{n'}\sum_{l=t_1-t}^{n'-1}\Big(1+\gamma_1-\frac{1}{2\gamma_1}|\boldsymbol{s}'_{1,l}|^2-\frac{1}{2}\big(\boldsymbol{s}'_{1,l+t-t_1}+\boldsymbol{z}_{1,l+t}-\boldsymbol{s}'_{1,l}\big)^2\Big)>\frac{\gamma_1}{3}-\frac{\epsilon}{\log e}\Big),\end{aligned}\tag{I.7}$$

where we are assuming that $\epsilon<\frac{\gamma_1}{3}\log e$. Define

$$\boldsymbol{w}_l=\begin{cases}1-\frac{1}{1+\frac{\gamma_1}{2}}\big(\frac{1}{2\gamma_1}|\boldsymbol{s}'_{1,l}|^2-\frac{1}{2}\big(\boldsymbol{z}_{1,l+t}-\boldsymbol{s}'_{1,l}\big)^2\big) & 0\leq l\leq t_1-t-1\\ 1-\frac{1}{1+\gamma_1}\big(\frac{1}{2\gamma_1}|\boldsymbol{s}'_{1,l}|^2-\frac{1}{2}\big(\boldsymbol{s}'_{1,l+t-t_1}+\boldsymbol{z}_{1,l+t}-\boldsymbol{s}'_{1,l}\big)^2\big) & t_1-t\leq l\leq n'-1\end{cases}.\tag{I.8}$$

Then we can write (I.7) as

$$\begin{aligned}&\mathbb{P}\Big(\big((\boldsymbol{s}'_{1,l})_{l=0}^{n'-1},(\boldsymbol{y}_{1,l})_{l=t}^{t+n'-1}\big)\in A_\epsilon^{(n')}[p^{(1)}]\Big)\\&\leq\mathbb{P}\Big(\frac{1}{n'}\sum_{l=0}^{t_1-t-1}(1+\frac{\gamma_1}{2})\boldsymbol{w}_l+\frac{1}{n'}\sum_{l=t_1-t}^{n'-1}(1+\gamma_1)\boldsymbol{w}_l>\frac{\gamma_1}{3}-\frac{\epsilon}{\log e}\Big)\\&\leq\mathbb{P}\Big(\frac{1}{n'}\sum_{l=0}^{n'-1}\boldsymbol{w}_l>\frac{1}{1+\gamma_1}\big(\frac{\gamma_1}{3}-\frac{\epsilon}{\log e}\big)\Big),\end{aligned}\tag{I.9}$$

where in the last step we have used the fact that $\frac{1}{n'}\sum_{l=0}^{t_1-t-1}\big(1+\frac{\gamma_1}{2}\big)\boldsymbol{w}_l+\frac{1}{n'}\sum_{l=t_1-t}^{n'-1}(1+\gamma_1)\boldsymbol{w}_l<(1+\gamma_1)\times\frac{1}{n'}\sum_{l=0}^{n'-1}\boldsymbol{w}_l$. Note that $\boldsymbol{w}_l$ have zero mean and finite variance and $\boldsymbol{w}_l\leq 1$ for all $0\leq l\leq n'-1$. However, in contrast with the previous case where we had $t\leq t_1-n'$, one can not apply Bernstein's Lemma because $\boldsymbol{w}_l$ are no longer independent random variables. In fact, any two $\boldsymbol{w}_l$ and $\boldsymbol{w}_{l'}$ are dependent if and only if $|l-l'|=t_1-t$. To circumvent this difficulty, we use a trick in Appendix 24B in [20]. We consider two cases:

* If t_1-t is odd, the terms with odd indices are independent. Similarly, the

[1] We have $\lim_{n\to\infty}(1-\frac{n'-1}{2n'})=\frac{1}{2}$ and hence, $1-\frac{n'-1}{2n'}>\frac{1}{3}$ for sufficiently large n.

terms with even indices are independent. Then

$$\begin{aligned}
&\mathbb{P}\Big(\frac{1}{n'}\sum_{l=0}^{n'-1}\boldsymbol{w}_l > \frac{1}{1+\gamma_1}\big(\frac{\gamma_1}{3}-\frac{\epsilon}{\log e}\big)\Big)\\
&\le \mathbb{P}\Bigg(\frac{1}{n'}\sum_{\substack{l=0\\ l \text{ even}}}^{n'-1}\boldsymbol{w}_l > \frac{1}{2(1+\gamma_1)}\big(\frac{\gamma_1}{3}-\frac{\epsilon}{\log e}\big)\Bigg)\\
&+\mathbb{P}\Bigg(\frac{1}{n'}\sum_{\substack{l=0\\ l \text{ odd}}}^{n'-1}\boldsymbol{w}_l > \frac{1}{2(1+\gamma_1)}\big(\frac{\gamma_1}{3}-\frac{\epsilon}{\log e}\big)\Bigg). \qquad \text{(I.10)}
\end{aligned}$$

At this point, similar to (I.3), one can apply Bernstein's inequality to conclude that each term on the right side of (I.10) is bounded from above by $e^{-\Theta(n')}$.

* If $t_1 - t$ is even, we need to partition the set of integers into two disjoint sets $\mathcal{I}$ and $\mathcal{J}$ such that the difference of any two element in each of these sets is not equal to $t_1 - t$. Such a partition is given in Appendix 24B in [20] or Appendix D in [25]. Then

$$\begin{aligned}
&\mathbb{P}\Big(\frac{1}{n'}\sum_{l=0}^{n'-1}\boldsymbol{w}_l > \frac{1}{1+\gamma_1}\big(\frac{\gamma_1}{3}-\frac{\epsilon}{\log e}\big)\Big)\\
&\le \mathbb{P}\Bigg(\frac{1}{n'}\sum_{\substack{l=0\\ l \in \mathcal{I}}}^{n'-1}\boldsymbol{w}_l > \frac{1}{1+\gamma_1}\big(\frac{\gamma_1}{3}-\frac{\epsilon}{\log e}\big)\Bigg)\\
&+\mathbb{P}\Bigg(\frac{1}{n'}\sum_{\substack{l=0\\ l \in \mathcal{J}}}^{n'-1}\boldsymbol{w}_l > \frac{1}{1+\gamma_1}\big(\frac{\gamma_1}{3}-\frac{\epsilon}{\log e}\big)\Bigg). \qquad \text{(I.11)}
\end{aligned}$$

As each of $\sum_{l\,\in\,\mathcal{I}}\boldsymbol{w}_l$ and $\sum_{l\,\in\,\mathcal{J}}\boldsymbol{w}_l$ are sums of independent random variables, it follows that each term on the right side of (I.11) is bounded from above by $e^{-\Theta(n')}$.

We conclude that whether $t \le t_1 - n'$ or $t_1 - n' + 1 \le t \le t_1 - 1$, each term in the first sum on the right side of (I.2) is bounded from above by $\Theta(1)e^{-\Theta(n')}$. Since there are t_1 terms in this sum, we get

$$\sum_{t=0}^{t_1-1} \mathbb{P}\left(\left((\boldsymbol{s}'_{1,l})_{l=0}^{n'-1}, (\boldsymbol{y}_{1,l})_{l=t}^{t+n'-1} \right) \in A_\epsilon^{(n')}[p^{(1)}] \right) \leq t_1 \Theta(1) e^{-\Theta(n')}. \tag{I.12}$$

- The term $\sum_{t=0}^{t_1-1} \mathbb{P}\big(\big((\boldsymbol{s}'_{2,l})_{l=0}^{n'-1}, (\boldsymbol{y}_{1,l})_{l=t}^{t+n'-1} \big) \in A_\epsilon^{(n')}[p^{(3)}] \big)$: The analysis in this case follows similar lines of reasoning in the previous case and is omitted. The result is that

$$\sum_{t=0}^{t_1-1} \mathbb{P}\left(\left((\boldsymbol{s}'_{2,l})_{l=0}^{n'-1}, (\boldsymbol{y}_{1,l})_{l=t}^{t+n'-1} \right) \in A_\epsilon^{(n')}[p^{(3)}] \right) \leq t_1 \Theta(1) e^{-\Theta(n')}. \tag{I.13}$$

By (I.1), (I.2), (I.12) and (I.13),

$$\mathbb{P}(\hat{\boldsymbol{t}}_1 \leq \boldsymbol{t}_1) \leq \Theta(1) e^{-\Theta(n')} \sum_{t_1 \geq 0} t_1 \mathbb{P}(\boldsymbol{t}_1 = t_1) = \Theta(1) \mathbb{E}[\boldsymbol{t}_1] e^{-\Theta(n')}. \tag{I.14}$$

But,

$$\begin{aligned} \mathbb{E}[\boldsymbol{t}_1] &= \mathbb{E}[\boldsymbol{\tau}_1^{(1)} + 1] \\ &= \mathbb{E}[\boldsymbol{\xi}_1^{(1)}] + \mathbb{E}[\lfloor n \boldsymbol{\nu}_1 \rfloor] + 1 \\ &\leq \mathbb{E}[\boldsymbol{\xi}_1^{(1)}] + \mathbb{E}[n \boldsymbol{\nu}_1] + 1 \\ &= \frac{\lfloor n \eta_1 \rfloor}{\lambda_1} + \frac{n}{2} + 1 \\ &= \Theta(n) \end{aligned} \tag{I.15}$$

By (I.14) and (I.15),

$$\mathbb{P}(\hat{\boldsymbol{t}}_1 < \boldsymbol{t}_1) \leq \Theta(n) e^{-\Theta(n')}, \tag{I.16}$$

as desired.

Appendix J

Proof of (3.45)

The proof follows similar lines of reasoning in (I.3). We explicitly write the constraint defining $A_\epsilon^{(m)}[p^{(3)}]$ in (2.20) as the set of all $(\vec{x}, \vec{y})$ such that $\left|\frac{1}{m}\sum_{l=1}^{m}\left(\frac{1}{2\gamma_2}x_l^2 + \frac{1}{2}\big(y_l - \sqrt{a_2}\,x_l\big)^2 - 1\right)\right| < \frac{\epsilon}{\log e}$. Then

$$
\begin{aligned}
&\mathbb{P}\Big(\big((\boldsymbol{s}'_{2,l})_{l=0}^{n'-1}, (\boldsymbol{y}_{1,l})_{l=t_1}^{t_1+n'-1}\big) \in A_\epsilon^{(n')}[p^{(3)}]\Big) \\
&\leq \mathbb{P}\Big(\Big|\frac{1}{n'}\sum_{l=0}^{n'-1}\Big(\frac{1}{2\gamma_2}|\boldsymbol{s}'_{2,l}|^2 + \frac{1}{2}(\boldsymbol{y}_{1,l+t_1} - \sqrt{a_2}\,\boldsymbol{s}'_{2,l})^2 - 1\Big)\Big| < \frac{\epsilon}{\log e}\Big) \\
&\leq \mathbb{P}\Big(\frac{1}{n'}\sum_{l=0}^{n'-1}\Big(\frac{1}{2\gamma_2}|\boldsymbol{s}'_{2,l}|^2 + \frac{1}{2}\big(\boldsymbol{y}_{1,l+t_1} - \sqrt{a_2}\,\boldsymbol{s}'_{2,l}\big)^2 - 1\Big) < \frac{\epsilon}{\log e}\Big) \\
&\overset{(a)}{=} \mathbb{P}\Big(\frac{1}{n'}\sum_{l=0}^{n'-1}\Big(\frac{1}{2\gamma_2}|\boldsymbol{s}'_{2,l}|^2 + \frac{1}{2}\big(\boldsymbol{s}'_{1,l} + \boldsymbol{z}_{1,l+t_1} - \sqrt{a_2}\,\boldsymbol{s}'_{2,l}\big)^2 - 1\Big) < \frac{\epsilon}{\log e}\Big) \\
&\overset{(b)}{\leq} \mathbb{P}\Big(\frac{1}{n'}\sum_{l=0}^{n'-1}\Big(1 + \frac{\gamma_1 + a_2\gamma_2}{2} - \frac{1}{2\gamma_2}|\boldsymbol{s}'_{1,l}|^2 - \frac{1}{2}\big(\boldsymbol{s}'_{1,l} + \boldsymbol{z}_{1,l+t_1} - \sqrt{a_2}\,\boldsymbol{s}'_{2,l}\big)^2\Big) > \frac{\gamma_1 + a_2\gamma_2}{2} - \frac{\epsilon}{\log e}\Big) \\
&\overset{(c)}{=} \mathbb{P}\Big(\frac{1}{n'}\sum_{l=0}^{n'-1}\Big(1 - \frac{\frac{1}{2\gamma_2}|\boldsymbol{s}'_{2,l}|^2 + \frac{1}{2}\big(\boldsymbol{s}'_{1,l} + \boldsymbol{z}_{1,l+t_1} - \sqrt{a_2}\,\boldsymbol{s}'_{2,l}\big)^2}{1 + \frac{\gamma_1 + a_2\gamma_2}{2}}\Big) > \frac{\frac{\gamma_1 + a_2\gamma_2}{2} - \frac{\epsilon}{\log e}}{1 + \frac{\gamma_1 + a_2\gamma_2}{2}}\Big) \\
&\overset{(d)}{\leq} e^{-\Theta(n')},
\end{aligned} \tag{J.1}
$$

where (a) is due to the fact that the signal at Rx 1 during time slots $t_1 \leq t \leq t_1 + n' - 1$ is $\boldsymbol{y}_t = \boldsymbol{s}'_{1,t} + \boldsymbol{z}_{1,t}$, in (b) we have added $\frac{\gamma_1 + a_2\gamma_2}{2}$ to both sides of the inequality in (a)

after multiplying both sides of the inequality by -1, in (c) both sides are divided by $1 + \frac{\gamma_1 + a_2\gamma_2}{2}$ and (d) is due to Lemma 2 (Bernstein's inequality) where it is assumed that $\epsilon < \frac{\gamma_1 + a_2\gamma_2}{2} \log e$. In fact, the random variables $\boldsymbol{w}_l = 1 - \frac{\frac{1}{2\gamma_2}|\boldsymbol{s}'_{1,l}|^2 + \frac{1}{2}(\boldsymbol{s}'_{1,l} + \boldsymbol{z}_{1,l+t_1} - \sqrt{a_2}\,\boldsymbol{s}'_{2,l})^2}{1 + \frac{\gamma_1 + a_2\gamma_2}{2}}$ are independent with zero mean and finite variance and $\boldsymbol{w}_l \leq 1$. Therefore, Bernstein's inequality applies.

Appendix K

Proof of Proposition 9

The j_1^{th} burst of Tx 1 overlaps with the j_2^{th} burst of Tx 2 if and only if one of the events

$$\mathcal{E}_n := \left\{ \boldsymbol{\tau}_1^{(j_1)} + 1 \leq \boldsymbol{\tau}_2^{(j_2)} + 1 \leq \boldsymbol{\tau}_1^{(j_1)} + n' + n_1 \right\} \tag{K.1}$$

or

$$\mathcal{F}_n := \left\{ \boldsymbol{\tau}_1^{(j_1)} + 1 \leq \boldsymbol{\tau}_2^{(j_2)} + n' + n_2 \leq \boldsymbol{\tau}_1^{(j_1)} + n' + n_1 \right\} \tag{K.2}$$

holds. Let us show that $\lim_{n\to\infty} \mathbb{P}(\mathcal{E}_n) = 0$ if and only if $j_2\mu_2 - j_1\mu_1 + \nu_2 - \nu_1 \in (0, \theta_1)$. Define

$$\boldsymbol{\rho}_n = \frac{1}{n}(\boldsymbol{\tau}_2^{(j_2)} - \boldsymbol{\tau}_1^{(j_1)}), \;\; \boldsymbol{\rho}'_n = \frac{1}{n}(\boldsymbol{\tau}_1^{(j_1)} - \boldsymbol{\tau}_2^{(j_2)} + n' + n_1 - 1). \tag{K.3}$$

Then

$$\mathcal{E}_n = \left\{ \boldsymbol{\rho}_n \geq 0, \boldsymbol{\rho}'_n \geq 0 \right\} = \left\{ \min\{\boldsymbol{\rho}_n, \boldsymbol{\rho}'_n\} \geq 0 \right\}. \tag{K.4}$$

Following similar arguments made in Appendix D,

$$\lim_{n\to\infty} \min\{\boldsymbol{\rho}_n, \boldsymbol{\rho}'_n\} = \min\{j_2\mu_2 - j_1\mu_1 + \nu_2 - \nu_1, j_1\mu_1 - j_2\mu_2 + \nu_1 - \nu_2 + \theta_1\}. \tag{K.5}$$

If $j_2\mu_2 - j_1\mu_1 + \nu_2 - \nu_1 \in (0, \theta_1)$, then $\min\{j_2\mu_2 - j_1\mu_1 + \nu_2 - \nu_1, j_1\mu_1 - j_2\mu_2 + \nu_1 - \nu_2 + \theta_1\} > 0$. Hence, $\lim_{n\to\infty} \mathbb{P}(\mathcal{E}_n) = 1$ by Lemma 4. Similarly, one can show that if $j_2\mu_2 - j_1\mu_1 + \nu_2 - \nu_1 \in (-\theta_2, \theta_1 - \theta_2)$, then $\lim_{n\to\infty} \mathbb{P}(\mathcal{F}_n) = 1$. It follows that if $j_2\mu_2 - j_1\mu_1 + \nu_2 - \nu_1 \in (0, \theta_1) \bigcup (-\theta_2, \theta_1 - \theta_2)$, then $\lim_{n\to\infty} \mathbb{P}(\mathcal{E}_n \bigcup \mathcal{F}_n) = 1$. Next, assume[1] $j_2\mu_2 - j_1\mu_1 + \nu_2 - \nu_1 \notin [0, \theta_1] \bigcup [-\theta_2, \theta_1 - \theta_2]$. Since $j_2\mu_2 - j_1\mu_1 + \nu_2 - \nu_1 \notin [0, \theta_1]$, then $\min\{j_2\mu_2 - j_1\mu_1 + \nu_2 - \nu_1, j_1\mu_1 - j_2\mu_2 + \nu_1 - \nu_2 + \theta_1\} < 0$ and we have $\lim_{n\to\infty} \mathbb{P}(\mathcal{E}_n) = 0$ by

[1]Recall that by (3.35), $j_2\mu_2 - j_1\mu_1 + \nu_2 - \nu_1 \notin \{0, \theta_1, -\theta_2, \theta_1 - \theta_2\}$.

Lemma 4. Similarly, $j_2\mu_2 - j_1\mu_1 + \nu_2 - \nu_1 \notin [-\theta_2, \theta_1 - \theta_2]$ results in $\lim_{n\to\infty} \mathbb{P}(\mathcal{F}_n) = 0$. But, $\mathbb{P}(\mathcal{E}_n \bigcup \mathcal{F}_n) \leq \mathbb{P}(\mathcal{E}_n) + \mathbb{P}(\mathcal{F}_n)$ and we get $\lim_{n\to\infty} \mathbb{P}(\mathcal{E}_n \bigcup \mathcal{F}_n) = 0$. Finally, the probability of the j_2^{th} burst of Tx 2 overlaping only with the preamble sequence in the j_1^{th} burst of Tx 1 vanishes as n grows due to Proposition 8. This completes the proof.

Appendix L

Proof of Proposition 10

Fix $\epsilon > 0$. We assume $i = 1$. Given the index j of the codeword of Tx 1, let both $\omega^- := \omega^-_{1,j}$ and $\omega^+ := \omega^+_{1,j}$ be nonzero, $\omega^- \neq \omega^+$ and $\omega_{1,j} = 0$. The proof can be easily extended to the cases $\omega^- = \omega^+$ or $\omega_{1,j} \geq 1$. Define the event $\mathcal{U}_n$ by

$$\begin{aligned}\mathcal{U}_n \;:=\; \Big\{ & \boldsymbol{\tau}_2^{(\omega^-)} + 1 < \boldsymbol{\tau}_1^{(j)} + n' + 1 \leq \boldsymbol{\tau}_2^{(\omega^-)} + n' + n_2 \\ & \leq \boldsymbol{\tau}_2^{(\omega^+)} + 1 \leq \boldsymbol{\tau}_1^{(j)} + n' + n_1 + \boldsymbol{\tau}_2^{(\omega^+)} + n' + n_2 \Big\}.\end{aligned} \tag{L.1}$$

The probability of error in decoding the j^{th} codeword of Tx 1 at Rx 1 is bounded as

$$\mathbb{P}(\text{error}) \leq \mathbb{P}(\text{error}, \mathcal{U}_n) + \mathbb{P}(\mathcal{U}_n^c) \leq \mathbb{P}(\text{error}, \mathcal{U}_n) + \epsilon, \tag{L.2}$$

where in the last step we have assumed n is large enough so that $\mathbb{P}(\mathcal{U}_n^c) \leq \epsilon$. This is due to Proposition 9 together with the fact that $\omega^-, \omega^+ \neq 0$. Under the event $\mathcal{U}_n$, error can happen in two ways. The first case is when at least one of (3.55), (3.56) or (3.57) is not satisfied for the actual transmitted codeword by Tx 1. We denote this error event by error_1. The second case is when all of (3.55), (3.56) and (3.57) are satisfied for a codeword that is different from the transmitted codeword by Tx 1. We denote this error event by error_2. Then

$$\mathbb{P}(\text{error}, \mathcal{U}_n) \leq \mathbb{P}(\text{error}_1, \mathcal{U}_n) + \mathbb{P}(\text{error}_2, \mathcal{U}_n). \tag{L.3}$$

Next, we address the two terms on the right side of (L.3) separately:

- The term $\mathbb{P}(\text{error}_1, \mathcal{U}_n)$: Here, we verify that (3.55) occurs with high probability for the actual transmitted codeword in the limit of large n. One can establish a similar result for (3.56) and (3.57) yielding $\lim_{n\to\infty} \mathbb{P}(\text{error}_1, \mathcal{U}_n) = 0$. Define the set $\widetilde{\mathcal{U}}_n$ by

$$\begin{aligned}\widetilde{\mathcal{U}}_n \;:=\; & \{(t_1,t_2,t_3)\in\mathbb{Z}^3 : t_2+1 < t_1+n'+1 \le t_2+n'+n_2 \\ & \qquad \le t_3+1 \le t_1+n'+n_1 < t_3+n'+n_2\}.\end{aligned} \tag{L.4}$$

Then

$$\mathcal{U}_n = \left\{(\boldsymbol{\tau}_1^{(j)}, \boldsymbol{\tau}_2^{(\omega^-)}, \boldsymbol{\tau}_2^{(\omega^+)}) \in \widetilde{\mathcal{U}}_n\right\}. \tag{L.5}$$

Assume $(\boldsymbol{s}_{1,l})_{l=0}^{n_1-1}$ is the j^{th} codeword sent by Tx 1. The probability that (3.55) does not occur for the actual transmitted codeword under $\mathcal{U}_n$ can be written as

$$\begin{aligned}&\mathbb{P}\left(\left((\boldsymbol{s}_{1,l})_{l=0}^{\boldsymbol{\tau}_2^{(\omega^-)}-\boldsymbol{\tau}_1^{(j)}+n_2-1}, (\boldsymbol{y}_{1,l})_{l=\boldsymbol{\tau}_1^{(j)}+n'+1}^{\boldsymbol{\tau}_2^{(\omega^-)}+n'+n_2}\right) \notin A_\epsilon^{(\boldsymbol{\tau}_2^{(\omega^-)}-\boldsymbol{\tau}_1^{(j)}+n_2)}[p^{(2)}], \mathcal{U}_n\right) \\ &= \sum_{(t_1,t_2,t_3)\in\widetilde{\mathcal{U}}_n} \mathbb{P}\left(\left((\boldsymbol{s}_{1,l})_{l=0}^{t_2-t_1+n_2-1}, (\boldsymbol{y}_{1,l})_{l=t_1+n'+1}^{t_2+n'+n_2}\right) \notin A_\epsilon^{(t_2-t_1+n_2)}[p^{(2)}]\right) \\ &\qquad\qquad \times \mathbb{P}\big((\boldsymbol{\tau}_1^{(j)}, \boldsymbol{\tau}_2^{(\omega^-)}, \boldsymbol{\tau}_2^{(\omega^+)}) = (t_1,t_2,t_3)\big).\end{aligned} \tag{L.6}$$

For any $(t_1,t_2,t_3)\in\widetilde{\mathcal{U}}_n$,

$$\begin{aligned}&\mathbb{P}\left(\left((\boldsymbol{s}_{1,l})_{l=0}^{t_2-t_1+n_2-1}, (\boldsymbol{y}_{1,l})_{l=t_1+n'+1}^{t_2+n'+n_2}\right) \notin A_\epsilon^{(t_2-t_1+n_2)}[p^{(2)}]\right) \\ &\le \mathbb{P}\left(\left|\frac{1}{t_2-t_1+n_2}\sum_{l=0}^{t_2-t_1+n_2-1} \log p^{(2)}(\boldsymbol{s}_{1,l}, \boldsymbol{y}_{1,l+t_1+n'+1}) + h(p^{(2)})\right| > \epsilon\right) \\ &\quad + \mathbb{P}\left(\left|\frac{1}{t_2-t_1+n_2}\sum_{l=0}^{t_2-t_1+n_2-1} \log p_1^{(2)}(\boldsymbol{s}_{1,l}) + h(p_1^{(2)})\right| > \epsilon\right) \\ &\quad + \mathbb{P}\left(\left|\frac{1}{t_2-t_1+n_2}\sum_{l=0}^{t_2-t_1+n_2-1} \log p_2^{(2)}(\boldsymbol{y}_{1,l+t_1+n'+1}) + h(p_2^{(2)})\right| > \epsilon\right),\end{aligned} \tag{L.7}$$

where $p_1^{(2)}$ and $p_2^{(2)}$ are the first and second marginals of $p^{(2)}$, respectively. The three terms on the right side of (L.7) can be treated similarly. Here, we only study the first term. Let us write

$$
\begin{aligned}
&\mathbb{P}\Big(\Big|\frac{1}{t_2-t_1+n_2}\sum_{l=0}^{t_2-t_1+n_2-1}\log p^{(2)}(\boldsymbol{s}_{1,l},\boldsymbol{y}_{1,l+t_1+n'+1})+h(p^{(2)})\Big|>\epsilon\Big)\\
&\leq \mathbb{P}\Big(\frac{1}{t_2-t_1+n_2}\sum_{l=0}^{t_2-t_1+n_2-1}\log p^{(2)}(\boldsymbol{s}_{1,l},\boldsymbol{y}_{1,l+t_1+n'+1})+h(p^{(2)})>\epsilon\Big)\\
&\quad+\mathbb{P}\Big(\frac{1}{t_2-t_1+n_2}\sum_{l=0}^{t_2-t_1+n_2-1}\log p^{(2)}(\boldsymbol{s}_{1,l},\boldsymbol{y}_{1,l+t_1+n'+1})+h(p^{(2)})<-\epsilon\Big).
\end{aligned}
\tag{L.8}
$$

The random variables $\log p^{(2)}(\boldsymbol{s}_{1,l},\boldsymbol{y}_{1,l+t_1+n'+1})$ for $0\leq l\leq t_2-t_1+n_2-1$ are independent and identically distributed with expectation $-h(p^{(2)})$. Using Chernoff's bounding technique [28] and for $r>0$, we can find an upper bound on the first term on the right side of (L.8) as

$$
\begin{aligned}
&\mathbb{P}\Big(\frac{1}{t_2-t_1+n_2}\sum_{l=0}^{t_2-t_1+n_2-1}\log p^{(2)}(\boldsymbol{s}_{1,l},\boldsymbol{y}_{1,l+t_1+n'+1})+h(p^{(2)})>\epsilon\Big)\\
&\leq 2^{-r(t_2-t_1+n_2)(\epsilon-h(p^{(2)}))}\left(\mathbb{E}\left[2^{r\log p^{(2)}(\boldsymbol{s}_{1,0},\boldsymbol{y}_{1,t_1+n'+1})}\right]\right)^{t_2-t_1+n_2}\\
&=2^{-r(t_2-t_1+n_2)(\epsilon-h(p^{(2)}))}\left(\mathbb{E}\left[\left(p^{(2)}(\boldsymbol{s}_{1,0},\boldsymbol{y}_{1,t_1+n'+1})\right)^r\right]\right)^{t_2-t_1+n_2}.
\end{aligned}
\tag{L.9}
$$

For notational simplicity and with a slight abuse of notation, let us write $\boldsymbol{y}_{1,t_1+n'+1}=\boldsymbol{s}_{1,0}+\sqrt{a_2}\,\boldsymbol{s}_2+\boldsymbol{z}_1$ where $\boldsymbol{s}_2\sim \mathrm{N}(0,\gamma_2)$ is a symbol of the $\omega^{-\ th}$ transmitted codeword by Tx 2 and $\boldsymbol{z}_1:=\boldsymbol{z}_{1,t_1+n'+1}\sim\mathrm{N}(0,1)$. We have

$$
\begin{aligned}
p^{(2)}(\boldsymbol{s}_{1,0},\boldsymbol{y}_{1,t_1+n'+1}) &= \mathrm{g}(\boldsymbol{s}_{1,0};\gamma_1)\mathrm{g}(\boldsymbol{y}_{1,t_1+n'+1}-\boldsymbol{s}_{1,0};1+a_2\gamma_2)\\
&= \mathrm{g}(\boldsymbol{s}_{1,0};\gamma_1)\mathrm{g}(\sqrt{a_2}\,\boldsymbol{s}_2+\boldsymbol{z}_1;1+a_2\gamma_2)
\end{aligned}
\tag{L.10}
$$

and

$$\begin{aligned}\mathbb{E}\left[\left(p^{(2)}(\boldsymbol{s}_{1,0},\boldsymbol{y}_{1,t_1+n'+1})\right)^r\right] &= \mathbb{E}\left[(\mathrm{g}(\boldsymbol{s}_{1,0};\gamma_1))^r\,(\mathrm{g}(\sqrt{a_2}\,\boldsymbol{s}_2+\boldsymbol{z}_1;1+a_2\gamma_2))^r\right]\\ &= \mathbb{E}\left[(\mathrm{g}(\boldsymbol{s}_{1,0};\gamma_1))^r\right]\mathbb{E}\left[(\mathrm{g}(\sqrt{a_2}\,\boldsymbol{s}_2+\boldsymbol{z}_1;1+a_2\gamma_2))^r\right]\\ &= \frac{1}{(1+r)\left(2\pi\sqrt{\gamma_1(1+a_2\gamma_2)}\right)^r}\\ &= \frac{e^r}{(1+r)2^{rh(p^{(2)})}}, \qquad \text{(L.11)}\end{aligned}$$

where the penultimate step is due to the fact that for $\boldsymbol{x}\sim \mathrm{N}(0,1)$ and any $u>-\frac{1}{2}$, we have $\mathbb{E}[e^{-u\boldsymbol{x}^2}]=\frac{1}{\sqrt{1+2u}}$ and the last step is due to $h(p^{(2)})=\log(2\pi e\sqrt{\gamma_1(1+a_2\gamma_2)})$. By (L.9) and (L.11),

$$\mathbb{P}\Big(\frac{1}{t_2-t_1+n_2}\sum_{l=0}^{t_2-t_1+n_2-1}\log p^{(2)}(\boldsymbol{s}_{1,l},\boldsymbol{y}_{1,l+t_1+n'+1})+h(p^{(2)})>\epsilon\Big)\le e^{-(t_2-t_1+n_2)u(r)}, \qquad \text{(L.12)}$$

where

$$u(r):=\ln(1+r)-r(1-\epsilon\ln 2),\ \ r>0. \qquad \text{(L.13)}$$

Following a similar approach that led us to (L.12), one can show that the second term on the right side of (L.8) is bounded from above as

$$\mathbb{P}\Big(\frac{1}{t_2-t_1+n_2}\sum_{l=0}^{t_2-t_1+n_2-1}\log p^{(2)}(\boldsymbol{s}_{1,l},\boldsymbol{y}_{1,l+t_1+n'+1})+h(p^{(2)})<-\epsilon\Big)\le e^{-(t_2-t_1+n_2)v(r)} \qquad \text{(L.14)}$$

where

$$v(r):=\ln(1-r)+r(1+\epsilon\ln 2),\ \ 0<r<1. \qquad \text{(L.15)}$$

Regardless of the value of $\epsilon>0$, there always exists an $0<r_0<1$ such that $u(r),v(r)>0$ for $0<r<r_0$. It is understood that we take r inside $(0,r_0)$. It is easy to see that for any $0<r<1$, $v(r)<u(r)$.[1] Using this fact together with (L.8),

[1] We have $v(r)<u(r)$ if and only if $w(r):=\ln(1+r)-\ln(1-r)-2r>0$. Note that $w(0)=0$ and $\frac{\mathrm{d}w}{\mathrm{d}r}=\frac{2r^2}{1-r^2}>0$ for any $0<r<1$. Therefore, $w(r)>0$ for any $0<r<1$ by the mean value theorem.

(L.12) and (L.14),

$$\mathbb{P}\left(\left|\frac{1}{t_2-t_1+n_2}\sum_{l=0}^{t_2-t_1+n_2-1}\log p^{(2)}(\boldsymbol{s}_{1,l},\boldsymbol{y}_{1,l+t_1+n'+1})+h(p^{(2)})\right|>\epsilon\right)\leq 2e^{-(t_2-t_1+n_2)v(r)}. \tag{L.16}$$

It can be shown similarly that the second and third terms on the right side of (L.7) are bounded from above by $2e^{-(t_2-t_1+n_2)v(r)}$. Therefore,

$$\mathbb{P}\left(\left((\boldsymbol{s}_{1,l})_{l=0}^{t_2-t_1+n_2-1},(\boldsymbol{y}_{1,l})_{l=t_1+n'+1}^{t_2+n'+n_2}\right)\notin A_\epsilon^{(t_2-t_1+n_2)}[p^{(2)}]\right)\leq 6e^{-(t_2-t_1+n_2)v(r)}. \tag{L.17}$$

Using (L.17) in (L.6),

$$\begin{aligned}
&\mathbb{P}\Big(\big((\boldsymbol{s}_{1,l})_{l=0}^{\boldsymbol{\tau}_2^{(\omega^-)}-\boldsymbol{\tau}_1^{(j)}+n_2-1},(\boldsymbol{y}_{1,l})_{l=\boldsymbol{\tau}_1^{(j)}+n'+1}^{\boldsymbol{\tau}_2^{(\omega^-)}+n'+n_2}\big)\notin A_\epsilon^{(\boldsymbol{\tau}_2^{(\omega^-)}-\boldsymbol{\tau}_1^{(j)}+n_2)}[p^{(2)}],\,\mathcal{U}_n\Big)\\
&\leq\sum_{(t_1,t_2,t_3)\in\widetilde{\mathcal{U}}_n}6e^{-(t_2-t_1+n_2)v(r)}\mathbb{P}((\boldsymbol{\tau}_1^{(j)},\boldsymbol{\tau}_2^{(\omega^-)},\boldsymbol{\tau}_2^{(\omega^+)})=(t_1,t_2,t_3))\\
&\overset{(a)}{\leq}\sum_{(t_1,t_2,t_3)\in\mathbb{Z}^3}6e^{-(t_2-t_1+n_2)v(r)}\mathbb{P}((\boldsymbol{\tau}_1^{(j)},\boldsymbol{\tau}_2^{(\omega^-)},\boldsymbol{\tau}_2^{(\omega^+)})=(t_1,t_2,t_3))\\
&=6\mathbb{E}\Big[e^{-(\boldsymbol{\tau}_2^{(\omega^-)}-\boldsymbol{\tau}_1^{(j)}+n_2)v(r)}\Big]\\
&=6e^{-(\lfloor n\nu_2\rfloor-\lfloor n\nu_1\rfloor+n_2)v(r)}\mathbb{E}\Big[e^{-(\boldsymbol{\xi}_2^{(\omega^-)}-\boldsymbol{\xi}_1^{(j)})v(r)}\Big]\\
&=6e^{-(\lfloor n\nu_2\rfloor-\lfloor n\nu_1\rfloor+n_2)v(r)}\mathbb{E}\Big[e^{-v(r)\boldsymbol{\xi}_2^{(\omega^-)}}\Big]\mathbb{E}\Big[e^{v(r)\boldsymbol{\xi}_1^{(j)}}\Big],
\end{aligned} \tag{L.18}$$

where in (a) we have removed the constraint $(t_1,t_2,t_3)\in\widetilde{\mathcal{U}}_n$ and the last step is due to independence of $\boldsymbol{\xi}_2^{(\omega^-)}$ and $\boldsymbol{\xi}_1^{(j)}$. Recalling the expression for the moment generating function of a negative Binomial random variable[2], we get

$$\mathbb{E}\Big[e^{-v(r)\boldsymbol{\xi}_2^{(\omega^-)}}\Big]=\left(\frac{q_2e^{-v(r)}}{1-(1-q_2)e^{-v(r)}}\right)^{\frac{\omega^-\lfloor n\eta_2\rfloor}{k_2}} \tag{L.19}$$

[2]The moment generating function of $\boldsymbol{x}\sim\mathrm{NB}(n,p)$ is given by $\mathbb{E}[e^{t\boldsymbol{x}}]=\left(\frac{pe^t}{1-(1-p)e^t}\right)^n$ for $t<-\ln(1-p)$.

and

$$\mathbb{E}\Big[e^{v(r)\boldsymbol{\xi}_1^{(j)}}\Big] = \left(\frac{q_1 e^{v(r)}}{1-(1-q_1)e^{v(r)}}\right)^{\frac{j\lfloor n\eta_1\rfloor}{k_1}}, \tag{L.20}$$

where (L.20) holds as long as $v(r) < -\ln(1-q)$. Since $\lim_{r\to 0^+} v(r) = 0$, one can make sure the constraint $v(r) < -\ln(1-q)$ holds by choosing r small enough. By (L.18), (L.19) and (L.20),

$$\begin{aligned}&\mathbb{P}\Big(\big((\boldsymbol{s}_{1,l})_{l=0}^{\boldsymbol{\tau}_2^{(\omega^-)}-\boldsymbol{\tau}_1^{(j)}+n_2-1}, (\boldsymbol{y}_{1,l})_{l=\boldsymbol{\tau}_1^{(j)}+n'+1}^{\boldsymbol{\tau}_2^{(\omega^-)}+n'+n_2}\big) \notin A_\epsilon^{(\boldsymbol{\tau}_2^{(\omega^-)}-\boldsymbol{\tau}_1^{(j)}+n_2)}[p^{(2)}], \mathcal{U}_n\Big)\\ &\le 6e^{-(\lfloor n\nu_2\rfloor-\lfloor n\nu_1\rfloor+n_2)v(r)}\left(\frac{q_2e^{-v(r)}}{1-(1-q_2)e^{-v(r)}}\right)^{\frac{\omega^-\lfloor n\eta_2\rfloor}{k_2}}\left(\frac{q_1e^{v(r)}}{1-(1-q_1)e^{v(r)}}\right)^{\frac{j\lfloor n\eta_1\rfloor}{k_1}}.\end{aligned} \tag{L.21}$$

Using the identity $\ln\frac{ae^x}{1-(1-a)e^x} = \frac{x}{a} + o(x)$ for $0 < a < 1$, one can write (L.21) as

$$\begin{aligned}&\mathbb{P}\Big(\big((\boldsymbol{s}_{1,l})_{l=0}^{\boldsymbol{\tau}_2^{(\omega^-)}-\boldsymbol{\tau}_1^{(j)}+n_2-1}, (\boldsymbol{y}_{1,l})_{l=\boldsymbol{\tau}_1^{(j)}+n'+1}^{\boldsymbol{\tau}_2^{(\omega^-)}+n'+n_2}\big) \notin A_\epsilon^{(\boldsymbol{\tau}_2^{(\omega^-)}-\boldsymbol{\tau}_1^{(j)}+n_2)}[p^{(2)}], \mathcal{U}_n\Big)\\ &\le 6e^{-(\lfloor n\nu_2\rfloor-\lfloor n\nu_1\rfloor+n_2)v(r)}e^{\frac{\omega^-\lfloor n\eta_2\rfloor}{k_2}(-\frac{v(r)}{q_2}+o(v(r)))}e^{\frac{j\lfloor n\eta_1\rfloor}{k_1}(\frac{v(r)}{q_1}+o(v(r)))}\\ &= e^{-nv(r)f(n)},\end{aligned} \tag{L.22}$$

where

$$f(n) := \frac{1}{n}\left(\frac{\omega^-\lfloor n\eta_2\rfloor}{\lambda_2} - \frac{j\lfloor n\eta_1\rfloor}{\lambda_1} + \lfloor n\nu_2\rfloor - \lfloor n\nu_1\rfloor + n_2 - (\lfloor n\eta_1\rfloor + \lfloor n\eta_2\rfloor)\frac{o(v(r))}{v(r)}\right). \tag{L.23}$$

But, $\lim_{n\to\infty} f(n) = \omega^-\mu_2 - j\mu_1 + \nu_2 - \nu_1 + \theta_2 + (\eta_1+\eta_2)\frac{o(v(r))}{v(r)}$. By definition, ω^- satisfies $\omega^-\mu_2 - j\mu_1 + \nu_2 - \nu_1 + \theta_2 > 0$. Since $\frac{o(v(r))}{v(r)}$ can be made arbitrarily small by choosing r small enough, we conclude that $\lim_{n\to\infty} f(n) > 0$. This together with (L.22) implies that $e^{-nv(r)f(n)}$ decays exponentially with n as desired.

- The term $\mathbb{P}(\text{error}_2, \mathcal{U}_n)$: Let $\delta > 0$ and define

$$\mathcal{V}_n := \left\{\max\left\{\left|\frac{\boldsymbol{\tau}_2^{(\omega^-)}}{n} - (\omega^-\mu_2+\nu_2)\right|, \left|\frac{\boldsymbol{\tau}_2^{(\omega^+)}}{n} - (\omega^+\mu_2+\nu_2)\right|\right\} < \delta\right\}. \tag{L.24}$$

Also define $\widetilde{\mathcal{V}}_n$ by

$$\widetilde{\mathcal{V}}_n := \left\{(t_2, t_3) \in \mathbb{Z}^2 : \max\left\{\left|\frac{t_2}{n} - (\omega^- \mu_2 + \nu_2)\right|, \left|\frac{t_3}{n} - (\omega^+ \mu_2 + \nu_2)\right|\right\} < \delta\right\}. \tag{L.25}$$

We can write

$$\mathbb{P}(\text{error}_2, \mathcal{U}_n) \leq \mathbb{P}(\text{error}_2, \mathcal{U}_n, \mathcal{V}_n) + \mathbb{P}(\mathcal{V}_n^c). \tag{L.26}$$

By SLLN, $\lim_{n\to\infty} \frac{\boldsymbol{\tau}_2^{(\omega^-)}}{n} = \omega^- \mu_2 + \nu_2$. Therefore, $\frac{\boldsymbol{\tau}_2^{(\omega^-)}}{n}$ also converges to $\omega^- \mu_2 + \nu_2$ in probability and one can select n large enough so that

$$\mathbb{P}\Big(\Big|\frac{\boldsymbol{\tau}_2^{(\omega^-)}}{n} - (\omega^- \mu_2 + \nu_2)\Big| \geq \delta\Big) < \epsilon/3. \tag{L.27}$$

Similarly,

$$\mathbb{P}\Big(\Big|\frac{\boldsymbol{\tau}_2^{(\omega^+)}}{n} - (\omega^+ \mu_2 + \nu_2)\Big| \geq \delta\Big) < \epsilon/3 \tag{L.28}$$

holds for large enough n. It follows that if n is sufficiently large, then $\mathbb{P}(\mathcal{V}_n^c) < \epsilon$. To find an upper bound on $\mathbb{P}(\text{error}_2, \mathcal{U}_n, \mathcal{V}_n)$, let us label the messages of Tx 1 as message 1 to message $2^{\lfloor n\eta_1 \rfloor}$. Assume the j^{th} transmitted message of Tx 1 is message 1 and $(\boldsymbol{s}_{1,l})_{l=0}^{n_1-1}$ is the codeword of user 1 corresponding to message 2. Recall that $\mathbb{P}(\text{error}_2, \mathcal{U}_n, \mathcal{V}_n)$ is the probability of the event that a codeword different from the transmitted codeword satisfies (3.55), (3.56) and (3.57). Then

$$\mathbb{P}(\text{error}_2, \mathcal{U}_n, \mathcal{V}_n) \leq 2^{\lfloor n\eta_1 \rfloor} \mathbb{P}\left(\mathcal{E} \bigcap \mathcal{F} \bigcap \mathcal{G} \bigcap \mathcal{U}_n \bigcap \mathcal{V}_n\right), \tag{L.29}$$

where

$$\mathcal{E} = \left\{\left((\boldsymbol{s}_{1,l})_{l=0}^{\boldsymbol{\tau}_2^{(\omega^-)} - \boldsymbol{\tau}_1^{(j)} + n_2 - 1}, (\boldsymbol{y}_{1,l})_{l=\boldsymbol{\tau}_1^{(j)} + n' + 1}^{\boldsymbol{\tau}_2^{(\omega^-)} + n' + n_2}\right) \in A_\epsilon^{(\boldsymbol{\tau}_2^{(\omega^-)} - \boldsymbol{\tau}_1^{(j)} + n_2)}[p^{(2)}]\right\}, \tag{L.30}$$

$$\mathcal{F} = \left\{\left((\boldsymbol{s}_{1,l})_{l=\boldsymbol{\tau}_2^{(\omega^-)} - \boldsymbol{\tau}_1^{(j)} + n_2}^{\boldsymbol{\tau}_2^{(\omega^+)} - \boldsymbol{\tau}_1^{(j)} - n' - 1}, (\boldsymbol{y}_{1,l})_{l=\boldsymbol{\tau}_2^{(\omega^-)} + n' + n_2 + 1}^{\boldsymbol{\tau}_2^{(\omega^+)}}\right) \in A_\epsilon^{(\boldsymbol{\tau}_2^{(\omega^+)} - \boldsymbol{\tau}_2^{(\omega^-)} - n' - n_2)}[p^{(1)}]\right\} \tag{L.31}$$

and

$$\mathcal{G} = \left\{ \left((\boldsymbol{s}_{1,l})_{l=\boldsymbol{\tau}_2^{(\omega^+)}-\boldsymbol{\tau}_1^{(j)}-n'}^{n_1-1}, (\boldsymbol{y}_{1,l})_{l=\boldsymbol{\tau}_2^{(\omega^+)}+1}^{\boldsymbol{\tau}_1^{(j)}+n'+n_1} \right) \in A_\epsilon^{(\boldsymbol{\tau}_1^{(j)}-\boldsymbol{\tau}_2^{(\omega^+)}+n'+n_1)}[p^{(2)}] \right\}. \quad \text{(L.32)}$$

Recalling the definition of $\mathcal{V}_n$ in (L.4),

$$\begin{aligned}
&\mathbb{P}\left(\mathcal{E} \bigcap \mathcal{F} \bigcap \mathcal{G} \bigcap \mathcal{U}_n \bigcap \mathcal{V}_n\right) \\
&= \sum_{\substack{(t_1,t_2,t_3)\in\widetilde{\mathcal{U}}_n \\ (t_2,t_3)\in\widetilde{\mathcal{V}}_n}} \mathbb{P}\left(\mathcal{E} \bigcap \mathcal{F} \bigcap \mathcal{G} \,\middle|\, (\boldsymbol{\tau}_1^{(j)}, \boldsymbol{\tau}_2^{(\omega^-)}, \boldsymbol{\tau}_2^{(\omega^+)}) = (t_1, t_2, t_3)\right) \\
&\qquad \times \mathbb{P}\big((\boldsymbol{\tau}_1^{(j)}, \boldsymbol{\tau}_2^{(\omega^-)}, \boldsymbol{\tau}_2^{(\omega^+)}) = (t_1, t_2, t_3)\big). \qquad \text{(L.33)}
\end{aligned}$$

For any $(t_1, t_2, t_3) \in \widetilde{\mathcal{U}}_n$,

$$\mathbb{P}\left(\mathcal{E} \bigcap \mathcal{F} \bigcap \mathcal{G} \,\middle|\, (\boldsymbol{\tau}_1^{(j)}, \boldsymbol{\tau}_2^{(\omega^-)}, \boldsymbol{\tau}_2^{(\omega^+)}) = (t_1, t_2, t_3)\right) = \mathbb{P}(\mathcal{E}_{t_1,t_2})\mathbb{P}(\mathcal{F}_{t_1,t_2,t_3})\mathbb{P}(\mathcal{G}_{t_1,t_3}), \quad \text{(L.34)}$$

where

$$\mathcal{E}_{t_1,t_2} = \left\{ \left((\boldsymbol{s}_{1,l})_{l=0}^{t_2-t_1+n_2-1}, (\boldsymbol{y}_{1,l})_{l=t_1+n'+1}^{t_2+n'+n_2} \right) \in A_\epsilon^{(t_2-t_1+n_2)}[p^{(2)}] \right\}, \quad \text{(L.35)}$$

$$\mathcal{F}_{t_1,t_2,t_3} = \left\{ \left((\boldsymbol{s}_{1,l})_{l=t_2-t_1+n_2}^{t_3-t_1-n'-1}, (\boldsymbol{y}_{1,l})_{l=t_2+n'+n_2+1}^{t_3} \right) \in A_\epsilon^{(t_3-t_2-n'-n_2)}[p^{(1)}] \right\} \quad \text{(L.36)}$$

and

$$\mathcal{G}_{t_1,t_3} = \left\{ \left((\boldsymbol{s}_{1,l})_{l=t_3-t_1-n'}^{n_1-1}, (\boldsymbol{y}_{1,l})_{l=t_3+1}^{t_1+n'+n_1} \right) \in A_\epsilon^{(t_1-t_3+n'+n_1)}[p^{(2)}] \right\}. \quad \text{(L.37)}$$

The reason behind (L.34) is that fixing $(\boldsymbol{\tau}_1^{(j)}, \boldsymbol{\tau}_2^{(\omega^-)}, \boldsymbol{\tau}_2^{(\omega^+)}) = (t_1, t_2, t_3)$, the events $\mathcal{E}$, $\mathcal{F}$ and $\mathcal{G}$ are independent as they depend on non-overlapping segments of the sequences $(\boldsymbol{s}_{1,l})_{l=0}^{n_1-1}$ and $(\boldsymbol{y}_{1,l})_{l=t_1+n'+1}^{t_1+n'+n_1}$. Using the standard properties of jointly typical sequences [28], we have

$$\mathbb{P}(\mathcal{E}_{t_1,t_2}) \leq 2^{-(t_2-t_1+n_2)(\kappa_1'-3\epsilon)}, \quad \text{(L.38)}$$

$$\mathbb{P}(\mathcal{F}_{t_1,t_2,t_3}) \leq 2^{-(t_3-t_2-n'-n_2)(\kappa_1-3\epsilon)} \tag{L.39}$$

and

$$\mathbb{P}(\mathcal{G}_{t_1,t_3}) \leq 2^{-(t_1-t_3+n'+n_1)(\kappa_1'-3\epsilon)}. \tag{L.40}$$

By (L.33), (L.34), (L.38), (L.39) and (L.40),

$$\begin{aligned} &\mathbb{P}\left(\mathcal{E}\bigcap\mathcal{F}\bigcap\mathcal{G}\bigcap\mathcal{U}_n\bigcap\mathcal{V}_n\right) \\ &\leq 2^{-n'(\kappa_1'-\kappa_1)}2^{-n_2(\kappa_1'-\kappa_1)}2^{-n_1(\kappa_1'-3\epsilon)} \\ &\qquad\times \sum_{\substack{(t_1,t_2,t_3)\in\widetilde{\mathcal{U}}_n \\ (t_2,t_3)\in\widetilde{\mathcal{V}}_n}} 2^{-(t_3-t_2)(\kappa_1-\kappa_1')}\mathbb{P}\big((\boldsymbol{\tau}_1^{(j)},\boldsymbol{\tau}_2^{(\omega^-)},\boldsymbol{\tau}_2^{(\omega^+)}) = (t_1,t_2,t_3)\big). \end{aligned} \tag{L.41}$$

We can write

$$\begin{aligned} &\sum_{\substack{(t_1,t_2,t_3)\in\widetilde{\mathcal{U}}_n \\ (t_2,t_3)\in\widetilde{\mathcal{V}}_n}} 2^{-(t_3-t_2)(\kappa_1-\kappa_1')}\mathbb{P}\big((\boldsymbol{\tau}_1^{(j)},\boldsymbol{\tau}_2^{(\omega^-)},\boldsymbol{\tau}_2^{(\omega^+)}) = (t_1,t_2,t_3)\big) \\ &\overset{(a)}{\leq} 2^{-n(\omega^+\mu_2+\nu_2-\delta-\omega^-\mu_2-\nu_2-\delta)(\kappa_1-\kappa_1')} \sum_{\substack{(t_1,t_2,t_3)\in\widetilde{\mathcal{U}}_n \\ (t_2,t_3)\in\widetilde{\mathcal{V}}_n}} \mathbb{P}\big((\boldsymbol{\tau}_1^{(j)},\boldsymbol{\tau}_2^{(\omega^-)},\boldsymbol{\tau}_2^{(\omega^+)}) = (t_1,t_2,t_3)\big) \\ &\overset{(b)}{\leq} 2^{-n(\mu_2-2\delta)(\kappa_1-\kappa_1')}, \end{aligned} \tag{L.42}$$

where (a) is due to the fact that if $(t_2,t_3)\in\widetilde{\mathcal{V}}_n$, then $t_2 \leq n(\omega^-\mu_2+\nu_2+\delta)$ and $t_3 \geq n(\omega^+\mu_2+\nu_2-\delta)$ and (b) is due to $\sum_{(t_1,t_2,t_3)\in\widetilde{\mathcal{U}}_n,\,(t_2,t_3)\in\widetilde{\mathcal{V}}_n}\mathbb{P}\big((\boldsymbol{\tau}_1^{(j)},\boldsymbol{\tau}_2^{(\omega^-)},\boldsymbol{\tau}_2^{(\omega^+)}) = (t_1,t_2,t_3)\big) \leq 1$ and the fact that $\omega^+-\omega^-=1$. By (L.29), (L.41) and (L.42),

$$\mathbb{P}(\mathrm{error}_2,\mathcal{U}_n,\mathcal{V}_n) \leq 2^{-nf(n)}, \tag{L.43}$$

where

$$f(n) := (\mu_2-2\delta)(\kappa_1-\kappa_1') + \frac{n_2}{n}(\kappa_1'-\kappa_1) + \frac{n_1}{n}(\kappa_1'-3\epsilon) + \frac{n'}{n}(\kappa_1'-\kappa_1) - \frac{\lfloor n\eta_1\rfloor}{n}. \tag{L.44}$$

We have $\lim_{n\to\infty} f(n) = \mu_2(\kappa_1 - \kappa_1') + \theta_2(\kappa_1' - \kappa_1) + \theta_1\kappa_1' - \eta_1 - 2(\kappa_1 - \kappa_1')\delta - 3\theta_1\epsilon$. By (3.58), $\mu_2(\kappa_1 - \kappa_1') + \theta_2(\kappa_1' - \kappa_1) + \theta_1\kappa_1' - \eta_1 > 0$. Therefore, $\lim_{n\to\infty} f(n) > 0$ for sufficiently small δ and ϵ and $\mathbb{P}(\text{error}_2, \mathcal{U}_n, \mathcal{V}_n)$ decays exponentially with n as desired.

Appendix M

Proof of Proposition 11

Given the index j of the codeword of Tx 1, we assume $\omega^- \neq 0$, $\omega^+ = 0$ and $\omega_{1,j} = 0$. The proof can be easily extended for arbitrary $\omega_{1,j} \geq 1$. Define $\mathcal{U}_n$ by

$$\mathcal{U}_n := \left\{ \boldsymbol{\tau}_2^{(\omega^-)} + 1 < \boldsymbol{\tau}_1^{(j)} + n' + 1 \leq \boldsymbol{\tau}_2^{(\omega^-)} + n' + n_2 \leq \boldsymbol{\tau}_1^{(j)} + n' + n_1 < \boldsymbol{\tau}_2^{\omega^- + 1} + 1 \right\}. \tag{M.1}$$

Also, let

$$\widetilde{\mathcal{U}}_n := \left\{ (t_1, t_2, t_3) \in \mathbb{Z}^3 : t_2 + 1 < t_1 + n' + 1 \leq t_2 + n' + n_2 \leq t_1 + n' + n_1 < t_3 + 1 \right\} \tag{M.2}$$

The probability of error in decoding the j^{th} codeword of Tx 1 at Rx 1 is bounded as

$$\mathbb{P}(\text{error}) \leq \mathbb{P}(\text{error}, \mathcal{U}_n) + \mathbb{P}(\mathcal{U}_n^c) \leq \mathbb{P}(\text{error}, \mathcal{U}_n) + \epsilon, \tag{M.3}$$

where in the last step we have assumed n is large enough so that $\mathbb{P}(\mathcal{U}_n^c) \leq \epsilon$. This follows by Proposition 9 together with the facts that $\omega^- \neq 0$, $\omega^+ = 0$ and $\omega_{1,j} = 0$. Under the event $\mathcal{U}_n$, error can happen in two possible ways. The first case is when at least one of (3.66) or (3.67) is not satisfied for the actual transmitted codeword by Tx 1. We denote this error event by error_1. The second case is when both (3.66) and (3.67) are satisfied for a codeword that is different from the transmitted codeword by Tx 1. We denote this error event by error_2. Then

$$\mathbb{P}(\text{error}, \mathcal{U}_n) \leq \mathbb{P}(\text{error}_1, \mathcal{U}_n) + \mathbb{P}(\text{error}_2, \mathcal{U}_n). \tag{M.4}$$

Analysis of $\mathbb{P}(\text{error}_1, \mathcal{U}_n)$ is quite similar to the one offered in Appendix G. Here, we only address $\mathbb{P}(\text{error}_2, \mathcal{U}_n)$. Let $\delta > 0$ and define

$$\mathcal{V}_n := \left\{ \max\left\{ \left| \frac{\boldsymbol{\tau}_1^{(j)}}{n} - (j\mu_1 + \nu_1) \right|, \left| \frac{\boldsymbol{\tau}_2^{(\omega^-)}}{n} - (\omega^- \mu_2 + \nu_2) \right| \right\} < \delta \right\}. \tag{M.5}$$

Also, let

$$\widetilde{\mathcal{V}}_n := \left\{ (t_1, t_2) \in \mathbb{Z}^2 : \max\left\{ \left| \frac{t_1}{n} - (j\mu_1 + \nu_1) \right|, \left| \frac{t_2}{n} - (\omega^- \mu_2 + \nu_2) \right| \right\} < \delta \right\}. \tag{M.6}$$

We can write

$$\mathbb{P}(\text{error}_2, \mathcal{U}_n) \leq \mathbb{P}(\text{error}_2, \mathcal{U}_n, \mathcal{V}_n) + \mathbb{P}(\mathcal{V}_n^c) \leq \mathbb{P}(\text{error}_2, \mathcal{U}_n, \mathcal{V}_n) + \epsilon, \tag{M.7}$$

where the last step we are assuming that n is large enough such that $\mathbb{P}(\mathcal{V}_n^c) \leq \epsilon$ following a similar reasoning in (L.27) in Appendix F. Let us label the messages of Tx 1 as message 1 to message $2^{\lfloor n\eta_1 \rfloor}$. Assume the j^{th} transmitted message of Tx 1 is message 1 and $(\boldsymbol{s}_{1,l})_{l=0}^{n_1-1}$ is the codeword corresponding to message 2. Then

$$\mathbb{P}(\text{error}_2, \mathcal{U}_n, \mathcal{V}_n) \leq 2^{\lfloor n\eta_1 \rfloor} \mathbb{P}\left(\mathcal{E} \bigcap \mathcal{F} \bigcap \mathcal{U}_n \bigcap \mathcal{V}_n \right), \tag{M.8}$$

where

$$\mathcal{E} = \left\{ \left((\boldsymbol{s}_{1,l})_{l=0}^{\boldsymbol{\tau}_2^{(\omega^-)} - \boldsymbol{\tau}_1^{(j)} + n_2 - 1}, (\boldsymbol{y}_{1,l})_{l=\boldsymbol{\tau}_1^{(j)} + n' + 1}^{\boldsymbol{\tau}_2^{(\omega^-)} + n' + n_2} \right) \in A_\epsilon^{(\boldsymbol{\tau}_2^{(\omega^-)} - \boldsymbol{\tau}_1^{(j)} + n_2)}[p^{(2)}] \right\} \tag{M.9}$$

and

$$\mathcal{F} = \left\{ \left((\boldsymbol{s}_{1,l})_{l=\boldsymbol{\tau}_2^{(\omega^-)} - \boldsymbol{\tau}_1^{(j)} + n_2}^{n_1 - 1}, (\boldsymbol{y}_{1,l})_{l=\boldsymbol{\tau}_2^{(\omega^-)} + n' + n_2 + 1}^{\boldsymbol{\tau}_1^{(j)} + n' + n_1} \right) \in A_\epsilon^{(\boldsymbol{\tau}_1^{(j)} - \boldsymbol{\tau}_2^{(\omega^-)} + n_1 - n_2)}[p^{(1)}] \right\}. \tag{M.10}$$

Then

$$\mathbb{P}\left(\mathcal{E} \bigcap \mathcal{F} \bigcap \mathcal{U}_n \bigcap \mathcal{V}_n \right) = \sum_{\substack{(t_1,t_2,t_3) \in \widetilde{\mathcal{U}}_n \\ (t_1,t_2) \in \widetilde{\mathcal{V}}_n}} \mathbb{P}\left(\mathcal{E} \bigcap \mathcal{F} \,\middle|\, (\boldsymbol{\tau}_1^{(j)}, \boldsymbol{\tau}_2^{(\omega^-)}, \boldsymbol{\tau}_2^{(\omega^- + 1)}) = (t_1, t_2, t_3) \right) \times \mathbb{P}\left((\boldsymbol{\tau}_1^{(j)}, \boldsymbol{\tau}_2^{(\omega^-)}, \boldsymbol{\tau}_2^{(\omega^- + 1)}) = (t_1, t_2, t_3) \right) \tag{M.11}$$

For any $(t_1, t_2, t_3) \in \widetilde{\mathcal{U}}_n$,

$$\mathbb{P}\Big(\mathcal{E} \bigcap \mathcal{F} \,\big|\, (\boldsymbol{\tau}_1^{(j)}, \boldsymbol{\tau}_2^{(\omega^-)}, \boldsymbol{\tau}_2^{(\omega^-+1)}) = (t_1, t_2, t_3)\Big) = \mathbb{P}(\mathcal{E}_{t_1,t_2})\mathbb{P}(\mathcal{F}_{t_1,t_2}), \tag{M.12}$$

where

$$\mathcal{E}_{t_1,t_2} = \left\{ \left((\boldsymbol{s}_{1,l})_{l=0}^{t_2-t_1+n_2-1}, (\boldsymbol{y}_{1,l})_{l=t_1+n'+1}^{t_2+n'+n_2} \right) \in A_\epsilon^{(t_2-t_1+n_2)}[p^{(2)}] \right\} \tag{M.13}$$

and

$$\mathcal{F} = \left\{ \left((\boldsymbol{s}_{1,l})_{l=t_2-t_1+n_2}^{n_1-1}, (\boldsymbol{y}_{1,l})_{l=t_2+n'+n_2+1}^{t_1+n'+n_1} \right) \in A_\epsilon^{(t_1-t_2+n_1-n_2)}[p^{(1)}] \right\}. \tag{M.14}$$

Using the standard properties of jointly typical sequences [28], we have

$$\mathbb{P}(\mathcal{E}_{t_1,t_2}) \leq 2^{-(t_2-t_1+n_2)(\kappa_1'-3\epsilon)} \tag{M.15}$$

and

$$\mathbb{P}(\mathcal{F}_{t_1,t_2}) \leq 2^{-(t_1-t_2+n_1-n_2)(\kappa_1-3\epsilon)}. \tag{M.16}$$

By (M.11), (M.12), (M.15) and (M.16),

$$\begin{aligned} \mathbb{P}\left(\mathcal{E} \bigcap \mathcal{F} \bigcap \mathcal{U}_n \bigcap \mathcal{V}_n\right) &\leq 2^{-n_1(\kappa_1-3\epsilon)} 2^{-n_2(\kappa_1'-\kappa_1)} \\ &\times \sum_{\substack{(t_1,t_2,t_3)\in\widetilde{\mathcal{U}}_n \\ (t_1,t_2)\in\widetilde{\mathcal{V}}_n}} 2^{-(t_1-t_2)(\kappa_1-\kappa_1')} \mathbb{P}\big((\boldsymbol{\tau}_1^{(j)}, \boldsymbol{\tau}_2^{(j^-)}, \boldsymbol{\tau}_2^{(\omega^-+1)}) = (t_1, t_2, t_3)\big). \end{aligned} \tag{M.17}$$

For any $(t_1, t_2) \in \widetilde{\mathcal{V}}_n$, we have $t_1 \geq n(j\mu_1 + \nu_1 - \delta)$ and $t_2 \leq n(\omega^-\mu_2 + \nu_2 + \delta)$. Using these bounds in (M.17), we get

$$\mathbb{P}\left(\mathcal{E} \bigcap \mathcal{F} \bigcap \mathcal{U}_n \bigcap \mathcal{V}_n\right) \leq 2^{-n_1(\kappa_1-3\epsilon)} 2^{-n_2(\kappa_1'-\kappa_1)} 2^{-n(j\mu_1-\omega^-\mu_2+\nu_1-\nu_2-2\delta)(\kappa_1-\kappa_1')}. \tag{M.18}$$

By (M.8) and (M.18),

$$\mathbb{P}(\text{error}_2, \mathcal{U}_n, \mathcal{V}_n) \leq 2^{-nf(n)}, \tag{M.19}$$

where

$$f(n) = (j\mu_1 - \omega^- \mu_2 + \nu_1 - \nu_2 - 2\delta)(\kappa_1 - \kappa_1') + \frac{n_1}{n}(\kappa_1 - 3\epsilon) + \frac{n_2}{n}(\kappa_1' - \kappa_1) - \frac{\lfloor n\eta_1 \rfloor}{n}. \tag{M.20}$$

We have $\lim_{n\to\infty} f(n) = (j\mu_1 - \omega^- \mu_2 + \nu_1 - \nu_2)(\kappa_1 - \kappa_1') + \theta_1\kappa_1 + \theta_2(\kappa_1' - \kappa_1) - \eta_1 - 2(\kappa_1 - \kappa_1')\delta - 3\theta_1\epsilon$. By (3.64), $(j\mu_1 - \omega^- \mu_2 + \nu_1 - \nu_2)(\kappa_1 - \kappa_1') + \theta_1\kappa_1 + \theta_2(\kappa_1' - \kappa_1) - \eta_1 > 0$. Therefore, $\lim_{n\to\infty} f(n) > 0$ for sufficiently small δ and ϵ and $\mathbb{P}(\text{error}_2, \mathcal{U}_n, \mathcal{V}_n)$ decays exponentially with n as desired.

Appendix N

We need the following lemma:

Lemma 5. *Let m, n be positive integers. The number of non-decreasing sequences of length n whose entries are among the numbers $1, 2, \cdots, m$ is $\binom{m+n-1}{n}$.*

Proof. Define the sets $\mathcal{A}$ and $\mathcal{B}$ by

$$\mathcal{A} := \left\{ (x_1, \cdots, x_{n+1}) : \begin{array}{c} x_1, \cdots, x_{n+1} \in \mathbb{Z} \\ x_1 \geq 1, x_2, \cdots, x_{n+1} \geq 0 \\ x_1 + \cdots + x_{n+1} = m \end{array} \right\} \tag{N.1}$$

and

$$\mathcal{B} := \Big\{ (y_1, \cdots, y_n) : y_1, \cdots, y_n \in \mathbb{Z}, 1 \leq y_1 \leq y_2 \leq \cdots \leq y_n \leq m \Big\}, \tag{N.2}$$

respectively. Define the map

$$\begin{gathered} f : \mathcal{A} \to \mathcal{B} \\ (x_1, \cdots, x_{n+1}) \mapsto (x_1, x_1 + x_2, x_1 + x_2 + x_3, \cdots, x_1 + \cdots + x_n). \end{gathered} \tag{N.3}$$

The codomain of the map f is as promised because for any $(x_1, \cdots, x_{n+1}) \in \mathcal{A}$, $1 \leq x_1 \leq x_1 + x_2 \leq x_1 + x_2 + x_3 \leq \cdots \leq x_1 + \cdots + x_n \leq m$. We make the following observations:

- The map f is one-to-one. In fact, let $(x_1, \cdots, x_{n+1}), (x'_1, \cdots, x'_{n+1}) \in \mathcal{A}$ and assume

$f(x_1, \cdots, x_{n+1}) = f(x'_1, \cdots, x'_{n+1})$. Then

$$\begin{aligned} x_1 &= x'_1 \\ x_1 + x_2 &= x'_1 + x'_2 \\ &\vdots \\ x_1 + \cdots + x_n &= x'_1 + \cdots + x'_n \end{aligned} \quad . \tag{N.4}$$

It follows that $x_i = x'_i$ for any $1 \leq i \leq n$ and hence, $x_{n+1} = m - \sum_{i=1}^{n} x_i = m - \sum_{i=1}^{n} x'_i = x'_{n+1}$. We conclude that $(x_1, \cdots, x_{n+1}) = (x'_1, \cdots, x'_{n+1})$.

- The map f is onto. To see this, let $(y_1, \cdots, y_n) \in \mathcal{B}$. Define $x_1 := y_1$, $x_i := y_i - y_{i-1}$ for any $2 \leq i \leq n$ and $x_{n+1} = m - \sum_{i=1}^{n} x_i$. Then it is easy to see that $(x_1, \cdots, x_{n+1}) \in \mathcal{A}$ and $f(x_1, \cdots, x_{n+1}) = (y_1, \cdots, y_n)$.

It follows that f is a bijection between $\mathcal{A}$ and $\mathcal{B}$ and hence, $|\mathcal{A}| = |\mathcal{B}|$. But, $|\mathcal{A}|$ is exactly the number of solutions for the tuples $(z_1, \cdots, z_{n+1})$ where $z_1 := x_1, z_2 := x_2 + 1 \cdots, z_{n+1} := x_{n+1} + 1$ are positive integers and $z_1 + \cdots + z_{n+1} = m + n$. This number is known to be $\binom{m+n-1}{n}$. □

Note that $\{(j, u_j, v_j) : 1 \leq j \leq N_2\}$ is a state if and only if

$$1 \leq i_1 \leq j_1 \leq i_2 \leq j_2 \leq \cdots \leq i_{N_2-1} \leq j_{N_2-1} \leq i_{N_2} \leq j_{N_2} \leq 2N_1 + 1, \tag{N.5}$$

i.e., $(i_1, j_1, i_2, j_2, \cdots, i_{N_2}, j_{N_2})$ is a non-decreasing sequence of integers whose entries are among the numbers $1, 2, \cdots, 2N_1 + 1$. By Lemma 5, the number of such sequences is $\binom{2N_1+2N_2}{2N_2}$.

Appendix O

Proof of Proposition 15

As shown earlier in (3.138), we only need to assume $j^* = 1$. Then $\mathscr{A}_1 = \{R > \lambda : \alpha < \mu < \alpha + \theta\} = \big(\lambda, \big(1 + \frac{\alpha}{\theta}\big)\lambda\big)$, $\mathscr{C}_1 = \{R > \lambda : \mu > \alpha + \theta\} = \big(\big(1 + \frac{\alpha}{\theta}\big)\lambda, \infty\big)$ and $\mathscr{B}_1 = \mathscr{D}_1 = \emptyset$. For notational simplicity, we show κ_γ and κ'_γ by κ and κ', respectively. It is beneficial to our presentation to write $\mathcal{P}(x, y; \gamma)$ in (3.116) as

$$\mathcal{P}(x, y; \gamma) = \widetilde{\mathcal{P}}(x, y; \gamma) \bigcap \mathcal{U}(\gamma), \tag{O.1}$$

where

$$\widetilde{\mathcal{P}}(x, y; \gamma) := \Big\{R_c > \lambda : \big(1 - \frac{x}{\lambda}(\kappa - \kappa')\big)R_c < \kappa' - (\kappa - \kappa')\frac{y}{\theta}\Big\} \tag{O.2}$$

and

$$\mathcal{U}(\gamma) := \Big\{R_c > \lambda : \gamma \leq \big(\frac{1}{N} + \frac{R_c}{\lambda}\big)P\Big\}. \tag{O.3}$$

We have

$$\mathcal{R}_1 = (\mathcal{R}_{\text{sym}} \bigcap \mathscr{A}_1) \bigcup (\mathcal{R}_{\text{sym}} \bigcap \mathscr{C}_1), \tag{O.4}$$

where

$$\mathcal{R}_{\text{sym}} \bigcap \mathscr{A}_1 = \bigcup_{\gamma \geq 0} \Big(\mathcal{U}(\gamma) \bigcap \widetilde{\mathcal{P}}(1, \theta; \gamma) \bigcap \widetilde{\mathcal{P}}(0, -\alpha; \gamma) \bigcap \mathscr{A}_1\Big) \tag{O.5}$$

and

$$\mathcal{R}_{\mathrm{sym}} \bigcap \mathscr{C}_1 \;=\; \bigcup_{\gamma \geq 0} \Big(\mathcal{U}(\gamma) \bigcap \widetilde{\mathcal{P}}(0,-\alpha;\gamma) \bigcap \mathscr{C}_1 \Big), \tag{O.6}$$

by (3.132) and (3.134), respectively. We have

$$\widetilde{\mathcal{P}}(0,-\alpha;\gamma) \bigcap \mathscr{C}_1 = \Big((1+\frac{\alpha}{\theta})\lambda, \kappa' + \frac{\alpha}{\theta}(\kappa-\kappa') \Big). \tag{O.7}$$

To compute the right side of (O.5), it is enough to note that

$$\begin{aligned} \widetilde{\mathcal{P}}(1,\theta;\gamma) \;&=\; \Big\{ R > \lambda : \big(1 - \frac{1}{\lambda}(\kappa-\kappa')\big) R < 2\kappa' - \kappa \Big\} \\ &=\; \begin{cases} \big(\lambda, \frac{2\kappa'-\kappa}{1-\frac{1}{\lambda}(\kappa-\kappa')}\big) & \kappa-\kappa' < \lambda < \kappa' \\ \Big(\max\Big\{\lambda, \frac{2\kappa'-\kappa}{1-\frac{1}{\lambda}(\kappa-\kappa')}\Big\}, \infty \Big) & \kappa \geq 2\kappa',\ \ \kappa-\kappa' > \lambda \\ (\lambda,\infty) & \kappa \leq 2\kappa',\ \ \kappa-\kappa' > \lambda \\ \emptyset & \text{otherwise} \end{cases} \end{aligned} \tag{O.8}$$

and

$$\widetilde{\mathcal{P}}(0,-\alpha;\gamma) \bigcap \mathscr{A}_1 = \Big(\lambda, \min\Big\{ (1+\frac{\alpha}{\theta})\lambda, \kappa' + \frac{\alpha}{\theta}(\kappa-\kappa') \Big\} \Big). \tag{O.9}$$

Having (O.8) and (O.9), we can find $\mathcal{R}_{\mathrm{sym}} \bigcap \mathscr{A}_1$ in (O.5) which together with $\mathcal{R}_{\mathrm{sym}} \bigcap \mathscr{C}_1$ in (O.6) and (O.7) complete the description of $\mathcal{R}_1$ in (O.4). To simplify our computations, let us consider two cases:

1. Assume
$$(1+\frac{\alpha}{\theta})\lambda < \kappa' + \frac{\alpha}{\theta}(\kappa-\kappa'). \tag{O.10}$$
Using this in (O.9), we see that $\widetilde{\mathcal{P}}(0,-\alpha;\gamma) \bigcap \mathscr{A}_1 = \big(\lambda, \big(1+\frac{\alpha}{\theta}\big)\lambda\big)$ and we get
$$\begin{aligned} &\widetilde{\mathcal{P}}(1,\theta;\gamma) \bigcap \widetilde{\mathcal{P}}(0,-\alpha;\gamma) \bigcap \mathscr{A}_1 \\ &= \begin{cases} \big(\lambda, \min\big\{(1+\frac{\alpha}{\theta})\lambda, \frac{2\kappa'-\kappa}{1-\frac{1}{\lambda}(\kappa-\kappa')}\big)\big\} & \kappa-\kappa' < \lambda < \kappa' \\ \Big(\max\Big\{\lambda, \frac{2\kappa'-\kappa}{1-\frac{1}{\lambda}(\kappa-\kappa')}\Big\}, (1+\frac{\alpha}{\theta})\lambda \Big) & \kappa \geq 2\kappa',\ \ \kappa-\kappa' > \lambda \\ (\lambda, (1+\frac{\alpha}{\theta})\lambda) & \kappa \leq 2\kappa',\ \ \kappa-\kappa' > \lambda \\ \emptyset & \text{otherwise} \end{cases} . \end{aligned} \tag{O.11}$$

Simple algebra shows that

$$\begin{aligned}&\big(\lambda-(\kappa-\kappa')\big)\Big(\kappa'+\frac{\alpha}{\theta}(\kappa-\kappa')-\frac{2\kappa'-\kappa}{1-\frac{1}{\lambda}(\kappa-\kappa')}\Big)\\&=(\kappa-\kappa')\big((1+\frac{\alpha}{\theta})\lambda-(\kappa'+\frac{\alpha}{\theta}(\kappa-\kappa'))\big).\end{aligned}\tag{O.12}$$

By (O.10) the right side of (O.12) is negative. Therefore,

$$\kappa-\kappa'<\lambda\implies\kappa'+\frac{\alpha}{\theta}(\kappa-\kappa')<\frac{2\kappa'-\kappa}{1-\frac{1}{\lambda}(\kappa-\kappa')}\implies\big(1+\frac{\alpha}{\theta}\big)\lambda<\frac{2\kappa'-\kappa}{1-\frac{1}{\lambda}(\kappa-\kappa')},\tag{O.13}$$

where in the last step we use (O.10). Therefore, the interval in the first row of (O.11) becomes $(\lambda,(1+\frac{\alpha}{\theta})\lambda)$, i.e.,

$$\begin{aligned}&\widetilde{\mathcal{P}}(1,\theta;\gamma)\bigcap\widetilde{\mathcal{P}}(0,-\alpha;\gamma)\bigcap\mathscr{A}_1\\&=\begin{cases}\big(\lambda,(1+\frac{\alpha}{\theta})\lambda\big) & \kappa-\kappa'<\lambda<\kappa'\\ \Big(\max\Big\{\lambda,\frac{2\kappa'-\kappa}{1-\frac{1}{\lambda}(\kappa-\kappa')}\Big\},(1+\frac{\alpha}{\theta})\lambda\Big) & \kappa\geq 2\kappa',\ \kappa-\kappa'>\lambda\\ (\lambda,(1+\frac{\alpha}{\theta})\lambda) & \kappa\leq 2\kappa',\ \kappa-\kappa'>\lambda\\ \emptyset & \text{otherwise}\end{cases}.\end{aligned}\tag{O.14}$$

It is notable that in the second line in (O.14) it is always true that $\frac{2\kappa'-\kappa}{1-\frac{1}{\lambda}(\kappa-\kappa')}<(1+\frac{\alpha}{\theta})\lambda$ provided $\kappa-\kappa'>\lambda$ and hence, the interval $\big(\max\big\{\lambda,\frac{2\kappa'-\kappa}{1-\frac{1}{\lambda}(\kappa-\kappa')}\big\},(1+\frac{\alpha}{\theta})\lambda\big)$ is nonempty.[1]

2. Assume

$$\big(1+\frac{\alpha}{\theta}\big)\lambda\geq\kappa'+\frac{\alpha}{\theta}(\kappa-\kappa').\tag{O.15}$$

Then

$$\widetilde{\mathcal{P}}(0,-\alpha;\gamma)\bigcap\mathscr{C}_1=\emptyset.\tag{O.16}$$

By (O.9) and (O.15), $\widetilde{\mathcal{P}}(0,-\alpha;\gamma)\bigcap\mathscr{A}_1=\big(\lambda,\kappa'+\frac{\alpha}{\theta}(\kappa-\kappa')\big)$. Using this together

[1]Multiplying both sides of $\frac{2\kappa'-\kappa}{1-\frac{1}{\lambda}(\kappa-\kappa')}<(1+\frac{\alpha}{\theta})\lambda$ by the negative number $1-\frac{1}{\lambda}(\kappa-\kappa')$ yields $\big(1+\frac{\alpha}{\theta}\big)\lambda<\kappa'+\frac{\alpha}{\theta}(\kappa-\kappa')$ which is our assumption in (O.10).

with (O.5) and (O.8),

$$\widetilde{\mathcal{P}}(1,\theta;\gamma)\bigcap\widetilde{\mathcal{P}}(0,-\alpha;\gamma)\bigcap\mathscr{A}_1 = \begin{cases} \left(\lambda, \min\left\{\frac{2\kappa'-\kappa}{1-\frac{1}{\lambda}(\kappa-\kappa')}, \kappa'+\frac{\alpha}{\theta}(\kappa-\kappa')\right\}\right) & \kappa-\kappa'<\lambda<\kappa' \\ \left(\max\left\{\lambda, \frac{2\kappa'-\kappa}{1-\frac{1}{\lambda}(\kappa-\kappa')}\right\}, \kappa'+\frac{\alpha}{\theta}(\kappa-\kappa')\right) & \kappa\geq 2\kappa',\ \kappa-\kappa'>\lambda, \\ (\lambda, \kappa'+\frac{\alpha}{\theta}(\kappa-\kappa')) & \kappa\leq 2\kappa',\ \kappa-\kappa'>\lambda \\ \emptyset & \text{otherwise} \end{cases}. \quad (O.17)$$

The right side of (O.12) is nonnegative due to (O.15). Therefore,

$$\kappa-\kappa'<\lambda \implies \frac{2\kappa'-\kappa}{1-\frac{1}{\lambda}(\kappa-\kappa')} \leq \kappa'+\frac{\alpha}{\theta}(\kappa-\kappa') \quad (O.18)$$

and

$$\kappa-\kappa'>\lambda \implies \frac{2\kappa'-\kappa}{1-\frac{1}{\lambda}(\kappa-\kappa')} \geq \kappa'+\frac{\alpha}{\theta}(\kappa-\kappa'). \quad (O.19)$$

Let us make the following observations:

- By (O.18), the interval in the first row of the definition of $\mathcal{R}_{\text{sym}}\bigcap\mathscr{A}_1$ in (O.17) becomes $\left(\lambda, \frac{2\kappa'-\kappa}{1-\frac{1}{\lambda}(\kappa-\kappa')}\right)$.
- By (O.19), the interval in the second row of (O.17) is empty.
- Combining $\kappa-\kappa'>\lambda$ with (O.15), we get $\lambda>\kappa'$ and hence, $\kappa>2\kappa'$. This shows that the constraints in the third row of (O.17) are not compatible with (O.15).

Based on these observations, we can simplify (O.17) as

$$\widetilde{\mathcal{P}}(1,\theta;\gamma)\bigcap\widetilde{\mathcal{P}}(0,-\alpha;\gamma)\bigcap\mathscr{A}_1 = \begin{cases} \left(\lambda, \frac{2\kappa'-\kappa}{1-\frac{1}{\lambda}(\kappa-\kappa')}\right) & \kappa-\kappa'<\lambda<\kappa' \\ \emptyset & \text{otherwise} \end{cases}. \quad (O.20)$$

Define γ_0 as the value of γ that solves $\kappa'=\kappa-\kappa'$. We consider two cases:

1. Let $\lambda<\kappa'_{\gamma_0}$. This situation is shown in panel (a) of Fig. O.1. The solutions for γ in $\kappa'_\gamma=\lambda$ and $\kappa_\gamma-\kappa'_\gamma=\lambda$ are shown by γ_1 and γ_3, respectively. It is only for values of γ in the interval (γ_1,γ_3) that $\kappa_\gamma-\kappa'_\gamma<\lambda<\kappa'_\gamma$. Moreover, the equation

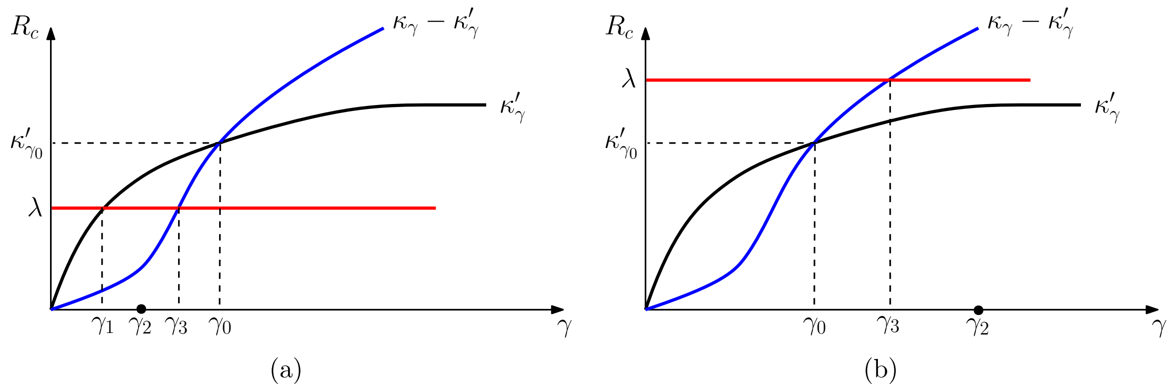


Figure O.1: Plots of κ'_γ (black curve) and $\kappa_\gamma - \kappa'_\gamma$ (blue curve) as function of γ. The constant level λ is shown in red. The value of γ for which $\kappa'_\gamma = \kappa_\gamma - \kappa'_\gamma$ is denoted by γ_0. Panel (a) shows a scenario where $\lambda < \kappa'_{\gamma_0}$. The solutions for γ in $\kappa'_\gamma = \lambda$ and $\kappa_\gamma - \kappa'_\gamma = \lambda$ are shown by γ_1 and γ_3, respectively. It is only for values of γ in the interval (γ_1, γ_3) that $\kappa_\gamma - \kappa'_\gamma < \lambda < \kappa'_\gamma$. Moreover, the equation $\kappa'_\gamma + \frac{\alpha}{\theta}\kappa_\gamma - \kappa'_\gamma = (1 + \frac{\alpha}{\theta})$ is solved for $\gamma = \gamma_2$ where $\gamma_1 < \gamma_2 < \gamma_3$. Panel (b) shows a scenario where $\lambda > \kappa'_{\gamma_0}$. The conditions $\kappa_\gamma - \kappa'_\gamma < \lambda < \kappa'_\gamma$ no longer hold.

$\kappa'_\gamma + \frac{\alpha}{\theta}(\kappa_\gamma - \kappa'_\gamma) = (1 + \frac{\alpha}{\theta})\lambda$ is solved for $\gamma = \gamma_2$ where $\gamma_1 < \gamma_2 < \gamma_3$.[2] We have

$$\widetilde{\mathcal{P}}(0, -\alpha; \gamma) \bigcap \mathscr{C}_1 = \begin{cases} \emptyset & \gamma \leq \gamma_2 \\ \big((1 + \frac{\alpha}{\theta})\lambda, \kappa' + \frac{\alpha}{\theta}(\kappa - \kappa')\big) & \gamma > \gamma_2 \end{cases} \tag{O.21}$$

and

$$\widetilde{\mathcal{P}}(1, \theta; \gamma) \bigcap \widetilde{\mathcal{P}}(0, -\alpha; \gamma) \bigcap \mathscr{A}_1 = \begin{cases} \emptyset & \gamma \leq \gamma_1 \\ \big(\lambda, \frac{2\kappa' - \kappa}{1 - \frac{1}{\lambda}(\kappa - \kappa')}\big) & \gamma_1 < \gamma < \gamma_2 \\ \big(\lambda, (1 + \frac{\alpha}{\theta})\lambda\big) & \gamma > \gamma_2 \end{cases} \tag{O.22}$$

2. Let $\lambda \geq \kappa'_{\gamma_0}$. This situation is shown in panel (b) of Fig. O.1. The conditions $\kappa - \kappa' < \lambda < \kappa'$ no longer hold. Moreover, $\gamma_3 < \gamma_2 < \gamma_1$ where we let $\gamma_1 = \infty$ if it

[2] if $\gamma < \gamma_1$, then κ' and $\kappa - \kappa'$ are both smaller than λ and if $\gamma > \gamma_3$, then κ' and $\kappa - \kappa'$ are both larger than λ. In either case, $\kappa'_\gamma + \frac{\alpha}{\theta}(\kappa_\gamma - \kappa'_\gamma) \neq (1 + \frac{\alpha}{\theta})\lambda$.

does not exist. The set $\widetilde{\mathcal{P}}(0,-\alpha;\gamma)\bigcap\mathscr{C}_1$ is given by (O.21) and

$$\widetilde{\mathcal{P}}(1,\theta;\gamma)\bigcap\widetilde{\mathcal{P}}(0,-\alpha;\gamma)\bigcap\mathscr{A}_1=\begin{cases}\emptyset & \gamma\le\gamma_2\\ \left(\frac{2\kappa'-\kappa}{1-\frac{1}{\lambda}(\kappa-\kappa')},(1+\frac{\alpha}{\theta})\lambda\right) & \gamma_2<\gamma\le\gamma_1\\ \left(\lambda,(1+\frac{\alpha}{\theta})\lambda\right) & \gamma>\gamma_1\end{cases}. \tag{O.23}$$

By (O.4), (O.5) and (O.6),

$$\mathcal{R}_{\text{sym}}=\bigcup_{\gamma\ge 0}\left(\mathcal{U}(\gamma)\bigcap\left(\left(\widetilde{\mathcal{P}}(1,\theta;\gamma)\bigcap\widetilde{\mathcal{P}}(0,-\alpha;\gamma)\bigcap\mathscr{A}_1\right)\bigcup\left(\widetilde{\mathcal{P}}(0,-\alpha;\gamma)\bigcap\mathscr{C}_1\right)\right)\right). \tag{O.24}$$

Using the expressions in (O.21), (O.22) and (O.23),

$$\left(\widetilde{\mathcal{P}}(1,\theta;\gamma)\bigcap\widetilde{\mathcal{P}}(0,-\alpha;\gamma)\bigcap\mathscr{A}_1\right)\bigcup\left(\widetilde{\mathcal{P}}(0,-\alpha;\gamma)\bigcap\mathscr{C}_1\right)=\left(f(\gamma),g(\gamma)\right), \tag{O.25}$$

where f and g are defined in (3.139) and (3.140) under the assumption that $\lambda<\kappa'_{\gamma_0}$ and $\lambda\ge\kappa'_{\gamma_0}$, respectively.