Lattice Codes for Secure Communication and Secret Key Generation

A Thesis

Submitted for the Degree of

Doctor of Philosophy

in the **Faculty of Engineering**

by

Shashank Vatedka

under the Guidance of

Navin Kashyap

Electrical Communication Engineering
Indian Institute of Science
Bangalore – 560 012, INDIA

MAY 2016

TO

MY PARENTS

If numbers aren't beautiful, I don't know what is.

— Paul Erdős

Acknowledgments

I consider myself extremely fortunate to have been advised by Prof. Navin Kashyap. He was always available for discussions and I would inevitably come out these meetings learning more than I had known before. He has always been patient and accommodating with me, and I have gained many experiences in the course of my PhD. I cannot thank him enough for that. I have tried to imbibe myself with some of his "goodness" properties, but it seems to require $N \to \infty$.

I am indebted to IISc for providing a wonderful environment for pursuing research and all the on-campus facilities. I am also thankful to Tata Consultancy Services for a generous fellowship that supported me during my PhD. I am also extremely thankful to all the Professors who taught me at IISc.

I must thank Prof. Andrew Thangaraj for various technical discussions that formed the "seeds" of this thesis. The work on secure bidirectional relaying was done in collaboration with him. Thanks are also due to Prof. Manjunath Krishnapur for many discussions and pointers. I would like to thank Prof. Gilles Zemor for insights on LDA lattices, and Manuj Mukherjee for discussions on the secret key generation problem. I am deeply grateful to Prof. Pascal Vontobel for hosting me for a summer at the Institute of Network Coding, and to Prof. Sidharth Jaggi for ideas that led to the work on concatenated lattices.

I am also thankful to Mr. Srinivasa Murthy, Suvarna and Veena madam for help with many administrative matters.

My experience at IISc was special because of the company of many wonderful friends. Thanks to Winston, Kamal, Avinash, Ananya, the rest of the *A-mess gang*, and all my other friends at IISc for their support and friendship. Also thanks to Birenjith, my constant travel companion outside (and within) IISc. Special thanks to my lab-mates Manuj and Shivkumar, with whom I shared several *moments*. With *comrade* Manuj, I had many useful discussions, both technical and non technical. Most memorable were the discussions with Shiv on all topics under the sun, which I must add, were very *constructive*.

My *Journey to the east* was fruitful because of the company and help of many wonderful people at Hong Kong. Thanks to my dear friends Eric Chan and Mao Ying for taking care of this lost soul and for the many bowls of noodles.

I am deeply indebted to my high school teacher, Manjunath *sir*, who first made me realize that mathematics is beautiful. I am also grateful to Dr. Vijaya Krishna, who inspired me and encouraged me to pursue a PhD.

None of all this would have been possible without the love and support of my parents, to whom I have the deepest gratitude.

Statement of Originality

I hereby declare that this thesis is my own work and has not been submitted in any form for another degree or diploma at any university or other institute of higher education.

I certify that to the best of my knowledge, the intellectual content of this thesis is the product of my own work and that all the assistance received in preparing this thesis and sources have been acknowledged.

Publications based on this Thesis

Journal

J1 S. Vatedka, N. Kashyap, and A. Thangaraj, "Secure compute-and-forward in a bidirectional relay," *IEEE Transactions on Information Theory,* vol. 51, no. 5, pp. 2531–2556, May 2015.

J2 S. Vatedka and N. Kashyap, "Some goodness properties of LDA lattices," submitted, *Problems of Information Transmission*, Dec. 2015.

Conference

C1 N. Kashyap, V. Shashank and A. Thangaraj, "Secure computation in a bidirectional relay," in *Proc. 2012 IEEE International Symposium on Information Theory*, Cambridge, MA, 2012, pp. 1162-1166.

C2 S. Vatedka and N. Kashyap, "Lattice coding for strongly secure compute-and-forward in a bidirectional relay, in *Proc. 2013 IEEE International Symposium on Information Theory*, Istanbul, Turkey, 2013, pp. 2775–2779.

C3 S. Vatedka and N. Kashyap, "Nested lattice codes for secure bidirectional relaying with asymmetric channel gains," in *Proc. 2015 IEEE Information Theory Workshop (Invited)*, Jerusalem, Israel, 2015, pp. 1–5.

C4 S. Vatedka and N. Kashyap, "Some goodness properties of LDA lattices," in *Proc. 2015 IEEE Information Theory Workshop*, Jerusalem, Israel, 2015, pp. 1–5.

C5 S. Vatedka and N. Kashyap, “A capacity-achieving coding scheme for the AWGN channel with polynomial encoding and decoding complexity,” in *Proc. 2016 National Conference on Communications*, Guwahati, India, 2016, pp. 36–41.

C6 S. Vatedka and N. Kashyap, “A Lattice Coding Scheme for Secret Key Generation from Gaussian Markov Tree Sources”, accepted, *2016 IEEE International Symposium on Information Theory*, Barcelona, Spain, 2016.

Preface

The work in Chapter 3 was done in collaboration with my advisor Prof. Navin Kashyap, and Prof. Andrew Thangaraj (IIT-Madras, Chennai, India). The results in the remaining chapters were obtained in collaboration with my advisor Prof. Navin Kashyap.

The work in Chapter 3 was published in C1, and subsequently in the journal version J1. The work in Chapter 4 was published in C2 and the journal version J1. Results in Chapter 5 were published in C3, and an online version with added proofs can be found in [97]. The work in Chapter 6 was published in C4, and the journal version J2 is currently under review. Chapter 7 was published in C5, and updated version with new results can be found in [101]. The work in Chapter 8 was published in C6.

Abstract

In this work, we study two problems in information-theoretic security. Firstly, we study a wireless network where two nodes want to securely exchange messages via an honest-but-curious bidirectional relay. There is no direct link between the user nodes, and all communication must take place through the relay. The relay behaves like a passive eavesdropper, but otherwise follows the protocol it is assigned. Our objective is to design a scheme where the user nodes can reliably exchange messages such that the relay gets no information about the individual messages. We first describe a perfectly secure scheme using nested lattices, and show that our scheme achieves secrecy regardless of the distribution of the additive noise, and even if this distribution is unknown to the user nodes. Our scheme is explicit, in the sense that for any pair of nested lattices, we give the distribution used for randomization at the encoders to guarantee security. We then give a strongly secure lattice coding scheme, and we characterize the performance of both these schemes in the presence of Gaussian noise. We then extend our perfectly-secure and strongly-secure schemes to obtain a protocol that guarantees end-to-end secrecy in a multihop line network. We also briefly study the robustness of our bidirectional relaying schemes to channel imperfections.

In the second problem, we consider the scenario where multiple terminals have access to private correlated Gaussian sources and a public noiseless communication channel. The objective is to generate a group secret key using their sources and public communication in a way that an eavesdropper having access to the public communication can obtain no information about the key. We give a nested lattice-based protocol for generating strongly secure secret keys from independent and identically distributed copies of the correlated *random variables. Under certain assumptions on the joint distribution of the sources, we derive achievable secret key rates.*

The tools used in designing protocols for both these problems are nested lattice codes, which have been widely used in several problems of communication and security. In this thesis, we also study lattice constructions that permit polynomial-time encoding and decoding. In this regard, we first look at a class of lattices obtained from low-density parity-check (LDPC) codes, called Low-density Construction-A (LDA) lattices. We show that high-dimensional LDA lattices have several "goodness" properties that are desirable in many problems of communication and security. We also present a new class of low-complexity lattice coding schemes that achieve the capacity of the AWGN channel. Codes in this class are obtained by concatenating an inner Construction-A lattice code with an outer Reed-Solomon code or an expander code. We show that this class of codes can achieve the capacity of the AWGN channel with polynomial encoding and decoding complexities. Furthermore, the probability of error decays exponentially in the blocklength for a fixed transmission rate R that is strictly less than the capacity. To the best of our knowledge, this is the first capacity-achieving coding scheme for the AWGN channel which has an exponentially decaying probability of error and polynomial encoding/decoding complexities.

Contents

List of Tables

List of Figures

Keywords

Information theory, nested lattice codes, physical-layer security, secret key generation, codes for secrecy, bidirectional relay, compute-and-forward, LDPC codes, expander graphs, concatenated codes.

Notation

Sets

$\mathbb{R}$	The set of real numbers
$\mathbb{R}^+$	The set of nonnegative real numbers
$\mathbb{Z}$	The set of integers
$\mathbb{Z}^+$	The set of nonnegative integers
$\mathbb{F}_p$	Finite field with p elements
$\mathbb{Z}_p$	The set of integers modulo p
$\mathrm{vol}(\mathcal{S})$	Volume of the set $\mathcal{S}$
$\bigtimes_{i=1}^{m} A_i$	Cartesian product of the sets $A_1, \ldots, A_m$

Vectors and matrices

$\mathcal{S}^n$	Set of column vectors of length n with entries from $\mathcal{S}$
$\mathbf{x}, \mathbf{y}, \mathbf{u}, \mathbf{v}, \mathbf{w}, \ldots$	Vectors
$\mathsf{A}, \mathsf{B}, \mathsf{G}, \ldots$	Matrices
A^T	Transpose of A
I_n	$n \times n$ identity matrix
$\lVert\mathbf{x}\rVert$	ℓ^2-norm of $\mathbf{x}$
$\mathrm{Supp}(\mathbf{x})$	Support of the vector $\mathbf{x}$, i.e., the set of indices corresponding to nonzero entries of $\mathbf{x}$

Random Variables and Events

pmf	Probability mass function
pdf	Probability density function
iid	Independent and identically distributed
$U, V, X, Y, Z, \ldots$	Random variables
$\mathbf{U}, \mathbf{V}, \mathbf{W}, \mathbf{Z}, \ldots$	Random vectors
$\Pr[A]$	Probability of event A
$\mathbb{E}[Z]$	Expectation of the random variable Z
$H(X)$	Entropy of the random variable X
$H(X\|Y)$	Conditional entropy of X given Y
$I(X;Y)$	Mutual information between X and Y
$X \perp\!\!\!\perp Y$	X and Y are independent random variables
$X \sim \mathcal{N}(a, \sigma^2)$	X is a Gaussian random variable with mean a and variance σ^2

Sequences

$f(n) = O(g(n))$	$\exists c > 0$ such that $f(n) < cg(n)$ for all sufficiently large n
$f(n) = \Omega(g(n))$	$\exists c > 0$ such that $g(n) < cf(n)$ for all sufficiently large n
$f(n) = o(g(n))$	$f(n)/g(n) \to 0$ as $n \to \infty$
$f(n) = o_n(1)$	$f(n) \to 0$ as $n \to \infty$

Graphs

$\mathcal{G} = (\mathcal{V}, \mathcal{E})$	Graph $\mathcal{G}$ with vertex set $\mathcal{V}$ and edge set $\mathcal{E}$
$\mathcal{G} = ((\mathcal{L}, \mathcal{R}), \mathcal{E})$	Bipartite graph $\mathcal{G}$ with vertex set $\mathcal{L} \cup \mathcal{R}$ and edge set $\mathcal{E}$. Here, $\mathcal{L}$ denotes the set of left vertices, and $\mathcal{R}$ denotes the set of right vertices.
$\mathtt{u}, \mathtt{v}, \ldots$	Vertices
$N(\mathtt{u})$	Neighbourhood of $\mathtt{u}$, i.e., the set of all vertices $\mathtt{v}$ such that $(\mathtt{u}, \mathtt{v})$ is an edge
$\mathcal{A}, \mathcal{B}, \ldots$	Subsets of $\mathcal{V}$
$N(\mathcal{A})$	Neighbourhood of $\mathcal{A}$, i.e., $\cup_{\mathtt{u} \in \mathcal{A}} N(\mathtt{u})$

Lattices

Λ, Λ_0	Lattices
$\Lambda^{(n)}, \Lambda_0^{(n)}$	Lattices in $\mathbb{R}^n$.
$Q_\Lambda(\mathbf{x})$	Lattice point (in Λ) closest to $\mathbf{x}$
$[\mathbf{x}] \bmod \Lambda$	$\mathbf{x} - Q_\Lambda(\mathbf{x})$
$\mathcal{V}(\Lambda)$	Fundamental Voronoi region of Λ
$r_{\text{cov}}(\Lambda)$	Covering radius of Λ
$r_{\text{pack}}(\Lambda)$	Packing radius of Λ
$r_{\text{eff}}(\Lambda)$	Effective radius of Λ
$\text{vol}\Lambda$ or $\det \Lambda$	Volume of the fundamental Voronoi region of Λ

Chapter 1

Introduction

This thesis mainly studies lattice codes for secure and reliable communication, and secret key generation. In recent years, we have seen several advancements in wireless communications. The fourth generation (4G) of wireless systems has already been commercially deployed, and research is underway to develop the next generation (5G) of wireless technologies [12, 105]. Although work in these areas is still in a preliminary stage, it is clear that the number of wireless devices will increase in the next few years, and securing communications in these systems will be a challenging task.

With the deployment of dense wireless networks, wireless communication would more closely resemble the internet, with our messages passing through various relay nodes before arriving at their destination. Mitigating interference would be a challenging problem, as would be the problem of keeping our messages from the prying eyes of potentially adversarial relays. Active research in mitigating interference has spawned an entire area of physical-layer network coding [55] to provide high-rate communication in wireless networks.

The bidirectional/two-way relay [75] is a three-node wireless network consisting of two nodes that want to exchange messages via an intermediate relay. The bidirectional relay can be viewed as a basic building block for larger wireless networks, and poses several challenges in itself. For a good part of this thesis, we will study the problem of how two parties can securely exchange messages in a bidirectional relay. To this end, we will develop coding schemes, and as we progress, we will attempt to improve these schemes by developing computationally efficient codes. Towards the end, we explore another problem where multiple parties having access to correlated data wish to generate secret keys using insecure public communication.

1.1 Information-Theoretic Security

A rigorous mathematical framework for secrecy systems was put forth by Shannon in [85] that used ideas from his work on information theory [84]. His work gave rise to what is now called information-theoretic security. Traditional cryptography assumed that the adversary, from whom the message is to be kept secure, has limited computational resources. Even today, most of the commonly used cryptosystems operate under the assumption that there are no known polynomial-time algorithms to solve certain computational problems. However, with the development of new and powerful computers, and advancements in quantum computing, many of these cryptographic protocols could soon be broken. In his paper, Shannon proposed an alternative approach by assuming that the adversary had unlimited time and computational resources. A key assumption was that the message to be secured can be modeled as the output of a random source. Both the message and the adversary's observations are random variables, and Shannon defined a cryptographic protocol to be perfectly secure if the message and the adversary's observations are statistically independent.

A fundamental characteristic of communication channels is that they are noisy, and this property can be used to deliver security to legitimate users. The first to use this property was Wyner [109], who studied the discrete memoryless wiretap channel. In this model, a source wants to communicate with a legitimate receiver through a discrete memoryless channel in presence of an eavesdropper. However, it is assumed that the channel between the source and the eavesdropper is degraded with respect to the source-receiver channel. Wyner proposed equivocation (the conditional entropy of the message given the eavesdropper's observation) as a measure of secrecy, and characterized the set of achievable rate-equivocation pairs in this setting. This was further extended to broadcast channels with confidential messages by Csiszár and Körner [17], and to the Gaussian wiretap channel by Leung-Yan-Cheong and Hellman [52]. These works also introduced the notion of *secrecy capacity*, or the maximum achievable transmission rates for reliable communication while satisfying the secrecy constraint. Since then, this field of physical-layer security has exploded, and several exciting developments have occurred in recent years — see, for example, [10, 11, 53].

1.1.1 When do We Say that a Protocol is Secure?

Traditionally, the following are the most commonly used measures used to judge the strength of a secure protocol:

- *Perfect secrecy* asks for statistical independence between the messages, M and the eavesdropper's observations, W, i.e., $M \perp\!\!\!\perp W$. However, this is usually too strong a requirement to be realized in practice. Hence, we typically relax the secrecy requirement.

- *Weak secrecy* asks for the mutual information rate, or the mutual information between M and W normalized by the number of channel uses (n), to be vanishingly small in n. This was the standard metric used in earlier works [109, 17, 52].

- *Strong secrecy* asks for the mutual information between M and W to be vanishingly small in n, i.e., $I(M;W) \to 0$ as $n \to \infty$. The weak secrecy constraint is in general not a good enough measure of secrecy, and we usually demand our protocols to satisfy strong secrecy. In many cases however, one can convert a weakly-secure scheme to a strongly-secure scheme with some additional preprocessing [65, 10, 11].

Like the problem of secure communication through noisy channels, a parallel line of work studied the problem of secret key generation from correlated sources. In this problem, multiple terminals have access to correlated random variables and must make use of insecure public communication to generate a secret key. The key so obtained must be independent or "almost independent" of the public communication. This problem was first studied by Maurer [64] and later by Ahlswede and Csiszár [1], and Csiszár and

Narayan [18, 19], among many others. A comprehensive survey of the developments in the areas of secure communication can be found in the book by Bloch and Barros [11].

Recent work has focused on designing coding schemes for information-theoretic security. Notable results include applications of low-density parity-check (LDPC) codes for wiretap channels [89, 88], and recent polynomial-time secrecy capacity-achieving protocols for wiretap channels using universal hash functions [8, 92] and polar codes [62, 38]. Another interesting development is the use of lattice codes [112] to design secure protocols for Gaussian channels. They have been used with success to achieve the secrecy capacity of the Gaussian wiretap channel [58], and design near-optimal schemes for bidirectional and co-operative relay channels [40, 41, 43]. They have also been used for secret key generation from correlated Gaussian sources [71, 57].

1.2 Lattice Codes

The theory of lattices goes back to at least the 18th century, when Gauss and Lagrange used lattices to prove results in number theory. A key figure in this area was Minkowski, who made remarkable contributions to lattices and the geometry of numbers [67]. Among the early motivations to study lattices was the problem of generating efficient sphere packings and coverings. In the sphere packing problem, we want to find an arrangement of non-intersecting spheres of a given radius that maximizes the average number of spheres packed per unit volume. On the other hand, the covering problem asks for an optimal covering of space by spheres of a given radius, that minimizes the average number of spheres per unit volume. Applications of lattices in packing and covering has been well-studied (see the classic book by Rogers [78] for more details), and it has been shown that lattices yield optimal coverings and packings in high dimensions [14, 78, 112]. Lattices have also been used to design optimal vector quantizers [36, 37]. More recently, they have found numerous applications in communications [112] and cryptography [66].

1.2.1 Communication over Wireless Channels

In the context of reliable communication over Gaussian channels, nested lattice codes have proved to be a very useful tool. It was shown by de Buda [20, 56], and later by Urbanke and Rimoldi [93] that lattice codes can achieve capacity of the additive white Gaussian noise (AWGN) channel with maximum likelihood (ML) decoding. For a long time, it was conjectured [59] that lattice codes cannot achieve the capacity of the AWGN channel with closest lattice-point (CLP) decoding (also called lattice decoding). This was finally resolved by Erez and Zamir [28], who showed that nested Construction-A lattice codes with CLP decoding can indeed achieve capacity. They have also been shown to achieve the capacity of the dirty-paper channel [30]. Inspired by these results, they have been applied to design protocols for reliable communication over wireless Gaussian networks. They have been used with much success for the interference channel [13, 103], the Gaussian bidirectional relay channel [107, 69], and generalized to the problem of physical layer network coding [3, 69] for multiuser Gaussian channels. Nested lattice coding has also been used for security in wiretap channels [6, 58] and bidirectional relay networks [41, 95]. For a more comprehensive treatment of lattices and their applications in communication problems, see the book by Zamir [112].

Constructing high dimensional lattices that have good structural properties is a problem that has been studied for a long time. Poltyrev [74] studied lattices in the context of coding for reliable transmission over the AWGN channel without power constraints, and showed that there exist lattices which are “good” for AWGN channel coding, i.e., achieve a vanishingly small probability of error with lattice decoding for all sufficiently small values of the noise variance.

Finding lattices with good structural properties is of particular importance in designing lattice codes that use nested lattice shaping for power-constrained Gaussian channels. A poorly designed shaping region leads to loss in transmission rates. Erez and Zamir [28] showed that using nested lattice codes, where the fine lattices are good for AWGN channel coding and the coarse lattices are good for mean squared error (MSE) quantization, we can achieve the capacity of the AWGN channel. Furthermore, the rates guaranteed by [107, 69] for bidirectional relaying and the compute-and-forward protocol are achievable using nested lattices that satisfy the aforementioned properties.

Instead of studying arbitrary lattices, it is easier to study lattices that have a special structure, i.e., lattices constructed by taking a linear code of length n over some prime field, and “lifting” it to $\mathbb{R}^n$. There are several such constructions of lattices from linear codes [14, Chapter 5]. One simple technique is Construction A, where the lattice is obtained by tessellating the codewords of the linear code (now viewed as points in $\mathbb{R}^n$) across the Euclidean space. It was shown by Erez et al. [29] that if we pick a random linear code and “lift” it using Construction A, then the resulting lattice is asymptotically good for covering, packing, MSE quantization, and AWGN channel coding with high probability[1].

1.2.2 The Quest for Good Lattice Codes with Low Decoding Complexity

While Construction-A lattices have several desirable properties, the problem with general Construction-A lattices is the complexity of closest lattice-point decoding. There is no known polynomial-time algorithm for decoding arbitrary lattice codes. In recent years, there have been constructions of lattices with low decoding complexity. Sadeghi et al. [81] proposed a construction based on LDPC codes using a technique called Construction D’ [14] to lift the LDPC code to a lattice. They analyzed the decoding complexity and showed using simulations that for error probabilities of roughly 10^{-6}, these lattices can achieve rates close to the Shannon limit. Sakzad et al. [82] proposed a construction using nested turbo codes and Construction D [14, Section 8.1, Chapter 8]. Simulations showed that these codes achieve rates close to capacity for an error probability of 10^{-6} [83].

Sommer et al. [86] introduced low density lattice codes, whose dual lattice had a sparse generator matrix. They proposed an iterative decoding scheme and analyzed the convergence of the decoder. Using simulations, they showed that for certain rates, the minimum volume-to-noise ratio required to achieve a probability of error of 10^{-6} with

[1] We will discuss Construction-A lattices and their properties in Chapter 2.

blocklength 10^5 is within 0.5 dB of the Poltyrev limit. However, none of these codes have been theoretically proven to possess any of the “goodness” properties that we mentioned earlier. Recently, two lattice constructions have been shown to achieve the capacity of the AWGN channel. Yan et al. [110, 111] introduced a construction using nested polar codes and Construction D. They gave a polynomial-time algorithm for encoding and decoding, and showed that this scheme is capacity achieving. Another construction using nonbinary LDPC codes and Construction A was proposed by di Pietro et al. [22]. These are called low-density Construction-A (LDA) lattices, and we will discuss them in detail in Chapter 6. Although LDA lattices permit low-complexity belief propagation decoders, these lattices have only been shown to achieve capacity using (exponential-time) CLP decoding.

1.3 Contributions and Outline of This Thesis

We have looked at four different problems (in different amounts of detail) in this thesis.

1.3.1 Secure Bidirectional Relaying

The first problem that we study is that of secure bidirectional relaying in the presence of an honest-but-curious eavesdropper. In this problem, two nodes, that we call the user nodes, want to exchange messages via a bidirectional relay. The bidirectional relay is a basic building block for larger, more complicated wireless networks, and studying this simple three-terminal network yields useful insights to design protocols for general networks. We assume that the relay follows the protocol that we assign to it, but we do not want the relay to obtain any information about the messages exchanged; Hence the name honest-but-curious. A similar problem was earlier studied by He and Yener [40, 41], and they gave weakly secure [40] and strongly secure protocols [41] for this problem.

Perfect Secrecy

In this thesis, we propose a perfectly-secure lattice coding scheme. This is the first perfectly secure coding scheme for secure communication in an honest-but-curious relay. Our scheme is fully explicit, in the sense that given *any* nested lattice pair in $\mathbb{R}^n$, we give the probability mass function (pmf) used for randomization at the encoders to guarantee perfect security. In Chapter 3, we formally introduce this problem, and describe our scheme. We first study the one-dimensional case before proceeding to the general setting. We show that perfect secrecy cannot be obtained if the codebook is restricted to have finite size (in fact, we show that the distribution of a random codeword cannot be light-tailed). In other words, there does not exist a perfectly secure scheme that satisfies a maximum power constraint. However, we can satisfy an average power constraint at the user nodes.

We use the Poisson summation formula from Fourier analysis to design our perfectly secure scheme. Specifically, the pmf used for randomization is obtained by sampling and marginalizing a density function whose characteristic function is compactly supported within the fundamental Voronoi region of the dual of the coarse lattice. We then restrict ourselves to those densities whose characteristic function is supported within a ball of radius equal to the packing radius of the dual of the coarse lattice. For this class of coding schemes, an earlier result of Ehm et al. [27] enables us to find the minimum average transmit power for a given pair of nested lattices.

We also study achievable rates under an average power constraint, and show that for large values of the signal-to-noise ratio, our scheme can achieve within $\log_2 2e$ bits of the achievable rate without secrecy constraints. To derive the rates achievable using our perfectly-secure scheme, we found the need to study the duals of Construction-A lattices. We cover the basics of lattice codes, and also present some of our novel findings on Construction-A lattices in Chapter 2.

Strong Secrecy

In Chapter 4, we relax the security requirement to strong secrecy, and give a scheme based on nested lattices and sampled Gaussian pmfs. The amount of information leaked to the relay is quantified in terms of the flatness factor of the lattice, which (loosely speaking) measures how close the quantization error of a Gaussian random vector is to being uniform. Under the assumption that the lattices satisfy certain “goodness” properties (it is known that such lattices indeed exist), we prove that our scheme achieves strong secrecy. We also study the achievable transmission rates under an average power constraint, and we show that this is within $\frac{1}{2}\log_2 2e$ bits of the achievable rate without secrecy constraints. We also extend the schemes for perfect and strong secrecy to a multi-hop line network.

Some Extensions

We then look at the robustness of our schemes to channel imperfections in Chapter 5. We operate under the assumption that the channel gains are not known to the user nodes, but known to the relay, and study what can be achieved in this scenario. We show that if the channel gains are rational, then perfect/strong secrecy can still be guaranteed, and derive achievable rates. However, we show that secure communication is not possible if the channel gains are irrational and the channels are noiseless.

Although we assume that the intermediate relay is “honest-but-curious”, we can ensure security even in some cases where the relay tries to modify the messages before forwarding. With some additional preprocessing along the lines of [41], we can ensure that the users can detect (with high probability) whether the relay has attempted to modify the messages. Once manipulation has been detected, the user nodes can retransmit the messages. In a larger network, the users may decide to flag the relay as an adversary and choose to exchange messages via a different relay.

1.3.2 “Good” Lattices with Low Decoding Complexity

In the first few chapters, we assume that the lattices used to design codes satisfy certain “goodness” properties. While prior work has shown that such lattices exist, finding explicit constructions that have polynomial-time encoders and decoders has proved to be a challenging task.

Geometric Properties of Low Density Construction-A Lattices

In Chapters 6 and 7, we study some practical lattice constructions that possess the properties that we desire. Chapter 6 shows that the LDA lattices proposed by di Pietro et al. [22] also satisfies several goodness properties, and thus proves their optimality in problems other than reliable communication over the point-to-point AWGN channel. Specifically, we will show that LDA lattices are good for packing, MSE quantization, and their duals are also good for packing. Following the work of di Pietro et al. [22], we assume that the factor graph of the LDPC code — from which the LDA lattice is constructed — is an expander graph. We extensively make use of the expansion properties of this graph, and derive lower bounds on the field size (over which the LDPC code is constructed) to guarantee these “goodness” properties.

Concatenation as a Method to Reduce Complexity

In Chapter 7, we take a different approach to design low-complexity lattice codes. We use ideas from Forney [34] and concatenate Construction-A lattices with nonbinary Reed-Solomon codes and expander codes. We show that this coding scheme can achieve the capacity of the AWGN channel with polynomial encoding and decoding complexities. We show that by concatenating an inner nested lattice code with an outer Reed-Solomon code, we can reduce the decoding complexity to be quadratic in the blocklength, N. Furthermore, concatenating with an outer expander code reduces the decoding complexity even further, to $O(N \log^2 N)$. We also make some remarks as to how this scheme can be used in other Gaussian channels.

1.3.3 Secret Key Generation

In Chapter 8, we study a problem of secret key generation from correlated Gaussian sources. Under the assumption of a quantization rate constraint at each terminal, we give a lattice-based scheme to generate a strongly secure key using public communication. Using the techniques developed in 7, we design a scheme whose overall computational complexity is polynomial in the number of samples (of the Gaussian source). For Markov tree sources, we derive expressions for the achievable secret key rate. We show that for certain classes of sources, our scheme achieves key capacity in the fine quantization limit. However, we also give examples where our scheme does not achieve capacity.

Finally, in Chapter 9, we conclude with some remarks and suggestions for future work.

Chapter 2

A Primer on Lattice Codes

In this chapter, we review some basic definitions and results on lattices, and also present some of our own results. A more detailed treatment of lattices can be found in [5, 14]. For a treatment of lattices with applications to compression and communication, see the book by Zamir [112].

We begin the chapter with some essential definitions and facts related to lattices. We will then describe some "goodness" properties that we want our lattices to satisfy, and present a class of lattices that satisfies these properties. These lattices are obtained from linear codes over prime fields by a process called Construction A. While the goodness properties of these lattices were well-known [29], their duals were not studied earlier. As we shall see in Chapter 3, having "good" lattices whose duals simultaneously satisfy some "goodness" properties is useful in proving some results. In this chapter, we show that there exist Construction-A lattices such that both the primal and dual lattices satisfy these "goodness" properties.

2.1 Lattices in $\mathbb{R}^n$

Let k, n be positive integers with $k \leq n$. Suppose $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k$ are linearly independent column vectors in $\mathbb{R}^n$. Then the set of all integer-linear combinations of the $\mathbf{u}_i$'s, $\Lambda = \{\sum_{i=1}^{k} a_i \mathbf{u}_i : a_i \in \mathbb{Z}, 1 \leq i \leq k\}$, is called a k-dimensional *lattice* in $\mathbb{R}^n$. It is easy to verify that Λ forms an abelian group under componentwise addition. The collection of vectors $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k\}$ is called a *basis* for the lattice Λ, and it is a standard fact that the basis of a lattice is not unique.

The $k \times n$ matrix $\mathsf{A} \triangleq [\mathbf{u}_1\ \mathbf{u}_2\ \cdots \mathbf{u}_k]^T$ is called a *generator matrix* of Λ, and we say that the vectors $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k$ generate[1] Λ. We write $\Lambda = \mathsf{A}^T\mathbb{Z}^k \triangleq \{\mathsf{A}^T\mathbf{x} : \mathbf{x} \in \mathbb{Z}^k\}$. If Λ is full-rank (i.e., Λ is a n-dimensional lattice in $\mathbb{R}^n$), then the *determinant* of Λ, denoted by $\det\Lambda$ is defined to be $|\det\mathsf{A}|$. Furthermore, $\det\Lambda$ does not depend on the generator matrix (see, e.g., [112, Chapter 2]). Unless mentioned otherwise, we will henceforth consider full-rank lattices in $\mathbb{R}^n$.

If Λ and Λ_0 are two lattices in $\mathbb{R}^n$ such that $\Lambda_0 \subset \Lambda$, then we say that Λ_0 is a *sublattice* of Λ, or that Λ_0 is *nested* within Λ. We call Λ_0 the *coarse lattice* and Λ the *fine lattice.* The number of cosets of Λ_0 in Λ is called the *index* or *nesting ratio* of Λ_0 in Λ, denoted by $|\Lambda/\Lambda_0|$. We also have $|\Lambda/\Lambda_0| = \det\Lambda_0/\det\Lambda$ [5, Theorem 5.2].

If A is a generator matrix of a lattice Λ, then $\Lambda^* \triangleq \{\mathsf{A}^{-1}\mathbf{z} : \mathbf{z} \in \mathbb{Z}^n\}$ is called the *dual lattice* of Λ. The dual lattice Λ^* is also equal to $\{\mathbf{x} : \sum_{i=1}^n x_i y_i \in \mathbb{Z}$ for every $\mathbf{y} \in \Lambda\}$[5]. The *Fourier dual* of Λ, denoted $\widehat{\Lambda}$, is defined as $2\pi\Lambda^*$.

For any $\mathbf{x} \in \mathbb{R}^n$, we define the nearest neighbour quantizer $Q_\Lambda(\mathbf{x}) \triangleq \arg\min_{\lambda\in\Lambda}\|\mathbf{x}-\lambda\|$ to be the function which maps $\mathbf{x}$ to the closest point in Λ, with ties resolved according to any fixed rule. The *fundamental Voronoi region* of Λ is defined as $\mathcal{V}(\Lambda) \triangleq \{\mathbf{y} : Q_\Lambda(\mathbf{y}) = \mathbf{0}\}$. The volume of the fundamental Voronoi region, $\mathrm{vol}(\mathcal{V}(\Lambda))$ is equal to $\det\Lambda$ [5, 14]. In this thesis, we use $\mathrm{vol}\Lambda \triangleq \mathrm{vol}(\mathcal{V}(\Lambda))$ and $\det\Lambda$ interchangeably.

For any $\mathbf{x} \in \mathbb{R}^n$, we define the modulo-$\Lambda$ operation as $[\mathbf{x}] \bmod \Lambda \triangleq \mathbf{x} - Q_\Lambda(\mathbf{x})$. In other words, $[\mathbf{x}] \bmod \Lambda$ gives the quantization error of the nearest neighbour quantizer $Q_\Lambda(\cdot)$. Figure 2.1 illustrates the $Q_\Lambda(\cdot)$ and the modulo-Λ operations.

The *covering radius* of Λ, denoted by $r_{\text{cov}}(\Lambda)$, is defined as the radius of the smallest closed ball in $\mathbb{R}^n$ centered at $\mathbf{0}$ which contains $\mathcal{V}(\Lambda)$. The *effective radius*, $r_{\text{eff}}(\Lambda)$, is defined as the radius of a ball in $\mathbb{R}^n$ having the same volume as that of $\mathcal{V}(\Lambda)$. The *packing radius* of Λ, $r_{\text{pack}}(\Lambda)$, is the radius of the largest open ball centered at $\mathbf{0}$ which is contained

[1] Note that the *rows* of the generator matrix, and not the columns, generate Λ.

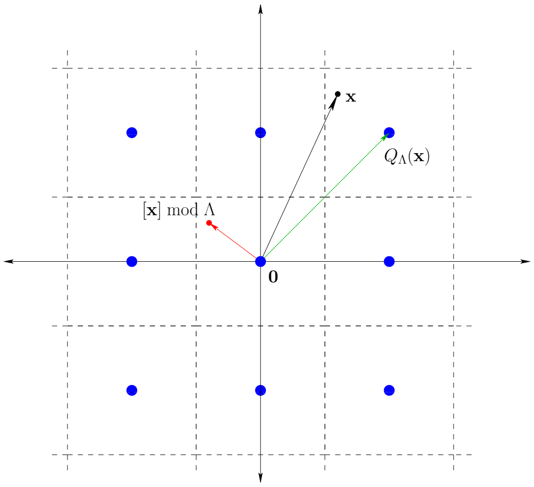


Figure 2.1: Illustrating the $Q_\Lambda(.)$ and the $[.] \bmod \Lambda$ operation for the $\mathbb{Z}^2$ lattice.

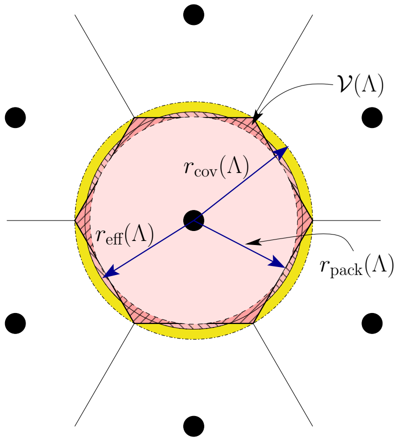

$\mathcal{V}(\Lambda)$
$r_{\text{cov}}(\Lambda)$
$r_{\text{eff}}(\Lambda)$
$r_{\text{pack}}(\Lambda)$

Figure 2.2: Illustrating the covering, packing and effective radii of the hexagonal lattice.

in $\mathcal{V}(\Lambda)$. Clearly, $r_{\text{cov}}(\Lambda) \geq r_{\text{eff}}(\Lambda) \geq r_{\text{pack}}(\Lambda)$. These parameters are illustrated for the hexagonal lattice in Fig. 2.2.

The *second moment per dimension* of Λ is defined as

$$\sigma^2(\Lambda) \triangleq \frac{1}{n \det \Lambda} \int_{\mathcal{V}(\Lambda)} \|\mathbf{y}\|^2 d\mathbf{y}, \tag{2.1}$$

which is equal to the second moment per dimension of a random vector uniformly distributed over $\mathcal{V}(\Lambda)$. The *normalized second moment per dimension* (also called the normalized moment of inertia) of a measurable set $\mathcal{S} \subset \mathbb{R}^n$ is defined as

$$G(\mathcal{S}) \triangleq \frac{1}{n\left[\mathrm{vol}(\mathcal{S})\right]^{1+2/n}} \int_{\mathcal{S}} \|\mathbf{y}\|^2 d\mathbf{y}. \tag{2.2}$$

The normalized second moment per dimension (NSM) of Λ is defined as the NSM of $\mathcal{V}(\Lambda)$.

$$G(\Lambda) = \frac{1}{(\det\Lambda)^{2/n}}\sigma^2(\Lambda) = \frac{1}{n\left[\det\Lambda\right]^{1+2/n}} \int_{\mathcal{V}(\Lambda)} \|\mathbf{y}\|^2 d\mathbf{y}. \tag{2.3}$$

2.2 "Goodness" Properties

Lattices have been used as a tool to solve several problems in mathematics, compression, communication and cryptography. In this section, we introduce some useful properties of sequences of lattices, and motivations for studying them.

2.2.1 Lattice coverings of space

One of the first problems where lattices were studied was how one can efficiently cover n-dimensional space using tessellations of a convex body. An interesting problem is the following: What is the best way to arrange identical n-dimensional spheres of a fixed radius so as to cover all of space, such that this arrangement minimizes the number of spheres placed per unit volume? This is called the sphere covering problem. One can obtain a *lattice covering* of space by placing these n-dimensional spheres such that the centers coincide with the lattice points. If the radius of each sphere is at least the covering radius of the lattice, then we can guarantee that all of space is covered by the spheres. For a given lattice covering, we can find the amount of "wastage", or the fraction of volume covered by more than one sphere. An equivalent quantity that captures the wastage is the *covering fraction*, defined as the ratio of the covering radius to the effective radius of the lattice. Clearly, this quantity is greater than one, but how small can this be? In particular, we are interested in high-dimensional lattices for which the covering fraction is as close to 1 as possible. We say that a sequence of lattices (indexed by the dimension, n) $\{\Lambda^{(n)}\}$, is *good for covering* if

$$\lim_{n\to\infty} \frac{r_{\text{cov}}(\Lambda^{(n)})}{r_{\text{eff}}(\Lambda^{(n)})} = 1. \tag{2.4}$$

It was shown by Rogers [78] that such lattices indeed exist. Lattices satisfying (2.4) are also called *covering-good* or *Rogers-good* lattices.

2.2.2 Lattice Packings

Another fundamental problem involving lattices is that of packing spheres in n-dimensional space. This problem asks for the most efficient way to arrange identical nonintersecting spheres in space so as to minimize the fraction of volume that is not covered by any sphere. As we did for the covering problem, we can place the spheres so that the centers coincide with the lattice points. As long as the radius of the spheres is less than the packing radius of the lattice, two spheres will not touch each other. We then want the lattice that minimizes the fraction of volume that is not covered by any sphere. An equivalent figure of merit for the efficiency of a lattice packing is the ratio of the packing radius to the effective radius of the lattice. In this regard, we say that a sequence of lattices $\{\Lambda^{(n)}\}$ is *good for packing* if

$$\lim_{n\to\infty} \frac{r_{\text{pack}}(\Lambda^{(n)})}{r_{\text{eff}}(\Lambda^{(n)})} \geq \frac{1}{2}. \tag{2.5}$$

We also call such lattices *packing-good* lattices.

2.2.3 Lattice Quantization

Lattices have been used to construct vector quantizers that minimize the mean-squared distortion. Gersho [36] conjectured that for a source which is uniformly distributed over a rectangular cell, the optimal fixed-length quantizer has a "regular" structure, i.e., it is obtained by tessellating some basic cell $\mathcal{S}$. If we assume that the conjecture is true, then for high quantization rates R, the mean-squared error (MSE) distortion is $G(\mathcal{S})2^{-2R}$, where $G(\mathcal{S})$ denotes the normalized second moment per dimension (NSM) of $\mathcal{S}$. It therefore makes sense to restrict our attention to lattice quantizers since they have a regular structure. Our objective would then be to look for a lattice quantizer with the least $G(\Lambda)$.

The NSM of any lattice is always lower bounded by that of a sphere, and the NSM of a sphere converges to $1/(2\pi e)$ as $n \to \infty$ (see, e.g., [29]). We say that a sequence of lattices $\{\Lambda^{(n)}\}$ is *good for mean-squared error (MSE) quantization* if

$$\lim_{n\to\infty} G(\Lambda^{(n)}) = \frac{1}{2\pi e}. \tag{2.6}$$

A useful fact is that if $\{\Lambda^{(n)}\}$ is good for covering, then $\{\Lambda^{(n)}\}$ is also good for MSE quantization [29].

2.2.4 Coding Without Restrictions for the AWGN Channel

Lattices have also been used to design good codes for the AWGN channel. Poltyrev [74] studied the problem of transmission of information in an AWGN channel without any input power constraints. In this problem, we want to design a code $\mathcal{C}$, which is simply an infinite collection of points in $\mathbb{R}^n$, such that *every* point in the constellation can be recovered reliably when transmitted across an AWGN channel. Suppose that some $\mathbf{x} \in \mathcal{C}$ was transmitted across the channel. Then, a minimum-distance decoder makes an error if the received vector is closer to some $\mathbf{x}' \in \mathcal{C}\backslash\{\mathbf{x}\}$, and we want the probability of this event (maximized over all possible $\mathbf{x} \in \mathcal{C}$) to vanish asymptotically in the blocklength. In other words,

$$\sup_{\mathbf{x}\in\mathcal{C}} \Pr\Big[\exists \mathbf{x}' \neq \mathbf{x} \text{ such that } \|\mathbf{z}\| > \|\mathbf{x}+\mathbf{z}-\mathbf{x}'\|\Big] \to 0 \text{ as } n \to \infty,$$

where $\mathbf{z}$ denotes AWGN with mean 0 and variance σ^2. If we restrict ourselves to lattice constellations, i.e., $\mathcal{C}$ is a lattice in $\mathbb{R}^n$, then the closest lattice point decoder is also the maximum likelihood decoder [74]. Moreover, studying the performance of the infinite lattice constellation also gives us a handle on the performance of a lattice code obtained by choosing a finite subset of points from the lattice (usually to satisfy a power constraint) with lattice decoding. If $\mathcal{C} = \Lambda$, then we have

$$\sup_{\mathbf{x}\in\mathcal{C}} \Pr[\exists \mathbf{x}' \neq \mathbf{x} \text{ such that } \|\mathbf{z}\| > \|\mathbf{x}+\mathbf{z}-\mathbf{x}'\|] = \Pr[\mathbf{z} \notin \mathcal{V}(\Lambda)].$$

Let $\mathbf{Z}$ be a zero-mean n-dimensional white Gaussian vector having second moment per dimension equal to σ^2. Let

$$\mu \triangleq \frac{\left(\mathrm{vol}\Lambda^{(n)}\right)^{2/n}}{\sigma^2}.$$

Then we say that $\{\Lambda^{(n)}\}$ is *good for AWGN channel coding* if the probability that $\mathbf{Z}$ lies outside the fundamental Voronoi region of $\Lambda^{(n)}$ is upper bounded as

$$\Pr[\mathbf{Z} \notin \mathcal{V}(\Lambda^{(n)})] \le e^{-n\left(E_U(\mu)-o_n(1)\right)}$$

for all σ^2 that satisfy $\mu \ge 2\pi e$. Here, $E_U(\cdot)$, called the *Poltyrev exponent*, is defined as follows:

$$E_U(\mu) = \begin{cases} \frac{\mu}{16\pi e}, & 8\pi e \le \mu \\ \frac{1}{2}\ln\frac{\mu}{8\pi}, & 4\pi e \le \mu \le 8\pi e \\ \frac{\mu}{4\pi e} - \frac{1}{2}\ln\frac{\mu}{2\pi}, & 2\pi e \le \mu \le 4\pi e \end{cases} \tag{2.7}$$

Suppose that we use a subcollection of points from $\Lambda^{(n)}$ as the codebook for transmission over an AWGN channel. Then, as long as

$$\frac{\left(\mathrm{vol}\Lambda^{(n)}\right)^{2/n}}{\sigma^2} \ge 2\pi e,$$

the probability that a lattice decoder decodes to a lattice point other than the one that was transmitted, decays to zero exponentially in the dimension n, with the exponent given by (2.7).

2.2.5 Flatness Factor and Secrecy Goodness

For $\mathbf{x}, \mathbf{c} \in \mathbb{R}^n$ and $\theta > 0$, we define

$$g_{\theta,\mathbf{c}}(\mathbf{x}) \triangleq \frac{1}{(2\pi\theta^2)^{n/2}} e^{-\frac{\|\mathbf{x}-\mathbf{c}\|^2}{2\theta^2}}.$$

For any lattice Λ in $\mathbb{R}^n$, and any $\theta > 0$, the *flatness factor* $\epsilon_\Lambda(\theta)$ is defined as [58, 7]

$$\epsilon_\Lambda(\theta) \triangleq \frac{\max_{\mathbf{x}\in\mathcal{V}(\Lambda)} \left| \left(\sum_{\lambda\in\Lambda} g_{\theta,\lambda}(\mathbf{x})\right) - (1/\det\Lambda)\right|}{(1/\det\Lambda)}. \tag{2.8}$$

If $\mathbf{Z}$ is a Gaussian vector with mean zero and variance θ^2, then $\epsilon_\Lambda(\theta)$ tells us how "far" the random vector $[\mathbf{Z}] \bmod \Lambda$ is from being uniformly distributed over $\mathcal{V}(\Lambda)$. This is a useful parameter in certain problems of communication and security [58, 7]. A useful property of the flatness factor is that it is a monotonic function of θ: for $a > b > 0$, and any lattice Λ, we have $\epsilon_\Lambda(a) \le \epsilon_\Lambda(b)$ [58, Remark 2]. Following [58], we define a sequence of lattices $\{\Lambda^{(n)}\}$ to be *secrecy-good* if

$$\epsilon_{\Lambda^{(n)}}(\theta) \le 2^{-\Omega(n)} \text{ for all } \theta \text{ such that } \frac{(\det(\Lambda^{(n)}))^{2/n}}{2\pi\theta^2} < 1.$$

2.2.6 Remarks on Goodness Properties

An interesting property of the goodness properties is that they are invariant to rotation and scaling by a constant. Specifically, if $\{\Lambda^{(n)}\}$ is a sequence of lattices which is good for covering/packing/MSE quantization/AWGN/secrecy, and $\alpha > 0$ is an arbitrary constant, then $\{\alpha\Lambda^{(n)}\}$ is also good for covering/packing/MSE quantization/AWGN/secrecy. Furthermore, if A is an $n \times n$ matrix having orthonormal columns, then $\{\mathsf{A}\Lambda^{(n)}\}$ is also good for covering/packing/MSE quantization/AWGN/secrecy.

2.3 Construction A

Let n and k be positive integers with $k \le n$, and let q be a prime number. Let $\mathbb{F}_q$ denote the field of integers modulo q.

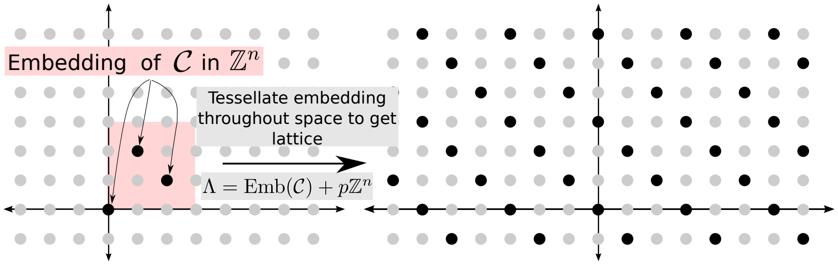

Embedding of $\mathcal{C}$ in $\mathbb{Z}^n$
Tessellate embedding throughout space to get lattice
$\Lambda = \text{Emb}(\mathcal{C}) + p\mathbb{Z}^n$

Figure 2.3: Construction A

1. Choose a $k \times n$ matrix G uniformly at random over $\mathbb{F}_q^{k\times n}$. This is done by choosing each element of G (there are nk of them) uniformly over $\mathbb{F}_q$, and independently of the other entries. Note that G need not be full-rank. However, the probability that G is full-rank goes to 1 as $(n-k)$ tends to ∞ [29]. The linear code over $\mathbb{F}_q$ generated by G is denoted by $\mathcal{C}(\mathsf{G})$.

2. Apply Construction A on the code generated above. This is done as follows:

   (c_1) The codebook generated using G is $\mathcal{C}(\mathsf{G}) = \{(\mathsf{G}^T\mathbf{y}) \bmod q : \mathbf{y} \in \mathbb{F}_q^k\}$.

   (c_2) Let Φ denote the natural embedding of $\mathbb{F}_q^n$ in $\mathbb{Z}^n$ (equivalently, in $\mathbb{R}^n$). The codebook is then scaled so that the codeword coordinates are restricted to lie within the n-dimensional unit cube: $\mathcal{C}' = (1/q)\Phi(\mathcal{C}(\mathsf{G})) = \{(1/q)\Phi(\mathbf{x}) : \mathbf{x} \in \mathcal{C}(\mathsf{G})\}$. It is easy to see that $\mathcal{C}'$ is a subset of the vertices of the rectangular grid of side $1/q$ sitting within the unit cube.

   (c_3) The lattice is obtained by repeating these points over the entire space, $\mathbb{R}^n$, i.e., $\widetilde{\Lambda}(\mathcal{C}) = \mathcal{C}' + \mathbb{Z}^n \triangleq \{\mathbf{c} + \mathbf{x} : \mathbf{c} \in \mathcal{C}', \mathbf{x} \in \mathbb{Z}^n\}$.

Following the terminology used in [29], we will henceforth call the above as the (n, k, q) *ensemble*. Also, from the construction, it is clear that $\mathbb{Z}^n$ is a sublattice of $\widetilde{\Lambda}(\mathcal{C})$. More detail regarding Construction-A lattices can be found in [14, 112].

Let Λ be a full-rank lattice chosen from the (n, k, q) ensemble. From the construction, we can see that Λ has $\mathbb{Z}^n$ as a sublattice. The nesting ratio of the $(\Lambda, \mathbb{Z}^n)$ nested lattice pair is q^k. This is also equal to the ratio of the volume of the fundamental Voronoi region of $\mathbb{Z}^n$ to that of Λ. The volume of $\mathcal{V}(\Lambda)$ is in turn equal to the volume of a n-dimensional ball of radius $r_{\text{eff}}(\Lambda)$, and hence,

$$q^k = \frac{\Gamma\left(\frac{n}{2}+1\right)}{\pi^{n/2}\left(r_{\text{eff}}(\Lambda)\right)^n}. \tag{2.9}$$

Rearranging, we get

$$r_{\text{eff}}(\Lambda) = \left(\frac{\Gamma\left(\frac{n}{2}+1\right)}{\pi^{n/2}q^k}\right)^{1/n}. \tag{2.10}$$

2.4 "Good" Construction-A Lattices having "Good" Duals

We have the following theorem by Erez et al. [29], which says that if the parameters k and q are selected appropriately, then almost all lattices in a (n, k, q) ensemble satisfy the "goodness" properties described earlier.

Theorem 2.4.1 ([29], Theorem 5)**.** *Let $0 < r < \frac{1}{4}$ be chosen arbitrarily. Let $\{\Lambda^{(n)}\}$ be a sequence of lattices, with each $\Lambda^{(n)}$ selected uniformly at random from an (n, k, q) ensemble, such that*

(L1) $k \leq \beta n$ for some $0 < \beta < 1$, but k grows faster than $\log^2 n$, and

(L2) q is chosen so that $r_{\text{eff}}(\Lambda^{(n)})$, as given by (2.10), satisfies $r < r_{\text{eff}}(\Lambda^{(n)}) < 2r$.

Then, the sequence of lattices $\{\Lambda^{(n)}\}$ is simultaneously good for covering, packing and MSE quantization, with probability approaching 1 as n tends to infinity. If, in addition, we have

(L3) $\beta < 1/2$,

then the sequence of lattices is also good for AWGN channel coding with probability tending to 1 as $n \to \infty$.

We will show the following result, which says that the duals of randomly chosen Construction-A lattices also satisfy the desired "goodness" properties.

Theorem 2.4.2 ("Good" dual lattices)**.** *Let $\{\Lambda^{(n)}\}$ be a sequence of lattices, with each $\Lambda^{(n)}$ chosen uniformly at random from an (n, k, q) ensemble. Suppose that*

(D1) $n - k \leq \beta n$ for some $0 < \beta < 1$, but $n - k$ grows faster than $\log^2 n$, and

(D2) q is chosen so that $r_{\text{eff}}(\Lambda^{(n)*})$, given by

$$r_{\text{eff}}(\Lambda^{(n)*}) = \left(\frac{\Gamma\left(\frac{n}{2}+1\right)}{\pi^{n/2}q^{n-k}}\right)^{1/n} \tag{2.11}$$

satisfies $r < r_{\text{eff}}(\Lambda^{(n)}) < 2r$ for some $0 < r < \frac{1}{4}$.

Then, the sequence of Fourier dual lattices $\{\widehat{\Lambda}^{(n)}\}$ is simultaneously good for covering, packing and MSE quantization, with probability approaching 1 as n tends to infinity. If, in addition, we have $\beta < 1/2$, then $\{\widehat{\Lambda}^{(n)}\}$ is also good for AWGN channel coding with probability tending to 1 as $n \to \infty$.

In proving Theorem 2.4.2, we use some properties of the duals of Construction-A lattices. Recall from Section 2.1 that if A is a generator matrix of a lattice Λ, then the dual lattice of Λ, denoted by Λ^*, is the set of all integer linear combinations of the rows of $(\mathsf{A}^{-1})^T$. It turns out that the dual of a Construction-A lattice is also a Construction-A lattice, as we shall see very soon. First, let us uncover some basic properties of the generator matrix of a Construction-A lattice.

Lemma 2.4.3. *Suppose that G is a $k \times n$ generator matrix of an (n, k) linear code $\mathcal{C}$ over $\mathbb{F}_q$, q being prime, and G having the form*

$$\mathsf{G} = \begin{bmatrix} \mathsf{I}_k & \mathsf{B} \end{bmatrix},$$

where I_k denotes the $k \times k$ identity matrix. Let $\Lambda(\mathcal{C})$ be the lattice obtained by employing Construction A on the code $\mathcal{C}$. Then, the matrix

$$\mathsf{A} = \frac{1}{q}\begin{bmatrix} \mathsf{I}_k & \mathsf{B} \\ 0 & q\mathsf{I}_{(n-k)} \end{bmatrix} \tag{2.12}$$

is a generator matrix for the Construction-A lattice $\Lambda(\mathcal{C})$.

Proof. We want to show that $\mathsf{A}^T\mathbb{Z}^n \triangleq \{\mathsf{A}^T\mathbf{y} : \mathbf{y} \in \mathbb{Z}^n\} = \Lambda(\mathcal{C})$. We first prove that $\mathsf{A}^T\mathbb{Z}^n \subseteq \Lambda(\mathcal{C})$. By definition, $\Lambda(\mathcal{C}) = \{\mathbf{x} \in \mathbb{R}^n : (q\mathbf{x}) \bmod q \in \mathcal{C}\}$. Therefore, it is enough to show that $(q\mathsf{A}^T\mathbf{z}) \bmod q \in \mathcal{C}$ for every $\mathbf{z} \in \mathbb{Z}^n$. Fix a $\mathbf{z} \in \mathbb{Z}^n$. Then,

$$\begin{aligned}(q\mathsf{A}^T\mathbf{z}) \bmod q &= \left(\begin{bmatrix} \mathsf{I}_k & \mathsf{B} \\ 0 & q\mathsf{I}_{(n-k)} \end{bmatrix}^T [z_1\, z_2 \ldots\, z_n]^T\right) \bmod q \\ &= \Big(\begin{bmatrix} \mathsf{I}_k & \mathsf{B} \end{bmatrix}^T [z_1\, \ldots\, z_k]^T \\ &\quad + \begin{bmatrix} 0 & q\mathsf{I}_{(n-k)} \end{bmatrix}^T [z_{k+1}\, \ldots\, z_n]^T\Big) \bmod q \\ &= \left(\begin{bmatrix} \mathsf{I}_k & \mathsf{B} \end{bmatrix}^T [z_1\, \ldots\, z_k]^T\right) \bmod q \\ &= (\mathsf{G}^T\widehat{\mathbf{z}}) \bmod q \in \mathcal{C}. \end{aligned} \tag{2.13}$$

For the converse, define $\mathcal{C}' = \{\frac{1}{q}\mathbf{c} : \mathbf{c} \in \mathcal{C}\}$. Then, $\Lambda(\mathcal{C}) = \mathcal{C}' + \mathbb{Z}^n \triangleq \{\mathbf{c}+\mathbf{z} : \mathbf{c} \in \mathcal{C}', \mathbf{z} \in \mathbb{Z}^n\}$. The set $\mathsf{A}^T\mathbb{Z}^n$ forms a group under (componentwise) addition. Hence, it is sufficient to show that $\mathcal{C}' \subseteq \mathsf{A}^T\mathbb{Z}^n$, and $\mathbb{Z}^n \subseteq \mathsf{A}^T\mathbb{Z}^n$. Fix an arbitrary $\mathbf{c} \in \mathcal{C}$. Let $\mathbf{c}' = \frac{1}{q}\mathbf{c}$. By definition, there exists a $\mathbf{x} \in \mathbb{Z}_q^k$ such that

$$\begin{aligned}\mathbf{c} &= \left(\begin{bmatrix} \mathsf{I}_k & \mathsf{B} \end{bmatrix}^T \mathbf{x}\right) \bmod q \\ &= \begin{bmatrix} \mathbf{x} \\ \mathsf{B}^T\mathbf{x} \end{bmatrix} \bmod q = \begin{bmatrix} \mathbf{x} \\ (\mathsf{B}^T\mathbf{x}) \bmod q \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{x} \\ \mathsf{B}^T\mathbf{x} \end{bmatrix} - q \begin{bmatrix} 0 \\ \mathbf{z}' \end{bmatrix}, \end{aligned} \tag{2.14}$$

for some $\mathbf{z}' \in \mathbb{Z}^{n-k}$. Therefore,

$$\mathbf{c} = \begin{bmatrix} \mathsf{I}_k & \mathsf{B} \\ 0 & q\mathsf{I}_{(n-k)} \end{bmatrix}^T \begin{bmatrix} \mathbf{x} \\ -\mathbf{z}' \end{bmatrix}.$$

Hence, there exists

$$\mathbf{z} = \begin{bmatrix} \mathbf{x} \\ \mathbf{z}' \end{bmatrix} \in \mathbb{Z}^n$$

so that $\mathbf{c}' = \mathsf{A}^T\mathbf{z}$. Therefore, we can say that $\mathcal{C}' \subseteq \mathsf{A}^T\mathbb{Z}^n$. Next, consider $\mathbf{z} \in \mathbb{Z}^n$. Let us define

$$\mathsf{A}^* \triangleq \begin{bmatrix} q\mathsf{I}_k & \mathsf{0} \\ -\mathsf{B}^T & \mathsf{I}_{(n-k)} \end{bmatrix}.$$

Note that $\mathsf{A}^T\mathsf{A}^* = \mathsf{I}_n$, the $n \times n$ identity matrix. Let $\mathbf{z}' = \mathsf{A}^*\mathbf{z} \in \mathbb{Z}^n$. Then, $\mathsf{A}^T\mathbf{z}' = \mathsf{A}^T(\mathsf{A}^*\mathbf{z}) = (\mathsf{A}^T\mathsf{A}^*)\mathbf{z} = \mathbf{z}$. Hence, we can say that for every $\mathbf{z} \in \mathbb{Z}^n$, there exists a $\mathbf{z}' \in \mathbb{Z}^n$ so that $\mathbf{z} = \mathsf{A}^T\mathbf{z}$, and hence $\mathbb{Z}^n \subseteq \mathsf{A}^T\mathbb{Z}^n$, thus concluding the proof. □

It can be shown in a similar manner that if G has the form

$$\mathsf{G} = \begin{bmatrix} \mathsf{B} & \mathsf{I}_k \end{bmatrix},$$

then,

$$\mathsf{A} = \frac{1}{q}\begin{bmatrix} \mathsf{B} & \mathsf{I}_k \\ q\mathsf{I}_{(n-k)} & \mathsf{0} \end{bmatrix}$$

is a generator matrix for $\Lambda(\mathcal{C})$.

Lemma 2.4.4. *Let $\mathcal{C}$, G, $\Lambda(\mathcal{C})$ be as in Lemma 2.4.3. Then, the dual of $\Lambda(\mathcal{C})$, denoted by $\Lambda^*(\mathcal{C})$, has generator matrix*

$$\mathsf{A}^* = \begin{bmatrix} q\mathsf{I}_k & \mathsf{0} \\ -\mathsf{B}^T & \mathsf{I}_{(n-k)} \end{bmatrix}. \tag{2.15}$$

Therefore, $\Lambda^*(\mathcal{C}) = q\Lambda(\mathcal{C}^{\perp})$, where $\mathcal{C}^{\perp}$ denotes the dual code of $\mathcal{C}$.

Proof. Consider,

$$\begin{aligned} \mathsf{A}(\mathsf{A}^*)^T &= \frac{1}{q}\begin{bmatrix} \mathsf{I}_k & \mathsf{B} \\ 0 & q\mathsf{I}_{(n-k)} \end{bmatrix}\begin{bmatrix} q\mathsf{I}_k & -\mathsf{B} \\ 0 & \mathsf{I}_{(n-k)} \end{bmatrix} \\ &= \frac{1}{q}\begin{bmatrix} q\mathsf{I}_k & 0 \\ 0 & q\mathsf{I}_{(n-k)} \end{bmatrix} \\ &= \mathsf{I}_n. \end{aligned}$$

Similarly, $(\mathsf{A}^*)^T\mathsf{A} = \mathsf{I}_n$.

Since a permutation of the rows of a generator matrix of a lattice also yields a valid generator matrix for the same lattice,

$$\mathsf{A}_1^* = \begin{bmatrix} -\mathsf{B}^T & \mathsf{I}_{(n-k)} \\ q\mathsf{I}_k & 0 \end{bmatrix}$$

is also a generator matrix for $\Lambda^*(\mathcal{C})$. If $\mathcal{C}^\perp$ denotes the dual code of $\mathcal{C}$, then $\mathcal{C}^\perp$ has a generator matrix [79]

$$\mathsf{G} = \begin{bmatrix} -\mathsf{B}^T & \mathsf{I}_{(n-k)} \end{bmatrix}.$$

Therefore, we can conclude that the dual lattice, $\Lambda^*(\mathcal{C})$ is a q-scaled version of the lattice obtained by applying Construction A to $\mathcal{C}^\perp$, i.e., $\Lambda^*(\mathcal{C}) = q\Lambda(\mathcal{C}^\perp)$. □

We now have all the tools required to prove that randomly chosen Construction-A lattices have "good" duals.

2.4.1 Proof of Theorem 2.4.2

Let us fix an $\epsilon > 0$. From Lemma 2.4.3 and Lemma 2.4.4, we know that for any Construction-A lattice $\Lambda(\mathcal{C})$, we have $\Lambda^*(\mathcal{C}) = q\Lambda(\mathcal{C}^\perp)$. If we choose an $(n-k)\times n$ parity check matrix H uniformly at random, then Theorem 2.4.1 tells us that $\Lambda(\mathcal{C}^\perp)$ is good for covering, packing, MSE quantization and AWGN channel coding with probability tending to 1 as $n \to \infty$. Since we are choosing the $k \times n$ generator matrix (and not the parity check matrix) uniformly at random, we cannot directly apply Theorem 2.4.1.

Let $\mathcal{A}_G$ denote the set of all linear codes over $\mathbb{F}_q$ generated by all possible (not necessarily full-rank) $k \times n$ matrices with entries from $\mathbb{F}_q$. Similarly, let $\mathcal{A}_P$ denote the set of all linear codes corresponding to all possible (not necessarily full-rank) $(n-k) \times n$ parity check matrices. Furthermore, let $\mathcal{F}$ denote the set of all linear codes in $\mathcal{A}_G$ (or $\mathcal{A}_P$) having dimension k. These are the linear codes corresponding to full-rank generator (and parity check) matrices. It is easy to see that $\mathcal{F} = \mathcal{A}_G \cap \mathcal{A}_P$. We also define $\mathcal{D}$ as the set of all those codes in $\mathcal{A}_P$ corresponding to "good" dual lattices, i.e., all those codes $\mathcal{C}$ such that

$$\frac{r_{\text{pack}}(\Lambda(\mathcal{C}^{\perp}))}{r_{\text{eff}}(\Lambda(\mathcal{C}^{\perp}))} \geq \frac{1}{2} - \epsilon, \quad \frac{r_{\text{cov}}(\Lambda(\mathcal{C}^{\perp}))}{r_{\text{eff}}(\Lambda(\mathcal{C}^{\perp}))} \leq 1 - \epsilon,$$

$$G(\Lambda(\mathcal{C}^{\perp})) \leq \frac{1}{2\pi e} + \epsilon, \quad \text{and}$$

$$\int_{\mathbf{x} \notin \mathcal{V}(\Lambda(\mathcal{C}^{\perp}))} \frac{1}{(2\pi\sigma^2)^{n/2}} e^{-\|\mathbf{x}\|^2/(2\sigma^2)} d\mathbf{x} < \epsilon \quad \text{for all } \sigma^2 < \frac{(\text{vol}(\Lambda(\mathcal{C}^{\perp})))^{2/n}}{2\pi e}.$$

For $\mathcal{B} \in \{\mathcal{A}_G, \mathcal{D}, \mathcal{F}\}$, let $\mathcal{M}_G(\mathcal{B})$ denote the set of all $k \times n$ generator matrices (not necessarily full-rank) corresponding to the codes in $\mathcal{B}$. Likewise, for $\mathcal{B} \in \{\mathcal{A}_P, \mathcal{D}, \mathcal{F}\}$, let $\mathcal{M}_P(\mathcal{B})$ denote the set of all $(n-k) \times n$ parity check matrices corresponding to the codes in $\mathcal{B}$. In other words, $\mathcal{M}_G(\mathcal{A}_G)$ and $\mathcal{M}_P(\mathcal{A}_P)$ respectively denote the set of all $k \times n$ and $(n-k) \times n$ matrices over $\mathbb{F}_q$. The sets $\mathcal{M}_G(\mathcal{F})$ and $\mathcal{M}_P(\mathcal{F})$ respectively denote the collection of all full-rank $k \times n$ and $(n-k) \times n$ matrices over $\mathbb{F}_q$.

If the hypotheses of Theorem 2.4.2 are satisfied, then using Theorem 2.4.1, we can say that the fraction of "good parity check matrices" tends to 1, or

$$\frac{|\mathcal{M}_P(\mathcal{D})|}{|\mathcal{M}_P(\mathcal{A}_P)|} = 1 - o_n(1). \tag{2.16}$$

To prove Theorem 2.4.2, we want to show that the number of "good generator matrices" (in the sense that the resulting lattice is in $\mathcal{D}$) is large, or

$$\frac{|\mathcal{M}_G(\mathcal{D})|}{|\mathcal{M}_G(\mathcal{A}_G)|} \to 1 \text{ as } n \to \infty. \tag{2.17}$$

For our choice of parameters, a randomly chosen generator matrix (or parity check matrix) has full rank with high probability. In other words,

$$\frac{|\mathcal{M}_G(\mathcal{F})|}{|\mathcal{M}_G(\mathcal{A}_G)|} = 1 - o_n(1) \text{ and } \frac{|\mathcal{M}_P(\mathcal{F})|}{|\mathcal{M}_P(\mathcal{A}_P)|} = 1 - o_n(1). \tag{2.18}$$

Using (2.16) and (2.18) and a simple application of the union bound, we obtain

$$\frac{|\mathcal{M}_P(\mathcal{D}\cap\mathcal{F})|}{|\mathcal{M}_P(\mathcal{A}_P)|} = 1 - o_n(1). \tag{2.19}$$

We have,

$$\frac{|\mathcal{M}_G(\mathcal{D})|}{|\mathcal{M}_G(\mathcal{A}_G)|} \geq \frac{|\mathcal{M}_G(\mathcal{D}\cap\mathcal{F})|}{|\mathcal{M}_G(\mathcal{A}_G)|} \geq \frac{|\mathcal{M}_G(\mathcal{D}\cap\mathcal{F})|}{|\mathcal{M}_G(\mathcal{A}_G)|}\frac{|\mathcal{M}_P(\mathcal{F})|}{|\mathcal{M}_P(\mathcal{A}_P)|}. \tag{2.20}$$

However, every code in $\mathcal{F}$ has the same number of distinct generator matrices, and this number is equal to $|\mathcal{M}_G(\mathcal{F})|/|\mathcal{F}|$. Likewise, every code in $\mathcal{F}$ has $|\mathcal{M}_P(\mathcal{F})|/|\mathcal{F}|$ distinct parity-check matrices. Therefore,

$$\begin{aligned}
|\mathcal{M}_G(\mathcal{D}\cap\mathcal{F})||\mathcal{M}_P(\mathcal{F})| &= \left(|\mathcal{D}\cap\mathcal{F}|\frac{|\mathcal{M}_G(\mathcal{F})|}{|\mathcal{F}|}\right)|\mathcal{M}_P(\mathcal{F})| \\
&= \left(|\mathcal{D}\cap\mathcal{F}|\frac{|\mathcal{M}_P(\mathcal{F})|}{|\mathcal{F}|}\right)|\mathcal{M}_G(\mathcal{F})| \\
&= |\mathcal{M}_P(\mathcal{D}\cap\mathcal{F})||\mathcal{M}_G(\mathcal{F})|. \qquad (2.21)
\end{aligned}$$

Substituting this in (2.4.1), and then using (2.18), we get

$$\frac{|\mathcal{M}_G(\mathcal{D})|}{|\mathcal{M}_G(\mathcal{A}_G)|} \geq \frac{|\mathcal{M}_P(\mathcal{D}\cap\mathcal{F})|}{|\mathcal{M}_P(\mathcal{A}_P)|}\frac{|\mathcal{M}_G(\mathcal{F})|}{|\mathcal{M}_G(\mathcal{A}_P)|} = 1 - o_n(1).$$

This completes the proof of the theorem. □

2.5 Nested Construction-A Lattices

In all our problems, we will use codes obtained from nested Construction-A lattices. Our construction of the $(\Lambda^{(n)}, \Lambda_0^{(n)})$ nested lattice codes is based on [28, 69]. Choose a sequence of coarse lattices $\{\Lambda_0^{(n)}\}$, each $\Lambda_0^{(n)}$ selected uniformly at random from the

(n, k, q) ensemble, where k and q may be functions of n chosen beforehand. Let $\mathsf{A}^{(n)}$ be a lattice generator matrix of $\Lambda_0^{(n)}$, for $n = 1, 2, ...$ For this choice of $\{\Lambda_0^{(n)}\}$, we construct another ensemble of lattices from which we pick the sequence of fine lattices $\{\Lambda^{(n)}\}$. This consists of two steps:

(f_1) Choose a sequence of lattices, $\{\widetilde{\Lambda}_f^{(n)}\}$, with each $\widetilde{\Lambda}_f^{(n)}$ coming from the (n, k_1, q_1) ensemble of Construction-A lattices. The parameters k_1 and q_1 could be possibly different from k and q. As mentioned earlier, $\widetilde{\Lambda}_f^{(n)}$ contains $\mathbb{Z}^n$ as a sublattice. If the generator matrix of $\widetilde{\Lambda}_f^{(n)}$ has full rank, then the number of cosets of $\mathbb{Z}^n$ in $\widetilde{\Lambda}_f^{(n)}$ is $q_1^{k_1}$.

(f_2) The lattice $\widetilde{\Lambda}_f^{(n)}$ is subjected to a linear transformation by the matrix $(\mathsf{A}^{(n)})^T$, to get $\Lambda^{(n)} = (\mathsf{A}^{(n)})^T \widetilde{\Lambda}_f^{(n)} \triangleq \left\{ (\mathsf{A}^{(n)})^T \mathbf{y} : \mathbf{y} \in \widetilde{\Lambda}_f^{(n)} \right\}$. Since $\mathbb{Z}^n$ is nested within $\widetilde{\Lambda}_f^{(n)}$, we have that $\Lambda_0^{(n)}$ is a sublattice of $\Lambda^{(n)}$.

We will call this ensemble of $(\Lambda^{(n)}, \Lambda_0^{(n)})$ pairs as the (n, k, q, k_1, q_1) *ensemble.* A $(\Lambda^{(n)}, \Lambda_0^{(n)})$ nested lattice code is typically the set of all points of $\Lambda^{(n)}$ that lie within the fundamental Voronoi region of $\Lambda_0^{(n)}$. The *rate* of the code is

$$R^{(n)} \triangleq \frac{1}{n} \log_2 |\Lambda^{(n)} \cap \mathcal{V}(\Lambda_0^{(n)})| = \frac{1}{n} \log_2 |\Lambda^{(n)}/\Lambda_0^{(n)}| = \frac{k_1}{n} \log_2(q_1). \tag{2.22}$$

By choosing q_1 and k_1 appropriately, the rate can be varied.

Our objective is to show the existence of a sequence $\{\Lambda^{(n)}, \Lambda_0^{(n)}\}$ of nested lattice pairs, each $(\Lambda^{(n)}, \Lambda_0^{(n)})$ pair coming from a (d, k, q, k_1, q_1) ensemble, that satisfy the following properties:

(GL_1) The sequence of coarse lattices, $\{\Lambda_0^{(n)}\}$, are good for covering, packing, MSE quantization and AWGN channel coding.

(GL_2) The sequence of dual lattices, $\{\widehat{\Lambda}_0^{(n)}\}$, are good for packing.

(GL_3) The sequence of fine lattices, $\{\Lambda^{(n)}\}$, are good for covering, MSE quantization, and AWGN channel coding.

In order to obtain nested lattices from the (n, k, q, k_1, q_1) ensemble that satisfy properties (GL_1)–(GL_3), we will restrict the choice of k, q, k_1, q_1. We will make sure the values of these parameters can be chosen so that, in spite of the restriction, the transmission rate, as given by (2.22), can be chosen freely.

2.5.1 Choice of k, q, k_1, q_1

We choose

$$k = \beta_0 n, \text{ and } k_1 = \beta_1 n, \tag{2.23}$$

for some $0 < \beta_0, \beta_1 < 1/2$, and q and q_1 are prime numbers chosen such that

$$\lim_{n\to\infty} \frac{n}{q_1} = 0, \text{ and } r^{(0)} < r_{\text{eff}}(\Lambda_0^{(n)}) < 2r^{(0)}, \tag{2.24}$$

for some $0 < r^{(0)} < 1/4$. It is possible to choose primes that satisfy the above conditions. Choosing q_1 to grow faster than n is sufficient to satisfy the first condition. To see that q can be chosen so as to satisfy the second constraint, substitute (2.9) in (2.24) ,

$$\frac{\Gamma\left(\frac{n}{2}+1\right)}{\pi^{n/2}\left(2r^{(0)}\right)^n} < q^k < \frac{\Gamma\left(\frac{n}{2}+1\right)}{\pi^{n/2}\left(r^{(0)}\right)^n}.$$

Since $k = \beta_0 n$, we can rewrite the above inequalities as follows:

$$\left(\frac{\Gamma\left(\frac{n}{2}+1\right)}{\pi^{n/2}\left(2r^{(0)}\right)^n}\right)^{\frac{1}{\beta_0 n}} < q < 2^{1/\beta_0}\left(\frac{\Gamma\left(\frac{n}{2}+1\right)}{\pi^{n/2}\left(2r^{(0)}\right)^n}\right)^{\frac{1}{\beta_0 n}}.$$

Let α denote the left hand side of the above inequality. It is enough to show that for every $0 < \beta_0 < 1$ and all sufficiently large α, there exists a prime q which satisfies $\alpha < q < 2^{1/\beta_0}\alpha$. For any $\alpha > \frac{2^{-1+1/\beta_0}+1}{2^{-1+1/\beta_0}-1}$, there exists an integer m such that $\alpha < m \leq \alpha+1$ and $2m < 2^{1/\beta_0}\alpha$ (since for $\alpha > \frac{2^{-1+1/\beta_0}+1}{2^{-1+1/\beta_0}-1}$, we have $2(\alpha+1) < 2^{1/\beta_0}\alpha - 1$). By Bertrand's postulate (see e.g., [76]), for every positive integer m, there exists a prime number between m and $2m$, and therefore, we can choose a prime q satisfying (2.24).

Theorem 2.5.1. *Let* $(\Lambda^{(n)}, \Lambda_0^{(n)})$ *be a nested lattice pair chosen uniformly at random from* *the* (n, k, q, k_1, q_1) *ensemble, with the parameters* k, q, k_1, q_1 *chosen so as to satisfy (2.23) and (2.24). Then, the probability that* $(\Lambda^{(n)}, \Lambda_0^{(n)})$ *satisfies* (G_1)*–*(G_3) *tends to one as* n *approaches infinity.*

Proof. Using Theorem 2.4.1 and Theorem 2.4.2, we can argue that if $k = \beta_0 n$ for some $\beta_0 < 1/2$, then a randomly picked sequence of lattices is good for covering, packing and AWGN coding, together with the property that its dual is good for packing and covering with probability going to 1 as $n \to \infty$. We would like to remark that to prove packing goodness of random Construction-A lattices, condition (D2) is not necessary, and the parameters k, q can be chosen so as to satisfy the hypotheses of Theorem 2.4.1 and Theorem 2.4.2. It was also shown in [69, Appendix B] (also see [28]) that if the coarse lattices are good for covering and AWGN channel coding, then as long as $n/q_1 \to 0$ as $n \to \infty$, the probability that a uniformly chosen sequence of fine lattices is good for AWGN channel coding tends to 1 as $n \to \infty$. Furthermore, it was shown in [49] that the fine lattices are also good for covering and MSE quantization. This completes the proof. □

Chapter 3

Perfectly Secure Bidirectional Relaying

3.1 Introduction

We now introduce the first problem that is addressed in this thesis. Consider a network having three nodes, denoted by A, B and R, as shown in Fig. 3.1. The nodes A and B, henceforth called the user nodes, wish to exchange information with each other. However, there is no direct link between the user nodes, and there are only links between the user nodes and R. The node R acts as a bidirectional relay between A and B, and facilitates communication between them. All nodes are assumed to operate in half-duplex mode (they cannot transmit and receive simultaneously), and all links between nodes are wireless (unit channel gain) additive white Gaussian noise (AWGN) channels. Bidirectional relaying in such settings has been studied extensively in the recent literature [3, 69, 75, 107, 116].

We use the compute-and-forward framework proposed in [69, 107] for bidirectional relaying, and we briefly describe a binary version for completeness and clarity. Suppose that A and B possess bits X and Y, respectively. We will assume that X and Y are generated independently and uniformly at random. The goal in bidirectional relaying is to transmit X to B and Y to A through R. To achieve this goal, a compute-and-forward protocol takes place in two phases as shown in Fig. 3.2:

(1) the (Gaussian) multiple access phase or the MAC phase, where the user nodes simultaneously transmit to the relay, and

(2) the broadcast phase, where the relay transmits to the user nodes.

In the MAC phase, the user nodes A and B independently modulate their bits X and Y into real-valued symbols U and V, respectively. The relay receives an instance of a random variable W, that can be modeled as

$$W = U + V + Z, \tag{3.1}$$

where it is assumed that the links A → R and B → R have unit gain, Z denotes additive white Gaussian noise independent of U and V, and communication is assumed to be synchronized. Using W, the relay computes the XOR of the two message bits, i.e., $X \oplus Y$, and in the broadcast phase, encodes it into a real symbol which is transmitted to the two users over a broadcast channel. Note that A and B can recover Y and X, respectively, from $X \oplus Y$.

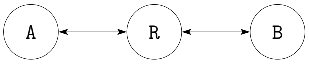


Figure 3.1: Bidirectional relay.

In the compute-and-forward bidirectional relaying problem described above, we study the scenario where an additional secrecy constraint is imposed on the relay R. Specifically, we require that, in the MAC phase, the relay remain ignorant of the individual bits X and Y, while still being able to compute the XOR $X \oplus Y$ reliably. The relay is assumed to be "honest-but-curious": it behaves like a passive eavesdropper, but otherwise helps in the exchange of messages. We study the problem under two secrecy constraints: perfect secrecy, which we describe next, and strong secrecy, which we describe further below. *Perfect secrecy* refers to the requirement that the relay be fully ignorant of the individual bits, i.e., that the random variables $U + V$, X, and Y be pairwise independent. More generally, the user nodes encode the messages X and Y into n-dimensional real vectors

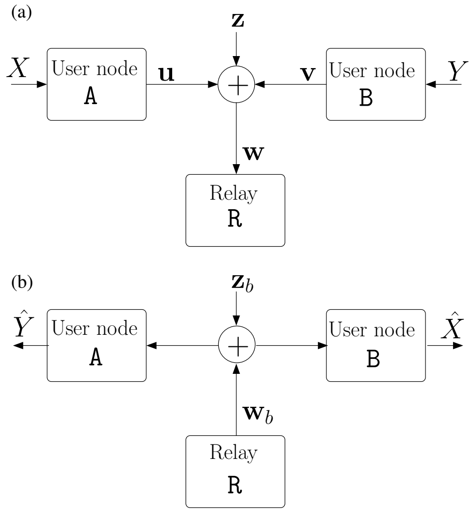


Figure 3.2: Bidirectional relaying: (a) MAC phase, (b) Broadcast phase.

U and **V** respectively, and we require $\mathbf{U}+\mathbf{V}$ to be statistically independent of each individual message. The problem of secure bidirectional relaying in the presence of an untrusted relay under a perfect secrecy constraint has not been studied prior to this work, and this is a major contribution of this thesis.

We propose a coding scheme for secure bidirectional relaying that uses a pair of nested lattices $(\Lambda^{(n)}, \Lambda_0^{(n)})$, with $\Lambda_0^{(n)} \subset \Lambda^{(n)}$. In our scheme, the messages are mapped to the cosets of the *coarse lattice* $\Lambda_0^{(n)}$ in the *fine lattice* $\Lambda^{(n)}$. Given a message (say, the jth coset, Λ_j) at the user node, the output of the encoder is a random point chosen from that coset according to a distribution p_j. This distribution is constructed using a well-chosen density function f on $\mathbb{R}^n$. Specifically, p_j is obtained by sampling and normalizing f over Λ_j. We will show that if the characteristic function of f is supported within the fundamental Voronoi region of the Fourier dual of $\Lambda_0^{(n)}$, then it is possible to achieve perfect secrecy. Our coding scheme for security is explicit, in that given *any* pair of nested lattices, we precisely specify the distributions p_j that must be used to obtain independence between $\mathbf{U}+\mathbf{V}$ and the individual messages. We then study the average transmit power and achievable rates for reliable and secure communication. We will show that a transmission rate of $\left[\frac{1}{2}\log_2 \frac{\mathcal{P}}{\sigma^2} - \log_2 2e\right]^+$ is achievable with perfect secrecy, where $[x]^+$ denotes $\max\{x, 0\}$. However, to achieve these rates, we assume that the nested lattices satisfy certain “goodness” properties. We will also discuss some constructions of “good” lattices that permit low decoding complexity in Chapters 6 and 7.

In the next chapter, we will relax the secrecy constraint, and only demand that the mutual information between $\mathbf{U}+\mathbf{V}$ and the individual messages be arbitrarily small for large block lengths, a requirement that is referred to as *strong secrecy* [65]. We again use a nested-lattice coding scheme, but now the distributions p_j are obtained by sampling and normalizing a Gaussian function, instead of a density having a compactly supported characteristic function. The idea of using probability mass functions (pmfs) obtained by sampling Gaussians was used [58] in the context of the Gaussian wiretap channel, and we will make use of the techniques developed there. Using this scheme, we will show that a rate of $\left[\frac{1}{2}\log_2\left(\frac{1}{2}+\frac{\mathcal{P}}{\sigma^2}\right)-\frac{1}{2}\log_2 2e\right]^+$ is achievable.

We show that our schemes can achieve secrecy even in the absence of noise, and that the addition of noise cannot leak any extra information to the relay. This allows us to develop the solution in two parts: first, we give coding schemes based on nested lattices that achieve secrecy over a noiseless channel. Then, we require the lattices to satisfy certain additional “goodness” properties in order to have reliable decoding in the presence of noise. The signal (codeword) transmitted by each user acts as a jamming signal for the other user’s message, and this helps achieve secrecy. In our scheme, the channel noise is not used to increase confidentiality, unlike the Gaussian wiretap channel [58] where an increase in the noise variance on the eavesdropper’s link can be used to achieve higher transmission rates. It may be possible to harness the additive noise in the MAC phase to obtain higher achievable rates, but we do not pursue this in the present work. However, our approach does offer an advantage: since our scheme guarantees secrecy in the absence of noise, the security properties continue to hold even when channel noise is present, and this is true *irrespective* of the noise distribution. Indeed, our scheme provides secrecy even if the channel noise follows an unknown probability distribution, a property that is in general not satisfied by coding schemes for wiretap channels. We only require the noise to be additive and independent of the transmitted codewords.

It is worth emphasizing the basic idea behind the construction of encoders in our coding schemes. Given a pair of nested lattices, the user nodes send points from the fine lattice in the nested lattice pair according to a pmf obtained by sampling a well-chosen density function at the fine lattice points. The choice of the density function determines the level of security that is achievable.

3.1.1 Related Work

In prior work, the problem of secure bidirectional relaying in the presence of an untrusted relay was studied by He and Yener in [40], who showed that the mutual information rate, defined to be $\frac{1}{n}\mathcal{I}(X;\mathbf{U}+\mathbf{V})=\frac{1}{n}\mathcal{I}(Y;\mathbf{U}+\mathbf{V})$ goes to zero for large blocklengths n. They later studied the problem under a strong secrecy constraint in [41], and gave a scheme based on nested lattice codes and universal hash functions. Using probabilistic arguments, they showed the existence of linear hash functions for randomization at the encoders that achieve strong secrecy. In both scenarios, they showed that a rate of $\left[\frac{1}{2}\log_2\left(\frac{1}{2}+\frac{\mathcal{P}}{\sigma^2}\right)-1\right]^+$ is achievable. The achievable rates guaranteed by our strongly secure scheme is slightly lower than that obtained in [41]. However, our scheme avoids the use of hash functions, and given a pair of nested lattices that satisfy certain "goodness" properties[1], we give an explicit probability distribution for randomization at the encoders that can be used to obtain strong secrecy.

More recently, Karpuk and Chorti [47] have given a perfectly secure coding scheme with QAM signaling. Their code satisfies a maximum power constraint, and they have derived bounds on achievable rates. However, under an average power constraint, our scheme can provide better achievable rates. Moreover, their scheme requires stronger assumptions on the sizes of the signal sets of terminals A and B. Furthermore, their scheme can guarantee security to only one of the two terminals, i.e., either $X \perp\!\!\!\perp U+V$ or $Y \perp\!\!\!\perp U+V$, but not both.

As remarked in Chapter 1, the idea of using nested lattice codes for secure communication is not new. They have been proposed for secure communication in other scenarios,

[1] Unfortunately, there are no known explicit constructions of lattices that satisfy these properties, but only existence results based on probabilistic arguments.

particularly the Gaussian wiretap channel (see e.g., [6, 58, 72]). They have also been used in interference networks [2], and for secret key generation using correlated Gaussian sources [71, 57].

3.2 Description of the Problem

The general set-up is as follows: two user nodes, denoted by A and B, possess messages taking values independently and uniformly in a finite set. For the purposes of computation at the relay, the messages are mapped into random variables X and Y taking values in a finite Abelian group $\mathbb{G}^{(n)}$, where the choice of $\mathbb{G}^{(n)}$ is left to the system designer. The mapping is such that the random variables X and Y remain uniformly distributed over $\mathbb{G}^{(n)}$, and we will see later that this distribution helps in achieving secrecy. The addition operation in the group $\mathbb{G}^{(n)}$ is denoted $\oplus$. The encoder at node A maps the given message X into a random n-dimensional real vector $\mathbf{U}$. In a similar fashion, the encoder at B maps the message Y to a random vector $\mathbf{V}$. The user nodes transmit their respective vectors to the relay simultaneously, and at the end of the MAC phase, the relay obtains

$$\mathbf{W} = \mathbf{U} + \mathbf{V} + \mathbf{Z}, \tag{3.2}$$

where $\mathbf{Z}$ is a Gaussian random vector with zero mean and covariance matrix $\sigma^2 \mathsf{I}_n$, where $+$ denotes componentwise real addition. The coding scheme at each user node must ensure that the relay can recover $X \oplus Y$ reliably from $\mathbf{W}$, and *perfect secrecy:* The mutual information between $\mathbf{W}$ and each individual message is exactly zero[2], i.e., $\mathcal{I}(\mathbf{W}; X) = \mathcal{I}(\mathbf{W}; Y) = 0$.

We in fact impose a slightly stronger security criterion than the one mentioned above. Even in the absence of noise, the mutual information between $\mathbf{U} + \mathbf{V}$ and each individual message must be exactly zero. Since the additive noise is independent of everything else, $X \to \mathbf{U}+\mathbf{V} \to \mathbf{U}+\mathbf{V}+\mathbf{Z}$ forms a Markov chain, and using the data processing inequality, $\mathcal{I}(X; \mathbf{U}+\mathbf{V}+\mathbf{Z}) \leq \mathcal{I}(X; \mathbf{U}+\mathbf{V})$. Likewise, $\mathcal{I}(Y; \mathbf{U}+\mathbf{V}+\mathbf{Z}) \leq \mathcal{I}(Y; \mathbf{U}+\mathbf{V})$. Therefore,

[2] Equivalently, we want $\mathbf{W} \perp\!\!\!\perp X$ and $\mathbf{W} \perp\!\!\!\perp Y$.

any scheme that achieves perfect secrecy in the absence of noise will also achieve perfect secrecy in a noisy channel.

The messages must also be protected from corruption by the additive noise in the multiple access phase. Since the messages are uniformly distributed over $\mathbb{G}^{(n)}$, $\frac{1}{n}\log_2|\mathbb{G}^{(n)}|$ gives the average number of bits of information sent to the relay by each user node in one channel use in the MAC phase. Our aim will be to ensure secure computation of $X \oplus Y$ at the highest possible rate (which we define to be $\frac{1}{n}\log_2|\mathbb{G}^{(n)}|$) for a given power constraint at the user nodes. To formalize these notions, we have the following definition:

Definition 1. *For a positive integer n, an $(n, M^{(n)})$ code for the MAC phase of the bidirectional relay channel with user nodes* `A`*,* `B` *and relay* `R` *consists of the following:*

1. ***Messages:*** *Nodes* `A` *and* `B` *possess messages X and Y, respectively, drawn independently and uniformly from a finite Abelian group $\mathbb{G}^{(n)}$ with $M^{(n)} = |\mathbb{G}^{(n)}|$ elements.*

2. ***Codebook:*** *The codebook, denoted by $\mathcal{C}$, is a discrete subset of $\mathbb{R}^n$, not necessarily finite. The elements of $\mathcal{C}$ are called codewords. The codebook consists of all those vectors that are allowed to be transmitted by the user nodes to the relay.*

3. ***Encoders:*** *The encoder at each node is a randomized mapping from $\mathbb{G}^{(n)}$ to $\mathbb{R}^n$, specified by the distributions $p_{\mathbf{U}|X}(\mathbf{u}|x) = \Pr[\mathbf{U} = \mathbf{u}|X = x]$ and $p_{\mathbf{V}|Y}(\mathbf{v}|y) = \Pr[\mathbf{V} = \mathbf{v}|Y = y]$ for all $\mathbf{u}, \mathbf{v} \in \mathcal{C}$ and $x, y \in \mathbb{G}^{(n)}$. At node* `A`*, given a message $x \in \mathbb{G}^{(n)}$ as input, the encoder outputs a codeword $\mathbf{u} \in \mathcal{C}$ at random, according to $p_{\mathbf{U}|X}(\mathbf{u}|x)$. Similarly, at node* `B`*, with y as input, the encoder outputs $\mathbf{v} \in \mathcal{C}$ according to $p_{\mathbf{V}|Y}(\mathbf{v}|y)$. The messages x and y are encoded independently.*

   The rate *of the code is defined to be*

$$R^{(n)} = \frac{\log_2 M^{(n)}}{n}. \tag{3.3}$$

   The code has an average transmit power per dimension *defined as*

$$P^{(n)} = \frac{1}{n}\mathbb{E}\|\mathbf{U}\|^2 = \frac{1}{n}\mathbb{E}\|\mathbf{V}\|^2. \tag{3.4}$$

4. ***Decoder:*** *The relay* R *receives a vector* $\mathbf{W} \in \mathbb{R}^n$ *as given in (3.2). The decoder,* $\mathcal{D}^{(n)} : \mathbb{R}^n \to \mathbb{G}^{(n)}$ *maps the received vector to an element of the set of messages. The average probability of error of the decoder is defined as*

$$\xi^{(n)} \triangleq \mathbb{E}\Big[\Pr[\mathcal{D}^{(n)}(\mathbf{W}) \neq X \oplus Y]\Big],$$

where $\mathbb{E}$ *denotes expectation over the messages,* X, Y*, and over the encoders (*$\mathbf{U}, \mathbf{V}$ *given* X, Y*).*

3.3 Perfect Secrecy

Recall that we have the following requirements for secure compute-and-forward:

(S1) $(\mathbf{U}, X) \perp\!\!\!\perp (\mathbf{V}, Y)$.

(S2) $(\mathbf{U}+\mathbf{V}) \perp\!\!\!\perp X$ and $(\mathbf{U}+\mathbf{V}) \perp\!\!\!\perp Y$.

(S3) $\mathbf{U}+\mathbf{V}$ almost surely determines $X \oplus Y$.

If conditions (S1)–(S3) are satisfied, the relay has no means of finding the individual messages. Property (S3) ensures that the relay can decode $X \oplus Y$, which can then be encoded/modulated for further transmission over the broadcast channel. On reception of the broadcast message, since user A (resp. B) knows X (resp. Y), it can recover Y (resp. X).

If the relay only had access to $X \oplus Y$ instead of $\mathbf{U}+\mathbf{V}$, the problem of secure communication would have been trivial due to the uniformity and independence of X and Y. However, the relay receives the real sum of $\mathbf{U}$ and $\mathbf{V}$, which makes the problem harder. For example, suppose that $n = 1$, and $\mathbb{G}^{(1)} = \mathbb{Z}_2$, the group of integers modulo 2. Consider the coding scheme $\mathbf{U} = X$, and $\mathbf{V} = Y$. Then, in the absence of noise, whenever $\mathbf{U}+\mathbf{V} = 0$ or $\mathbf{U}+\mathbf{V} = 2$, the relay can determine both X and Y.

The performance of a coding scheme is generally evaluated in terms of the average transmit power, and the transmission rate. To make these notions formal, we define achievable power-rate pairs as follows.

Definition 2. *A power-rate pair* $(\mathcal{P}, \mathcal{R})$ *is* achievable with perfect secrecy *if there exists a sequence of* $(n, M^{(n)})$ *codes such that for every* $\delta > 0$*,*

- *conditions (S1)–(S3) are satisfied for all* n*,*

and for all sufficiently large n*,*

- *the transmission rate,* $R^{(n)}$*, is greater than* $\mathcal{R} - \delta$*;*
- *the average transmit power per dimension,* $P^{(n)}$*, is less than* $\mathcal{P} + \delta$*; and*
- *the average probability of decoding error,* $\xi^{(n)}$*, is less than* δ*.*

The objective of the next couple of sections will be to prove the following result.

Theorem 3.3.1. *A power-rate pair of*

$$\left(\mathcal{P}, \left[\frac{1}{2}\log_2\frac{\mathcal{P}}{\sigma^2} - \log_2(2e)\right]^+\right)$$

is achievable with perfect secrecy in the MAC phase of the bidirectional relay.

3.4 Perfect secrecy: The Noiseless Setting

To get a clear picture as to how secure communication can be achieved, we first describe the binary case. The messages X and Y are chosen independently and uniformly at random from $\{0, 1\}$, or equivalently, the set of integers modulo-2 ($\mathbb{G} = \mathbb{Z}_2$). They are modulated to U and V, respectively, which take values in $\mathbb{R}$. Studying the one-dimensional case will give us the intuition needed to tackle the general case, and we will see that the techniques developed here extend quite naturally to the n-dimensional setting.

We will give examples of distributions on U and V that permit secure computation defined by properties (S1)–(S3). This is somewhat surprising since we cannot have non-degenerate real-valued random variables U, V that satisfy $(U+V) \perp\!\!\!\perp U$ and $(U+V) \perp\!\!\!\perp V$, as shown in the following proposition:

Proposition 3.4.1. *Let U and V be independent real-valued random variables, and let $+$ denote addition over $\mathbb{R}$. Then, we have $(U+V) \perp\!\!\!\perp U$ and $(U+V) \perp\!\!\!\perp V$ iff U and V are constant a.s. (i.e., there exist $a, b \in \mathbb{R}$ such that $\Pr[U=a] = \Pr[V=b] = 1$).*

Proof. The "if" part is trivial, so let us prove the "only if" part. Let $W = U+V$, so that by assumption, U, V and W are pairwise independent. Let φ_U, φ_V and φ_W denote the characteristic functions of U, V and W, respectively. In particular, $\varphi_W = \varphi_U \varphi_V$. From $U = W - V$, we also have that $\varphi_U = \varphi_W \overline{\varphi_V}$, where $\overline{\varphi_V}$ denotes the complex conjugate of φ_V. Putting the two equalities together, we obtain $\varphi_U = \varphi_U |\varphi_V|^2$. To be precise, $\varphi_U(t) = \varphi_U(t)|\varphi_V(t)|^2$ for all $t \in \mathbb{R}$.

Now, characteristic functions are continuous and take the value 1 at $t = 0$. Hence, φ_U is non-zero within the interval $[-\delta, \delta]$ for some $\delta > 0$. Thus, $|\varphi_V(t)| = 1$ for all $t \in [-\delta, \delta]$. By a basic property of characteristic functions (see Lemma 4 of Section XV.1 in [33]), this implies that there exists $b \in \mathbb{R}$ such that $\varphi_V(t) = e^{ibt}$ for all $t \in \mathbb{R}$, thus proving that $V = b$ with probability 1.

A similar argument using $V = W - U$ shows that U is also constant with probability 1. □

3.4.1 Secure Computation of XOR at the Relay

In this section, X and Y are independent and identically distributed (iid) uniform binary random variables (rvs), and $X \oplus Y$ denotes their modulo-2 sum (XOR). We describe a construction of integer-valued rvs U and V satisfying the properties (S1)–(S3).

Conditions on PMFs and Characteristic Functions

We first derive conditions under which integer-valued rvs U and V can satisfy the desired properties. We introduce some notation: for $k \in \mathbb{Z}$, let $p_U(k) = \Pr[U = k]$, $p_V(k) = \Pr[V = k]$, and for $a \in \{0, 1\}$, let $p_{U|a}(k) = \Pr[U = k \mid X = a]$, $p_{V|a}(k) = \Pr[V = k \mid Y = a]$. Thus, $p_U = (1/2)(p_{U|0} + p_{U|1})$ and $p_V = (1/2)(p_{V|0} + p_{V|1})$.

Property (S1) is equivalent to requiring that the joint probability mass function (pmf) of (U, V, X, Y) be expressible as

$$p_{UVXY}(k, l, a, b) = (1/2)(1/2)p_{U|a}(k)p_{V|b}(l) \tag{3.5}$$

for $k, l \in \mathbb{Z}$ and $a, b \in \{0, 1\}$. Next, we look at (S3). Without the requirement that $U + V \perp\!\!\!\perp X$ and $U + V \perp\!\!\!\perp Y$, it is trivial to define U and V such that (S3) is satisfied: for example, take $U = X$ and $V = Y$. More generally, property (S3) is satisfied by any U, V such that

$$\begin{aligned} p_{U|0}(k) = p_{V|0}(k) = 0 \ \text{ for all odd } k \in \mathbb{Z}, \\ p_{U|1}(k) = p_{V|1}(k) = 0 \ \text{ for all even } k \in \mathbb{Z}. \end{aligned} \tag{3.6}$$

Finally, we turn our attention to (S2). We want $(U+V) \perp\!\!\!\perp X$ and $(U+V) \perp\!\!\!\perp Y$. Let us define, for $k \in \mathbb{Z}$, $p_{U+V}(k) = \Pr[U+V = k]$, and for $a \in \{0, 1\}$, $p_{U+V|X=a}(k) = \Pr[U+V = k \mid X = a]$ and $p_{U+V|Y=a}(k) = \Pr[U + V = k \mid Y = a]$. Assuming $(U, X) \perp\!\!\!\perp (V, Y)$, we have $p_{U+V} = p_U * p_V$, $p_{U+V|X=a} = p_{U|a} * p_V$, and $p_{U+V|Y=a} = p_U * p_{V|a}$, where $*$ denotes the convolution operation. Thus, when $(U, X) \perp\!\!\!\perp (V, Y)$, (S2) holds iff

$$p_U * p_V = p_{U|a} * p_V = p_U * p_{V|a} \ \text{ for } a \in \{0, 1\}. \tag{3.7}$$

It helps to view this in the Fourier domain. Let φ_U, φ_V, $\varphi_{U|a}$ etc. denote the respective characteristic functions of the pmfs p_U, p_V, $p_{U|a}$ etc. — for example, $\varphi_{U|a}(t) = \sum_{k\in\mathbb{Z}} p_{U|a}(k)e^{ikt}$. Then, (3.7) is equivalent to

$$\varphi_U\varphi_V = \varphi_{U|a}\varphi_V = \varphi_U\varphi_{V|a} \ \text{ for } a \in \{0, 1\}. \tag{3.8}$$

Note that $\varphi_U = (1/2)(\varphi_{U|0} + \varphi_{U|1})$ and $\varphi_V = (1/2)(\varphi_{V|0} + \varphi_{V|1})$. Hence, (3.8) should be viewed as a requirement on the conditional pmfs $p_{U|a}$ and $p_{V|a}$, $a \in \{0, 1\}$.

In summary, we have the following lemma.

Lemma 3.4.2. *Suppose that the conditional pmfs $p_{U|a}$ and $p_{V|a}$, $a \in \{0, 1\}$, satisfy (3.6) and (3.8). Then, the rvs U, V, X, Y with joint pmf given by (3.5) have properties (S1)–(S3).*

The observations made up to this point also allow us to prove the following negative result.[3]

Proposition 3.4.3. *Properties (S1)–(S3) cannot be satisfied by integer-valued rvs U, V that are finitely supported.*

Proof. Suppose that U and V are finitely supported $\mathbb{Z}$-valued rvs. Then, $\varphi_U(t)$ and $\varphi_V(t)$ are finite linear combinations of some exponentials $e^{ik_1 t}, \ldots, e^{ik_n t}$. Equivalently, the real and imaginary parts of φ_U and φ_V are trigonometric polynomials. Thus, either φ_U (resp. φ_V) is identically zero, or it has a discrete set of zeros. The former is impossible as $\varphi_U(0) = \varphi_V(0) = 1$. Now, suppose that (S1) and (S2) are satisfied, which means that (3.8) must hold. The equality $\varphi_U \varphi_V = \varphi_U \varphi_{V|a}$ in (3.8) implies that $\varphi_{V|a}(t) = \varphi_V(t)$ for all t such that $\varphi_U(t) \neq 0$. But since $\varphi_U(t)$ has a discrete set of zeros, continuity of characteristic functions in fact implies that $\varphi_{V|a}(t) = \varphi_V(t)$ for all t. An analogous argument shows that $\varphi_{U|a}(t) = \varphi_U(t)$ for all t. Hence, $U \perp\!\!\!\perp X$ and $V \perp\!\!\!\perp Y$. From this, and (S1), we obtain that $U + V \perp\!\!\!\perp X \oplus Y$, thus precluding (S3). □

Practical communication systems generally have a maximum power constraint, which means that we would like to have U, V be finitely supported. But from Proposition 3.4.3, we see that it is not possible to have finitely supported U, V that permit secure computation of the XOR at the relay. Therefore, in order to ensure secure computation, we will have to relax the power constraint to an *average power constraint* on the user nodes. This means that we require finite-variance, integer-valued random variables U, V, with infinite support, that satisfy properties (S1)–(S3), or equivalently, the hypotheses of Lemma 3.4.2.

We now give a construction of U, V that satisfy the hypotheses of Lemma 3.4.2. We will choose a density function whose characteristic function is compactly supported. The random variables U and V are chosen according to a distribution obtained by sampling and appropriately normalizing this density function. To study this in more detail, we rely upon methods and results from Fourier analysis. The key tool we need is the Poisson

[3] In fact, a stronger negative result can be shown — see Proposition 3.4.8.

summation formula, which we briefly recall here. Our description is based largely on Section XIX.5 in [33].

3.4.2 The Poisson Summation Formula

Fix a positive integer n, and let Λ be a full-rank lattice in $\mathbb{R}^n$. Recall from Chapter 2 that $\widehat{\Lambda}$ denotes the Fourier dual of Λ.

Let $\psi : \mathbb{R}^n \to \mathbb{C}$ be the characteristic function of a $\mathbb{R}^n$-valued random variable, such that $\int_{\mathbb{R}^n} |\psi(\mathbf{t})|\, d\mathbf{t} < \infty$. In particular, ψ is continuous and $\psi(\mathbf{0}) = 1$. Since ψ is absolutely integrable, the random variable has a continuous density $f : \mathbb{R}^n \to \mathbb{R}^+$. The Poisson summation formula can be expressed as follows: for any $\mathbf{s} \in \mathbb{R}^n$, we have for all $\boldsymbol{\zeta} \in \mathbb{R}^n$,

$$\sum_{\mathbf{x}\in\widehat{\Lambda}} \psi(\boldsymbol{\zeta}+\mathbf{x})\, e^{-i\langle \mathbf{x},\, \mathbf{s}\rangle} = (\det\Lambda) \sum_{\mathbf{k}\in\Lambda} f(\mathbf{k}+\mathbf{s})\, e^{i\langle \mathbf{k}+\mathbf{s},\, \boldsymbol{\zeta}\rangle}, \tag{3.9}$$

provided that the series on the left converges to a continuous function $\Psi(\boldsymbol{\zeta})$. It should be pointed out that texts in Fourier analysis typically state the Poisson summation formula for an arbitrary L^1 function f, and would then require that f and ψ decay sufficiently quickly — see e.g., [87, Chapter VII, Corollary 2.6] or [5, Eq. (17.1.2)] — for (3.9) to hold. However, as argued by Feller in proving the formula in the one-dimensional setting [33, Chapter XIX, equation (5.9)], in the special case of a non-negative L^1 function f, it is sufficient to assume that the left-hand side (LHS) of (3.9) converges to a continuous function $\Psi(\boldsymbol{\zeta})$.

Note that $\Psi(\mathbf{0}) = (\det\Lambda) \sum_{\mathbf{k}\in\Lambda} f(\mathbf{k}+\mathbf{s})$, which is a non-negative quantity. If $\Psi(\mathbf{0}) \neq 0$, then dividing both sides of (3.9) by $\Psi(\mathbf{0})$ yields the important fact that $\Psi(\boldsymbol{\zeta})/\Psi(\mathbf{0})$ is the characteristic function of a discrete random variable supported within the set $\Lambda+\mathbf{s} \triangleq \{\mathbf{k}+\mathbf{s} : \mathbf{k} \in \Lambda\}$, the probability mass at the point $\mathbf{k}+\mathbf{s}$ being equal to $f(\mathbf{k}+\mathbf{s})/\sum_{\boldsymbol{\ell}\in\Lambda} f(\boldsymbol{\ell}+\mathbf{s})$.

A special case of interest is when ψ is compactly supported; specifically, it is supported within the fundamental Voronoi region of $\widehat{\Lambda}$: $\psi(\mathbf{t}) = 0$ for all $\mathbf{t} \notin \mathcal{V}(\widehat{\Lambda})$. In this case, we can readily show that the series on the LHS of (3.9) converges to a continuous function Ψ. Indeed, if we define $\widetilde{\psi}(\mathbf{t}) \triangleq \psi(\mathbf{t})e^{-i\langle \mathbf{t},\mathbf{s}\rangle}$, then the series on the LHS of (3.9) may be written as $\Psi(\boldsymbol{\zeta}) \triangleq e^{i\langle\boldsymbol{\zeta},\mathbf{s}\rangle}\widetilde{\Psi}(\boldsymbol{\zeta})$, where

$$\widetilde{\Psi}(\boldsymbol{\zeta}) \triangleq \sum_{\mathbf{n}\in\widehat{\Lambda}} \widetilde{\psi}(\boldsymbol{\zeta}+\mathbf{n}).$$

Now, recall that ψ, being a characteristic function, is continuous on $\mathbb{R}^n$; hence, so is $\widetilde{\psi}$. Also, by assumption, ψ is supported within $\mathcal{V}(\widehat{\Lambda})$; hence, so is $\widetilde{\psi}$. In particular, by continuity, $\widetilde{\psi}$ must be 0 on the boundary of $\mathcal{V}(\widehat{\Lambda})$; therefore, the supports of $\widetilde{\psi}(\cdot)$ and $\widetilde{\psi}(\cdot+\mathbf{n})$ do not intersect for any non-zero $\mathbf{n} \in \widehat{\Lambda}$. From this, we infer that $\widetilde{\Psi}$, which is formed by the superposition of continuous functions with disjoint supports, must be continuous. Hence, we can conclude that $\Psi(\boldsymbol{\zeta}) = e^{i\langle\boldsymbol{\zeta},\mathbf{s}\rangle}\widetilde{\Psi}(\boldsymbol{\zeta})$ is a continuous function.

Moreover, it is clear that $\Psi(\mathbf{0}) = \psi(\mathbf{0})$, and since ψ is a characteristic function, $\psi(\mathbf{0}) = 1$. As explained above, this shows that Ψ is the characteristic function of a discrete random vector supported within $\Lambda + \mathbf{s}$. In fact, by plugging in $\boldsymbol{\zeta} = \mathbf{0}$ in (3.9) we obtain that $\Psi(\mathbf{0}) = (\det\Lambda)\sum_{\mathbf{k}\in\Lambda} f(\mathbf{k}+\mathbf{s})$, which shows that $\sum_{\mathbf{k}\in\Lambda} f(\mathbf{k}+\mathbf{s}) = 1/(\det\Lambda)$. For future reference, we summarize this in the form of a proposition.

Proposition 3.4.4. *Let Λ be a full-rank lattice in $\mathbb{R}^n$. Let $\psi : \mathbb{R}^n \to \mathbb{C}$ be a characteristic function such that $\psi(\mathbf{t}) = 0$ for all $\mathbf{t} \notin \mathcal{V}(\widehat{\Lambda})$, and let $f : \mathbb{R}^n \to \mathbb{R}^+$ be the corresponding probability density function. Then, for any $\mathbf{s} \in \mathbb{R}^n$, the function $\Psi : \mathbb{R}^n \to \mathbb{C}$ defined by*

$$\Psi(\boldsymbol{\zeta}) = \sum_{\mathbf{n}\in\widehat{\Lambda}} \psi(\boldsymbol{\zeta}+\mathbf{n})\, e^{-i\langle\mathbf{n},\,\mathbf{s}\rangle}$$

is the characteristic function of a random vector supported within the set $\Lambda + \mathbf{s} \triangleq \{\mathbf{k}+\mathbf{s} : \mathbf{k} \in \Lambda\}$. The probability mass at the point $\mathbf{k}+\mathbf{s}$ is equal to $(\det\Lambda)\, f(\mathbf{k}+\mathbf{s})$.

It should be noted that compactly supported characteristic functions do indeed exist — see e.g., [33, Section XV.2, Table 1], [27], [80]. We also give an explicit construction in Example 1 in Section 3.4.3.

Applying Proposition 3.4.4 to the one-dimensional lattice $T\mathbb{Z} = \{kT : k \in \mathbb{Z}\}$, with $T > 0$, we obtain the corollary below.

Corollary 3.4.5. *Let ψ be a characteristic function of a real-valued random variable*

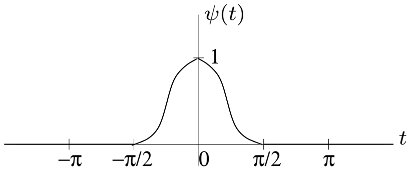


Figure 3.3: A generic characteristic function supported on $[-\pi/2, \pi/2]$.

such that $\psi(t) = 0$ whenever $|t| \geq \pi/T$ for some $T > 0$, and let f be the corresponding probability density function. Then, for any $s \in \mathbb{R}$, the function $\Psi : \mathbb{R} \to \mathbb{C}$ defined by

$$\Psi(\zeta) = \sum_{n=-\infty}^{\infty} \psi(\zeta + 2n\pi/T)\, e^{-is(2n\pi/T)}$$

is the characteristic function of a discrete random variable supported within the set $\{kT + s : k \in \mathbb{Z}\}$. The probability mass at the point $kT + s$ is equal to $Tf(kT + s)$.

This corollary plays a central role in the construction described next.

3.4.3 Construction of $\mathbb{Z}$-Valued RVs Satisfying (S1)–(S3)

We now describe the construction of integer-valued rvs that satisfy (S1)–(S3). Let ψ be a characteristic function (of a continuous rv X) with the properties that

(C1) $\psi(t) = 0$ for $|t| \geq \pi/2$, and

(C2) $\psi(t)$ is real and non-negative for all $t \in \mathbb{R}$.[4]

A generic such ψ is depicted in Fig. 3.3; we give a specific example a little later in this section. Since ψ is real-valued, it must be an even function: $\psi(-t) = \psi(t)$ for all $t \in \mathbb{R}$. Also, $\psi(0) = 1$. Moreover, since ψ is integrable over $\mathbb{R}$, by the Fourier inversion formula, the rv X has a continuous density f. Note that Corollary 3.4.5 holds for $T \leq 2$.

Let φ be the periodic function with period 2π that agrees with ψ on $[-\pi, \pi]$, as depicted in Fig. 3.4. Note that $\varphi(\zeta) = \sum_{n=-\infty}^{\infty} \psi(\zeta + 2\pi n)$. Thus, applying Corollary 3.4.5 with

[4] There is no loss of generality in imposing this requirement. Suppose that an rv X has characteristic function ψ, which is complex-valued in general. Let X_1, X_2 be iid rvs with the same distribution as X. Then, $X_1 - X_2$ has characteristic function $\psi\bar{\psi} = |\psi|^2$.

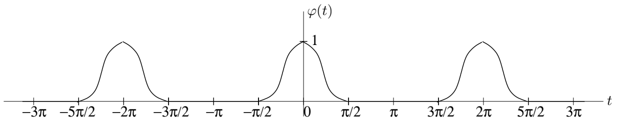


Figure 3.4: Period-2π extension of generic ψ from Fig. 3.3.

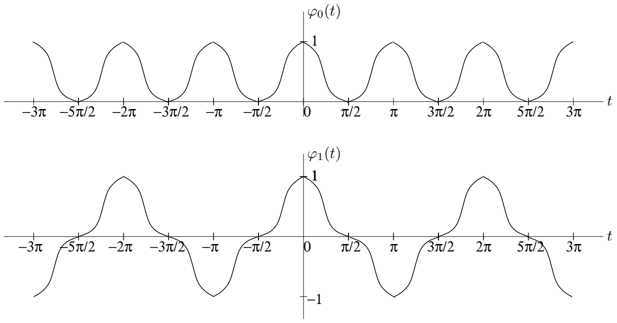


Figure 3.5: The periodic functions φ_0 and φ_1 derived from ψ.

$T = 1$ and $s = 0$, we find that φ is the characteristic function of an integer-valued rv, with pmf given by

$$p(k) = f(k) \text{ for all } k \in \mathbb{Z}. \tag{3.10}$$

Next, for $s = 0, 1$, define φ_s as follows: for $\zeta \in \mathbb{R}$,

$$\varphi_s(\zeta) = \sum_{n=-\infty}^{\infty} \psi(\zeta + n\pi) e^{-isn\pi}.$$

It is easily seen that φ_0 is the periodic extension of ψ with period π, i.e., φ_0 is the periodic function with period π that agrees with ψ on $[-\pi/2, \pi/2]$, as depicted at the top of Fig. 3.5 for a generic ψ shown in Fig. 3.3. On the other hand, φ_1 is periodic with period 2π: its graph is obtained from that of φ_0 by reflecting about the ζ-axis every second copy of ψ, as depicted at the bottom of Fig. 3.5.

Applying Corollary 3.4.5 with $T = 2$ and $s \in \{0, 1\}$, we get that φ_0 and φ_1 are characteristic functions of rvs supported within the even and odd integers, respectively.

The pmf corresponding to φ_0 is given by

$$p_0(k) = \begin{cases} 2f(k) & \text{if } k \text{ is an even integer} \\ 0 & \text{otherwise.} \end{cases} \tag{3.11}$$

and that corresponding to φ_1 is

$$p_1(k) = \begin{cases} 2f(k) & \text{if } k \text{ is an odd integer} \\ 0 & \text{otherwise.} \end{cases} \tag{3.12}$$

From (3.10)–(3.12), we have $p(k) = \frac{1}{2}(p_0(k) + p_1(k))$ for all $k \in \mathbb{Z}$.

Finally, note that since $\varphi_0(t)$ and $\varphi_1(t)$ differ from $\varphi(t)$ only when $\varphi(t) = 0$, we have

$$\varphi^2 = \varphi\varphi_0 = \varphi\varphi_1. \tag{3.13}$$

With these facts in hand, we can describe the construction of $\mathbb{Z}$-valued rvs U and V satisfying properties (S1)–(S3). Set $p_{U|0} = p_{V|0} = p_0$ and $p_{U|1} = p_{V|1} = p_1$. This implies that $p_U = p_V = p$, where p is as defined in (3.10). Clearly, (3.6) holds. To verify (3.8), note that, by virtue of (3.13), we have for $a \in \{0, 1\}$,

$$\varphi_U\varphi_V = \varphi^2 = \varphi\varphi_a.$$

But, by construction, $\varphi_U\varphi_{V|a} = \varphi_V\varphi_{U|a} = \varphi\varphi_a$. Therefore, by Lemma 3.4.2, the rvs (U, V, X, Y) with joint pmf given by (3.5) have the properties (S1)–(S3).

Recall from the discussion following Proposition 3.4.3 that we need the rvs U and V to have finite variance. To ensure this, we use the fact [33, pp. 512–513] that a probability distribution F with characteristic function χ has finite variance iff χ is twice differentiable; in this case, $\chi'(0) = i\mu$ and $\chi''(0) = -\mu_2$, where μ and μ_2 are the mean and second moment of F. Thus, the rvs U and V (with pmf p as above) have finite variance iff the characteristic function φ is twice differentiable. In this case, as φ is real, so is $\varphi'(0)$, which implies that U and V have zero mean. Hence, their variances are equal to their second moments, and so, $\mathrm{Var}(U) = \mathrm{Var}(V) = -\varphi''(0)$. By construction, φ is twice differentiable iff ψ is twice differentiable and $\varphi''(0) = \psi''(0)$. We summarize our construction of the rvs U and V in the following theorem.

Theorem 3.4.6. *Let X, Y be iid Bernoulli$(1/2)$ rvs. Suppose that we are given a probability density function $f : \mathbb{R} \to \mathbb{R}^+$ with a non-negative real characteristic function ψ such that $\psi(t) = 0$ for $|t| \geq \pi/2$. Set $p_{U|0} = p_{V|0} = p_0$ and $p_{U|1} = p_{V|1} = p_1$, where p_0 and p_1 are as in (3.11) and 3.12). Then, the resulting $\mathbb{Z}$-valued rvs U and V satisfy properties (S1)–(S3). Additionally, the rvs U and V have finite variance iff ψ is twice differentiable, in which case the variance equals $-\psi''(0)$.*

Based on Theorem 3.4.6, secure computation of XOR at the relay works as follows: the nodes `A` and `B` modulate their bits independently to an integer k, with probability $p_0(k)$ (from (3.11)) if the bit is 0, or with probability $p_1(k)$ (from (3.12)) if the bit is 1. The probability distributions can be chosen such that the modulated symbols have finite average power. The average transmit power is equal to the variance of the modulated random variable, which is $-\psi''(0)$, and a handle on this can be obtained by choosing ψ carefully. The relay receives the sum of the two integers, which is independent of the individual bits X and Y (of `A` and `B` respectively). However, the XOR of the two bits can be recovered at `R` with probability 1. This is done by simply mapping the received integer W to 1, if W is odd, and 0 if W is even. To gain a better understanding of the construction of the rvs, let us see an example.

Example 1. *Consider the density (from [33, Section XV.2, Table 1])*

$$f(x) = \begin{cases} \frac{1}{2\pi} & \text{if } x = 0 \\ \frac{1-\cos x}{\pi x^2} & \text{if } x \neq 0 \end{cases} \tag{3.14}$$

which has characteristic function

$$\widehat{f}(t) = \max\{0, 1 - |t|\} \tag{3.15}$$

The function $\widehat{f}$ is plotted in Fig. 3.6. In particular, $\widehat{f}(t) = 0$ for $|t| \geq 1$.

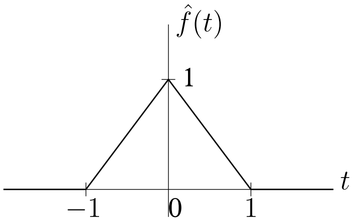


Figure 3.6: $\widehat{f}(t) = \max\{0, 1 - |t|\}$.

The function $\widehat{f}$ is compactly supported but it is not differentiable at 0*. This can be rectified by considering instead $g = \widehat{f} * \widehat{f}$, where $*$ denotes convolution, which can be explicitly computed to be*

$$g(t) = (\widehat{f} * \widehat{f})(t) = \begin{cases} \frac{1}{2}|t|^3 - t^2 + \frac{2}{3} & \text{if } |t| \le 1 \\ \frac{1}{6}(2 - |t|)^3 & \text{if } 1 \le |t| \le 2 \\ 0 & \text{otherwise.} \end{cases} \cdot \tag{3.16}$$

Claim 1. *Define $h(x) \triangleq (3\pi^2/4)\,[f(\pi x/4)]^2$, with f as in (3.14). Then, h is a probability density function whose characteristic function is given by*

$$\psi(t) = \tfrac{3}{2}\, g(\tfrac{4t}{\pi}),$$

where g is as in (3.16). Furthermore, ψ is non-negative with $\psi(t) = 0$ for $|t| \ge \pi/2$, and that ψ is twice differentiable, with $\psi''(0) = -48/\pi^2$.

Using Claim 1, we see that rvs U and V can indeed be constructed as in Theorem 3.4.6 with $\mathrm{var}(U) = \mathrm{var}(V) = 48/\pi^2$.

Proof of Claim 1

Note first that $\widehat{f}$ defined in (3.15) is also a probability density function — it is non-negative and its integral over $(-\infty, \infty)$ is 1. By Fourier inversion, its characteristic function is $2\pi f$. Therefore, $g = \widehat{f} * \widehat{f}$ is a density with characteristic function $4\pi^2 f^2$.

Now, f^2 is integrable since $(\widehat{f})^2$ is integrable (see corollary to Theorem 3 of Section XV.3 of [33]). Hence, $\tilde{h}(x) = f^2(x)/(\int_{-\infty}^{\infty} f^2(y)\,dy)$ is a probability density function. The integral in the denominator can be explicitly evaluated by means of the Plancherel identity:

$$\int_{-\infty}^{\infty} f^2(y)\,dy = \frac{1}{2\pi}\int_{-\infty}^{\infty} [\widehat{f}(t)]^2\,dt = \frac{1}{2\pi}\,g(0) = \frac{1}{3\pi},$$

the last equality following from (3.16). Thus, $\tilde{h}(x) = 3\pi f^2(x)$.

From the fact that $4\pi^2 f^2$ is the characteristic function of g, it follows by Fourier inversion that $\tilde{h}$ has characteristic function given by $\tilde{\psi}(t) = \frac{3}{2}\,g(t)$. Hence, $h(x) = (\pi/4)\tilde{h}(\pi x/4)$ is a density function with characteristic function $\tilde{\psi}(4t/\pi)$, which is precisely $\psi(t)$.

Since $g(t)$ is zero for $|t| > 2$, we have that $\psi(t) = \frac{3}{2}g(4t/\pi)$ is zero for $|t| > \pi/2$. It can also be easily verified that $\psi''(0) = -48/\pi^2$. □

Remark 3.4.7. *It is even possible to construct compactly supported C^∞ characteristic functions. Constructions of such functions are given in [80]. In fact, [80] constructs compactly supported characteristic functions ψ such that the corresponding density functions f are even functions satisfying $\lim_{x\to\infty} x^m f(x) = 0$ for all $m > 0$. This implies that all the absolute moments $\int_{-\infty}^{\infty} |x|^m f(x)\,dx$ exist, and hence, ψ is a C^∞ function (see [33, p. 512]). If such a characteristic function ψ is used in the construction described in Theorem 3.4.6, then the resulting $\mathbb{Z}$-valued rvs U, V will have pmfs $p_U(k), p_V(k)$ whose tails decay faster than any polynomial in k. To be precise, $\lim_{k\to\infty} k^m p_U(k) = \lim_{k\to\infty} k^m p_V(k) = 0$ for any $m > 0$.*

The above remark shows that we can have $\mathbb{Z}$-valued rvs U, V satisfying properties (S1)–(S3), with pmfs decaying faster than any polynomial. However, the rate of decay cannot be much faster than that. Indeed, it is not possible to construct $\mathbb{Z}$-valued rvs with exponentially decaying pmfs that satisfy properties (S1)–(S3). Define a pmf $p(k)$, $k \in \mathbb{Z}$, to be *light-tailed* if there are positive constants C and λ such that $p(k) \le C\lambda^{-|k|}$ for all sufficiently large $|k|$.

Proposition 3.4.8. *Properties (S1)–(S3) cannot be satisfied by integer-valued rvs U, V having light-tailed pmfs.*

Proof.[5] Suppose that U, V are $\mathbb{Z}$-valued rvs satisfying (S1) and (S2). Using $\varphi_U = (1/2)(\varphi_{U|0} + \varphi_{U|1})$ and $\varphi_V = (1/2)(\varphi_{V|0} + \varphi_{V|1})$ in (3.8), we readily obtain

$$\varphi_{U|0}^2 = \varphi_{U|1}^2 \text{ and } \varphi_{V|0}^2 = \varphi_{V|1}^2. \tag{3.17}$$

If U, V have light-tailed pmfs, then $p_{U|a}$ and $p_{V|a}$, $a \in \{0, 1\}$, must also be light-tailed, since $p_{U|a} \leq 2p_U$ and $p_{V|a} \leq 2p_V$. The key observation is that the characteristic function of a light-tailed pmf is real-analytic, i.e., it has a power series expansion $\sum_{n=0}^{\infty} c_n t^n$, with $c_n \in \mathbb{C}$, that is valid for all $t \in \mathbb{R}$ [61, Chapter 7]. Thus, $\varphi_{U|a}$ and $\varphi_{V|a}$, for $a \in \{0, 1\}$, are real-analytic. It follows by comparing power series coefficients, that if functions g and h are real-analytic and $g^2 = h^2$, then either $g = h$ or $g = -h$. Applying this to (3.17), we find that $\varphi_{U|0} = \pm\varphi_{U|1}$, and similarly for V. In fact, since φ_U and φ_V cannot be identically 0, we actually have $\varphi_{U|0} = \varphi_{U|1} = \varphi_U$, and similarly for V. This implies that $U \perp\!\!\!\perp X$ and $V \perp\!\!\!\perp Y$. From this, and (S1), we obtain that $U + V \perp\!\!\!\perp X \oplus Y$, thus precluding (S3).

3.4.4 Extension to Finite Abelian Groups

A close look at the modulations in the previous section reveals the following structure: we had a fine lattice $\Lambda = \mathbb{Z}$ and a coarse lattice $\Lambda_0 = 2\mathbb{Z}$, with the quotient group Λ/Λ_0, consisting of the two cosets $2\mathbb{Z}$ and $1 + 2\mathbb{Z}$, making up the probabilistically-chosen modulation alphabet. Given a message $X \in \Lambda/\Lambda_0$, the encoder outputs a random point from the coset X according to a carefully chosen probability distribution. Note that the quotient group in this case is isomorphic to $\mathbb{Z}_2$, and this enables recovery of the XOR of the bits (addition in $\mathbb{Z}_2$) from integer addition of the transmitted symbols modulo the coarse lattice. Also, the choice of the probability distribution (from Theorem 3.4.6) ensures that the choice of coset at each transmitter is independent of the integer sum at the relay. We shall extend the construction described in the previous subsection to n dimensions, thereby obtaining a scheme that satisfies properties (S1)–(S3).

[5] This proof was conveyed to the authors by Manjunath Krishnapur.

Now, any finite Abelian group $\mathbb{G}$ can be expressed as the quotient group Λ/Λ_0 for some pair of nested lattices $\Lambda_0 \subseteq \Lambda$. Indeed, any such $\mathbb{G}$ is isomorphic to a direct sum of cyclic groups: $\mathbb{G} \cong \mathbb{Z}_{N_1} \oplus \mathbb{Z}_{N_2} \oplus \cdots \oplus \mathbb{Z}_{N_k}$ for some positive integers $N_1, N_2, \ldots, N_k$ [44, Theorem 2.14.1]. Here, $\mathbb{Z}_{N_j}$ denotes the group of integers modulo-N_j. Taking $\Lambda = \mathbb{Z}^n$ and $\Lambda_0 = \mathsf{A}^T \mathbb{Z}^n$, where A is the diagonal matrix diag$(N_1, N_2, \ldots, N_k)$, we have $\mathbb{G} \cong \Lambda/\Lambda_0$. So, the finite Abelian group case is equivalent to considering the quotient group, i.e., the group of cosets, of a coarse lattice Λ_0 within a fine lattice Λ. These lattices may be taken to be full-rank lattices in $\mathbb{R}^n$.

As an example, let $N \geq 2$ be an integer, and let $\mathbb{Z}_N = \{0, 1, \ldots, N-1\}$ denote the set of integers modulo N. Let X, Y be iid random variables uniformly distributed over $\mathbb{Z}_N$, and let $X \oplus Y$ now denote their modulo-N sum. Similar to the binary case discussed so far, given a non-negative real characteristic function ψ such that $\psi(t) = 0$ for $|t| \geq \pi/N$, we can construct $\mathbb{Z}$-valued random variables U, V, jointly distributed with X, Y, for which properties (S1)–(S3) hold. In this case, the finite Abelian group can be taken as the group of cosets of the coarse lattice $N\mathbb{Z}$ within the fine lattice $\mathbb{Z}$, which is isomorphic to $\mathbb{Z}_N$.

Let Λ_0 be a sublattice of Λ of index M (i.e., the number of cosets of Λ_0 in Λ is M). List the cosets of Λ_0 in Λ as $\Lambda_0, \Lambda_1, \ldots, \Lambda_{M-1}$, which constitute the quotient group $\mathbb{G} = \Lambda/\Lambda_0$. As before, $\oplus$ denotes addition within $\mathbb{G}$.

Consider rvs X, Y uniformly distributed over $\mathbb{G}$. We wish to construct random vectors $\mathbf{U}, \mathbf{V}$ taking values in Λ, having the properties (S1)–(S3). The following theorem shows that this is possible.

Theorem 3.4.9. *Suppose that $\psi : \mathbb{R}^n \to \mathbb{R}^+$ is the characteristic function of a probability density function $f : \mathbb{R}^n \to \mathbb{R}^+$, such that $\psi(\mathbf{t}) = 0$ for $\mathbf{t} \notin \mathcal{V}(\widehat{\Lambda}_0)$, where $\widehat{\Lambda}_0$ is the Fourier dual of Λ_0. For $j = 0, 1, \ldots, M-1$, define the pmf p_j as follows:*

$$p_j(\mathbf{k}) = \begin{cases} |\text{det}\Lambda_0| f(\mathbf{k}) & \textit{if } \mathbf{k} \in \Lambda_j \\ 0 & \textit{otherwise.} \end{cases} \tag{3.18}$$

Finally, define a random vector $\mathbf{U}$ (resp. $\mathbf{V}$) jointly distributed with X (resp. Y) as follows:

if $X = \Lambda_j$ (resp. $Y = \Lambda_j$), $\mathbf{U}$ (resp. $\mathbf{V}$) is a random point from Λ_j picked according to the distribution p_j. Then, the resulting Λ-valued random vectors $\mathbf{U}, \mathbf{V}$ satisfy properties (S1)–(S3). Additionally, $\mathbb{E}\|\mathbf{U}\|^2$ and $\mathbb{E}\|\mathbf{V}\|^2$ are finite iff ψ is twice differentiable at $\mathbf{0}$, in which case $\mathbb{E}\|\mathbf{U}\|^2 = \mathbb{E}\|\mathbf{V}\|^2 = -\Delta\psi(\mathbf{0})$, where $\Delta = \sum_{j=1}^{n} \partial_j^2$ is the Laplacian operator.

As with Theorem 3.4.6 and XOR, the above theorem allows for secure computation at the relay of the group operation $X \oplus Y$. The theorem is proved using Proposition 3.4.4, in a manner completely analogous to Theorem 3.4.6.

Proof of Theorem 3.4.9

We are given an index-M sublattice Λ_0 of the lattice Λ. Recall from Section 2.1 that $(\det\Lambda_0)/(\det\Lambda) = M$. Let $\Lambda_0, \Lambda_1, \ldots, \Lambda_{M-1}$ denote the M cosets of Λ_0 in Λ. These constitute the elements of the quotient group $\mathbb{G} = \Lambda/\Lambda_0$.

Suppose that X, Y are iid random variables, each uniformly distributed over $\mathbb{G}$. For each $j \in \{0, 1, \ldots M-1\}$, let p_j be a pmf supported within the coset Λ_j, so that $p_j(\mathbf{k}) = 0$ for $\mathbf{k} \notin \Lambda_j$. We define a random vector $\mathbf{U}$ (resp. $\mathbf{V}$) jointly distributed with X (resp. Y) as follows: if $X = \Lambda_j$ (resp. $Y = \Lambda_j$), $\mathbf{U}$ (resp. $\mathbf{V}$) is a random point from Λ_j picked according to the distribution p_j. Then, $\mathbf{U}$ and $\mathbf{V}$ are identically distributed with $p_{\mathbf{U}} = p_{\mathbf{V}} = \frac{1}{M}\sum_{i=0}^{M-1} p_i$. Let $\varphi_{\mathbf{U}}$, $\varphi_{\mathbf{V}}$ and φ_j, $j = 0, 1, \ldots, M-1$, be the characteristic functions corresponding to $p_{\mathbf{U}}$, $p_{\mathbf{V}}$ and p_j, $j = 0, 1, \ldots, M-1$, respectively. We have the following straightforward generalization of Lemma 3.4.2.

Lemma 3.4.10. *Suppose that $\varphi_{\mathbf{U}}\varphi_{\mathbf{V}} = \varphi_j\varphi_{\mathbf{V}} = \varphi_{\mathbf{U}}\varphi_j$ for $j = 0, 1, \ldots, M-1$. Then, the random variables $(\mathbf{U}, \mathbf{V}, X, Y)$ with joint pmf given by*

$$p_{\mathbf{U}\mathbf{V}XY}(\mathbf{k}, \mathbf{l}, \Lambda_i, \Lambda_j) = (1/M)(1/M)p_i(\mathbf{k})p_j(\mathbf{l}) \quad \text{for} \quad \mathbf{k}, \mathbf{l} \in \Lambda \text{ and } \Lambda_i, \Lambda_j \in \mathbb{G} \tag{3.19}$$

have properties (S1)–(S3).

We will now construct the characteristic functions φ_j that satisfy the above lemma. Let f be the (continuous) probability density function corresponding to the compactly supported characteristic function ψ in the hypothesis of Theorem 3.4.9. The function f can be retrieved from ψ by Fourier inversion:

$$\begin{aligned} f(\mathbf{x}) &= \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} \psi(\mathbf{t}) e^{-i\langle \mathbf{t}, \mathbf{x} \rangle}\, d\mathbf{t} \\ &= \frac{1}{(2\pi)^n} \int_{\mathcal{V}(\widehat{\Lambda}_0)} \psi(\mathbf{t}) e^{-i\langle \mathbf{t}, \mathbf{x} \rangle}\, d\mathbf{t}. \end{aligned} \tag{3.20}$$

Note that each coset Λ_j can be expressed as $\mathbf{u}_j + \Lambda_0$ for some $\mathbf{u}_j \in \Lambda$. We set

$$\varphi_j(\boldsymbol{\zeta}) = \sum_{\mathbf{y} \in \widehat{\Lambda}_0} \psi(\boldsymbol{\zeta} + \mathbf{y})\, e^{-i\langle \mathbf{y},\, \mathbf{u}_j \rangle} \tag{3.21}$$

for all $\boldsymbol{\zeta} \in \mathbb{R}^n$. Then, by Proposition 3.4.4, we have that p_j is supported within Λ_j, and

$$p_j(\mathbf{k}) = (\det \Lambda_0)\, f(\mathbf{k}) \text{ for all } \mathbf{k} \in \Lambda_j. \tag{3.22}$$

Finally, define

$$\varphi(\boldsymbol{\zeta}) = \sum_{\mathbf{n} \in \widehat{\Lambda}} \psi(\boldsymbol{\zeta} + \mathbf{n}) \tag{3.23}$$

for all $\boldsymbol{\zeta} \in \mathbb{R}^n$.

We make two claims:

(i) $\varphi^2 = \varphi \varphi_j$ for $j = 0, 1, \ldots, M-1$;

(ii) $\varphi = \varphi_{\mathbf{U}} = \varphi_{\mathbf{V}}$.

Given these claims, by Lemma 3.4.10, the random vectors $\mathbf{U}, \mathbf{V}$ satisfy the properties (S1)–(S3).

Both claims follow from the fact that $\widehat{\Lambda}$ is a sublattice of $\widehat{\Lambda}_0$. (If a lattice Γ contains a sublattice Γ_0, then the dual Γ^* is a sublattice of Γ_0^*.) To see (i), we re-write (3.23) as

$$\varphi(\boldsymbol{\zeta}) = \sum_{\mathbf{s} \in \widehat{\Lambda}} \psi(\boldsymbol{\zeta} + \mathbf{s})\, e^{-i\langle \mathbf{s},\, \mathbf{u}_j \rangle}. \tag{3.24}$$

This is possible because, for $\mathbf{s} \in \widehat{\Lambda} = 2\pi\Lambda^*$ and $\mathbf{u}_j \in \Lambda$, we have $e^{-i\langle \mathbf{s}, \mathbf{u}_j \rangle} = 1$. Comparing (3.21) and (3.24), and noting that ψ is supported within $\mathcal{V}(\widehat{\Lambda}_0)$, it is evident that $\text{supp}(\varphi) \triangleq \{\boldsymbol{\zeta} : \varphi(\boldsymbol{\zeta}) \neq 0\}$ is contained in $\text{supp}(\varphi_j) \triangleq \{\boldsymbol{\zeta} : \varphi_j(\boldsymbol{\zeta}) \neq 0\}$. Furthermore, for all $\boldsymbol{\zeta} \in \text{supp}(\varphi)$, we have $\varphi(\boldsymbol{\zeta}) = \varphi_j(\boldsymbol{\zeta})$. Claim (i) directly follows from this.

For Claim (ii), we note that $\mathcal{V}(\widehat{\Lambda}_0) \subseteq \mathcal{V}(\widehat{\Lambda})$, since $\widehat{\Lambda}$ is a sublattice of $\widehat{\Lambda}_0$. Hence, we can apply Proposition 3.4.4 to deduce that φ is the characteristic function of a pmf p supported within Λ, with

$$p(\mathbf{k}) = (\det\Lambda)\, f(\mathbf{k}) \text{ for all } \mathbf{k} \in \Lambda.$$

Thus, from (3.22) and the fact that $(\det\Lambda_0)/(\det\Lambda) = M$, we see that $p = \frac{1}{M}\sum_{j=0}^{M-1} p_j$. In other words, $p = p_{\mathbf{U}} = p_{\mathbf{V}}$, which proves Claim (ii).

It remains to prove the statements concerning finiteness of $\mathbb{E}\|\mathbf{U}\|^2$ and $\mathbb{E}\|\mathbf{V}\|^2$. Theorem 1 in [108] shows that these moments are finite iff φ is twice differentiable at $\mathbf{0}$ (i.e., all second-order partial derivatives exist at $\mathbf{0}$). From (3.23), we see that φ agrees with ψ in a small neighbourhood around $\mathbf{0}$; hence, φ is twice differentiable at $\mathbf{0}$ iff ψ is twice differentiable at $\mathbf{0}$.

Assuming that ψ has all second-order partial derivatives at $\mathbf{0}$, we must show that $\mathbb{E}\|\mathbf{U}\|^2 = \mathbb{E}\|\mathbf{V}\|^2 = -\Delta\psi(\mathbf{0})$. Since $\mathbf{U}$ and $\mathbf{V}$ are identically distributed, it is enough to show that $\mathbb{E}\|\mathbf{U}\|^2 = -\Delta\psi(\mathbf{0})$. Write $\mathbf{U} = (U_1, \ldots, U_n)$, so that $\|\mathbf{U}\|^2 = U_1^2 + \cdots + U_n^2$. We want to show that $\mathbb{E}[U_j^2] = -\frac{\partial^2}{\partial t_j^2}\psi(\mathbf{0})$, for $j = 1, \ldots, n$. For notational simplicity, we show this for $j = 1$. Note that the characteristic function of U_1 is given by $\varphi_{U_1}(t_1) = \varphi_{\mathbf{U}}(t_1, 0, \ldots, 0)$. As argued prior to the statement of Theorem 3.4.6 in Section 3.4.3, $\mathbb{E}[U_1^2] = -\varphi_{U_1}''(0)$. Now, $\varphi_{U_1}''(0) = \frac{\partial^2}{\partial t_1^2}\varphi_{\mathbf{U}}(0, 0, \ldots, 0)$. From (3.23), we have that $\varphi_{\mathbf{U}} = \psi$ in a small neighbourhood around $\mathbf{0} = (0, 0, \ldots, 0)$. Therefore, $\frac{\partial^2}{\partial t_1^2}\varphi_{\mathbf{U}}(\mathbf{0}) = \frac{\partial^2}{\partial t_1^2}\psi(\mathbf{0})$, and hence, $\mathbb{E}[U_1^2] = -\frac{\partial^2}{\partial t_1^2}\psi(\mathbf{0})$, as desired.

This concludes the proof of Theorem 3.4.9. □

Constructing compactly supported twice-differentiable (or even C^∞) characteristic functions $\psi : \mathbb{R}^n \to \mathbb{R}^+$, $n \geq 1$, is straightforward, given our previous constructions

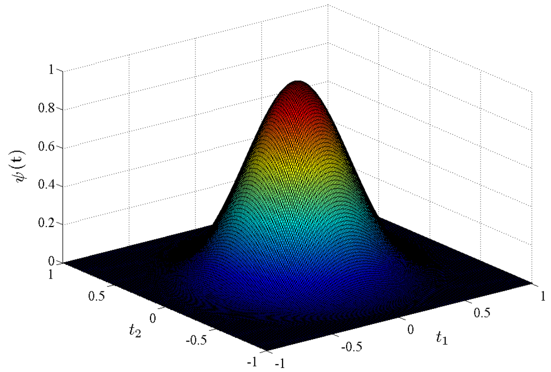


Figure 3.7: Example of a characteristic function supported within $\mathcal{V}(2\mathbb{Z}^2)$.

of such functions from $\mathbb{R}$ to $\mathbb{R}^+$. Suppose that for $i = 1, 2, \ldots, n$, $\psi_i : \mathbb{R} \to \mathbb{R}^+$ is the characteristic function of a random variable X_i, such that $\psi_i(t) = 0$ for $|t| \geq \lambda_i$, with $\lambda_i > 0$, and $X_1, X_2, \ldots, X_n$ are mutually independent. Then, $\psi(t_1, \ldots, t_n) = \prod_{i=1}^{n} \psi_i(t_i)$ is the characteristic function of the random vector $\mathbf{X} = (X_1, \ldots, X_n)$. Note that ψ is compactly supported: $\psi(\mathbf{t}) = 0$ for $\mathbf{t} \notin \prod_{i=1}^{n}(-\lambda_i, \lambda_i)$. Moreover, if the ψ_is are twice-differentiable (or C^∞) for all i, then so is ψ. Constructions other than product constructions are also in abundance; see e.g., [27], [80] and Theorem 3.4.11 below. A smooth, compactly supported characteristic function in $\mathbb{R}^2$ is depicted in Fig. 3.7.

3.4.5 Minimizing the Average Transmit Power

Our objective is to design codes (as defined in Definition 1) for secure computation at the relay. With the construction described in Theorem 3.4.9, the rate of the code depends on the number of cosets, M, of Λ_0 in Λ. For a given average power constraint, the system designer is usually faced with the task of maximizing the rate. Equivalently, for a given rate, the average transmit power must be kept as small as possible. The transmit power is equal to the second moment of $\mathbf{U}$ (or $\mathbf{V}$). Therefore, while any characteristic function ψ supported within $\mathcal{V}(\widehat{\Lambda}_0)$ suffices for the construction of Theorem 3.4.9, we must use a ψ for which $-\Delta\psi(\mathbf{0})$ is the least among such ψ's. This would yield random vectors $\mathbf{U}$ and $\mathbf{V}$ of least second moment (and hence least transmit power), and having the desired properties.

It is evident that by simply scaling the nested lattice pair, the average transmit power may be made as small as required. Suppose that the random vectors $\mathbf{U}$ and $\mathbf{V}$, distributed over a fine lattice Λ, have second moment P. Then, for any $\alpha > 0$, the random vectors $\mathbf{U}' = \alpha\mathbf{U}$ and $\mathbf{V}' = \alpha\mathbf{V}$, distributed over $\alpha\Lambda \triangleq \{\alpha\mathbf{z} : \mathbf{z} \in \Lambda\}$ have second moment $\alpha^2 P$. Choosing a small enough α would suffice to satisfy the power constraint. However, as we will see in the following sections, when we have to deal with the additive noise in the MAC channel, it is not possible to scale down the lattice arbitrarily if the probability of error is to be made small. Also, for a given (fixed) coarse lattice, it turns out that the second moment (which depends solely on the choice of ψ) cannot be made arbitrarily small. Indeed, the following result, adapted from [27], gives a precise and complete answer to the question of how small $-\Delta\psi(\mathbf{0})$ can be for a characteristic function ψ supported within a ball of radius ρ in $\mathbb{R}^n$.

Theorem 3.4.11 ([27], Theorem 5.1)**.** *Fix a $\rho > 0$. If ψ is a characteristic function of a random vector distributed over $\mathbb{R}^n$ such that $\psi(\mathbf{t}) = 0$ for $\|\mathbf{t}\| \geq \rho$, then*

$$-\Delta\psi(\mathbf{0}) \geq \frac{4}{\rho^2}\, j^2_{\frac{n-2}{2}}, \tag{3.25}$$

with equality iff $\psi(\mathbf{t}) = \widetilde{\psi}(\mathbf{t}/\rho)$ for $\widetilde{\psi} = \omega_n \tilde{*} \omega_n$. Here, j_k denotes the first positive zero of the Bessel function J_k. Also, $\omega_n(\mathbf{t}) = \gamma_n\, \Omega_n(2\|\mathbf{t}\| j_{\frac{n-2}{2}})$ for $\|\mathbf{t}\| \leq 1/2$ and $\omega_n(\mathbf{t}) = 0$ for $\|\mathbf{t}\| > 1/2$, and

$$\omega_n \tilde{*} \omega_n(\mathbf{t}) = \int \omega_n(\boldsymbol{\tau})\overline{\omega_n(\mathbf{t}+\boldsymbol{\tau})}\, n\boldsymbol{\tau}$$

denotes the folded-over self convolution of ω_n, with $\overline{\omega_n(\mathbf{t})}$ denoting the complex conjugate of $\omega_n(\mathbf{t})$. Furthermore, for $t \in \mathbb{R}$,

$$\Omega_n(t) = \Gamma(n/2)\Big(\frac{2}{t}\Big)^{\frac{n-2}{2}} J_{\frac{n-2}{2}}(t)$$

and

$$\gamma_n^2 = \frac{4 j^{n-2}_{\frac{n-2}{2}}}{\pi^{n/2}\Gamma(n/2) J^2_{\frac{n}{2}}(j_{\frac{n-2}{2}})},$$

where $\Gamma(\cdot)$ denotes the Gamma function. The density f corresponding to the minimum-variance ψ is given by $f(\mathbf{x}) = \rho^n \tilde{f}(\rho\mathbf{x})$, where

$$\tilde{f}(\mathbf{x}) = c_n \left(\frac{\Omega_n(\|\mathbf{x}\|/2)}{j^2_{\frac{n-2}{2}} - (\|\mathbf{x}\|/2)^2} \right)^2, \tag{3.26}$$

where

$$c_n = \frac{4 j^2_{\frac{n-2}{2}}}{4^n \pi^{n/2} \Gamma(n/2)}.$$

Observe that Theorem 3.4.9 is true for any nested lattice pair (Λ, Λ_0). As long as $\psi(\mathbf{t})$ is a characteristic function supported within $\mathcal{V}(\widehat{\Lambda}_0)$, we have an encoding scheme that satisfies (S1)–(S3). If we restrict ψ to be supported within a ball of radius ρ, which is contained within $\mathcal{V}(\widehat{\Lambda}_0)$, then Theorem 3.4.11 gives us a suitable candidate for ψ that can be used to obtain perfect secrecy. Since we are interested in minimizing the transmission power, we can choose ρ to be as large as $r_{\text{pack}}(\widehat{\Lambda}_0)$, where $r_{\text{pack}}(\widehat{\Lambda}_0)$ denotes the packing radius of $\widehat{\Lambda}_0$. Hence, we now have a coding scheme that achieves perfect secrecy for any arbitrary nested lattice pair. This is rather interesting, since earlier work on weak and strong secrecy using lattices [40, 41, 58] invariably required that the nested lattices satisfy certain goodness properties. Therefore, ours is an explicit scheme which specifies, for any nested lattice pair, a distribution to be used for randomization at the encoder in order to obtain perfect secrecy. In particular, our randomization scheme can also be used in conjunction with "practical" lattice coding schemes (e.g., [24, 86, 110]) that have low decoding complexity.

3.5 The Gaussian Noise Setting

Given any nested lattice pair, we now have a scheme whereby the relay can compute $X \oplus Y$ from $\mathbf{U} + \mathbf{V}$, but cannot determine X or Y separately. We next consider the scenario where the symbols received by the relay are corrupted by noise, and prove the achievability of the power-rate pairs described in Theorem 3.3.1. Recall that in the MAC phase, the relay receives

$$\mathbf{W} = \mathbf{U} + \mathbf{V} + \mathbf{Z},$$

where $\mathbf{Z}$ is zero-mean iid Gaussian noise with variance σ^2. The coding scheme that we use is largely based on the work in [28, 69], and is described below.

3.5.1 Coding Scheme for Perfect Secrecy

We describe here a sequence of $(n, M^{(n)})$ (recall Definition 1) codes that achieve perfect secrecy.

Code: A $(\Lambda^{(n)}, \Lambda_0^{(n)})$ *nested lattice code* consists of a pair of full-rank nested lattices $\Lambda_0^{(n)} \subseteq \Lambda^{(n)}$ in $\mathbb{R}^n$. The messages are chosen from the group $\mathbb{G}^{(n)} = \Lambda^{(n)}/\Lambda_0^{(n)}$, whose $M^{(n)} \triangleq |\Lambda^{(n)}/\Lambda_0^{(n)}|$ elements are listed as $\Lambda_0, \Lambda_1, \ldots, \Lambda_{M^{(n)}-1}$.

Encoding: We have messages X, Y at nodes $\mathtt{A}, \mathtt{B}$ that are independent rvs, uniformly distributed over $\mathbb{G}^{(n)}$. We first pick a characteristic function ψ supported within $\mathcal{V}(\widehat{\Lambda}_0^{(n)})$, as needed in Theorem 3.4.9. We impose the restriction that ψ be supported within a ball centered at $\mathbf{0}$ with radius equal to the packing radius, $r_{\text{pack}}(\widehat{\Lambda}_0^{(n)})$, of the dual lattice $\widehat{\Lambda}_0^{(n)}$. Recall that the packing radius is, by definition, the largest radius of a ball centered at $\mathbf{0}$ that is contained within $\mathcal{V}(\widehat{\Lambda}_0^{(n)})$. So, if $\psi(\mathbf{t}) = 0$ for $\|\mathbf{t}\| \geq r_{\text{pack}}(\widehat{\Lambda}_0^{(n)})$, then $\psi(\mathbf{t})$ is certainly supported within $\mathcal{V}(\widehat{\Lambda}_0^{(n)})$. If $X = \Lambda_j$, node $\mathtt{A}$ transmits a random vector $\mathbf{U} \in \Lambda_j$ picked according to the distribution p_j of Theorem 3.4.9. Similarly, if $Y = \Lambda_k$, node $\mathtt{B}$ transmits a random vector $\mathbf{V} \in \Lambda_k$ picked according to the distribution p_k. The rate of transmission from $\mathtt{A}$ or $\mathtt{B}$ is $R^{(n)} = \frac{1}{n}\log_2 M^{(n)}$. The average transmit power per dimension at each node is $P^{(n)} = \frac{-\Delta\psi(\mathbf{0})}{n}$, as in Theorem 3.4.9.

From Theorem 3.4.11, we see that an average transmit power per dimension as low as

$$P^{(n)} = \frac{4j^2_{\frac{n-2}{2}}}{n\left(r_{\text{pack}}(\widehat{\Lambda}_0^{(n)})\right)^2}, \tag{3.27}$$

is achievable by a suitable choice of ψ. It was shown in [90] (see also [32]) that the first positive zero of the Bessel function J_k can be written as $j_k = k + bk^{1/3} + \mathcal{O}(k^{-1/3})$, where

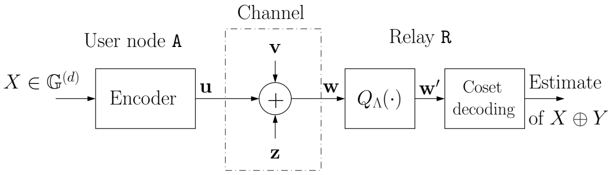


Figure 3.8: The operations performed by the user nodes and the relay.

b is a constant independent of k. Therefore,

$$P^{(n)} = \frac{n}{{r_{\text{pack}}}^2(\widehat{\Lambda}_0^{(n)})}(1 + o_n(1)), \tag{3.28}$$

where $o_n(1) \to 0$ as $n \to \infty$, is achievable by a suitable choice of ψ using Theorem 3.4.11.
Decoding: The relay $\texttt{R}$ receives $\mathbf{W} = \mathbf{U} + \mathbf{V} + \mathbf{Z}$, where $\mathbf{Z}$ is a Gaussian noise vector with n independent $\mathcal{N}(0, \sigma^2)$ components, which are all independent of $\mathbf{U}$ and $\mathbf{V}$. The relay estimates $\Lambda_j \oplus \Lambda_k$ to be the coset of $\Lambda_0^{(n)}$ represented by $Q_{\Lambda^{(n)}}(\mathbf{W})$, the closest vector to $\mathbf{W}$ in the lattice $\Lambda^{(n)}$. The decoder mapping is denoted by $\mathcal{D}(\cdot)$.
Security: Since the noise $\mathbf{Z}$ is independent of everything else, Theorem 3.4.9 shows that $\mathbf{W}$ is independent of the individual messages X, Y. Hence, even in the noisy setting, perfect security continues to be guaranteed at the relay for any choice of the nested lattice code. It is worth reiterating that perfect secrecy can be guaranteed irrespective of the noise $\mathbf{Z}$. The distribution of $\mathbf{Z}$ only determines the reliability of decoding, which in turn influences the power-rate pairs achievable with perfect secrecy.
Reliability and achievable power-rate pairs: Let $\xi^{(n)}$ denote the average probability that $Q_\Lambda(\mathbf{W})$ is different from the coset to which $\mathbf{U}+\mathbf{V}$ belongs. From Definition 2, a pair $(\mathcal{P}, \mathcal{R})$ is achievable if for every $\delta > 0$, there exists a sequence of nested lattice codes $(\Lambda^{(n)}, \Lambda_0^{(n)})$ for which the following hold for sufficiently large n: $R^{(n)} > \mathcal{R} - \delta$, $P^{(n)} < \mathcal{P} + \delta$ and $\xi^{(n)} < \delta$.

For a given nested lattice pair, Theorem 3.4.11 gives us the minimum average transmit power per dimension that guarantees perfect secrecy (subject to the condition that the characteristic function is supported within a ball of radius $r_{\text{pack}}(\widehat{\Lambda}_0^{(n)})$), and the pmf p_j that achieves the minimum. The choice of the nested lattices affects the reliability of decoding $X \oplus Y$ at the relay, and consequently determines achievable transmission rates.

To guarantee secure and reliable computation at the relay, we restrict the class of nested lattice pairs $(\Lambda^{(n)}, \Lambda_0^{(n)})$ to those which satisfy the following "goodness" properties:

(G_1) The sequence of coarse lattices, $\{\Lambda_0^{(n)}\}$, is good for covering and AWGN channel coding.

(G_2) The sequence of dual lattices, $\{\widehat{\Lambda}_0^{(n)}\}$, is good for packing.

(G_3) The sequence of fine lattices, $\{\Lambda^{(n)}\}$, is good for AWGN channel coding.

Unlike prior work on nested lattices [2, 28, 69, 71] which only required $\{\Lambda_0^{(n)}\}$ and $\{\Lambda^{(n)}\}$ to satisfy properties (G_1) and (G_3) above, we have the additional requirement that the sequence of Fourier duals, $\{\widehat{\Lambda}_0^{(n)}\}$ must be good for packing. We know from Section 2.5 that lattices satisfying (G_1)–(G_3) indeed exist, and we will use this fact to find achievable rates in the next section.

3.5.2 Achievable Rates

We now find achievable transmission rates for reliable and secure computation of $X \oplus Y$ at the relay. The analysis closely follows that in [28, 69, 70]. As defined in Section 3.5.1, let $\mathcal{D}(\mathbf{W})$ be the estimate of $X \oplus Y$ made by the relay; to be precise, $\mathcal{D}(\mathbf{W})$ is the coset of $\Lambda_0^{(n)}$ to which $Q_{\Lambda^{(n)}}(\mathbf{W})$ belongs. This is the same as the coset represented by $Q_{\Lambda^{(n)}}([\mathbf{W}] \bmod \Lambda_0^{(n)})$.

Each lattice point in $\Lambda^{(n)} \cap \mathcal{V}(\Lambda_0^{(n)})$ is a coset representative for a coset of $\Lambda_0^{(n)}$ in $\Lambda^{(n)}$. This is illustrated in Fig. 3.9. Suppose that Λ_j and Λ_k are the cosets which represent the messages X and Y, respectively. Let $\mathbf{X} = [\mathbf{U}] \bmod \Lambda_0^{(n)}$ and $\mathbf{Y} = [\mathbf{V}] \bmod \Lambda_0^{(n)}$ be the coset representatives of Λ_j and Λ_k, respectively. Then, $\Lambda_j \oplus \Lambda_k$ has $[\mathbf{X} + \mathbf{Y}] \bmod \Lambda_0^{(n)}$ as its representative. Therefore, the estimate $\mathcal{D}(\mathbf{W})$ has $\widehat{\mathbf{W}} = [Q_{\Lambda^{(n)}}(\mathbf{W})] \bmod \Lambda_0^{(n)}$ as its coset representative. This is equal to $\widehat{\mathbf{W}} = [Q_{\Lambda^{(n)}}([\mathbf{W}] \bmod \Lambda_0^{(n)})] \bmod \Lambda_0^{(n)}$. Let $\widetilde{\mathbf{W}} =$

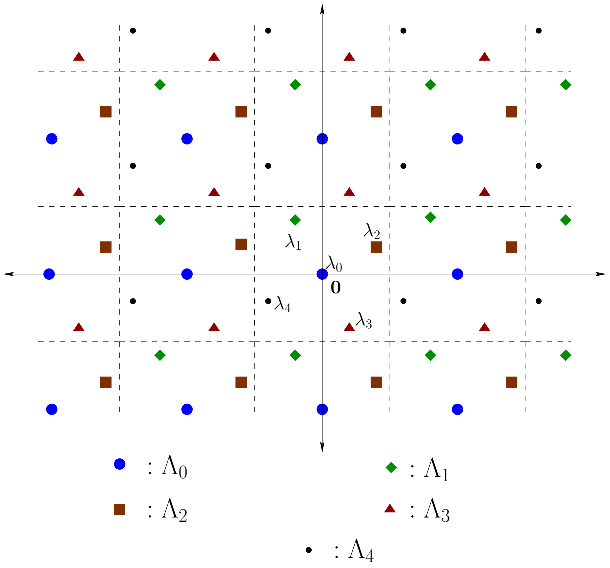


Figure 3.9: Different cosets of Λ_0 in Λ. The coset representative of Λ_j within $\mathcal{V}(\Lambda_0)$ is λ_j.

$[\mathbf{W}] \bmod \Lambda_0^{(n)}$. Then, $\widehat{\mathbf{W}} = [Q_{\Lambda^{(n)}}(\widetilde{\mathbf{W}})] \bmod \Lambda_0^{(n)}$. As a consequence of the transmitter-receiver operations, the "effective" channel from $\mathbf{X}, \mathbf{Y}$ to $\widetilde{\mathbf{W}}$ can be written as follows [69]:

$$\begin{aligned}\widetilde{\mathbf{W}} &= [\mathbf{U} + \mathbf{V} + \mathbf{Z}] \bmod \Lambda_0^{(n)} \\ &= \left[\left([\mathbf{U} + \mathbf{V}] \bmod \Lambda_0^{(n)}\right) + \mathbf{Z}\right] \bmod \Lambda_0^{(n)} \\ &= \left[\left([\mathbf{X} + \mathbf{Y}] \bmod \Lambda_0^{(n)}\right) + \mathbf{Z}\right] \bmod \Lambda_0^{(n)}.\end{aligned}$$

A channel of the form $\mathbf{W} = [\mathbf{X} + \mathbf{N}] \bmod \Lambda_0^{(n)}$, where $\mathbf{N}$ denotes the noise vector, is called a $\Lambda_0^{(n)}$-modulo lattice additive noise ($\Lambda_0^{(n)}$-MLAN) channel [28]. The random vector $\widetilde{\mathbf{W}}$ behaves like the output of a point-to-point transmission over a $\Lambda_0^{(n)}$-MLAN channel, with the transmitted vector being $[\mathbf{X} + \mathbf{Y}] \bmod \Lambda_0^{(n)}$. Looking from $\widetilde{\mathbf{W}}$, the "effective" channel is a $\Lambda_0^{(n)}$-MLAN channel, and the relay has to decode $[\mathbf{X} + \mathbf{Y}] \bmod \Lambda_0^{(n)}$ reliably from $\widetilde{\mathbf{W}}$. This is illustrated in Fig. 3.10. We will use the properties of the $\Lambda_0^{(n)}$-MLAN channel to determine achievable rate regions for our coding scheme.

We choose a sequence of nested lattice pairs that satisfy (G_1)–(G_3), with each nested lattice pair coming from a (n, k, q, k_1, q_1) ensemble, where k, q, k_1 and q_1 satisfy (2.23) and (2.24). Using the coding scheme of Section 3.5.1, we can achieve perfect secrecy. The

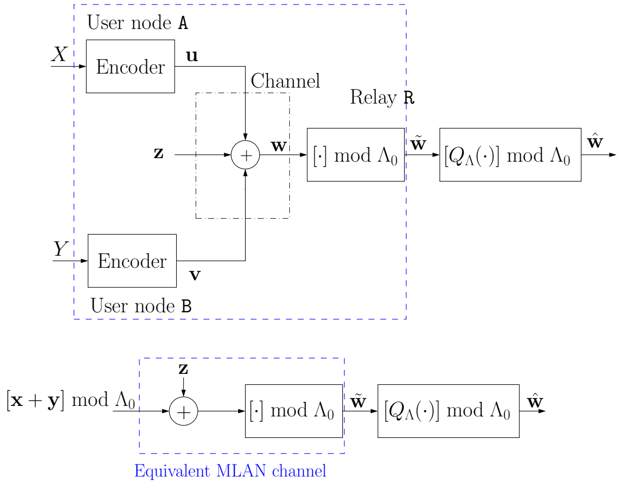


Figure 3.10: MAC phase of the bidirectional relay and equivalent MLAN channel representation.

proposition below provides us with the means of determining the rates achievable with this coding scheme.

Proposition 3.5.1. *Let $P > 0$ be a constant, and $\{\Lambda^{(n)}, \Lambda_0^{(n)}\}$ be a sequence of nested lattice pairs that satisfy (G_1)–(G_3), and scaled so as to satisfy $r_{\text{eff}}(\Lambda_0^{(n)}) = \sqrt{nP}$. Then, using the coding scheme of Section 3.5.1 with this sequence of nested lattice pairs, any rate less than $\frac{1}{2}\log_2\left(\frac{P}{\sigma^2}\right)$ is achievable with perfect secrecy.*

The proposition can be proved along the same lines as [28, Theorem 4]; we omit the details.

3.5.3 Relating Achievable Rates to Transmit Power

From (3.28), we know that as long as the average transmit power per dimension is less than $\left(\frac{n}{{r_{\text{pack}}}^2(\widehat{\Lambda}_0^{(n)})}\right)(1+o_n(1))$, we can guarantee perfect secrecy at the relay. From Proposition 3.5.1, we see that as long as the transmission rate is less than $\frac{1}{2}\log_2({r_{\text{eff}}}^2(\Lambda_0^{(n)})/(n\sigma^2))$, the relay can reliably compute $X \oplus Y$ from $\mathbf{W}$. In order to achieve positive rates, we need $r_{\text{eff}}(\Lambda_0^{(n)})$ to grow at least as fast as $\sqrt{n}$, i.e., $r_{\text{eff}}(\Lambda_0^{(n)}) = \Omega(\sqrt{n})$. Furthermore, to satisfy an average power constraint, we require $r_{\text{pack}}(\widehat{\Lambda}_0^{(n)}) = \Omega(\sqrt{n})$. The rate is an increasing function of $r_{\text{eff}}(\Lambda_0^{(n)})$, and the average transmit power per dimension is a decreasing function of $r_{\text{pack}}(\widehat{\Lambda}_0^{(n)})$. Since we want to maximize the rate for a given power constraint, we would like both $r_{\text{eff}}(\Lambda_0^{(n)})$ and $r_{\text{pack}}(\widehat{\Lambda}_0^{(n)})$ to be as large as possible. However, for any lattice $\Lambda_0^{(n)}$, we have $r_{\text{cov}}(\Lambda_0^{(n)})r_{\text{pack}}(\widehat{\Lambda}_0^{(n)}) \leq \pi n$ [5, Theorem 18.3], and since $r_{\text{eff}}(\Lambda_0^{(n)}) \leq r_{\text{cov}}(\Lambda_0^{(n)})$, we get $r_{\text{eff}}(\Lambda_0^{(n)})r_{\text{pack}}(\widehat{\Lambda}_0^{(n)}) \leq \pi n$. Hence, to obtain positive rates and at the same time satisfy the power constraint, both $r_{\text{eff}}(\Lambda_0^{(n)})$ and $r_{\text{pack}}(\widehat{\Lambda}_0^{(n)})$ must grow roughly as $\sqrt{n}$. Therefore, we seek lattices satisfying properties (G_1)–(G_3), for which the product $r_{\text{eff}}(\Lambda_0^{(n)})r_{\text{pack}}(\widehat{\Lambda}_0^{(n)})$ is close to the upper bound of πn.

For a sequence of Construction-A coarse lattices satisfying (G_1) and (G_2), we can find an asymptotic lower bound for $(1/n)r_{\text{eff}}(\Lambda_0^{(n)})r_{\text{pack}}(\widehat{\Lambda}_0^{(n)})$,[6] as the following theorem shows.

Lemma 3.5.2. *Let $\{\Lambda_0^{(n)}\}$ be a sequence of coarse lattices, with each $\Lambda_0^{(n)}$ chosen from an (n, k, q) ensemble and k, q satisfying (2.23) and (2.24). If $\{\Lambda_0^{(n)}\}$ satisfies conditions (G_1)–(G_2), then,*

$$\lim_{d\to\infty} \frac{r_{\text{eff}}(\Lambda_0^{(n)})r_{\text{pack}}(\widehat{\Lambda}_0^{(n)})}{d} \geq \frac{1}{2e}. \tag{3.29}$$

Proof. For ease of notation, denote by r_{eff}, the effective radius of $\Lambda_0^{(n)}$. The index, n, in r_{eff} has been dropped but it must be understood that this is a function of n. Let $\mathcal{C}^{(n)}$ denote the (n, k) code over $\mathbb{Z}_q$ that is used to generate the coarse lattice. Using (2.10),

$$\begin{aligned} q^k &= \frac{\Gamma(n/2+1)}{\pi^{n/2}{r_{\text{eff}}}^n} \\ &= \sqrt{n\pi}\left(\frac{n}{2\pi e {r_{\text{eff}}}^2}\right)^{n/2}(1+o_n(1)), \end{aligned} \tag{3.30}$$

where the second step uses Stirling's approximation, and $o_n(1)$ is a term that approaches 0 as $n \to \infty$. From (2.23), $k = \beta_0 n$ for some $0 < \beta_0 < 1/2$. Substituting this in the above,

[6] The product $r_{\text{eff}}(\Lambda_0^{(n)})r_{\text{pack}}(\widehat{\Lambda}_0^{(n)})$ is invariant to scaling of $\Lambda_0^{(n)}$. This is because, for a constant $\alpha > 0$, $r_{\text{eff}}(\alpha\Lambda_0^{(n)}) = \alpha r_{\text{eff}}(\Lambda_0^{(n)})$, and if $\Lambda' = \alpha\Lambda_0^{(n)}$, then the Fourier dual of Λ' is $(1/\alpha)\widehat{\Lambda}_0^{(n)}$.

and raising both sides to the power $1/n$, we get

$$\begin{aligned} q^{\beta_0} &= (n\pi)^{\frac{1}{2n}} \left(\frac{n}{2\pi e {r_{\text{eff}}}^2}\right)^{1/2} (1+o_n(1))^{1/n} \\ &= (n\pi)^{\frac{1}{2n}} \frac{\sqrt{n}}{\sqrt{2\pi e}\, r_{\text{eff}}} (1+o_n(1)). \end{aligned} \tag{3.31}$$

Let $\Lambda_0^{(n)*}$ denote the dual of $\Lambda_0^{(n)}$, and ${r_{\text{eff}}}^*$ denote the effective radius of $\Lambda_0^{(n)*}$. Let $\Lambda_0(\mathcal{C}^{(n)\perp})$ be the lattice obtained by applying Construction-A on the dual of $\mathcal{C}^{(n)}$, i.e., on $\mathcal{C}^{(n)\perp}$. As remarked in Section 2.4, $\Lambda_0(\mathcal{C}^{(n)\perp})$ comes from an $(n, n-k, q)$ ensemble. From Lemma 2.4.4, $\Lambda_0^{(n)*} = q\Lambda_0(\mathcal{C}^{(n)\perp})$. Therefore, $(1/q)\Lambda_0^{(n)*} = \Lambda_0(\mathcal{C}^{(n)\perp})$ will satisfy

$$q^{n-k} = \sqrt{n\pi} \left(\frac{n}{2\pi e \left(r_{\text{eff}}(\frac{1}{q}\Lambda_0^{(n)*}) \right)^2} \right)^{n/2} (1+o_n(1)),$$

where $o_n(1) \to 0$ as $n \to \infty$. But $r_{\text{eff}}(\frac{1}{q}\Lambda_0^{(n)*}) = \frac{1}{q}{r_{\text{eff}}}^*$, and hence, analogous to (3.31), we have

$$q^{n(1-\beta_0)} = \sqrt{n\pi} \left(\frac{n}{2\pi e (1/q)^2 ({r_{\text{eff}}}^*)^2} \right)^{n/2} (1+o_n(1)). \tag{3.32}$$

Rearranging,

$${r_{\text{eff}}}^* = (n\pi)^{\frac{1}{2n}} \frac{\sqrt{n} q^{\beta_0}}{\sqrt{2\pi e}} (1+o_n(1))^{1/n}. \tag{3.33}$$

Let the packing radius of $\Lambda_0^{(n)*}$ be $r_{\text{pack}}(\Lambda_0^{(n)*}) = \gamma(n){r_{\text{eff}}}^*$. From the definition of the packing radius, $\gamma(n) \leq 1$ for all n. Again, since the dual lattice is good for packing, $\lim_{n\to\infty} \gamma(n) \geq 1/2$. Also, since $o_n(1) \to 0$ as $n \to \infty$, we have $(1+o_n(1))^{1/n} = (1+o_n(1))$. Therefore, we have,

$$r_{\text{eff}}(\Lambda_0^{(n)}) r_{\text{pack}}(\Lambda_0^{(n)*}) = \gamma(n) r_{\text{eff}}(\Lambda_0^{(n)}) (n\pi)^{(1/2n)} \frac{\sqrt{n} q^{\beta_0}}{\sqrt{2\pi e}} (1+o_n(1)).$$

Substituting for q^{β_0} from (3.31) in the above equation, we get

$$\frac{r_{\text{eff}}(\Lambda_0^{(n)}) r_{\text{pack}}(\Lambda_0^{(n)*})}{n} = \gamma(n)(n\pi)^{(1/n)} \frac{1}{2\pi e}(1+o_n(1)). \tag{3.34}$$

Therefore, as $n \to \infty$, the above expression converges to a value greater than or equal to $1/4\pi e$. Using $r_{\text{pack}}(\widehat{\Lambda}_0^{(n)}) = 2\pi r_{\text{pack}}(\Lambda_0^{(n)*})$, we get Lemma 3.5.2. □

3.5.4 Proof of Theorem 3.3.1

Let us choose $r_{\text{eff}}(\Lambda_0^{(n)}) = \frac{1}{2e}\sqrt{n\mathcal{P}}$, for a constant $\mathcal{P} > 4e^2\sigma^2$. Fix a $\delta > 0$. Using Lemma 3.5.2, we see that

$$r_{\text{pack}}(\widehat{\Lambda}_0^{(n)}) \geq \frac{n}{2er_{\text{eff}}(\Lambda_0^{(n)})}(1 - o_d(1)) \geq \frac{\sqrt{n}}{\sqrt{\mathcal{P}}}(1 - o_n(1)). \tag{3.35}$$

From (3.28), we see that perfect secrecy can be achieved with an average power constraint as low as $P^{(n)} = \left(n/{r_{\text{pack}}}^2(\widehat{\Lambda}_0^{(n)})\right)(1 + o_n(1))$. Combining this and (3.35), perfect secrecy can be achieved with an average transmission power,

$$P^{(n)} < \mathcal{P} + \delta \tag{3.36}$$

for all sufficiently large n. From Proposition 3.5.1, we have seen that the average probability of error can be made to go down to zero as long as

$$R^{(n)} < \mathcal{R} \triangleq \frac{1}{2}\log_2 \frac{\mathcal{P}}{(2e)^2\sigma^2}. \tag{3.37}$$

Therefore, for every $\delta > 0$, we can choose a sequence of nested lattice codes such that for all sufficiently large n, we have $R^{(n)} > \mathcal{R} - \delta$, $P^{(n)} < \mathcal{P} + \delta$ and $\xi^{(n)} < \delta$. Hence, a power-rate pair of

$$\left(\mathcal{P}, \left[\frac{1}{2}\log_2 \frac{\mathcal{P}}{\sigma^2} - \log_2 2e\right]^+\right)$$

is achievable with perfect secrecy, concluding the proof of Theorem 3.3.1. □

Chapter 4

Strongly Secure Bidirectional Relaying

4.1 Strong Secrecy

Let us quickly summarize the coding scheme used to obtain perfect secrecy in the previous chapter. We chose a pair of nested lattices $(\Lambda^{(n)}, \Lambda_0^{(n)})$, where $\Lambda_0^{(n)} \subseteq \Lambda^{(n)}$, and the messages were chosen uniformly over the quotient group, $\Lambda^{(n)}/\Lambda_0^{(n)}$. Given a message (coset) Λ_j, the user node transmits a point from that coset according to a distribution p_j, which is obtained by sampling (and appropriately normalizing) a distribution f over $\mathbb{R}^n$, at precisely those lattice points that belong to the coset Λ_j. As long as the characteristic function corresponding to f is supported within the fundamental Voronoi region of the Fourier dual of $\Lambda_0^{(n)}$, perfect secrecy can be obtained. We imposed an additional constraint, that the characteristic function corresponding to f must be supported within a ball of radius $r_{\text{pack}}(\widehat{\Lambda}_0^{(n)})$. Using this, we obtained the minimum average transmit power, and found an achievable rate. In this chapter, we relax the security constraint to strong secrecy in order to achieve higher transmission rates.

Ideas from the perfectly secure scheme

A natural question that arises is what happens if we replace f in Theorem 3.4.9 by a density function for which the support of the characteristic function is larger than $\mathcal{V}(\widehat{\Lambda}_0^{(n)})$. Can we obtain different secrecy properties by simply changing the density f? Specifically, let $\psi(\mathbf{t})$ be a characteristic function which is supported within a ball of radius $\rho > r_{\text{pack}}(\widehat{\Lambda}_0^{(n)})$, and choose the characteristic function $\phi_{\mathbf{U}|X=\mathbf{x}}(\mathbf{t}) = \sum_{\mathbf{n}\in\widehat{\Lambda}_0^{(n)}} \psi(\mathbf{t}+\mathbf{n})e^{-i\langle \mathbf{n},\mathbf{x}\rangle}$. Clearly, we cannot expect perfect secrecy, but can we at least obtain strong secrecy? Let us take ψ to be the characteristic function of the minimum-variance distribution in (3.26), with the support of ψ chosen to be a ball of radius $\rho = \min\{r_{\text{eff}}(\widehat{\Lambda}_0^{(n)}), 2r_{\text{pack}}(\widehat{\Lambda}_0^{(n)})\}$.[1] Doing so would give us an improved rate of $\left[\frac{1}{2}\log_2\frac{\mathcal{P}}{\sigma^2} - \log_2 e\right]^+$.

We can show that the ℓ^2 norm of the difference between the joint pmf of $U+V$ and X, and the product of the marginals goes to zero exponentially in n. Knowing only that the ℓ^2 norm of the difference between $p_{\mathbf{U}+\mathbf{V},X}$ and $p_{\mathbf{U}+\mathbf{V}}p_X$ goes to zero as $n\to\infty$, we cannot conclude whether strong secrecy is obtained. In fact, by itself, the ℓ^2 norm is not a good measure of secrecy. In any case, we will use a different approach to obtaining strong secrecy, and show that an even higher transmission rate of $\left[\frac{1}{2}\log_2\left(\frac{1}{2}+\frac{\mathcal{P}}{\sigma^2}\right) - \frac{1}{2}\log_2 2e\right]^+$ is achievable.

Instead of using distributions with compactly supported characteristic functions, we will use a sampled Gaussian density for randomization at the encoders. Such a scheme was used in context of the wiretap channel in [58]. We will show that if a Gaussian pdf is used instead of a density f having a compactly supported characteristic function, then we can obtain strong secrecy. It is interesting to note that the same basic coding scheme, but with a different pdf used for randomization, can give different secrecy properties.

4.1.1 The Gaussian Density

We now introduce some notation that will be used in the sequel. Let Λ be a lattice in $\mathbb{R}^n$. For any $\mathbf{x}\in\mathbb{R}^n$, and any $\kappa>0$, we define $g_{\kappa,\mathbf{x}}(\cdot)$ to be the Gaussian density with mean

[1] If we have $\rho > 2r_{\text{pack}}(\widehat{\Lambda}_0^{(n)})$, then $\sum_{\mathbf{n}\in\widehat{\Lambda}_0^{(n)}} \psi(\mathbf{t}+\mathbf{n})e^{-i\langle \mathbf{n},\mathbf{x}\rangle}$ would have to be normalized to make it a characteristic function, and this makes the analysis more complicated.

$\mathbf{x}$ and covariance matrix $\kappa^2\mathsf{I}_n$, i.e., $\forall\mathbf{z}\in\mathbb{R}^n$,

$$g_{\kappa,\mathbf{x}}(\mathbf{z}) \triangleq \frac{1}{(2\pi\kappa^2)^{n/2}} e^{-\frac{\|\mathbf{z}-\mathbf{x}\|^2}{2\kappa^2}}. \tag{4.1}$$

We also define

$$g_{\kappa,\mathbf{x}}(\Lambda) \triangleq \sum_{\lambda\in\Lambda} g_{\kappa,\mathbf{x}}(\lambda). \tag{4.2}$$

We will use $g_\kappa(\mathbf{z})$ and $g_\kappa(\Lambda)$ to denote $g_{\kappa,\mathbf{0}}(\mathbf{z})$ and $g_{\kappa,\mathbf{0}}(\Lambda)$, respectively.

4.2 Coding Scheme for Strong Secrecy

Code: Following Section 3.5.1, we use a $(\Lambda^{(n)},\Lambda_0^{(n)})$ nested lattice code, with $\Lambda_0^{(n)}\subseteq\Lambda^{(n)}$. As before, the messages are chosen from $\mathbb{G}^{(n)}\triangleq\Lambda^{(n)}/\Lambda_0^{(n)}$, and $\oplus$ is the addition operation on $\mathbb{G}^{(n)}$. The $M^{(n)}\triangleq|\mathbb{G}^{(n)}|$ cosets of $\Lambda_0^{(n)}$ in $\Lambda^{(n)}$ are denoted by $\Lambda_0,\ldots,\Lambda_{M^{(n)}-1}$.
Encoding: For a coset Λ_j of $\Lambda_0^{(n)}$ in $\Lambda^{(n)}$, let λ_j denote its representative within $\mathcal{V}(\Lambda_0^{(n)})$ (see Fig. 3.9 for an illustration). Fix a $\kappa>0$. Corresponding to the message Λ_j, the user node transmits a random lattice point from Λ_j, according to the distribution

$$p_j(\mathbf{u}) = \begin{cases} \frac{g_\kappa(\mathbf{u})}{g_{\kappa,-\lambda_j}(\Lambda_0^{(n)})} & \text{if } \mathbf{u}\in\Lambda_j, \\ \mathbf{0} & \text{otherwise.} \end{cases} \tag{4.3}$$

Decoding: The relay computes the closest point in $\Lambda^{(n)}$ to the linear minimum mean-squared error (MMSE) estimate of the received vector, as in [58, 28, 69], and the output of the decoder is the coset to which this point belongs. Let $\alpha^*=\frac{2\kappa^2}{2\kappa^2+\sigma^2}$ be the linear MMSE coefficient, and $\widetilde{\mathbf{W}}=[\alpha^*\mathbf{W}]\bmod\Lambda_0^{(n)}$. The estimate of $X\oplus Y$, denoted by $\mathcal{D}(\mathbf{W})$, is then the coset to which $Q_{\Lambda^{(n)}}(\widetilde{\mathbf{W}})$ belongs.
Achievable power-rate pair: A power-rate pair of $(\mathcal{P},\mathcal{R})$ is achievable if there exists a sequence of $(\Lambda^{(n)},\Lambda_0^{(n)})$ nested lattice codes such that for every $\delta>0$ and for all sufficiently large n,

- the average transmit power per dimension is less than $\mathcal{P}+\delta$:

$$P^{(n)} \triangleq \frac{1}{n}\mathbb{E}\|\mathbf{U}\|^2 = \frac{1}{n}\mathbb{E}\|\mathbf{V}\|^2 < \mathcal{P}+\delta;$$

- the transmission rate is greater than $\mathcal{R}-\delta$:

$$R^{(n)} \triangleq \frac{1}{n}\log_2 M^{(n)} > \mathcal{R}-\delta;$$

- the average probability of decoding $X \oplus Y$ incorrectly from $\mathbf{W}$ is less than δ; and

- the mutual information between each message and $\mathbf{U}+\mathbf{V}$ is less than δ:

$$\mathcal{I}(X;\mathbf{U}+\mathbf{V}) = \mathcal{I}(Y;\mathbf{U}+\mathbf{V}) < \delta.$$

We will show that

Theorem 4.2.1. *A power-rate pair of*

$$\left(\mathcal{P}, \left[\frac{1}{2}\log_2\left(\frac{1}{2}+\frac{\mathcal{P}}{\sigma^2}\right) - \frac{1}{2}\log_2 2e\right]^+\right)$$

can be achieved with strong secrecy using the coding scheme of Section 4.2.

We will prove this theorem in two parts. In Section 4.3, we will show that the coding scheme outlined earlier in this section achieves strong secrecy in the absence of noise. In Section 4.4, we will impose certain restrictions on the nested lattice pairs and show that the rate guaranteed by Theorem 4.2.1 is achievable.

4.3 Strong Secrecy in the Absence of Noise

We will first prove that the scheme described in the previous section achieves strong secrecy in the absence of noise. Let us establish some more notation. Let $p_{\mathbf{U}+\mathbf{V}}(\cdot)$ denote the distribution of $\mathbf{U}+\mathbf{V}$, and for any $\mathbf{x} \in \Lambda^{(n)} \cap \mathcal{V}(\Lambda_0^{(n)})$, let $p_{\mathbf{U}+\mathbf{V}|\mathbf{x}}(\cdot)$ denote the distribution of $\mathbf{U}+\mathbf{V}$ conditioned on the event that X is the coset to which $\mathbf{x}$ belongs. We will show that for every $\mathbf{x}$ in $\Lambda^{(n)} \cap \mathcal{V}(\Lambda_0^{(n)})$ the *variational distance* (also called the total variation distance) between $p_{\mathbf{U}+\mathbf{V}}$ and $p_{\mathbf{U}+\mathbf{V}|\mathbf{x}}(\cdot)$, defined as[2]

$$\mathbb{V}(p_{\mathbf{U}+\mathbf{V}}, p_{\mathbf{U}+\mathbf{V}|\mathbf{x}}) \triangleq \sum_{\mathbf{w}\in\Lambda^{(n)}} |p_{\mathbf{U}+\mathbf{V}}(\mathbf{w}) - p_{\mathbf{U}+\mathbf{V}|\mathbf{x}}(\mathbf{w})|, \tag{4.4}$$

goes to zero exponentially in the dimension n. Therefore, the *average variational distance* between the joint pmf of $\mathbf{U}+\mathbf{V}$ and X, and the product of the marginals,

$$\overline{\mathbb{V}} \triangleq \sum_{\mathbf{x}\in\Lambda^{(n)}\cap\mathcal{V}(\Lambda_0^{(n)})} \frac{1}{|\mathbb{G}^{(n)}|}\mathbb{V}(p_{\mathbf{U}+\mathbf{V}}, p_{\mathbf{U}+\mathbf{V}|\mathbf{x}}),$$

also goes to zero exponentially in n. We can then use the following lemma, which relates the mutual information and the variational distance.

Lemma 4.3.1 ([19], Lemma 1)**.** *For $|\mathbb{G}^{(n)}| \geq 4$, we have*

$$\mathcal{I}(X;\mathbf{U}+\mathbf{V}) \leq \overline{\mathbb{V}}\left(\log_2|\mathbb{G}^{(n)}| - \log_2(\overline{\mathbb{V}})\right). \tag{4.5}$$

We are mainly interested in the case where the asymptotic rate ($\lim_{n\to\infty}\frac{1}{n}\log_2|\mathbb{G}^{(n)}|$) is finite[3]. In this scenario, $|\mathbb{G}^{(n)}|$ grows exponentially in n and it is sufficient to have $\overline{\mathbb{V}}$ going to zero as $o(1/n)$ to ensure that $\mathcal{I}(X;\mathbf{U}+\mathbf{V})$ goes to zero. We will in fact show that $\overline{\mathbb{V}}$ can be made to go to zero exponentially in n, which will guarantee that the mutual information also decays exponentially in n. In order to have $\overline{\mathbb{V}}$ going to zero exponentially in n, we will require the coarse and fine lattices to satisfy certain properties.

Recall from Chapter 2 that for any lattice Λ in $\mathbb{R}^n$, and any $\theta > 0$, the flatness factor

[2] For probability measures P_1 and P_2 defined on a discrete alphabet $\mathcal{X}$, the total variation distance between them is usually defined as $\mathbb{V}(P_1,P_2) \triangleq \sup_{A\subseteq\mathcal{X}}|P_1(A)-P_2(A)|$. This can be shown to be equal to $\frac{1}{2}\sum_{x\in\mathcal{X}}|P_1(x)-P_2(x)|$ (see e.g., [15, Section 11.6]). We have dropped the $\frac{1}{2}$ factor for simplicity.

[3] Otherwise, we would not be able to guarantee reliable decoding at the relay in the presence of noise.

$\epsilon_\Lambda(\theta)$ is defined as [58, 7]

$$\epsilon_\Lambda(\theta) \triangleq \frac{\max_{\mathbf{x}\in\mathcal{V}(\Lambda)} \left| \left(\sum_{\lambda\in\Lambda} g_{\theta,\lambda}(\mathbf{x})\right) - (1/\det\Lambda)\right|}{1/\det\Lambda}. \tag{4.6}$$

Moreover, a sequence of lattices $\{\Lambda^{(n)}\}$ is said to be *secrecy-good* if

$$\epsilon_{\Lambda^{(n)}}(\theta) \le 2^{-\Omega(d)} \text{ for all } \theta \text{ such that } \frac{(\det(\Lambda^{(n)}))^{2/d}}{2\pi\theta^2} < 1.$$

It was shown in [58] that there exist secrecy-good lattices that are also simultaneously good for covering, packing, AWGN channel coding..

Let us choose κ in (4.3) to be equal to $\sqrt{\mathcal{P}}$. We can bound the variational distance in terms of the flatness factor of the coarse lattice as follows:

Theorem 4.3.2. *Suppose that the sequence of nested lattice pairs $\{\Lambda^{(n)}, \Lambda_0^{(n)}\}$ satisfies*

$$\epsilon^{(n)} \triangleq \epsilon_{\Lambda_0^{(n)}}(\sqrt{\mathcal{P}/2}) < 1/2.$$

Then, for every $\mathbf{x} \in \Lambda^{(n)} \cap \mathcal{V}(\Lambda_0^{(n)})$, we have

$$\mathbb{V}(p_{\mathbf{U}+\mathbf{V}}, p_{\mathbf{U}+\mathbf{V}|\mathbf{x}}) \le 216 \cdot \epsilon^{(n)}. \tag{4.7}$$

We prove the above theorem in Section 4.3.1. The constant 216 in the above theorem can be improved, but we do not attempt to do so, as the exact constant is not important for our purposes.

The following result from [58, Section V-B] tells us that if the flatness factor of the coarse lattice goes to zero as $n \to \infty$, then the average transmit power converges to $\mathcal{P}$.

Lemma 4.3.3. *If the flatness factor $\epsilon_1 \triangleq \epsilon_{\Lambda_0^{(n)}}\left(\mathcal{P}\sqrt{1-1/(e\pi)}\right) < 1/2$, then,*

$$\left|\mathbb{E}\|\mathbf{U}\|^2 - n\mathcal{P}\right| = \left|\mathbb{E}\|\mathbf{V}\|^2 - n\mathcal{P}\right| \le \frac{2\pi\epsilon_1}{1-\epsilon_1}\mathcal{P}.$$

Since $\sqrt{1-1/(e\pi)} > 1/\sqrt{2}$, it is sufficient to have (by monotonicity of the flatness factor) $\epsilon_{\Lambda_0^{(n)}}(\sqrt{\mathcal{P}/2}) \to 0$ to satisfy the power constraint for all sufficiently large n. From Theorem 4.3.2 and Lemma 4.3.1, we see that strong secrecy can be obtained in the noiseless scenario.

4.3.1 Proof of Theorem 4.3.2

The following lemma from [58] will be used in the proof.

Lemma 4.3.4 ([58], Lemma 4)**.** *Let Λ be a lattice in $\mathbb{R}^n$. Then, for all $\mathbf{z} \in \mathbb{R}^n$, and $\kappa > 0$,*

$$\frac{1-\epsilon_\Lambda(\kappa)}{1+\epsilon_\Lambda(\kappa)} \leq \frac{g_{\kappa,\mathbf{z}}(\Lambda)}{g_\kappa(\Lambda)} \leq 1.$$

For ease of notation, we will suppress the index n in $\epsilon^{(n)}$, $\Lambda_0^{(n)}$ and $\Lambda^{(n)}$. We will find upper and lower bounds for $p_{\mathbf{U}+\mathbf{V}}(\mathbf{u})$ and $p_{\mathbf{U}+\mathbf{V}|\mathbf{x}}(\mathbf{u})$, and then use these to get an upper bound on the absolute value of the difference between the two.

For a message X chosen at node $\mathtt{A}$, let $\mathbf{x}$ be the coset representative of X from $\Lambda \cap \mathcal{V}(\Lambda_0)$. For any subset $S \subseteq \mathbb{R}^n$, let $\mathbf{1}_S(\cdot)$ denote the indicator function of S, i.e., $\mathbf{1}_S(\mathbf{u})$ is 1 if $\mathbf{u} \in S$, and 0 otherwise. From (4.3), with $\kappa = \sqrt{\mathcal{P}}$, we have

$$p_{\mathbf{U}|\mathbf{x}}(\mathbf{u}) = \frac{g_{\sqrt{\mathcal{P}}}(\mathbf{u})}{g_{\sqrt{\mathcal{P}},-\mathbf{x}}(\Lambda_0)} \mathbf{1}_{\Lambda_0+\mathbf{x}}(\mathbf{u}). \tag{4.8}$$

Let $\mathbb{G}_X \triangleq \Lambda \cap \mathcal{V}(\Lambda_0)$, and $M \triangleq |\mathbb{G}^{(n)}| = |\mathbb{G}_X|$. Since the messages are uniformly distributed,

$$p_{\mathbf{U}}(\mathbf{u}) = \sum_{\mathbf{x} \in \mathbb{G}_X} \frac{g_{\sqrt{\mathcal{P}}}(\mathbf{u})}{g_{\sqrt{\mathcal{P}},-\mathbf{x}}(\Lambda_0)} \frac{\mathbf{1}_{\Lambda_0+\mathbf{x}}(\mathbf{u})}{M}. \tag{4.9}$$

By monotonicity of the flatness factor, $\epsilon_{\Lambda_0}(\sqrt{\mathcal{P}}) < \epsilon_{\Lambda_0}(\sqrt{\mathcal{P}/2}) = \epsilon$, and using Lemma 4.3.4,

$$\frac{g_{\sqrt{\mathcal{P}}}(\mathbf{u})}{g_{\sqrt{\mathcal{P}}}(\Lambda_0)} \leq \frac{g_{\sqrt{\mathcal{P}}}(\mathbf{u})}{g_{\sqrt{\mathcal{P}},-\mathbf{x}}(\Lambda_0)} \leq \frac{g_{\sqrt{\mathcal{P}}}(\mathbf{u})}{g_{\sqrt{\mathcal{P}}}(\Lambda_0)} \frac{1+\epsilon}{1-\epsilon}.$$

Using this in (4.9), we get for $\mathbf{u} \in \Lambda$,

$$\frac{g_{\sqrt{\mathcal{P}}}(\mathbf{u})}{M g_{\sqrt{\mathcal{P}}}(\Lambda_0)} \leq p_{\mathbf{U}}(\mathbf{u}) \leq \frac{g_{\sqrt{\mathcal{P}}}(\mathbf{u})}{M g_{\sqrt{\mathcal{P}}}(\Lambda_0)} \frac{1+\epsilon}{1-\epsilon}. \tag{4.10}$$

We will require bounds on $g_{\sqrt{\mathcal{P}}}(\Lambda)$ in the proof. Rearranging the terms above,

$$\left(\frac{1-\epsilon}{1+\epsilon}\right) p_{\mathbf{U}}(\mathbf{u}) M g_{\sqrt{\mathcal{P}}}(\Lambda_0) \leq g_{\sqrt{\mathcal{P}}}(\mathbf{u}) \leq p_{\mathbf{U}}(\mathbf{u}) M g_{\sqrt{\mathcal{P}}}(\Lambda_0).$$

Since $p_{\mathbf{U}}$ is a pmf supported over Λ, and $\sum_{\mathbf{u}\in\Lambda} p_{\mathbf{U}}(\mathbf{u}) = 1$, we can get

$$\left(\frac{1-\epsilon}{1+\epsilon}\right) M g_{\sqrt{\mathcal{P}}}(\Lambda_0) \leq g_{\sqrt{\mathcal{P}}}(\Lambda) \leq M g_{\sqrt{\mathcal{P}}}(\Lambda_0). \tag{4.11}$$

It can be similarly verified that for any $\mathbf{a} \in \mathbb{R}^n$,

$$\left(\frac{1-\epsilon}{1+\epsilon}\right) M g_{\sqrt{\frac{\mathcal{P}}{2}},\mathbf{a}}(\Lambda_0) \leq g_{\sqrt{\frac{\mathcal{P}}{2}},\mathbf{a}}(\Lambda) \leq M g_{\sqrt{\frac{\mathcal{P}}{2}},\mathbf{a}}(\Lambda_0). \tag{4.12}$$

We establish some more notation for convenience. Let

$$\alpha(\mathbf{w}) \triangleq \frac{g_{\sqrt{2\mathcal{P}}}(\mathbf{w})}{M g_{\sqrt{\mathcal{P}}}(\Lambda_0)} \frac{g_{\sqrt{\frac{\mathcal{P}}{2}}}(\Lambda_0)}{g_{\sqrt{\mathcal{P}}}(\Lambda_0)}, \tag{4.13}$$

$$\beta(\mathbf{x}, \mathbf{w}) \triangleq \left(\frac{g_{\sqrt{\frac{\mathcal{P}}{2}},\frac{\mathbf{w}}{2}-\mathbf{x}}(\Lambda_0)}{g_{\sqrt{\frac{\mathcal{P}}{2}}}(\Lambda_0)}\right) \left(\frac{g_{\sqrt{\mathcal{P}},-\mathbf{x}}(\Lambda_0)}{g_{\sqrt{\mathcal{P}}}(\Lambda_0)}\right)^{-1}. \tag{4.14}$$

We can bound $p_{\mathbf{U}+\mathbf{V}|\mathbf{x}}$ and $p_{\mathbf{U}+\mathbf{V}}$ as follows.

Lemma 4.3.5. *For any lattice point $\mathbf{w} \in \Lambda$, and any $\mathbf{x} \in \mathbb{G}_X$, we have*

$$\left(\frac{1-\epsilon}{1+\epsilon}\right) \alpha(\mathbf{w}) \leq p_{\mathbf{U}+\mathbf{V}}(\mathbf{w}) \leq \left(\frac{1+\epsilon}{1-\epsilon}\right)^2 \alpha(\mathbf{w}) \tag{4.15}$$

$$\beta(\mathbf{x}, \mathbf{w}) \alpha(\mathbf{w}) \leq p_{\mathbf{U}+\mathbf{V}|\mathbf{x}}(\mathbf{w}) \leq \left(\frac{1+\epsilon}{1-\epsilon}\right) \beta(\mathbf{x}, \mathbf{w}) \alpha(\mathbf{w}). \tag{4.16}$$

Proof. Let $\mathbf{x}$ be any fine lattice point from $\mathbb{G}_X$. Then,

$$p_{\mathbf{U}+\mathbf{V}|\mathbf{x}}(\mathbf{w}) = \sum_{\mathbf{t}\in\Lambda_0+\mathbf{x}} p_{\mathbf{U}|\mathbf{x}}(\mathbf{t})p_{\mathbf{V}}(\mathbf{w}-\mathbf{t}).$$

Using (4.8) and (4.10) in the above equation, we obtain

$$\sum_{\mathbf{t}\in\Lambda_0+\mathbf{x}} \frac{g_{\sqrt{\mathcal{P}}}(\mathbf{t})}{g_{\sqrt{\mathcal{P}},-\mathbf{x}}(\Lambda_0)} \frac{g_{\sqrt{\mathcal{P}}}(\mathbf{w}-\mathbf{t})}{Mg_{\sqrt{\mathcal{P}}}(\Lambda_0)} \leq p_{\mathbf{U}+\mathbf{V}|\mathbf{x}}(\mathbf{w}) \leq \sum_{\mathbf{t}\in\Lambda_0+\mathbf{x}} \frac{g_{\sqrt{\mathcal{P}}}(\mathbf{t})}{g_{\sqrt{\mathcal{P}},-\mathbf{x}}(\Lambda_0)} \frac{g_{\sqrt{\mathcal{P}}}(\mathbf{w}-\mathbf{t})}{Mg_{\sqrt{\mathcal{P}}}(\Lambda_0)} \left(\frac{1+\epsilon}{1-\epsilon}\right). \tag{4.17}$$

Consider the term

$$\begin{aligned}
&\sum_{\mathbf{t}\in\Lambda_0+\mathbf{x}} \frac{g_{\sqrt{\mathcal{P}}}(\mathbf{t})}{g_{\sqrt{\mathcal{P}},-\mathbf{x}}(\Lambda_0)} \frac{g_{\sqrt{\mathcal{P}}}(\mathbf{w}-\mathbf{t})}{Mg_{\sqrt{\mathcal{P}}}(\Lambda_0)} \\
&= \frac{1}{g_{\sqrt{\mathcal{P}},-\mathbf{x}}(\Lambda_0)} \frac{1}{Mg_{\sqrt{\mathcal{P}}}(\Lambda_0)} \sum_{\mathbf{t}\in\Lambda_0+\mathbf{x}} \frac{e^{\left(-\frac{\|\mathbf{t}\|^2}{2\mathcal{P}}-\frac{\|\mathbf{t}-\mathbf{w}\|^2}{2\mathcal{P}}\right)}}{(2\pi\mathcal{P})^d} \\
&= \frac{1}{Mg_{\sqrt{\mathcal{P}}}(\Lambda_0)g_{\sqrt{\mathcal{P}},-\mathbf{x}}(\Lambda_0)} \sum_{\mathbf{t}\in\Lambda_0+\mathbf{x}} \frac{e^{\left(-\frac{\|\mathbf{w}\|^2}{4\mathcal{P}}-\frac{\|\mathbf{t}-\frac{\mathbf{w}}{2}\|^2}{\mathcal{P}}\right)}}{(2\pi\mathcal{P})^d} \\
&= \frac{g_{\sqrt{2\mathcal{P}}}(\mathbf{w})}{Mg_{\sqrt{\mathcal{P}}}(\Lambda_0)g_{\sqrt{\mathcal{P}},-\mathbf{x}}(\Lambda_0)} \sum_{\mathbf{t}\in\Lambda_0+\mathbf{x}} g_{\sqrt{\frac{\mathcal{P}}{2}},\frac{\mathbf{w}}{2}}(\mathbf{t}) \\
&= \frac{g_{\sqrt{2\mathcal{P}}}(\mathbf{w})}{Mg_{\sqrt{\mathcal{P}}}(\Lambda_0)} \frac{g_{\sqrt{\frac{\mathcal{P}}{2}},\frac{\mathbf{w}}{2}-\mathbf{x}}(\Lambda_0)}{g_{\sqrt{\mathcal{P}},-\mathbf{x}}(\Lambda_0)}.
\end{aligned} \tag{4.18}$$

Substituting this in (4.17), and writing this in terms of α and β, we obtain (4.16). Similarly, bounding both $p_{\mathbf{U}}$ and $p_{\mathbf{V}}$ from above and below using (4.10), proceeding as above, and finally using (4.12) to bound $g_{\sqrt{\frac{\mathcal{P}}{2}},\frac{\mathbf{w}}{2}}(\Lambda)$, we get (4.15). □

Observe that $\beta(\mathbf{x},\mathbf{w})$ in (4.14) is a ratio of two terms, both of which can be bounded using Lemma 4.3.4 to get

$$\left(\frac{1-\epsilon}{1+\epsilon}\right) \leq \beta(\mathbf{x},\mathbf{w}) \leq \left(\frac{1+\epsilon}{1-\epsilon}\right). \tag{4.19}$$

Let $\overline{p}_{\mathbf{U}+\mathbf{V}}$ and $\underline{p}_{\mathbf{U}+\mathbf{V}}$ respectively denote the upper and lower bounds for $p_{\mathbf{U}+\mathbf{V}}$ in

(4.15), and let $\overline{p}_{\mathbf{U}+\mathbf{V}|\mathbf{x}}$ and $\underline{p}_{\mathbf{U}+\mathbf{V}|\mathbf{x}}$ respectively denote the upper and lower bounds for $p_{\mathbf{U}+\mathbf{V}|\mathbf{x}}$ in (4.16). Then, we can say that $|p_{\mathbf{U}+\mathbf{V}|\mathbf{x}}(\mathbf{w}) - p_{\mathbf{U}+\mathbf{V}}(\mathbf{w})|$ is less than or equal to the maximum of $|\overline{p}_{\mathbf{U}+\mathbf{V}|\mathbf{x}}(\mathbf{w}) - \underline{p}_{\mathbf{U}+\mathbf{V}}(\mathbf{w})|$ and $|\underline{p}_{\mathbf{U}+\mathbf{V}|\mathbf{x}}(\mathbf{w}) - \overline{p}_{\mathbf{U}+\mathbf{V}}(\mathbf{w})|$.

Substituting for $|\overline{p}_{\mathbf{U}+\mathbf{V}|\mathbf{x}}(\mathbf{w}) - \underline{p}_{\mathbf{U}+\mathbf{V}}(\mathbf{w})|$, we get

$$|\overline{p}_{\mathbf{U}+\mathbf{V}|\mathbf{x}}(\mathbf{w}) - \underline{p}_{\mathbf{U}+\mathbf{V}}(\mathbf{w})| = \alpha(\mathbf{w})\left(\frac{1-\epsilon}{1+\epsilon}\right)\left|\left(\frac{1+\epsilon}{1-\epsilon}\right)^2 \beta(\mathbf{x},\mathbf{w}) - 1\right|. \tag{4.20}$$

However, from (4.19), we see that

$$1 < \left(\frac{1+\epsilon}{1-\epsilon}\right) \le \left(\frac{1+\epsilon}{1-\epsilon}\right)^2 \beta(\mathbf{x},\mathbf{w}) \le \left(\frac{1+\epsilon}{1-\epsilon}\right)^3,$$

and for $\epsilon \le 1/2$, we have $\left(\frac{1+\epsilon}{1-\epsilon}\right)^3 \le 1 + 64\epsilon$. Therefore,

$$|\overline{p}_{\mathbf{U}+\mathbf{V}|\mathbf{x}}(\mathbf{w}) - \underline{p}_{\mathbf{U}+\mathbf{V}}(\mathbf{w})| \le \alpha(\mathbf{w})\left(\frac{1-\epsilon}{1+\epsilon}\right) 64\epsilon. \tag{4.21}$$

Similarly, expressing $|\underline{p}_{\mathbf{U}+\mathbf{V}|\mathbf{x}}(\mathbf{w}) - \overline{p}_{\mathbf{U}+\mathbf{V}}(\mathbf{w})|$ in terms of α and β, and using the fact that $((1-\epsilon)/(1+\epsilon))^3 \ge 1 - 8\epsilon$ for $\epsilon < 1/2$, we get

$$|\underline{p}_{\mathbf{U}+\mathbf{V}|\mathbf{x}}(\mathbf{w}) - \overline{p}_{\mathbf{U}+\mathbf{V}}(\mathbf{w})| \le \alpha(\mathbf{w})\left(\frac{1+\epsilon}{1-\epsilon}\right)^2 8\epsilon. \tag{4.22}$$

Rearranging (4.15), and observing that $\sum_{\mathbf{w}\in\Lambda} p_{\mathbf{U}+\mathbf{V}}(\mathbf{w}) = 1$, we have

$$\left(\frac{1-\epsilon}{1+\epsilon}\right)^2 \le \sum_{\mathbf{w}\in\Lambda} \alpha(\mathbf{w}) \le \left(\frac{1+\epsilon}{1-\epsilon}\right). \tag{4.23}$$

Combining (4.21) and (4.22), and summing over $\mathbf{w}$, we get

$$\mathbb{V}(p_{\mathbf{U}+\mathbf{V}}, p_{\mathbf{U}+\mathbf{V}|\mathbf{x}}) \le \sum_{\mathbf{w}\in\Lambda} \alpha(\mathbf{w}) \max\left\{\left(\frac{1-\epsilon}{1+\epsilon}\right) 64\epsilon, \left(\frac{1+\epsilon}{1-\epsilon}\right)^2 8\epsilon\right\},$$

and using (4.23) to bound $\sum_{\mathbf{w}\in\Lambda}\alpha(\mathbf{w})$ from above, we get

$$\begin{aligned}\mathbb{V}(p_{\mathbf{U}+\mathbf{V}},p_{\mathbf{U}+\mathbf{V}|\mathbf{x}}) &\leq \max\left\{64\epsilon, \left(\frac{1+\epsilon}{1-\epsilon}\right)^3 8\epsilon\right\} \\ &\leq \max\left\{64\epsilon, 27\times 8\epsilon\right\},\end{aligned}$$

since $\epsilon \leq 1/2$. Therefore,

$$\mathbb{V}(p_{\mathbf{U}+\mathbf{V}},p_{\mathbf{U}+\mathbf{V}|\mathbf{x}}) \leq 216\epsilon,$$

thereby completing the proof. □

4.4 Strong Secrecy and Reliability of Decoding in the Presence of AWGN

We have shown in the previous section that the scheme in Section 4.2 achieves strong secrecy in the absence of noise. Since the noise $\mathbf{Z}$ is independent of everything else, we have strong secrecy in a noisy channel as well. To see why this is the case, observe that $X \rightarrow (\mathbf{U}+\mathbf{V}) \rightarrow (\mathbf{U}+\mathbf{V}+\mathbf{Z})$ forms a Markov chain. Using the data-processing inequality, we see that $\mathcal{I}(X;\mathbf{U}+\mathbf{V}+\mathbf{Z}) \leq \mathcal{I}(X;\mathbf{U}+\mathbf{V})$, verifying our claim. Note that the claim holds regardless of the probability distribution of the noise $\mathbf{Z}$. The fact that the noise is Gaussian will be used to determine achievable rates for reliable decoding of $X\oplus Y$ at the relay.

In this section, we will assume that the additive noise is Gaussian and show that the rate guaranteed by Theorem 4.2.1 is achievable.

We choose our sequence of nested lattices $\{\Lambda^{(n)},\Lambda_0^{(n)}\}$ so as to satisfy the following properties:

(L1) The sequence of coarse lattices, $\{\Lambda_0^{(n)}\}$, is good for covering, MSE quantization, and AWGN channel coding[4].

[4] For the definitions of lattices good for covering, MSE quantization, and AWGN channel coding, see Chapter 2.

(L2) The sequence of coarse lattices, $\{\Lambda_0^{(n)}\}$, is secrecy-good.

(L3) The sequence of fine lattices, $\{\Lambda^{(n)}\}$, is good for AWGN channel coding.

Using (44) in [58, Appendix II] and [58, Proposition 2], we can show that if Λ_0 is a lattice sampled uniformly at random from a (n, k, q) ensemble, where n, k, q satisfy (2.23) and (2.24), then for all sufficiently large n, we have $\mathbb{E}[\epsilon_{\Lambda_0}(\theta)] \leq 2\left(\frac{(\det(\Lambda_0))^{2/n}}{2\pi\theta^2}\right)^{n/2}$, which goes to zero exponentially in n as long as $\frac{(\det(\Lambda_0))^{2/n}}{2\pi\theta^2} < 1$. Using the Markov inequality, we can say that the probability of choosing a lattice whose flatness factor is less than $4\left(\frac{(\det(\Lambda_0))^{2/d}}{2\pi\theta^2}\right)^{d/2}$ is at least 1/2 for all sufficiently large n. From Theorem 2.4.1, we know that a randomly chosen nested lattice pair satisfies (L1) and (L3) with probability tending to 1 as $n \to \infty$. We can then use the union bound to conclude that a randomly chosen pair of nested lattices from the (n, k, q, k_1, q_1) ensemble satisfies (L1)–(L3) with probability at least 1/2 as $n \to \infty$.

We now work towards an estimate of the probability of error of decoding $X \oplus Y$ from $\mathbf{W}$. Recall that the relay computes $\widetilde{\mathbf{W}} = [\alpha^* \mathbf{W}] \bmod \Lambda_0^{(n)}$, where $\alpha^* = \frac{2\mathcal{P}}{2\mathcal{P}+\sigma^2}$, and the estimate of $X \oplus Y$ is the coset to which $Q_{\Lambda^{(n)}}(\widetilde{\mathbf{W}})$ belongs. The quantity $\widetilde{\mathbf{W}}$ can be written as

$$\begin{aligned}\widetilde{\mathbf{W}} &= [\alpha^*(\mathbf{U}+\mathbf{V}+\mathbf{Z})] \bmod \Lambda_0^{(n)} \\ &= [\mathbf{U}+\mathbf{V}-(1-\alpha^*)(\mathbf{U}+\mathbf{V})+\alpha^*\mathbf{Z}] \bmod \Lambda_0^{(n)} \\ &= \Big[[\mathbf{X}+\mathbf{Y}] \bmod \Lambda_0^{(n)} + \mathbf{Z}'\Big] \bmod \Lambda_0^{(n)}, \end{aligned} \tag{4.24}$$

where $\mathbf{Z}' = (\alpha^*-1)(\mathbf{U}+\mathbf{V})+\alpha^*\mathbf{Z}$ is the effective noise of the MLAN channel. Unlike in Section 3.5.2, $\mathbf{Z}'$ is not statistically independent of $[\mathbf{X}+\mathbf{Y}] \bmod \Lambda_0^{(n)}$. However, as we shall see, if the flatness factor of the coarse lattice is small, then the effective noise behaves like an almost independent Gaussian vector. Let $f_{\mathbf{Z}'|\mathbf{x},\mathbf{y}}$ denote the density function of $\mathbf{Z}'$ conditioned on $\mathbf{X} = \mathbf{x}$ and $\mathbf{Y} = \mathbf{y}$, and $f_{\mathbf{N}}$ denote the density function of a Gaussian random vector, $\mathbf{N}$, with mean $\mathbf{0}$ and covariance matrix $\big(2(1-\alpha^*)^2\mathcal{P} + (\alpha^*)^2\sigma^2\big)\mathsf{I}_n$. Given two density functions f_1 and f_2 over $\mathbb{R}^n$, the variational distance between f_1 and f_2, denoted by $\mathbb{V}(f_1, f_2)$, is defined as

$$\mathbb{V}(f_1, f_2) \triangleq \int_{\mathbf{x}\in\mathbb{R}^n} |f_1(\mathbf{x}) - f_2(\mathbf{x})|\, d\mathbf{x}.$$

Lemma 4.4.1 ([58], Lemma 8)**.** *Let Λ be a lattice in $\mathbb{R}^n$, $\mathbf{x} \in \mathbb{R}^n$, and $\sigma_1, \sigma_2 > 0$. Let $\mathbf{U}$ be a random vector supported on $\Lambda + \mathbf{x}$, having pmf $g_{\sigma_1}(\mathbf{u})/g_{\sigma_1,\mathbf{x}}(\Lambda)$. If $\mathbf{Z}$ is an iid Gaussian random vector with mean zero and variance σ_2^2, and $\epsilon_\Lambda\left(\frac{\sigma_1\sigma_2}{\sqrt{\sigma_1^2+\sigma_2^2}}\right) < 1/2$, then the density of $\mathbf{U} + \mathbf{Z}$, $f_{\mathbf{U}+\mathbf{Z}}$, satisfies*

$$\mathbb{V}(f_{\mathbf{U}+\mathbf{Z}}, g_{\sqrt{\sigma_1^2+\sigma_2^2}}) \leq 4\,\epsilon_\Lambda\left(\frac{\sigma_1\sigma_2}{\sqrt{\sigma_1^2+\sigma_2^2}}\right).$$

We now show the following.

Lemma 4.4.2. *If $\epsilon_{\Lambda_0^{(n)}}(\sqrt{\alpha^*\mathcal{P}}) < 1/2$, then for every $\mathbf{x}$ and $\mathbf{y}$ in $\mathbb{G}^{(n)}$,*

$$\mathbb{V}(f_{\mathbf{Z}'|\mathbf{x},\mathbf{y}}, f_{\mathbf{N}}) \leq 8\epsilon_{\Lambda_0^{(n)}}(\sqrt{\alpha^*\mathcal{P}}).$$

Proof. Recall that $\mathbf{Z}$ is the additive Gaussian noise vector in the MAC phase having mean zero and variance σ^2, and $\mathbf{Z}'$ denotes the additive noise in the effective MLAN channel, and is equal to $(\alpha^* - 1)(\mathbf{U} + \mathbf{V}) + \alpha^*\mathbf{Z}$. Let $\mathbf{N}$ denote a zero-mean Gaussian vector with covariance matrix $((1-\alpha^*)^2 2\mathcal{P} + (\alpha^*)^2\sigma^2)\mathsf{I}_d$, and $\mathbf{N}'$ denote a zero-mean Gaussian vector with covariance matrix $((1-\alpha^*)^2\mathcal{P} + (\alpha^*)^2\sigma^2)\mathsf{I}_d$. Let $f_{\mathbf{N}}$ and $f_{\mathbf{N}'}$ denote the densities of $\mathbf{N}$ and $\mathbf{N}'$ respectively, and $f_{\mathbf{U}'|\mathbf{x}}$ denote the density of $\mathbf{U}' \triangleq (\alpha^* - 1)\mathbf{U} + \mathbf{N}'$ conditioned on $\mathbf{X} = \mathbf{x}$. Let $f_{\mathbf{V}'|\mathbf{y}}$ denote the density function of $\mathbf{V}' \triangleq (\alpha^* - 1)\mathbf{V} + \alpha^*\mathbf{Z}$ conditioned on $\mathbf{Y} = \mathbf{y}$. Then, we can write

$$\mathbb{V}(f_{\mathbf{Z}'|\mathbf{x},\mathbf{y}}, f_{\mathbf{N}}) \leq \mathbb{V}(f_{\mathbf{Z}'|\mathbf{x},\mathbf{y}}, f_{\mathbf{U}'|\mathbf{x}}) + \mathbb{V}(f_{\mathbf{U}'|\mathbf{x}}, f_{\mathbf{N}}).$$

But

$$
\begin{aligned}
\mathbb{V}(f_{\mathbf{Z}'|\mathbf{x},\mathbf{y}}, f_{\mathbf{U}'|\mathbf{x}}) &= \int_{\mathbf{w}\in\mathbb{R}^n} |f_{Z'|\mathbf{x},\mathbf{y}}(\mathbf{w}) - f_{\mathbf{U}'|\mathbf{x}}(\mathbf{w})|\, d\mathbf{w} \\
&= \int_{\mathbf{w}\in\mathbb{R}^n} \left| \sum_{\mathbf{u}\in\Lambda_0^{(n)}+\mathbf{x}} p_{\mathbf{U}|\mathbf{x}}(\mathbf{u}) \Big(f_{\mathbf{V}'|\mathbf{y}}(\mathbf{w}-(\alpha^*-1)\mathbf{u}) - f_{\mathbf{N}'}(\mathbf{w}-(\alpha^*-1)\mathbf{u}) \Big) \right| d\mathbf{w} \\
&\le \sum_{\mathbf{u}\in\Lambda_0^{(n)}+\mathbf{x}} p_{\mathbf{U}|\mathbf{x}}(\mathbf{u}) \bigg(\int_{\mathbf{w}\in\mathbb{R}^n} \Big| f_{\mathbf{V}'|\mathbf{y}}(\mathbf{w}-(\alpha^*-1)\mathbf{u}) - f_{\mathbf{N}'}(\mathbf{w}-(\alpha^*-1)\mathbf{u}) \Big| \, d\mathbf{w} \bigg) \\
&= \sum_{\mathbf{u}\in\Lambda_0^{(n)}+\mathbf{x}} p_{\mathbf{U}|\mathbf{x}}(\mathbf{u}) \mathbb{V}(f_{\mathbf{V}'|\mathbf{y}}, f_{\mathbf{N}'}) \\
&= \mathbb{V}(f_{\mathbf{V}'|\mathbf{y}}, f_{\mathbf{N}'}).
\end{aligned}
$$

Therefore,

$$
\mathbb{V}(f_{\mathbf{Z}'|\mathbf{x},\mathbf{y}}, f_{\mathbf{N}}) \le \mathbb{V}(f_{\mathbf{V}'|\mathbf{y}}, f_{\mathbf{N}'}) + \mathbb{V}(f_{\mathbf{U}'|\mathbf{x}}, f_{\mathbf{N}}). \tag{4.25}
$$

Using Lemma 4.4.1 and the fact that for any constant $a > 0$, $\epsilon_{a\Lambda_0^{(n)}}(a\theta) = \epsilon_{\Lambda_0^{(n)}}(\theta)$ [58, Remark 4], we get

$$
\begin{aligned}
\mathbb{V}(f_{\mathbf{V}'|\mathbf{y}}, f_{\mathbf{N}'}) &\le 4\epsilon_{\Lambda_0^{(n)}} \left(\frac{\alpha^*\sqrt{\mathcal{P}\sigma^2}}{\sqrt{(1-\alpha^*)^2\mathcal{P} + (\alpha^*)^2\sigma^2}} \right) &(4.26) \\
&\le 4\epsilon_{\Lambda_0^{(n)}} \left(\frac{\alpha^*\sqrt{\mathcal{P}\sigma^2}}{\sqrt{(1-\alpha^*)^2 2\mathcal{P} + (\alpha^*)^2\sigma^2}} \right) &(4.27) \\
&= 4\epsilon_{\Lambda_0^{(n)}} \left(\sqrt{\alpha^*\mathcal{P}} \right), &(4.28)
\end{aligned}
$$

where (4.27) is by the monotonicity of the flatness factor. Equation (4.28) is then obtained by substituting $\alpha^* = 2\mathcal{P}/(2\mathcal{P}+\sigma^2)$ and simplifying. Using similar arguments, we can show that

$$
\mathbb{V}(f_{\mathbf{U}'|\mathbf{x}}, f_{\mathbf{N}}) \le 4\epsilon_{\Lambda_0^{(n)}} \left(\sqrt{\alpha^*\mathcal{P}} \right).
$$

Substituting in (4.25) completes the proof. □ □

4.4.1 Proof of Theorem 4.2.1

If $\mathbb{P}_1$ and $\mathbb{P}_2$ are probability measures on $\mathbb{R}^n$ having densities f_1 and f_2 respectively, then $\sup_{A\subset\mathbb{R}^n}|\mathbb{P}_1(A)-\mathbb{P}_2(A)| = \frac{1}{2}\mathbb{V}(f_1,f_2)$, where the supremum is taken over all measurable subsets of $\mathbb{R}^n$ (assuming that both P_1 and P_2 are defined on a common event space) [26, Section 7.7]. Using this and Lemma 4.4.2, the probability of error of the decoder can be bounded by

$$\begin{aligned}\xi^{(n)} &\leq \Pr\left[\mathbf{Z}'\notin\mathcal{V}(\Lambda^{(n)})\right]\\ &\leq \Pr\left[\mathbf{N}\notin\mathcal{V}(\Lambda^{(n)})\right] + 4\epsilon_{\Lambda_0^{(n)}}(\sqrt{\alpha^*\mathcal{P}}). \end{aligned} \tag{4.29}$$

The variance of $\mathbf{N}$ is equal to $\sigma_N^2 = 2(1-\alpha^*)^2\mathcal{P} + (\alpha^*)^2\sigma^2 = \frac{2\mathcal{P}\sigma^2}{2\mathcal{P}+\sigma^2}$. If the flatness factor $\epsilon_{\Lambda_0^{(n)}}(\sqrt{\alpha^*\mathcal{P}}) \to 0$ as $n\to\infty$, and the fine lattices are good for AWGN channel coding, then the probability of error at the relay goes to zero as long as $\frac{(\det(\Lambda^{(n)}))^{2/n}}{2\pi e\sigma_N^2} > 1$, or equivalently,

$$\frac{1}{|\mathbb{G}^{(n)}|^{2/n}}\frac{(\det(\Lambda_0^{(n)}))^{2/n}}{2\pi e\sigma_N^2} > 1.$$

In other words,

$$R^{(n)} = \frac{1}{n}\log_2|\mathbb{G}^{(n)}| < \frac{1}{2}\log_2\left(\frac{(\det(\Lambda_0^{(n)}))^{2/n}}{2\pi e\sigma_N^2}\right). \tag{4.30}$$

If we have $\alpha^* \geq 1/2$, then by monotonicity of the flatness factor, $\epsilon_{\Lambda_0^{(n)}}(\sqrt{\alpha^*\mathcal{P}}) \leq \epsilon_{\Lambda_0^{(n)}}(\sqrt{\mathcal{P}/2})$. This requires $\frac{2\mathcal{P}}{2\mathcal{P}+\sigma^2}\geq 1/2$, or $\mathcal{P}\geq\sigma^2/2$. Observe that having $\epsilon_{\Lambda_0^{(n)}}(\sqrt{\mathcal{P}/2}) \to 0$ has three important consequences: (a) strong secrecy, even in the absence of noise (Theorem 4.3.2); (b) the average transmit power converges to $\mathcal{P}$ (Lemma 4.3.3); and (c) the effective noise vector is "almost" independent of the message (Lemma 4.4.2).

Using (L2), in order to have the flatness factor $\epsilon_{\Lambda_0^{(n)}}(\sqrt{\mathcal{P}/2}) \to 0$, the coarse lattices must be scaled so that

$$\frac{\left(\det(\Lambda_0^{(n)})\right)^{2/n}}{2\pi(\mathcal{P}/2)} < 1. \tag{4.31}$$

Let us choose $\left(\det(\Lambda_0^{(n)})\right)^{2/n} = \pi\mathcal{P}-\delta$, for some arbitrary $\delta > 0$, so as to satisfy (4.31).

Substituting this in (4.30), we get that for $\mathcal{P} \geq \sigma^2/2$, as long as

$$R^{(n)} < \frac{1}{2}\log_2\left(\frac{\mathcal{P}-\delta/\pi}{2e\sigma_N^2}\right),$$

the probability of error of decoding $X \oplus Y$ at the relay, as well as the mutual information between the individual messages and $\mathbf{W}$, go to zero as $n \to \infty$. Substituting for σ_N^2, we complete the proof of Theorem 4.2.1. □

Remark 4.4.3. *In the perfect secrecy setting, we were only able to show that a rate of* $\frac{1}{2}\log_2\frac{\mathcal{P}}{\sigma^2} - \log_2(2e)$ *is achievable. We tried to show that using a linear MMSE estimator at the decoder (as we did in Section 4.2), a higher rate of* $\frac{1}{2}\log_2\left(\frac{1}{2}+\frac{\mathcal{P}}{\sigma^2}\right) - \log_2(2e)$ *is achievable. As in the strong-secrecy case, suppose that the relay computes* $\widetilde{\mathbf{W}} \triangleq [\alpha^*\mathbf{W}] \bmod \Lambda_0^{(n)}$, *where* $\alpha^* \triangleq (2\mathcal{P})/(2\mathcal{P}+\sigma^2)$ *is the MMSE coefficient. The effective noise vector,* $\mathbf{Z}_{\text{eff}} = -(1-\alpha^*)(\mathbf{U}+\mathbf{V}) + \alpha^*\mathbf{Z}$ *is not Gaussian, since* $\mathbf{U}$ *and* $\mathbf{V}$ *are not Gaussian. In order to find the probability of decoding error, we require an upper bound on the probability that* $\mathbf{Z}_{\text{eff}} \notin \mathcal{V}(\Lambda^{(n)})$, *which is not straightforward unlike in the Gaussian case. Consequently, we were not able to say whether lattice decoding achieves vanishingly small error probabilities in this situation.*

4.4.2 Related Work on Strong Secrecy and Handling Byzantine Adversaries

The strongly secure scheme proposed by He and Yener in [41] also used nested lattice codes as we have done here. They obtain strong secrecy using universal hash functions, and show the existence of a suitable linear hash function that ensures that the mutual information decays exponentially in n. Unlike [41], we have used a sampled Gaussian pmf for randomization at the encoder, and hence, for a given pair of nested lattices, we explicitly specify the distribution used for randomization. Even using our scheme, the mutual information goes down to zero exponentially in n. But unlike [41], which was valid under a maximum power constraint at each node, the codebook we use is unbounded, so our scheme can only satisfy an average power constraint. Also, the achievable rate in the scheme of He and Yener is slightly higher (by $\frac{1}{2}\log_2 \frac{e}{2}$ bits per channel use). On the other hand, the He-Yener randomization scheme uses hash functions whose existence is only guaranteed by a probabilistic argument, while our randomization scheme has the advantage of being specified by sampled Gaussian pmfs that can be given in explicit form. The scheme in [41] was coupled with an Algebraic Manipulation Detection (AMD) code [16] for Byzantine detection, and it was shown that the probability of a Byzantine attack being undetected could be made to decay to zero exponentially in n. We remark that our coding scheme can also be extended to this scenario, where it can be used as a replacement for the nested lattice code in [41].

4.5 Extensions: Multi-hop Line Network

The bidirectional relay can be viewed as a building block in many wireless networks. In particular, the problem of secure compute-and-forward can be extended to scenarios where we want secure relaying of messages from one point to another on a network with multiple honest-but-curious relays. As an example, we will extend our results to the multi-hop line network studied in [40, 42]. The structure of a multi-hop line network with $K+1$ hops is shown in Fig. 4.1. It consists of $K+2$ nodes: a source node, $\mathtt{S}$, a destination node, $\mathtt{D}$, and K relay nodes, $\mathtt{R}_1, \mathtt{R}_2, \ldots, \mathtt{R}_K$. It is assumed that all links are identical AWGN (mean zero, variance σ^2) wireless links. All nodes are half-duplex and can communicate only with their neighbours. Nodes broadcast their messages to their immediate neighbours.

The source wants to send N messages, $X_1, X_2, \ldots, X_N$, to the destination across the network of honest-but-curious relays. The messages are assumed to be independent and uniformly distributed over the set of all messages. It is assumed that the relays do not co-operate with each other, i.e., the information available at a relay is not shared with the other relays. As remarked by He and Yener in [40], this also takes care of the situation wherein the eavesdropper has access to one of the relays, but it is not known which relay has been compromised. We study this problem mainly under the strong secrecy constraint, but the arguments can be extended to the perfect secrecy scenario.

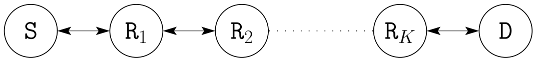


Figure 4.1: Multi-hop line network with $K+1$ hops.

He and Yener showed that their scheme [40, 42] achieves weak secrecy over the multi-hop line network, but their arguments cannot be directly extended for strong secrecy. We give a new proof that shows that our strongly secure scheme for the bidirectional relay can be used with the He and Yener co-operative jamming protocol to obtain strong secrecy in a multi-hop line network.[5]

4.5.1 The Communication Scheme

We use the co-operative jamming scheme proposed by He and Yener for relaying. The communication takes place in $2N+K$ phases, where each phase consists of n channel uses. Let us choose a sequence of $(\Lambda^{(n)}, \Lambda_0^{(n)})$ nested lattice pairs that satisfy properties (L1)–(L3). Each node in the network employs the encoding and decoding scheme described in Section 4.2. Let $\mathcal{D} : \mathbb{R}^n \to \mathbb{G}^{(n)}$ denote the decoder map of Section 4.2. Also, for any $X \in \mathbb{G}^{(n)}$, let $\mathcal{E}(X)$ denote the encoded form of X as in Section 4.2.

- Each relay node i ($i = 1, \ldots, K$) generates a *jamming signal,* J_i, which is chosen uniformly at random from $\mathbb{G}^{(n)}$, and independently of everything else. The destination generates N independent jamming signals, J_{K+l}, for $l = 1, 2, \ldots, N$, where N is the number of messages to be relayed.

- Let $\mathbf{W}_i[n]$ denote the n-dimensional vector received by the ith node in the nth phase, and let $\mathbf{V}_i[n]$ be the vector transmitted by the ith node in the nth phase.

An average power constraint is imposed at the nodes: $\frac{1}{n}\mathbb{E}\|\mathbf{V}_i[m]\|^2 \leq P^{(m)}$ for $i = 0, 1, \ldots, K+1$ and $m = 1, 2, \ldots, K+2N$.

[5] In fact, our proof shows that any strongly secure coding scheme for the bidirectional relay can be used to obtain strong secrecy in the multihop network. However, the achievable rate would depend on the coding scheme.

Since it takes $K+2N$ phases for sending N messages, the rate of the scheme is defined as

$$R_N^{(n)} \triangleq \frac{N}{n(K+2N)} \log_2 |\mathbb{G}^{(n)}|. \tag{4.32}$$

We say that a *power-rate pair* of $(\mathcal{P}, \mathcal{R})$ is *achievable for N-message transmission* with strong secrecy in a multi-hop line network with $K+1$ hops, if there exists a sequence of $(\Lambda^{(n)}, \Lambda_0^{(n)})$ nested lattice codes such that for every $\delta > 0$ and all sufficiently large n, we have

- $P^{(n)} < \mathcal{P} + \delta$;
- $R_N^{(n)} > \mathcal{R} - \delta$;
- the probability of the destination decoding $X_1, X_2, \dots, X_N$ incorrectly, $\xi^{(n)}$, is less than δ; and,
- for $k = 1, 2, \dots, K$, the mutual information between the N messages and all the variables available at the kth relay is less than δ, i.e.,

$$\mathcal{I}(X_1, \dots, X_N; J_k, \mathbf{W}_k[1], \dots, \mathbf{W}_k[2N+K]) < \delta.$$

We will describe the scheme for secure message relaying in the next subsection, and find achievable power-rate pairs. As the main result, letting the number of messages to go to infinity, we will show the following:

Theorem 4.5.1. *A power-rate pair of*

$$\left(\mathcal{P}, \left[\frac{1}{4}\log_2\left(\frac{1}{2} + \frac{\mathcal{P}}{\sigma^2}\right) - \frac{1}{4}\log_2 2e\right]^+\right)$$

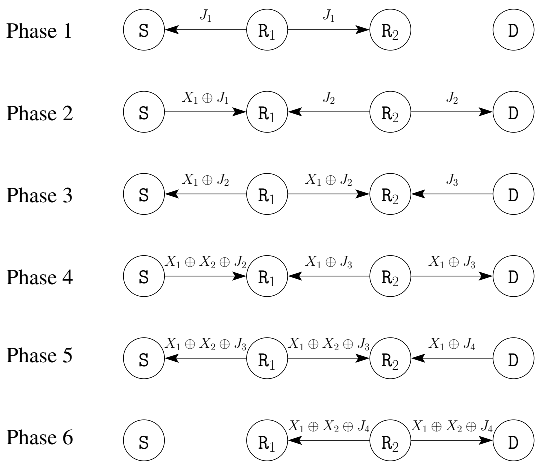


Figure 4.2: Secure relaying of two messages in a 3-hop relay network.

is achievable with strong secrecy[6]*, and a power-rate pair of*

$$\left(\mathcal{P}, \left[\frac{1}{4}\log_2\left(\frac{\mathcal{P}}{\sigma^2}\right) - \frac{1}{2}\log_2 2e\right]^+\right)$$

is achievable with perfect secrecy at the relay nodes in a multi-hop line network with $K+1$ *hops.*

Scheme of He and Yener for Multi-Hop Relaying

We now describe the scheme for secure relaying. A more detailed description can be found in [40]. The case where S wants to send two messages, X_1 and X_2, to the destination is illustrated for a network with two relays in Fig. 4.2. Only the messages (elements of $\mathbb{G}^{(n)}$) transmitted by each node are indicated in the figure, and it is assumed that actual transmitted vectors are the encoded versions of the messages indicated. The messages available at various nodes at the end of each phase are tabulated in Table 4.1. Let us use the notation $\oplus_{p=1}^{t} X_p$ to denote $X_1 \oplus X_2 \oplus \cdots \oplus X_t$.

- The ith node ($i = 0, 1, 2, \ldots, K+1$) transmits in the $(2t+i)$th phase, for $t = 0, 1, \ldots, N$.

[6] *If the scheme in [41] is used at each node, then the achievable rate with strong secrecy can be improved to* $\left[\frac{1}{4}\log_2\left(\frac{1}{2}+\frac{\mathcal{P}}{\sigma^2}\right)-\frac{1}{2}\right]^+$.

Phase	Messages available at node at the end of phase			
	$\mathtt{S}$	$\mathtt{R}_1$	$\mathtt{R}_2$	$\mathtt{D}$
0	$X_1,\ X_2$	J_1	J_2	$J_3,\ J_4$
1	$X_1,\ X_2,\ J_1$	J_1	$J_1,\ J_2$	$J_3,\ J_4$
2	$X_1,\ X_2,\ J_1$	$J_1,\ X_1 \oplus J_2$	$J_1,\ J_2$	$J_2,\ J_3,\ J_4$
3	$X_1,\ X_2,\ J_1,\ J_2$	$J_1,\ X_1 \oplus J_2$	$J_1,\ J_2,\ X_1 \oplus J_3$	$J_2,\ J_3,\ J_4$
4	$X_1,\ X_2,\ J_1,\ J_2$	$J_1,\ X_1 \oplus J_2,\ X_1 \oplus X_2 \oplus J_3$	$J_1,\ J_2,\ X_1 \oplus J_3$	$X_1,\ J_2,\ J_3,\ J_4$
5	$X_1,\ X_2,\ J_1,\ J_2,\ J_3$	$J_1,\ X_1 \oplus J_2,\ X_1 \oplus X_2 \oplus J_3$	$J_1,\ J_2,\ X_1 \oplus J_3,\ X_1 \oplus X_2 \oplus J_4$	$X_1,\ J_2,\ J_3,\ J_4$
6	$X_1,\ X_2,\ J_1,\ J_2,\ J_3$	$J_1,\ X_1 \oplus J_2,\ X_1 \oplus X_2 \oplus J_3,\ X_1 \oplus X_2 \oplus J_4$	$J_1,\ J_2,\ X_1 \oplus J_3,\ X_1 \oplus X_2 \oplus J_4$	$X_1,\ X_2,\ J_2,\ J_3,\ J_4$

Table 4.1: Messages available at various nodes at the end of each phase for the protocol in Fig. 4.2.

- In the $(2t+i)$th phase $(t = 0, 1, 2, \ldots, N)$, the ith node sends

$$\mathbf{V}_i[2t+i] = \mathcal{E}\big((\oplus_{p=1}^{t} X_p) \oplus J_{i+t}\big). \tag{4.33}$$

  This holds for all nodes, $i = 0, 1, \ldots, K+1$. The ith node evaluates $(\oplus_{p=1}^{t} X_p) \oplus J_{i+t}$ by subtracting the message transmitted by it in the $(2t+i-2)$nd phase from the message decoded in the $(2t+i-1)$st phase.

Since the destination knows $J_{K+1}, \ldots, J_{K+N}$, it can compute $\oplus_{p=1}^{t} X_p$ from $\mathcal{E}\big((\oplus_{p=1}^{t} X_p) \oplus J_{K+t}\big)$, for $t = 0, 1, \ldots N$, and hence, each of the messages X_l.

4.5.2 Secrecy

Let us assume that all links are noiseless. As argued at the end of Section 4.3, it is enough to show that strong secrecy is obtained in this situation. Let $\{X_p : p = 1, \ldots, N\}$ denote the set of i.i.d. messages to be sent to the destination. Let us fix a k from $\{1, 2, \ldots, K\}$.

In the $(2t+k-1)$st phase, the kth relay receives

$$\begin{aligned}\mathbf{W}_k[2t+k-1] &= \mathbf{V}_{k-1}[2t+k-1]+\mathbf{V}_{k+1}[2t+k-1] &(4.34)\\ &= \mathcal{E}\Big((\oplus_{p=1}^{t}X_p)\oplus J_{k+t-1}\Big)+\mathcal{E}\Big((\oplus_{p=1}^{t-1}X_p)\oplus J_{k+t}\Big), &(4.35)\end{aligned}$$

for $1 \le t \le N$, and $\mathbf{W}_k[k-1]=\mathcal{E}(J_{k-1})$. For $t=1,2,\ldots,N$, let us define

$$\Theta_{k,t} \triangleq \{J_k, J_{k-1}, \mathbf{W}_k[2m+k-1] : 1 \le m \le t\} \tag{4.36}$$

to be the set of all random variables available at the kth relay at the end of the $(2t+k-1)$st phase. We also define $\Theta_{k,0} \triangleq \{J_k, J_{k-1}\}$. Note that $\Theta_{k,t-1} \subset \Theta_{k,t}$ for $t=1,2,\ldots,N$, and $\Theta_{k,N}$ is the set of all random variables available at the kth relay at the end of all phases. We have to show that $\mathcal{I}(X_1,\ldots,X_N;\Theta_{k,N}) \to 0$ as $d \to \infty$.

Lemma 4.5.2. *Let $\epsilon^{(n)} \triangleq \epsilon_{\Lambda_0^{(n)}}(\sqrt{\mathcal{P}/2}) < 1/2$. Then, the total information available at the kth relay node at the end of all relaying phases can be bounded from above as follows:*

$$\mathcal{I}(X_1,\ldots,X_N;\Theta_{k,N}) \le N\epsilon^{(n)}\left(\log_2|\mathbb{G}^{(n)}| - \log_2\epsilon^{(n)}\right). \tag{4.37}$$

Since for our choice of nested lattices, $\epsilon^{(n)} \to 0$ exponentially in n, the mutual information $\mathcal{I}(X_1,\ldots,X_N;\Theta_{k,N})$ also goes to zero exponentially in n, thereby guaranteeing strong secrecy. Before proving Lemma 4.5.2, let us use this to complete the proof of Theorem 4.5.1.

4.5.3 Achievable Rate and Proof of Theorem 4.5.1

Using the union bound, one can show that for each N, the probability of the kth relay being in error in the ith phase goes to zero as $d \to \infty$ for all k and i. Using Theorem 4.2.1, we can say that a power-rate pair of $\Big(\mathcal{P}, \frac{N}{2(K+2N+1)}\left[\log_2\left(\frac{1}{2}+\frac{\mathcal{P}}{\sigma^2}\right)-\log_2 2e\right]^+\Big)$ is achievable for the transmission of N messages using this scheme. Letting the number of messages, N, go to infinity, we have the first part of Theorem 4.5.1. The second part of the theorem can be proved in a similar manner. □

4.5.4 Proof of Lemma 4.5.2

We want to show that $\mathcal{I}(X_1, \dots, X_N; \Theta_{k,N})$ is arbitrarily small for all sufficiently large n. Using the chain rule of mutual information, and making some observations about the conditional independence of these random variables, we will show that this quantity can be written as a sum of mutual information terms between the ith message, X_i, and the vector $\mathbf{W}_k[2i+k-1]$, conditioned on everything observed by the kth relay in the first $2i+k-2$ phases. We will then bound each of these mutual information terms from above by a quantity of the form $\mathcal{I}(X; \mathcal{E}(X) + \mathcal{E}(Y))$, so that we can invoke the results of Section 4.1 to conclude that each of these terms go to zero as $n \to \infty$. We would like to remark that the techniques used in this proof hold good for any coding scheme that achieves strong secrecy over the bidirectional relay, and in particular, the one in [41].

Making repeated use of the chain rule of mutual information, we see that

$$\begin{aligned}\mathcal{I}(X_1, \dots, X_N; \Theta_{k,N}) &= \sum_{t=1}^{N} \mathcal{I}(X_t; \Theta_{k,N} \big| X_1, \dots, X_{t-1}) \\ &= \sum_{t=1}^{N} \Big[\mathcal{I}(X_t; J_k, J_{k-1} \big| X_1, \dots, X_{t-1}) \\ &\qquad + \sum_{m=1}^{N} \mathcal{I}(X_t; \mathbf{W}_k[2m+k-1] \big| X_1, \dots, X_{t-1}, \Theta_{k,m-1}) \Big] \\ &= \sum_{t=1}^{N} \sum_{m=1}^{N} \mathcal{I}(X_t; \mathbf{W}_k[2m+k-1] \big| X_1, \dots, X_{t-1}, \Theta_{k,m-1}), \qquad (4.38)\end{aligned}$$

where the last step follows from the fact that $\mathcal{I}(X_t; J_k, J_{k-1} \big| X_1, \dots, X_{t-1}) = 0$ for $1 \leq t \leq N$, since the messages and the jamming signals are independent.

We will first show that many terms in the above summation are zero. We will make use of the fact that if X, Y, and Z are random variables distributed over a finite group $\mathbb{G}$, with X being uniformly distributed over $\mathbb{G}$ and independent of (Y, Z), then $X \oplus Y$ is uniformly distributed over $\mathbb{G}$ and independent of Z. Observe that for $m \in \{1, 2, \dots, N\}$, $\Theta_{k,m-1}$ consists of random variables which are all functions of $X_1, \dots, X_{m-1}$ and $J_{k-1}, \dots, J_{k+m-1}$, which are all independent of X_t for $m \leq t$ (even when conditioned on the first $l - 1 < t$ messages). Therefore,

Proposition 4.5.3. *Let $1 \leq t \leq N$, and $m, l \in \{1, 2, \dots, t\}$. Then, the message X_t is conditionally independent of $\Theta_{k,m-1}$ given $X_1, X_2, \dots, X_{l-1}$.*

Using a similar argument, we obtain

Proposition 4.5.4. *Let $1 \leq t < m \leq N$.The vector $\mathbf{W}_k[2m + k - 1]$ received by the kth relay in the $(2m + k - 1)$st phase is independent of $X_1, \dots, X_t$ and $\Theta_{k,m-1}$.*

We now evaluate the terms in (4.38). Using Proposition 4.5.3, we get

$$\mathcal{I}(X_t; \mathbf{W}_k[2m + k - 1] \big| X_1, \dots, X_{t-1}, \Theta_{k,m-1}) = 0 \tag{4.39}$$

for all $1 \leq m < t \leq N$. Similarly, using Proposition 4.5.4,

$$\mathcal{I}(X_t; \mathbf{W}_k[2m + k - 1] \big| X_1, \dots, X_{t-1}, \Theta_{k,m-1}) = 0 \tag{4.40}$$

for all $1 \leq t < m \leq N$. Therefore, (4.38) reduces to

$$\mathcal{I}(X_1, \dots, X_N; \Theta_{k,N}) = \sum_{t=1}^{N} \mathcal{I}(X_t; \mathbf{W}_k[2t + k - 1] \big| X_1, \dots, X_{t-1}, \Theta_{k,t-1}). \tag{4.41}$$

We can write the mutual information $\mathcal{I}(X_t; \mathbf{W}_k[2t + k - 1] \big| X_1, \dots, X_{t-1}, \Theta_{k,t-1})$ in terms of conditional entropies as

$$\begin{aligned}\mathcal{I}(X_t; \mathbf{W}_k[2t + k - 1] \big| X_1, \dots, X_{t-1}, \Theta_{k,t-1}) = {} & \mathcal{H}(\mathbf{W}_k[2t + k + 1] \big| X_1, \dots, X_{t-1}, \Theta_{k,t-1}) \\ & - \mathcal{H}(\mathbf{W}_k[2t + k + 1] \big| X_1, \dots, X_t, \Theta_{k,t-1}).\end{aligned} \tag{4.42}$$

Let us evaluate each of the terms on the right hand side. Consider the second term,

$$\mathcal{H}(\mathbf{W}_k[2t+k+1]\big|X_1,\ldots,X_t,\Theta_{k,t-1}) \geq \mathcal{H}(\mathbf{W}_k[2t+k+1]\big|X_1,\ldots,X_t,X,\Theta_{k,t-1}) \quad (4.43)$$

$$= \mathcal{H}(\mathbf{W}_k[2t+k+1]\big|X), \quad (4.44)$$

where $X \triangleq \oplus_{p=1}^{t} X_p \oplus J_{k+t-1}$. The inequality in (4.43) is true because conditioning reduces entropy. The equality in (4.44) requires some justification. Given X, the term $\mathbf{V}_{k-1}[2t+k-1]$ is independent of $X_1,\ldots,X_t,\Theta_{k,t-1}$. The jamming signal, J_{k+t} is independent of $\Theta_{k,t-1}$, all the first t messages, and X. Therefore, $\mathbf{V}_{k+1}[2t+k-1]$, and hence, $\mathbf{W}_k[2t+k-1]$ is also independent of $\Theta_{k,t-1}$, the first t messages and X, thus justifying (4.44).

Now, define $Y \triangleq \oplus_{p=1}^{t-1} X_p \oplus J_{k+t}$. Then, we have,

$$\mathcal{H}(\mathbf{W}_k[2t+k+1]\big|X_1,\ldots,X_t,\Theta_{k,t-1}) \geq \mathcal{H}(\mathcal{E}(X)+\mathcal{E}(Y)\big|X).$$

From Proposition 4.5.4, the first term of (4.42), $\mathcal{H}(\mathbf{W}_k[2t+k+1]\big|X_1,\ldots,X_{t-1},\Theta_{k,t-1}) = \mathcal{H}\left(\mathcal{E}(X)+\mathcal{E}(Y)\right)$. Therefore, $\mathcal{I}(X_t;\mathbf{W}_k[2t+k-1]\big|X_1,\ldots,X_{t-1},\Theta_{k,t-1})$ is bounded above by $\mathcal{I}(X;\mathcal{E}(X)+\mathcal{E}(Y))$, and the random variables X and $\widetilde{Y}$ are independent and uniformly distributed over $\mathbb{G}^{(n)}$. The lemma now follows by using Theorem 4.3.2 and Lemma 4.3.1 to bound this quantity. □

Chapter 5

Secure Bidirectional Relaying with Unequal Channel Gains

In this chapter, we extend the results of Chapters 3 and 4, and make an attempt to study the robustness of the schemes presented there. Let us briefly recall the general setup of compute-and-forward [69] for bidirectional relaying. User nodes `A` and `B` have messages X and Y respectively, which are assumed to be uniformly distributed over a finite abelian group $\mathbb{G}$. Let $\oplus$ denote the addition operation in $\mathbb{G}$. The messages $X, Y \in \mathbb{G}$ are mapped to n-dimensional real-valued codewords $\mathbf{U}$ and $\mathbf{V}$ respectively, and transmitted simultaneously to `R`, who receives

$$\mathbf{W} = h_1\mathbf{U} + h_2\mathbf{V} + \mathbf{Z}, \tag{5.1}$$

where $h_1, h_2 \in \mathbb{R}$, and $\mathbf{Z}$ is additive white Gaussian noise (AWGN) with variance σ^2. The relay computes an integer-linear combination of the messages, $k_1X \oplus k_2Y$, and forwards it to the user nodes. In the previous chapters, we assumed that $h_1 = h_2 = 1$, and chose $k_1 = k_2 = 1$. We will now relax the assumption on the channel gains, and say that h_1 and h_2 are arbitrary but fixed nonzero real numbers. Furthermore, we do not restrict k_1 and k_2 to be 1, but the choice of these integers will depend on the values of h_1 and h_2.

Nazer and Gastpar [69] gave a $(\Lambda^{(n)}, \Lambda_0^{(n)})$ nested lattice coding scheme for reliable computation of $k_1X \oplus k_2Y$ at the relay with $\mathbb{G} = \Lambda^{(n)}/\Lambda_0^{(n)}$, and we used a similar protocol in the previous chapters. If $\Lambda^{(n)}$ and $\Lambda_0^{(n)}$ are nested Construction-A lattices (as in [69], and Chapters 3 and 4), then there is a group isomorphism between $\Lambda^{(n)}/\Lambda_0^{(n)}$ and $\mathbb{F}_q^k$ (viewed as an additive group), for some prime q and positive integer k. For the rest of this chapter, we will assume that $\mathbb{G} = \mathbb{F}_q^k$. If q does not divide k_2 (resp. k_1) and none of the nodes make a decoding error, then $\mathtt{A}$ (resp. $\mathtt{B}$) can recover Y (resp. X) from $k_1X \oplus k_2Y$.

In Chapters 3 and 4, we demanded that the relay must not obtain any information about the individual messages. We addressed the problem under two measures of security:

(S1) *Perfect secrecy:* The received vector is independent of the individual messages, i.e., $\mathbf{W} \perp\!\!\!\perp X$ and $\mathbf{W} \perp\!\!\!\perp Y$.

(S2) *Strong secrecy:* The information leaked by $\mathbf{W}$ about the individual messages must be vanishingly small for large n, i.e., $\lim_{n\to\infty} I(X;\mathbf{W}) = \lim_{n\to\infty} I(Y;\mathbf{W}) = 0$

In the previous chapters, we assumed that the channel gains h_1, h_2 are equal to 1. If the channel gains are known at the user nodes, then they can equalize these channel gains so that the problem reduces to the form studied earlier. However, in a practical scenario, the user nodes may not know h_1 and h_2 exactly, since there is always an error in estimation of the channel gains. In this chapter, we assume that the user nodes *do not know* the values of the channel gains h_1 and h_2. We operate under a worst-case assumption and say that the relay knows h_1 and h_2 exactly. We want to know if it is still possible to achieve security in this (rather pessimistic) situation. We split the analysis into two parts: (1) the case when h_1/h_2 is rational, and (2) when h_1/h_2 is irrational. In Sections 5.1 and 5.2, we will find sufficient conditions to guarantee perfect/strong security in case (1). In Section 5.3, we show that no lattice-based coding scheme can guarantee secrecy in the latter case if we assume that the channel is noiseless.

If h_1/h_2 is rational, then we can express $h_1 = hk_1$ and $h_2 = hk_2$ for some real number h and co-prime integers k_1 and k_2. Therefore, in the first few sections, we will assume that the channel gains h_1 and h_2 are *co-prime integers, but are unknown* to both users.

The relay must compute $h_1X \oplus h_2Y$ and forward it to the users in the next phase. The relay also forwards h_1, h_2 to the users in this phase, which will enable them to recover the respective messages. It is therefore assumed that the users have no knowledge of the channel gains prior to the broadcast phase. We will also assume that q *does not divide* h_1 or h_2. For e.g., if q divides h_1, then $h_1X = 0$ for all $X \in \mathbb{F}_q^k$, and hence, $\mathtt{B}$ will never be able to recover X. We will mostly study the noiseless scenario, i.e., the relay receives $\mathbf{W} = h_1\mathbf{U} + h_2\mathbf{V}$, and find conditions under which our scheme achieves security. We will also briefly discuss achievable rates in presence of Gaussian noise, but without any proofs, since this is just a reapplication of the results of Chapter 3 and Chapter 4.

We reiterate that demanding security in the noiseless scenario is a much stronger condition. Since the additive noise $\mathbf{Z}$ is independent of everything else, $X \to h_1\mathbf{U}+h_2\mathbf{V} \to h_1\mathbf{U}+h_2\mathbf{V}+\mathbf{Z}$ forms a Markov chain, and hence, $I(X; h_1\mathbf{U}+h_2\mathbf{V}+\mathbf{Z}) \le I(X; h_1\mathbf{U}+h_2\mathbf{V})$. Therefore, any scheme that achieves perfect/strong secrecy in the noiseless setting also continues to achieve the same in presence of noise. Furthermore, such a scheme has the added advantage that security is achieved *irrespective of the distribution* on $\mathbf{Z}$, and even when this distribution is *unknown* to the users.

The chapter is organized as follows: The coding scheme is described in Section 5.0.1. The main results on perfect secrecy are presented in Section 5.1, with the main result summarized in Theorem 5.1.1. Strong secrecy is studied in Section 5.2, and Theorem 5.2.1 gives the main result. In Section 5.3, we discuss the case where the channel gains are not integral and co-prime, and conclude with some final remarks.

5.0.1 The coding scheme

Recall the basic coding scheme used in Chapter 3–4. It consists of a pair of nested lattices $(\Lambda^{(n)}, \Lambda_0^{(n)})$ in $\mathbb{R}^n$, where $\Lambda_0^{(n)} \subset \Lambda^{(n)}$, and a well chosen continuous pdf $f^{(n)}$ over $\mathbb{R}^n$.

- *Lattices:* $\Lambda^{(n)}$ and $\Lambda_0^{(n)}$ are nested *Construction-A* lattices [28] over $\mathbb{F}_q$ for a prime q. Specifically, let $\Lambda^{(n)}$ be constructed from an (n, k_1) linear code $\mathcal{C}$, and $\Lambda_0^{(n)}$ from an (n, k_0) linear code $\mathcal{C}_0$, with $\mathcal{C}_0 \subset \mathcal{C}$. If $k \triangleq k_1 - k_0$, then there exists a group isomorphism from $\Lambda^{(n)}/\Lambda_0^{(n)}$ to $\mathbb{F}_q^k$ [69]. However, these results hold for the more general case where $\Lambda^{(n)}$ and $\Lambda_0^{(n)}$ are arbitrary n-dimensional nested lattices.

- *Messages:* The messages are mapped bijectively into $\mathbb{G}^{(n)} \triangleq \Lambda^{(n)}/\Lambda_0^{(n)}$. Therefore, each message is identified by a coset of $\Lambda_0^{(n)}$ in $\Lambda^{(n)}$. We also have $M \triangleq |\mathbb{G}^{(n)}| = q^k$, and the rate of the code is $R = \frac{1}{n}\log_2 M = \frac{k}{n}\log_2 q$.

- *Encoding:* Given $X \in \mathbb{G}^{(n)}$, node $\mathtt{A}$ transmits a vector $\mathbf{u} \in \mathbb{R}^n$ with probability

$$p_{\mathbf{U}|X}(\mathbf{u}) = \begin{cases} \frac{f^{(n)}(\mathbf{u})}{\sum_{\mathbf{u}' \in X} f^{(n)}(\mathbf{u}')}, & \text{if } \mathbf{u} \in X \\ 0, & \text{otherwise.} \end{cases} \tag{5.2}$$

  Likewise, $\mathtt{B}$ transmits $\mathbf{v} \in Y$ with probability $p_{\mathbf{V}|Y}(\mathbf{v})$. The scheme can satisfy an average power constraint: $\frac{1}{n}\mathbb{E}\|\mathbf{U}\|^2 = \frac{1}{n}\mathbb{E}\|\mathbf{V}\|^2 \leq \mathcal{P}$.

- *Decoding:* The relay finds the closest point in $\Lambda^{(n)}$ to the received vector $\mathbf{w}$, and determines $h_1 X \oplus h_2 Y$ to be the coset to which this point belongs.

We are interested in two kinds of pdfs $f^{(n)}$ over $\mathbb{R}^n$:

- *Density with a compactly supported characteristic function for perfect secrecy:* Let ψ be the characteristic function corresponding to $f^{(n)}$. Let $\mathcal{R}(\psi)$ be the support of ψ, i.e., the region where ψ is nonzero. We will show that for certain values of (h_1, h_2), if $\mathcal{R}(\psi)$ is supported within a certain compact subset of $\mathbb{R}^n$, then perfect secrecy can be obtained.

- *The Gaussian density for strong secrecy:* For $\mathbf{x}, \mathbf{w} \in \mathbb{R}^n$ and $P > 0$, we define

$$g_{-\mathbf{x},\sqrt{P}}(\mathbf{w}) = \frac{1}{(2\pi P)^{n/2}} e^{-\frac{\|\mathbf{w}-\mathbf{x}\|^2}{2P}},$$

  and $g_{-\mathbf{x},\sqrt{P}}(\Lambda) = \sum_{\mathbf{w}\in\Lambda} g_{-\mathbf{x},\sqrt{P}}(\mathbf{w})$. For ease of notation, we will use $g_{\sqrt{P}}(\mathbf{w})$ and $g_{\sqrt{P}}(\Lambda)$ instead of $g_{\mathbf{0},\sqrt{P}}(\mathbf{w})$ and $g_{\mathbf{0},\sqrt{P}}(\Lambda)$ respectively. We will show that if $\Lambda_0^{(n)}$ satisfies certain properties, then with $f^{(n)} = g_{\sqrt{P}}$, we can obtain strong secrecy.

We say that a rate $\mathcal{R}$ is *achievable* with perfect (resp. strong) secrecy using our scheme if there exists a sequence of nested lattice pairs $\{(\Lambda^{(n)}, \Lambda_0^{(n)})\}$ and a sequence of density functions $\{f^{(n)}\}$ so that for every $\delta > 0$ and all sufficiently large n, the above coding scheme satisfies the following properties:

- $\mathbf{W} \perp\!\!\!\perp X$ and $\mathbf{W} \perp\!\!\!\perp Y$ (resp. $I(\mathbf{W};X) < \delta$ and $I(\mathbf{W};Y) < \delta$),
- the probability of error of decoding $h_1X \oplus h_2Y$ at the relay is less than δ, and
- the transmission rate $R = \frac{1}{n}\log_2 M > \mathcal{R} - \delta$.

For the rest of this chapter, we will drop the superscript n in $\Lambda^{(n)}$, $\Lambda_0^{(n)}$, etc. to simplify notation.

5.1 Perfect secrecy with integral channel gains

5.1.1 The noiseless case

In this section and the next, we assume that h_1 and h_2 are co-prime integers. Recall Proposition 3.4.4, which says that if f is a pdf over $\mathbb{R}^n$ such that the corresponding characteristic function, ψ, is compactly supported within $\mathcal{V}(\hat{\Lambda})$, then $\phi(\mathbf{t}) \triangleq \sum_{\mathbf{u}\in\hat{\Lambda}} \psi(\mathbf{t}+\mathbf{u})e^{-i\langle \mathbf{x},\mathbf{u}\rangle}$ is the characteristic function of a random vector supported within $\Lambda + \mathbf{x}$, and having pmf

$$p(\mathbf{u}) = \begin{cases} \det\Lambda \cdot f(\mathbf{u}) & \text{if } \mathbf{u} \in \Lambda + \mathbf{x} \\ 0 & \text{otherwise.} \end{cases}$$

In other words, if ψ is compactly supported within $\mathcal{V}(\widehat{\Lambda})$, then $\phi(\mathbf{t})$ is the characteristic function corresponding to the pmf obtained by sampling and normalizing f over $\Lambda + \mathbf{x}$.

Given message (coset) x, user $\mathtt{A}$ transmits a random point $\mathbf{U}$ in the coset x according to distribution $p_{\mathbf{U}|x}$ as given by (5.2), and given message y at $\mathtt{B}$, the user transmits $\mathbf{V}$ in the coset y according to distribution $p_{\mathbf{V}|y}(\mathbf{v})$. The density f from which these pmfs are sampled from is compactly supported within $\mathcal{R}(\psi)$. The following result gives sufficient conditions under which perfect security is achieved.

Theorem 5.1.1. *If the order of no non-zero element of Λ/Λ_0 divides h_1 or h_2, and the* *support of ψ, $\mathcal{R}(\psi)$, is contained within the interior of $\frac{2\mathcal{V}(\widehat{\Lambda}_0^{(n)})}{|h_1|+|h_2|}$, then $(h_1\mathbf{U}+h_2\mathbf{V}) \perp\!\!\!\perp X$ and $(h_1\mathbf{U}+h_2\mathbf{V}) \perp\!\!\!\perp Y$.*

If Λ and Λ_0 are Construction-A lattices obtained from linear codes over $\mathbb{F}_q$, then the order of no non-zero element of Λ/Λ_0 divides h_1 or h_2 iff q does not divide h_1 or h_2.

We can choose a characteristic function ψ which is supported within a ball of radius $r = \alpha r_{\text{pack}}(\widehat{\Lambda}_0)$ $(\alpha \leq 1)$, where $r_{\text{pack}}(\widehat{\Lambda}_0)$ denotes the packing radius of $\widehat{\Lambda}_0$. If $r < 2r_{\text{pack}}(\widehat{\Lambda}_0)/(|h_1|+|h_2|)$, then we certainly have $\mathcal{R}(\psi) \subset 2\mathcal{V}(\widehat{\Lambda}_0)/(|h_1|+|h_2|)$, which guarantees perfect secrecy. Therefore, perfect secrecy can be attained for all h_1, h_2 for which q does not divide either h_1 or h_2, and $2/(|h_1|+|h_2|) \geq \alpha$. An interesting point to note at this juncture is that the nested lattice pair does not have to satisfy any additional properties in order to obtain perfect secrecy. The above result holds for any pair of nested lattices, and for any value of n, unlike most results on secrecy which usually require the lattices to satisfy special properties and n to be sufficiently large.

Proof of Theorem 5.1.1

Fix any $x, y \in \mathbb{G}$. We want to show that $p_{h_1\mathbf{U}+h_2\mathbf{V}|x} = p_{h_1\mathbf{U}+h_2\mathbf{V}}$, and $p_{h_1\mathbf{U}+h_2\mathbf{V}|y} = p_{h_1\mathbf{U}+h_2\mathbf{V}}$. We only prove the first statement here, and the second can be proved analogously. Let ψ be the characteristic function corresponding to f, and $\phi_{h_1\mathbf{U}|x}$ be the characteristic function of $h_1\mathbf{U}$ conditioned on $X = x$. Furthermore, let $\phi_{h_1\mathbf{U}}$ and $\phi_{h_2\mathbf{V}}$ be the characteristic functions of $h_1\mathbf{U}$ and $h_2\mathbf{V}$ respectively. We will show that $\phi_{h_1\mathbf{U}|x}\phi_{h_2\mathbf{V}} = \phi_{h_1\mathbf{U}}\phi_{h_2\mathbf{V}}$. Let $\mathbf{x}$ be the coset representative of x within $\mathcal{V}(\Lambda_0)$. Using Proposition 3.4.4, we have

$$\phi_{h_1\mathbf{U}}(\mathbf{t}) = \sum_{\lambda\in\widehat{\Lambda}} \psi\left(\frac{\lambda+\mathbf{t}}{|h_1|}\right), \quad \phi_{h_2\mathbf{V}}(\mathbf{t}) = \sum_{\lambda\in\widehat{\Lambda}} \psi\left(\frac{\lambda+\mathbf{t}}{|h_2|}\right),$$

and

$$\phi_{h_1\mathbf{U}|x}(\mathbf{t}) = \sum_{\lambda\in\widehat{\Lambda}_0} \psi\left(\frac{\lambda+\mathbf{t}}{|h_1|}\right) e^{-i\langle\lambda,\mathbf{x}\rangle}.$$

Since $\Lambda_0 \subset \Lambda$, we have $\widehat{\Lambda} \subset \widehat{\Lambda}_0$. Using this, and the fact that $\langle \lambda, \mathbf{x} \rangle \in 2\pi\mathbb{Z}$ for $\lambda \in \widehat{\Lambda}$, we can write

$$\phi_{h_1\mathbf{U}|x}(\mathbf{t}) = \phi_{h_1\mathbf{U}}(\mathbf{t}) + \sum_{\lambda \in \widehat{\Lambda}_0 \backslash \widehat{\Lambda}} \psi\left(\frac{\lambda + \mathbf{t}}{|h_1|}\right) e^{-i\langle \lambda, \mathbf{x} \rangle} \tag{5.3}$$

Therefore, $\phi_{h_1\mathbf{U}|x}(\mathbf{t})\phi_{h_2\mathbf{V}}(\mathbf{t}) = \phi_{h_1\mathbf{U}}(\mathbf{t})\phi_{h_2\mathbf{V}}(\mathbf{t})$, is equivalent to

$$\phi_{h_2\mathbf{V}}(\mathbf{t}) \sum_{\lambda \in \widehat{\Lambda}_0 \backslash \widehat{\Lambda}} \psi\left(\frac{\lambda + \mathbf{t}}{|h_1|}\right) e^{-i\langle \lambda, \mathbf{x} \rangle} = 0,$$

or,

$$\sum_{\lambda' \in \widehat{\Lambda}} \psi\left(\frac{\lambda' + \mathbf{t}}{|h_2|}\right) \left(\sum_{\lambda \in \widehat{\Lambda}_0 \backslash \widehat{\Lambda}} \psi\left(\frac{\lambda + \mathbf{t}}{|h_1|}\right) e^{-i\langle \lambda, \mathbf{x} \rangle} \right) = 0.$$

It is enough to show that for every $\lambda_1 \in \widehat{\Lambda}_0 \backslash \widehat{\Lambda}$, $\lambda_2 \in \widehat{\Lambda}$, and $\mathbf{t} \in \mathbb{R}^n$,

$$\psi\left(\frac{\lambda_1 + \mathbf{t}}{|h_1|}\right) \psi\left(\frac{\lambda_2 + \mathbf{t}}{|h_2|}\right) = 0.$$

Observe that

$$\text{Supp}\left(\psi\left(\frac{\lambda_1 + \mathbf{t}}{|h_1|}\right)\right) = \frac{\mathcal{R}(\psi) - \lambda_1}{|h_1|},$$

and

$$\text{Supp}\left(\psi\left(\frac{\lambda_2 + \mathbf{t}}{|h_2|}\right)\right) = \frac{\mathcal{R}(\psi) - \lambda_2}{|h_2|}.$$

We will show that for every $\lambda_1 \in \widehat{\Lambda}_0 \backslash \widehat{\Lambda}$ and $\lambda_2 \in \widehat{\Lambda}$,

$$\text{Supp}\left(\psi\left(\frac{\lambda_2 + \mathbf{t}}{|h_2|}\right)\right) \bigcap \text{Supp}\left(\psi\left(\frac{\lambda_1 + \mathbf{t}}{|h_1|}\right)\right) = \{\},$$

or equivalently,

$$\left(\frac{\mathcal{R}(\psi) - \lambda_1}{|h_1|}\right) \bigcap \left(\frac{\mathcal{R}(\psi) - \lambda_2}{|h_2|}\right) = \{\},$$

where $\{\}$ denotes the empty set.

Let us assume the contrary, that there exist $\mathbf{t}_1, \mathbf{t}_2$ in $\mathcal{R}(\psi)$, $\lambda_1 \in \widehat{\Lambda}_0 \backslash \widehat{\Lambda}$ and $\lambda_2 \in \widehat{\Lambda}$

such that $\frac{\mathbf{t}_1-\lambda_1}{|h_1|} = \frac{\mathbf{t}_2-\lambda_2}{|h_2|}$. This can be rewritten as

$$|h_2|\mathbf{t}_1 - |h_1|\mathbf{t}_2 = |h_2|\lambda_1 - |h_1|\lambda_2. \tag{5.4}$$

Clearly, $|h_2|\mathbf{t}_1-|h_1|\mathbf{t}_2$ lies in $(|h_2|+|h_1|)\mathcal{R}(\psi)$, which is contained in the interior of $2\mathcal{V}(\widehat{\Lambda}_0)$. Since $|h_2|\lambda_1 - |h_1|\lambda_2 \in \widehat{\Lambda}_0$, the proof will be complete if we show that this is nonzero. To this end, we write $\lambda_1 = \lambda_1^{(0)} + \lambda_1^{(1)}$, where $\lambda_1^{(0)} \in \widehat{\Lambda}_0 \cap \mathcal{V}(\widehat{\Lambda})$, and $\lambda_1^{(1)} \in \widehat{\Lambda}$. Therefore, $|h_2|\lambda_1^{(1)} - |h_1|\lambda_2 \in \widehat{\Lambda}$. Since $\lambda_1 \in \widehat{\Lambda}_0 \backslash \widehat{\Lambda}$, we are assured that $\lambda_1^{(0)}$ is nonzero. Using the quotient group duality property of orthogonal subgroups, it can be shown that the quotient group $\widehat{\Lambda}_0/\widehat{\Lambda}$ is isomorphic to Λ/Λ_0 [35]. We are also told that the order of no non-zero element of Λ/Λ_0 divides h_1 or h_2. Therefore, the order of no non-zero element of $\widehat{\Lambda}_0/\widehat{\Lambda}$ divides h_1 or h_2. Hence, $[|h_2|\lambda_1^{(0)}] \bmod \widehat{\Lambda} \neq \mathbf{0}$. We can now say that $|h_2|\lambda_1-|h_1|\lambda_2 \in \widehat{\Lambda}_0 \backslash \widehat{\Lambda}$, which contradicts (5.4). This completes the proof of the theorem. □

5.1.2 Achievable rates in presence of Gaussian noise

We choose ψ to be a characteristic function supported within a ball of radius $r = \alpha r_{\text{pack}}(\widehat{\Lambda}_0)$, as discussed in Section 5.1.1. Since our scheme achieves secrecy in the noiseless case, it is also guaranteed to do so in the presence of AWGN. The average transmit power is given by $P = \frac{1}{n}\mathbb{E}\|\mathbf{U}\|^2$ (also $\frac{1}{n}\mathbb{E}\|\mathbf{V}\|^2$), where the expectation is over the distribution $p_{\mathbf{U}|X}$ and over X. Using Theorem 3.4.11, for a given Λ_0, the average transmit power can be made no less than $\frac{n}{r^2}(1 + o(1))$, where $o(1) \to 0$ as $n \to \infty$. Moreover, Theorem 3.4.11 also gives us the characteristic function that achieves this minimum. The following theorem can then be proved analogously to Theorem 3.3.1.

Theorem 5.1.2. *Let (Λ, Λ_0) be a pair of nested lattices such that Λ_0 is good for covering, $\widehat{\Lambda}_0$ is good for packing, and Λ is good for AWGN channel coding. Let ψ be supported within a ball of radius $r = \alpha r_{\text{pack}}(\widehat{\Lambda}_0)$. Then, a rate of $\frac{1}{2}\log\frac{\alpha^2 P}{\sigma^2} - \log(2e)$, is achievable with perfect secrecy as long as no non-zero element of Λ/Λ_0 has order which divides either h_1 or h_2, and $2/(|h_1| + |h_2|) \geq \alpha$.*

5.2 Strong secrecy with integral channel gains

5.2.1 The noiseless case

To obtain strong secrecy, we use the pmf obtained by sampling the Gaussian density, i.e., $f = g_{\sqrt{P}}$ in (5.2). Recall that for $\theta > 0$, the *flatness factor*, $\epsilon_\Lambda(\theta)$ is defined as

$$\epsilon_\Lambda(\theta) = \max_{\mathbf{x}\in\mathcal{V}(\Lambda)} \left|\det\Lambda \cdot g_{\mathbf{x},\theta}(\Lambda) - 1\right|.$$

This parameter will be used to bound the mutual information between the individual messages and $\mathbf{W}$. We recall the following properties of ϵ_Λ that will be useful in the remainder of the chapter:

- (Lemma 4.3.4) For every $\mathbf{z} \in \mathbb{R}^n$ and $\theta > 0$, we have
$$\frac{g_{\mathbf{z},\theta}(\Lambda)}{g_\theta(\Lambda)} \in \left[\frac{1-\epsilon_\Lambda(\theta)}{1+\epsilon_\Lambda(\theta)}, 1\right] \tag{5.5}$$
- ([58, Remark 2]) For every $\kappa \geq \theta$ and $a > 0$, we have $\epsilon_\Lambda(\theta) \geq \epsilon_\Lambda(\kappa)$, and $\epsilon_{a\Lambda}(a\theta) = \epsilon_\Lambda(\theta)$.

As in Chapter 4, will show that if a certain flatness factor of Λ_0 is asymptotically vanishing in n, then we can obtain strong secrecy. Specifically,

Theorem 5.2.1. *Let $\epsilon \triangleq \epsilon_{\Lambda_0}\left(\sqrt{\frac{P}{h_1^2+h_2^2}}\right)$. If $\epsilon < 1/(16e)$, and Λ/Λ_0 has no non-zero element whose order divides h_1 or h_2, then*

$$I(X; h_1\mathbf{U} + h_2\mathbf{V}) \leq \frac{16\epsilon}{3}\left(\log_2|\mathbb{G}| - \log_2\left(\frac{16\epsilon}{3}\right)\right).$$

If we have $\epsilon = o(1/n)$, then $\mathcal{I}(X; h_1\mathbf{U}+h_2\mathbf{V}) \to 0$ and $\mathcal{I}(Y; h_1\mathbf{U}+h_2\mathbf{V}) \to 0$ as $n \to \infty$, thereby guaranteeing strong secrecy. For secrecy good lattices, we have the flatness factor $\epsilon_{\Lambda_0}(\theta)$ going to zero *exponentially* in n for all θ that satisfies $\text{vol}(\mathcal{V}(\Lambda_0)) < 2\pi\theta^2$. Suppose we choose Λ_0 which is secrecy-good, and $\text{vol}(\mathcal{V}(\Lambda_0)) < 2\pi\alpha^2 P$ for some $\alpha < 1$. Then,

$I(X;\mathbf{W})$ and $I(Y;\mathbf{W})$ can be driven to zero exponentially in n for all co-prime h_1, h_2 that satisfy $1/(h_1^2+h_2^2) \geq \alpha^2$, thereby ensuring strong secrecy. Unlike the scheme of Section 5.1 which guaranteed perfect secrecy for any pair of nested Construction-A lattices, this scheme requires Λ_0 to be secrecy-good to obtain strong security. Before we prove Theorem 5.2.1, we state the following technical lemmas.

Lemma 5.2.2. *Let Λ be a lattice in $\mathbb{R}^n$, and k_1, k_2 be co-prime integers. Then, $\{k_1\mathbf{u}+k_2\mathbf{v} : \mathbf{u},\mathbf{v} \in \Lambda\} = \Lambda$.*

Proof. Clearly, $\{k_1\mathbf{u}+k_2\mathbf{v} : \mathbf{u},\mathbf{v} \in \Lambda\} \subseteq \Lambda$. The converse, $\Lambda \subseteq \{k_1\mathbf{u}+k_2\mathbf{v} : \mathbf{u},\mathbf{v} \in \Lambda\}$ can be proved using the fact that $\exists m, l \in \mathbb{Z}$ such that $k_1m+k_2l = 1$ if k_1, k_2 are co-prime, and $m\mathbf{x}, l\mathbf{x} \in \Lambda$ for $\mathbf{x} \in \Lambda$. □

Lemma 5.2.3. *Let k_1, k_2 be co-prime integers, and $\mathbf{w}_1, \mathbf{w}_2 \in \mathbb{R}^n$. If $\mathbf{w}_2 - \mathbf{w}_1 \notin \Lambda$, then $(k_1\Lambda + \mathbf{w}_1) \cap (k_2\Lambda + \mathbf{w}_2)$ is empty. Otherwise, there exist $\mathbf{u},\mathbf{v} \in \Lambda$ such that $\mathbf{w}_2 - \mathbf{w}_1 = k_1\mathbf{u} + k_2\mathbf{v}$, and $(k_1\Lambda + \mathbf{w}_1) \cap (k_2\Lambda + \mathbf{w}_2) = k_1k_2\Lambda + k_1\mathbf{u} + \mathbf{w}_1$.*

Proof. Define $\mathbf{w} = \mathbf{w}_2 - \mathbf{w}_1$. We can write $(k_1\Lambda+\mathbf{w}_1)\cap(k_2\Lambda+\mathbf{w}_2) = (k_1\Lambda\cap(k_2\Lambda+\mathbf{w}))+\mathbf{w}_1$. If $\mathbf{w} \notin \Lambda$, then clearly $(k_1\Lambda) \cap (k_2\Lambda + \mathbf{w}) = \{\}$.

Now suppose that $\mathbf{w} \in \Lambda$. We can write $\mathbf{w} = k_1\mathbf{u} + k_2\mathbf{v}$ for some $\mathbf{u},\mathbf{v} \in \Lambda$. Since $k_2\Lambda + \mathbf{w} = k_2\Lambda + k_1\mathbf{u}$, we have $k_1k_2\Lambda + k_1\mathbf{u} \subset k_2\Lambda + \mathbf{w}$. Since we also have $k_1k_2\Lambda + k_1\mathbf{u} \subset k_1\Lambda + k_1\mathbf{u}$, we can say that $(k_1k_2\Lambda + k_1\mathbf{u}) \subset (k_1\Lambda) \cap (k_2\Lambda + \mathbf{w})$. To complete the proof, we need to show that $(k_1\Lambda) \cap (k_2\Lambda + \mathbf{w}) \subset (k_1k_2\Lambda + k_1\mathbf{u})$.

For every $\lambda \in (k_1\Lambda) \cap (k_2\Lambda + \mathbf{w}) = (k_1\Lambda) \cap (k_2\Lambda + k_1\mathbf{u})$, there exist $\mathbf{x},\mathbf{y} \in \Lambda$ so that $\lambda = k_1\mathbf{x} = k_2\mathbf{y} + k_1\mathbf{u}$. In other words, $\lambda - k_1\mathbf{u} = k_1(\mathbf{x}-\mathbf{u}) = k_2\mathbf{y}$. Hence, $\lambda - k_1\mathbf{u} \in k_1\Lambda \cap k_2\Lambda$. We now claim that since k_1 and k_2 are co-prime integers, $k_1\Lambda \cap k_2\Lambda = k_1k_2\Lambda$. Clearly, $k_1k_2\Lambda \subset k_1\Lambda \cap k_2\Lambda$. Let G be a generator matrix for Λ. For every $\mathbf{x} \in k_1\Lambda \cap k_2\Lambda$, there exist $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{Z}^n$ so that $\mathbf{x} = k_1G\mathbf{x}_1 = k_2G\mathbf{x}_2$. In other words, $k_1\mathbf{x}_1 = k_2\mathbf{x}_2$, which implies that $\mathbf{x}_1 \in k_2\mathbb{Z}^n$, and $\mathbf{x}_2 \in k_1\mathbb{Z}^n$ since k_1, k_2 are co-prime. Hence, $\mathbf{x} \in k_1k_2\Lambda$, and $k_1\Lambda \cap k_2\Lambda \subset k_1k_2\Lambda$. Therefore, $\lambda - k_1\mathbf{u} \in k_1k_2\Lambda$, or $\lambda \in k_1k_2\Lambda + k_1\mathbf{u}$. Hence, $(k_1\Lambda) \cap (k_2\Lambda + \mathbf{w}) \subset (k_1k_2\Lambda + k_1\mathbf{u})$. This completes the proof. □

Fix any coset $x \in \mathbb{G}$. Let $\mathbf{x}$ be the unique coset representative of x in $\Lambda \cap \mathcal{V}(\Lambda_0)$, and $\mathbf{W} \triangleq h_1\mathbf{U} + h_2\mathbf{V}$. Recall from Chapter 4 that the *variational distance* between $p_{\mathbf{W}}$ and $p_{\mathbf{W}|\mathbf{x}}$ is defined as

$$\mathbb{V}(p_{\mathbf{W}}, p_{\mathbf{W}|\mathbf{x}}) = \sum_{\mathbf{w}\in\Lambda} |p_{\mathbf{W}}(\mathbf{w}) - p_{\mathbf{W}|\mathbf{x}}(\mathbf{w})|,$$

and the average variational distance is $\overline{\mathbb{V}} = \frac{1}{M}\sum_{\mathbf{x}\in\Lambda\cap\mathcal{V}(\Lambda_0)} \mathbb{V}(p_{\mathbf{W}}, p_{\mathbf{W}|\mathbf{x}})$. To prove the theorem, we will find an upper bound on the average variational distance, and then bound the mutual information using the average variational distance. Recall that $\epsilon = \epsilon_{\Lambda_0}\left(\sqrt{P/(h_1^2+h_2^2)}\right)$.

Lemma 5.2.4. *If $\epsilon < 1/2$, and Λ/Λ_0 has no non-zero element whose order divides h_1 or h_2, then for every $\mathbf{x} \in \Lambda \cap \mathcal{V}(\Lambda_0)$, we have*

$$\mathbb{V}(p_{\mathbf{h}_1 U + h_2\mathbf{V}|\mathbf{x}}, p_{h_1\mathbf{U}+h_2\mathbf{V}}) \leq 16\epsilon.$$

Proof. We have

$$p_{\mathbf{W}|\mathbf{x},\mathbf{y}}(\mathbf{w}) = \sum_{\mathbf{u}\in h_1\Lambda_0 + h_1\mathbf{x}} p_{h_1\mathbf{U}|\mathbf{x}}(\mathbf{u}) p_{h_2\mathbf{V}|\mathbf{y}}(\mathbf{w}-\mathbf{u}).$$

The supports of $p_{h_1\mathbf{U}|\mathbf{x}}$ and $p_{h_2\mathbf{V}|\mathbf{y}}$ are $h_1\Lambda_0 + h_1\mathbf{x}$ and $h_2\Lambda_0 + h_2\mathbf{y}$ respectively. Hence, $p_{h_1\mathbf{U}|\mathbf{x}}(\mathbf{u})p_{h_2\mathbf{V}|\mathbf{y}}(\mathbf{w}-\mathbf{u})$ is nonzero iff $\mathbf{u} \in (h_1\Lambda_0 + h_1\mathbf{x})$ and $\mathbf{w}-\mathbf{u} \in (h_2\Lambda_0 + h_2\mathbf{y})$, or equivalently, if $\mathbf{u} \in (h_1\Lambda_0 + h_1\mathbf{x}) \cap (h_2\Lambda_0 - h_2\mathbf{y} + \mathbf{w})$. Using Lemma 5.2.3, we have $(h_1\Lambda_0 + h_1\mathbf{x}) \cap (h_2\Lambda_0 - h_2\mathbf{y} + \mathbf{w}) = h_1h_2\Lambda_0 + h_1\mathbf{x} + h_2\mathbf{y} - \mathbf{w}$ if $\mathbf{w} \in \Lambda_0 + h_1\mathbf{x} + h_2\mathbf{y}$, and empty otherwise. We can therefore conclude that the support of $p_{\mathbf{W}|\mathbf{x},\mathbf{y}}$ is $\Lambda_0 + h_1\mathbf{x} + h_2\mathbf{y}$. Since the order of no element of Λ/Λ_0 divides h_2, we have $\cup_{\mathbf{y}}(\Lambda_0 + h_1\mathbf{x} + h_2\mathbf{y}) = \Lambda$. Therefore, the support of $p_{\mathbf{W}|\mathbf{x}}$ is Λ. Substituting for $p_{h_1\mathbf{U}|\mathbf{x}}$ and $p_{h_2\mathbf{V}|\mathbf{y}}$, we have for all $\mathbf{w} \in \Lambda$,

$$p_{\mathbf{W}|\mathbf{x}}(\mathbf{w}) = \sum_{\mathbf{y}\in\mathbb{G}} \sum_{\substack{\mathbf{u}\in h_1h_2\Lambda_0 + \\ h_1\mathbf{x}+h_2\mathbf{y}-\mathbf{w}}} \frac{e^{-\frac{\|\mathbf{u}\|^2}{2h_1^2P} - \frac{\|\mathbf{w}-\mathbf{u}\|^2}{2h_2^2P}}}{\xi} \tag{5.6}$$

where

$$\xi \triangleq M(2\pi h_1 h_2 P)^n g_{-h_1\mathbf{x},h_1\sqrt{P}}(h_1\Lambda_0) g_{-h_2\mathbf{y},h_2\sqrt{P}}(h_2\Lambda_0).$$

The remainder of the proof follows that of Theorem 4.2.1, and we only give an outline. A simple calculation tells us that

$$e^{-\frac{\|\mathbf{u}\|^2}{2h_1^2P}-\frac{\|\mathbf{w}-\mathbf{u}\|^2}{2h_2^2P}} = \exp\left(-\frac{\|\mathbf{w}\|^2}{2P(h_1^2+h_2^2)} - \frac{(h_1^2+h_2^2)}{2P(h_1^2h_2^2)}\left\|\mathbf{u}-\frac{h_1^2\mathbf{w}}{h_1^2+h_2^2}\right\|^2\right).$$

Let $h \triangleq h_1h_2/\sqrt{h_1^2+h_2^2}$, and $k \triangleq \sqrt{h_1^2+h_2^2}$. Using this and the above equation in (5.6), and simplifying, we get

$$\begin{aligned} p_{\mathbf{W}|\mathbf{x}}(\mathbf{w}) &= \sum_{\mathbf{y}\in\mathbb{G}} \sum_{\substack{\mathbf{u}\in h_1h_2\Lambda_0\\ +h_1\mathbf{x}+h_2\mathbf{y}-\mathbf{w}}} \frac{e^{\left(-\frac{\|\mathbf{w}\|^2}{2P(h_1^2+h_2^2)}-\frac{\|\mathbf{u}-h^2\mathbf{w}/h_2^2\|^2}{2h^2P}\right)}}{\xi} \\ &= e^{-\frac{\|\mathbf{w}\|^2}{2k^2P}} \sum_{\mathbf{y}\in\mathbb{G}} \sum_{\substack{\mathbf{u}\in h_1h_2\Lambda_0+h_1\mathbf{x}\\ +h_2\mathbf{y}-\mathbf{w}-h^2\mathbf{w}/h_2^2}} \frac{e^{-\frac{1}{2h^2P}\|\mathbf{u}\|^2}}{\xi} \end{aligned}$$

Let us define $\mathbf{t} \triangleq h_1\mathbf{x}+h_2\mathbf{y}-\mathbf{w}-(h^2/h_2^2)\mathbf{w}$. The above equation can be simplified to

$$p_{\mathbf{W}|\mathbf{x}}(\mathbf{w}) = \frac{1}{M}\sum_{\mathbf{y}} \frac{g_{k\sqrt{P}}(\mathbf{w})}{g_{-h_1\mathbf{x},h_1\sqrt{P}}(h_1\Lambda_0)} \frac{g_{-\mathbf{t},h\sqrt{P}}(h_1h_2\Lambda_0)}{g_{-h_2\mathbf{y},h_2\sqrt{P}}(h_2\Lambda_0)}$$

Using the properties of the flatness factor,

$$\epsilon_{h_1h_2\Lambda_0}\left(\sqrt{\frac{h_1^2h_2^2P}{h_1^2+h_2^2}}\right) = \epsilon_{\Lambda_0}\left(\sqrt{\frac{P}{h_1^2+h_2^2}}\right) = \epsilon. \tag{5.7}$$

Using (5.5),

$$\frac{1-\epsilon}{1+\epsilon} \le \frac{g_{-\mathbf{t},h\sqrt{P}}(h_1h_2\Lambda_0)}{g_{h\sqrt{P}}(h_1h_2\Lambda_0)} \le 1,$$

Similarly,

$$\frac{1-\epsilon_{\Lambda_0}(\sqrt{P})}{1+\epsilon_{\Lambda_0}(\sqrt{P})} \le \frac{g_{-h_1\mathbf{x},h_1\sqrt{P}}(h_1\Lambda_0)}{g_{h_1\sqrt{P}}(h_1\Lambda_0)} \le 1.$$

Since $\sqrt{h_1^2+h_2^2} > 1$, we have $\epsilon_{\Lambda_0}(\sqrt{P}) \le \epsilon$. Using this, and the fact that $(1-x)/(1+x)$

is a decreasing function of x, we have and

$$\frac{1-\epsilon}{1+\epsilon} \leq \frac{g_{-h_1\mathbf{x},h_1\sqrt{P}}(h_1\Lambda_0)}{g_{h_1\sqrt{P}}(h_1\Lambda_0)} \leq 1.$$

Let us define

$$p(\mathbf{w}) = \frac{1}{M}\sum_{\mathbf{y}} \frac{g_{k\sqrt{P}}(\mathbf{w})}{g_{h_1\sqrt{P}}(h_1\Lambda_0)} \frac{g_{h\sqrt{P}}(h_1h_2\Lambda_0)}{g_{-h_2\mathbf{y},h_2\sqrt{P}}(h_2\Lambda_0)},$$

which is a function independent of $\mathbf{x}$. We can therefore say that

$$\frac{1-\epsilon}{1+\epsilon}p(\mathbf{w}) \leq p_{\mathbf{W}|\mathbf{x}}(\mathbf{w}) \leq \frac{1+\epsilon}{1-\epsilon}p(\mathbf{w}). \tag{5.8}$$

Since $p(\mathbf{w})$ does not depend on $\mathbf{x}$, we can use the above to bound $p_{\mathbf{W}}(\mathbf{w}) = \frac{1}{M}\sum_{\mathbf{x}} p_{\mathbf{W}|\mathbf{x}}(\mathbf{w})$ in the same manner, and obtain $\sum_{\mathbf{w}\in\Lambda} |p_{\mathbf{W}|\mathbf{x}}(\mathbf{w}) - p_{\mathbf{W}}(\mathbf{w})| \leq \frac{4\epsilon}{(1-\epsilon)^2}$. Using the fact that $\epsilon < 1/2$, we get $\mathbb{V}(p_{\mathbf{W}|\mathbf{x}}, p_{\mathbf{W}}) \leq 16\epsilon$, thus completing the proof. □

Proof of Theorem 5.2.1

If $\epsilon < 1/2$, we have $\mathbb{V}(p_{\mathbf{W}|\mathbf{x}}, p_{\mathbf{W}}) \leq 16\epsilon$ from Lemma 5.2.4. Since this is true for every $\mathbf{x} \in \Lambda \cap \mathcal{V}(\Lambda_0)$, we also have $\overline{\mathbb{V}} \leq 16\epsilon$. We can then use [Lemma 1, [19]], which says that if $|\mathbb{G}| > 4$, then $I(\mathbf{W};X) \leq \overline{\mathbb{V}}(\log_2|\mathbb{G}| - \log_2\overline{\mathbb{V}})$.

Since $-x\log x$ is an increasing function of x for $x < 1/e$, we can use the upper bound of 16ϵ for $\overline{\mathbb{V}}$ if $\epsilon < 1/16e$. This completes the proof of the theorem. □

5.2.2 Achievable rates in presence of Gaussian noise

As remarked in the previous section, we choose Λ_0 so that the flatness factor $\epsilon_{\Lambda_0}(\alpha\sqrt{P})$ goes to zero exponentially in n, for some $\alpha \leq 1$. The following statement can be proved analogously to Theorem 4.2.1:

Theorem 5.2.5. *If Λ_0 is good for MSE quantization and secrecy-good, and Λ is good for AWGN channel coding, then the average transmit power converges to P, and any rate less than $\frac{1}{2}\log_2\frac{\alpha^2P}{\sigma^2} - \frac{1}{2}\log_2 e$ can be achieved with strong secrecy as long as the order of no non-zero element of Λ/Λ_0 divides h_1 or h_2, and $1/(h_1^2+h_2^2) \geq \alpha^2$.*

5.3 Discussion

So far, we studied the case where h_1 and h_2 were co-prime integers. This can easily be extended to the general case where h_1/h_2 is rational. We can express $h_1 = hk_1$ and $h_2 = hk_2$ for some $h \in \mathbb{R}$ and co-prime integers k_1 and k_2. Then, it is easy to show that perfect (resp. strong) secrecy can be obtained as long as the order of no non-zero element of Λ/Λ_0 divides k_1 or k_2, and $2/(|k_1|+|k_2|) \geq \alpha$ $\big($resp. $1/(k_1^2+k_2^2) \geq \alpha^2\big)$. Furthermore, the achievable rate is given by $\frac{1}{2}\log\frac{h^2\alpha^2 P}{\sigma^2} - \log(2e)$ $\Big($resp. $\frac{1}{2}\log_2\frac{h^2\alpha^2 P}{\sigma^2} - \frac{1}{2}\log_2 e\Big)$.

We now make the observation that if h_1 and h_2 are nonzero and h_1/h_2 is irrational, then the relay can uniquely recover the individual messages if the channel is noiseless.

Proposition 5.3.1. *Suppose that h_1, h_2 are nonzero, and h_1/h_2 is irrational. Let Λ be a full-rank lattice in $\mathbb{R}^n$. Then, for every $\mathbf{u}, \mathbf{v} \in \Lambda$, $\mathbf{w} = h_1\mathbf{u} + h_2\mathbf{v}$ uniquely determines $(\mathbf{u}, \mathbf{v})$.*

Proof. Consider any $\mathbf{u}_1, \mathbf{u}_2, \mathbf{v}_1, \mathbf{v}_2 \in \Lambda$ that satisfy $h_1\mathbf{u}_1 + h_2\mathbf{v}_1 = h_1\mathbf{u}_2 + h_2\mathbf{v}_2$. If A is a (full-rank) generator matrix of Λ, then we can write $\mathbf{u}_1 = \mathsf{A}^T\tilde{\mathbf{u}}_1$, $\mathbf{u}_2 = \mathsf{A}^T\tilde{\mathbf{u}}_2$, $\mathbf{v}_1 = \mathsf{A}^T\tilde{\mathbf{v}}_1$, and $\mathbf{v}_2 = \mathsf{A}^T\tilde{\mathbf{v}}_2$, where $\tilde{\mathbf{u}}_1, \tilde{\mathbf{u}}_2, \tilde{\mathbf{v}}_1$, and $\tilde{\mathbf{v}}_2$ belong to $\mathbb{Z}^n$. Therefore, $h_1(\tilde{\mathbf{u}}_1 - \tilde{\mathbf{u}}_2) = h_2(\tilde{\mathbf{v}}_2 - \tilde{\mathbf{v}}_1)$. Suppose $\mathbf{u}_1 \neq \mathbf{u}_2$. Then, there exists some $1 \leq i \leq n$ such that $\tilde{u}_1(i) \neq \tilde{u}_2(i)$. Rearranging $h_1(\tilde{u}_1(i) - \tilde{u}_2(i)) = h_2(\tilde{v}_2(i) - \tilde{v}_1(i))$, we get $\frac{h_1}{h_2} = \frac{\tilde{v}_2(i)-\tilde{v}_1(i)}{\tilde{u}_1(i)-\tilde{u}_2(i)}$. However, the right hand side is clearly a rational number, in contradiction to our hypothesis of h_1/h_2 being irrational. Therefore, $\mathbf{u}_1 = \mathbf{u}_2$. Similarly, $\mathbf{v}_1 = \mathbf{v}_2$. □

For our lattice-based scheme to achieve perfect/strong secrecy it is therefore necessary that h_1/h_2 is rational, in which case we can write $h_1 = hk_1$ and $h_2 = hk_2$ for some $h \in \mathbb{R}$ and co-prime integers k_1 and k_2. In addition to this, no element of Λ/Λ_0 can have order that divides k_1 or k_2 if we want to achieve security. While we have seen that the second requirement is sufficient to to guarantee perfect/strong secrecy, we also claim that it is also a *necessary* condition for perfect secrecy. To see why this is the case, recall that we want $p_{k_1\mathbf{U}+k_2\mathbf{V}|x} = p_{k_1\mathbf{U}+k_2\mathbf{V}}$ for all $x \in \Lambda/\Lambda_0$. For this, the supports of the two pmfs must be the same. While the support of $p_{k_1\mathbf{U}+k_2\mathbf{V}|x}$ is $k_1\Lambda_0 + k_2\Lambda + k_1 x$, the support of $p_{k_1\mathbf{U}+k_2\mathbf{V}}$ is $k_1\Lambda + k_2\Lambda = \Lambda$ (since $\gcd(k_1, k_2) = 1$). We can write $k_1\Lambda_0 + k_2\Lambda + k_1\mathbf{x} =$

$\cup_{y\in\Lambda\cap\mathcal{V}(\Lambda_0)}(k_1\Lambda_0 + k_2\Lambda_0 + k_1\mathbf{x} + k_2\mathbf{y}) = \cup_{y\in\Lambda\cap\mathcal{V}(\Lambda_0)}(\Lambda_0 + k_1\mathbf{x} + k_2\mathbf{y})$. If the order of some element of Λ/Λ_0 divides k_2, then we can argue using the pigeonhole principle that $\cup_{y\in\Lambda\cap\mathcal{V}(\Lambda_0)}(\Lambda_0 + k_1\mathbf{x} + k_2\mathbf{y}) \neq \Lambda$, and hence, perfect secrecy is not obtained. This justifies our claim.

The requirement of h_1/h_2 being rational to obtain security may appear discouraging for a practical scenario, where the channel gains are almost surely irrational. However, we must note that we have used a rather pessimistic model for the system. In practice, the user nodes do have a rough estimate of the channel gains, and the channel is noisy. While it may not be possible to achieve perfect security even in presence of noise when the channel gains are irrational unknown to the user nodes, we may hope to achieve strong secrecy. We observed that if we proceed along the lines of Lemma 5.2.4, strong secrecy can be achieved if the flatness factors $\epsilon_{\Lambda_0}\left(\sqrt{\frac{h_i^2P\sigma^2}{h_i^2P+\sigma^2}}\right) = o(1/n)$ for $i = 1, 2$. To achieve this, we could use a secrecy-good lattice scaled so that $\mathrm{vol}(\mathcal{V}(\Lambda_0)) < 2\pi\frac{h_i^2P\sigma^2}{h_i^2P+\sigma^2}$ for $i = 1, 2$. However, it turns out that this is in conflict with the requirement of reliable decoding of X and Y, for which we need $\mathrm{vol}(\mathcal{V}(\Lambda))$ to be greater than $2\pi e\frac{h_i^2P\sigma^2}{h_i^2P+\sigma^2}$. Hence, it seems that a different approach is required to tackle this problem.

Before concluding the chapter, we make a final remark. We saw that no scheme can guarantee security in the absence of noise when the channel gains are irrational. The assumptions that we made were very pessimistic in nature. In practice, the user nodes have some estimate of the channel gains, and the channel is noisy. However, we were unable to show that our scheme guarantees secure and reliable communication in such a setting. In the following subsection, we will outline a simple scheme with which strong secrecy can be obtained in the case where the channel is noisy and the users have a noisy estimate of the channel gains.

5.3.1 Co-operative jamming: Security using Gaussian jamming signals

We can use the following four-stage amplify-and-forward bidirectional relaying strategy: In the first phase, user `A` transmits its codeword $\mathbf{U}_1$, which is jammed by a Gaussian random vector $\mathbf{V}_1$ generated by B. The relay simply scales the received vector and sends it to B, who knows $\mathbf{V}_1$ and can recover $\mathbf{U}_1$. The channel from A to B can be modeled as a Gaussian wiretap channel, where R acts as the eavesdropper. Using a wiretap code [58] for $\mathbf{U}$, we can achieve strong secrecy. User B similarly uses a wiretap code to transmit its message to user A via R in the third and fourth phases.

A reasonable assumption to make is that the error in the estimation of h_1 and h_2 at both user nodes is at most δ. To keep things simple, let us assume that R simply forwards the received signal to the users without scaling. At the end of the second phase, B receives $h_1\mathbf{U}_1+h_2\mathbf{V}_1+\mathbf{Z}$, where $\mathbf{Z}=\mathbf{Z}_1+\mathbf{Z}_2$ is the sum of the noise vectors accumulated in the first two phases, and has variance $\sigma_1^2+\sigma_2^2$. Suppose that the estimates of h_1, h_2 made by B are h_1' and h_2' respectively. Due to the error in estimation, there would be a residual component of $\mathbf{V}$ remaining even after the jamming signal has been removed. Therefore, B "sees" an effective channel of $h_1'\mathbf{U}_1+\mathbf{Z}_B$, where the effective noise is $\mathbf{Z}_B=(h_1-h_1')\mathbf{U}_1+(h_2-h_2')\mathbf{V}_1+\mathbf{Z}$. On the other hand, R "sees" the effective channel $h_1\mathbf{U}_1+\mathbf{Z}'$, where $\mathbf{Z}'=\mathbf{Z}_1+h_2\mathbf{V}_1$. It can be shown that [58] using the lattice Gaussian distribution for randomization, i.e., $p_{\mathbf{U}_1|X}$ given by (5.2) with $f=g_{\sqrt{P}}$, a rate of $\frac{1}{4}\log_2\left(1+\frac{h_1^2P}{2\delta^2P+\sigma^2}\right)-\frac{1}{4}\log_2\left(1+\frac{h_1^2P}{h_2^2P+\sigma_1^2}\right)-\frac{1}{2}\log_2 e$ can be achieved by A with strong secrecy. In fact, the rate can be slightly improved by using a modulo-and-forward scheme [115] instead of the simple amplify-and-forward scheme for relaying.

Chapter 6

Goodness Properties of Low-Density Construction-A Lattices

As seen in the previous chapters, nested lattice codes play a very useful role in designing good codes for secure and reliable communication. We restricted ourselves to Construction-A lattices obtained from linear codes over $\mathbb{F}_p$, for large primes p, and observed that random Construction-A lattices have all the goodness properties that we look for in these problems.

The problem with general Construction-A lattices is the complexity of closest lattice-point decoding. There is no known polynomial-time algorithm for decoding Construction-A lattices obtained from arbitrary linear codes. A natural way of circumventing this is to restrict ourselves to LDPC codes to construct lattices. We can then use low-complexity belief propagation (BP) decoders instead of the closest lattice-point decoder which has exponential complexity. Such lattices, termed low-density Construction-A (LDA) lattices, were introduced by di Pietro et al. in [22]. Simulation results in [21, 91] showed that these lattices perform well with BP decoding. While there is no formal proof that these lattices are good under BP decoding, it was proved in [23] that LDA lattices are good for AWGN channel coding[1], and subsequently shown by di Pietro et al. [24, 21, 25] that nested LDA

[1] It is to be noted that the definition of AWGN goodness used by di Pietro et al. is a little different from what we defined in Chapter 2. They only ask for the probability of decoding error to go to zero as $n \to \infty$, and not necessarily at an exponential rate.

lattices achieve the capacity of the power constrained AWGN channel with closest lattice-point decoding. In this chapter, we show that LDA lattices have several other goodness properties. We will prove that a randomly chosen LDA lattice (whose parameters satisfy certain conditions) is good for packing and MSE quantization with probability tending to 1 as $n \to \infty$. In addition, we will show that the dual of a randomly chosen LDA lattice is good for packing with probability tending to 1 as $n \to \infty$. This means that the capacities of the AWGN channel and the dirty paper channel, the rates guaranteed by compute-and-forward framework [69], and the rates guaranteed in Chapter 3 for perfectly secure bidirectional relaying can all be achieved using nested LDA lattices (with closest lattice-point decoding). However, showing that the aforementioned results can all be achieved using belief propagation decoding still remains an open problem. Even though other AWGN-good lattice constructions that permit low-complexity decoding algorithms have been proposed [86, 110], this is the first instance where such a class of lattices have been shown to satisfy other goodness properties, and this is our main contribution.

The rest of the chapter is organized as follows: Section 6.2 describes the ensemble of lattices, and the main result is stated in Theorem 6.2.2. Some preliminary lemmas are stated in Section 6.3. This is then followed by results on the various goodness properties of lattices in the LDA ensemble. In Section 6.4, the goodness of these lattices for channel coding is described. This is followed by Section 6.5 on the packing goodness of LDA lattices. In Section 6.6, we discuss sufficient conditions for goodness of these lattices for MSE quantization. We then prove the goodness of the duals for packing in Section 6.7, and conclude with some final remarks in Section 6.8.

6.1 Basic Definitions

We also want to use lattices to design good codebooks for reliable transmission over additive noise channels. Classically, a lattice was defined to be good for AWGN channel coding if with high probability, the closest lattice-point decoder returned the actual lattice point that was transmitted over an AWGN channel. This notion was made slightly more general in [73], using the notion of semi norm-ergodic noise:

Definition 3 ([73])**.** *A sequence of random vectors $\{\mathbf{Z}^{(n)}\}$ (where $\mathbf{Z}^{(n)}$ is an n-dimensional random vector) having second moment per dimension $\sigma^2 \triangleq \frac{1}{n}\mathbb{E}[\|\mathbf{Z}^{(n)}\|^2]$ for all n, is said to be semi norm-ergodic if for every $\delta > 0$,*

$$\Pr[\mathbf{Z}^{(n)} \notin (\sqrt{(1+\delta)n\sigma^2})\mathcal{B}] \to 0 \text{ as } n \to \infty.$$

As remarked in [73], any zero-mean noise whose components are independent and identically distributed (iid) is semi norm-ergodic. We say that a sequence of lattices $\{\Lambda^{(n)}\}$ is *good for coding in presence of semi norm-ergodic noise* if for every sequence of semi norm-ergodic noise vectors $\{\mathbf{Z}^{(n)}\}$, with second moment per dimension equal to $\sigma^2 \triangleq \frac{1}{n}\mathbb{E}[\|\mathbf{Z}^{(n)}\|^2]$, the probability that the lattice point closest to $\mathbf{Z}^{(n)}$ is not $\mathbf{0}$ goes to zero as $n \to \infty$, i.e.,

$$\Pr[\mathbf{Z}^{(n)} \notin \mathcal{V}(\Lambda^{(n)})] \to 0 \text{ as } n \to \infty,$$

as long as $(\text{vol}(\Lambda^{(n)}))^{2/n} > 2\pi e\sigma^2$ for all sufficiently large n.

An LDPC code can be defined by its parity check matrix, or by the corresponding edge-labeled Tanner graph [77]. A (Δ_V, Δ_C)-regular bipartite graph $\mathcal{G} = ((\mathcal{L}, \mathcal{R}), \mathcal{E})$ is defined as an undirected bipartite graph with every left vertex (i.e., every vertex in $\mathcal{L}$) having degree Δ_V, and every right vertex (i.e., every vertex in $\mathcal{R}$) having degree Δ_C. The vertices in $\mathcal{L}$ are also called the variable nodes, and those in $\mathcal{R}$ are called parity check (or simply, check) nodes. If $\mathcal{A}$ is a subset of $\mathcal{L}$ (resp. $\mathcal{A}' \subset \mathcal{R}$), then $N(\mathcal{A})$ is the neighbourhood of $\mathcal{A}$, defined as $N(\mathcal{A}) \triangleq \{\mathtt{v} \in \mathcal{R} : (\mathtt{u}, \mathtt{v}) \in \mathcal{E} \text{ for some } \mathtt{u} \in \mathcal{A}\}$ (resp. $N(\mathcal{A}') \triangleq \{\mathtt{u} \in \mathcal{L} : (\mathtt{u}, \mathtt{v}) \in \mathcal{E} \text{ for some } \mathtt{v} \in \mathcal{A}'\}$).

6.2 The Ensemble of LDA Lattices

Throughout this chapter, λ and R are real numbers chosen so that $\lambda > 0$, and $1 > R > 0$. For $n \in \mathbb{Z}^+$, define $k \triangleq \lceil nR \rceil$. For each $n \in \mathbb{Z}^+$, let p (which is really a sequence indexed by n) be the smallest prime number greater than or equal to n^λ, and $\mathbb{F}_p$ denote the field of integers modulo p.

We study the constant-degree LDA ensemble introduced in [23, 21]. Specifically, let $\mathcal{G}$ denote a (Δ_V, Δ_C)-regular bipartite graph ($\Delta_V < \Delta_C$), with n variable nodes, $\frac{n\Delta_V}{\Delta_C}$ check nodes, and satisfying $R = 1-(\Delta_V/\Delta_C)$. Let V be the set of variable nodes, and C denote the set of parity check nodes. This graph $\mathcal{G}$ is the Tanner graph of a binary linear code with parity check matrix $\widehat{H}$. The matrix $\widehat{H}$ has entries from $\{0, 1\}$, and the (i, j)th entry is 1 if and only if there is an edge in $\mathcal{G}$ between the ith vertex in C and jth vertex in V. The graph $\mathcal{G}$ is required to satisfy certain expansion properties, which are stated in the definition below.

Definition 4 ([21], Definition 3.3). *Let A, α, B, β be positive real numbers satisfying $1 \leq \alpha < A$, and $\frac{1}{1-R} < \beta < \min\{\frac{2}{1-R}, B\}$. Let ϵ and ϑ be two small positive constants. The graph $\mathcal{G}$ is said to be (α, A, β, B)-good if it satisfies the following properties:*

(L1) $S \subset V$ and $|S| \leq \lceil \epsilon n \rceil \implies |N(S)| \geq A|S|$.

(L2) $S \subset V$ and $|S| \leq \left\lceil \frac{n(1-R)}{2\alpha} \right\rceil \implies |N(S)| \geq \alpha|S|$.

(R1) $T \subset C$ and $|T| \leq \vartheta n(1-R) \implies |N(T)| \geq B|T|$.

(R2) $T \subset C$ and $|T| \leq \frac{n(1-R)}{2} \implies |N(T)| \geq \beta|T|$.

The following lemma by di Pietro [21] asserts that a randomly chosen graph satisfies the above properties with high probability.

Lemma 6.2.1 ([21], Lemma 3.3). *Let $\mathcal{G}$ be chosen uniformly at random from the standard ensemble [77, Definition 3.15] of (Δ_V, Δ_C)-regular bipartite graphs with n variable nodes. Let ϵ and ϑ be two constants that satisfy*

$$0 < \epsilon < \frac{(1-R)(\Delta_V - A - 1)}{A(\Delta_V - 2 + R)}, \tag{6.1}$$

$$0 < \vartheta < \frac{\Delta_V - (B+1)(1-R)}{B(1-R)(\Delta_V - 2 + R)}. \tag{6.2}$$

If Δ_V satisfies

$$\begin{aligned}\Delta_V > \max\Bigg\{ & \frac{h_2\left(\frac{1-R}{2\alpha}\right)+1-R}{h_2\left(\frac{1-R}{2\alpha}\right)-\frac{1}{2}h_2\left(\frac{1-R}{\alpha}\right)}, R+2\alpha, A+1, \frac{h_2(\epsilon)+(1-R)h_2\left(\frac{A\epsilon}{1-R}\right)}{h_2(\epsilon)-\frac{A\epsilon}{1-R}h_2\left(\frac{1-R}{A}\right)}, \\ & \frac{1-R+h_2\left(\frac{\beta(1-R)}{2}\right)}{1-\frac{\beta(1-R)}{2}h_2\left(\frac{1}{\beta(1-R)}\right)}, \frac{(2+\beta R)(1-R)}{2-\beta(1-R)}, (1-R)(B+1), \\ & \frac{(1-R)h_2(\vartheta)+h_2(B\vartheta(1-R))}{h_2(\vartheta)-B\vartheta(1-R)h_2\left(\frac{1}{B(1-R)}\right)}\Bigg\}, \end{aligned} \tag{6.3}$$

then the probability that $\mathcal{G}$ is not (α, A, β, B)-good tends to zero as $n \to \infty$.

6.2.1 The $(\mathcal{G}, \lambda)$ LDA Ensemble

Let $\lambda > 0$, and $1 > R > 0$ be two constants, and $n \in \{1, 2, 3, \ldots\}$. Let p be the smallest prime number greater than n^λ.[2] Let $\Delta_C \triangleq \Delta_V/(1-R)$. Let us pick a (Δ_V, Δ_C)-regular bipartite graph $\mathcal{G}$ with n variable nodes. Throughout the chapter, we assume that the parameters of $\mathcal{G}$ satisfy the hypotheses of Lemma 6.2.1, and that $\mathcal{G}$ is (α, A, β, B)-good. Let $\widehat{H}$ denote the $n(1-R) \times n$ parity check matrix corresponding to the Tanner graph $\mathcal{G}$. In other words, the (i, j)th entry of $\widehat{H}$ is 1 if there is an edge between the ith parity check node and the jth variable node, and 0 otherwise. We describe the LDA ensemble obtained using the Tanner graph $\mathcal{G}$, which will henceforth be called the $(\mathcal{G}, \lambda)$ *LDA ensemble.*

We construct a new $n(1-R) \times n$ matrix, H, by replacing the 1's in $\widehat{H}$ with independent random variables uniformly distributed over $\mathbb{F}_p$. For $1 \le i \le n(1-R)$ and $1 \le j \le n$, let $h'_{i,j}$ be $n^2(1-R)$ iid random variables, each uniformly distributed over $\mathbb{F}_p$, and let $\widehat{h}_{i,j}$ be the (i, j)th entry of $\widehat{H}$. Then, the (i, j)th entry of H, denoted $h_{i,j}$, is given by $h_{i,j} = \widehat{h}_{i,j} h'_{i,j}$. Therefore, $h_{i,j}$ is equal to $h'_{i,j}$ if $\widehat{h}_{i,j}$ is 1, and zero otherwise. For example,

[2] In our proofs, we take $p = n^\lambda$, and $k = nR$ for convenience, but choosing p to be the smallest prime number greater than n^λ, and $k = \lceil nR \rceil$ will not change any of the results.

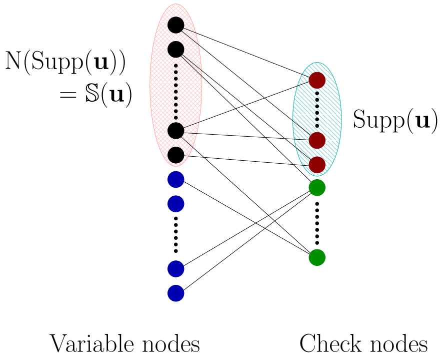


Figure 6.1: Nodes corresponding to Supp($\mathbf{u}$) and $\mathbb{S}(\mathbf{u})$.

if

$$\widehat{H} = \begin{pmatrix} 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \end{pmatrix},$$

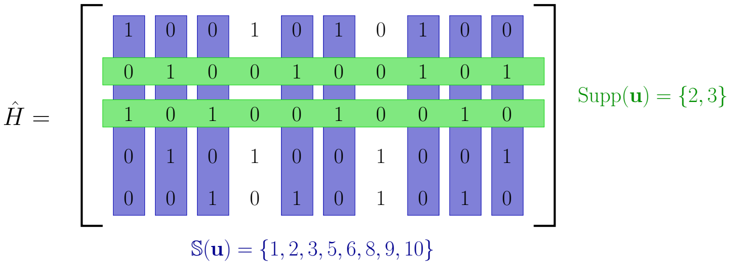
then

$$H = \begin{pmatrix} h'_{11} & h'_{12} & 0 & 0 & h'_{15} & 0 \\ 0 & h'_{22} & 0 & h'_{24} & 0 & h'_{26} \\ 0 & 0 & h'_{33} & h'_{34} & h'_{35} & 0 \\ h'_{41} & 0 & h'_{43} & 0 & 0 & h'_{46} \end{pmatrix}. \tag{6.4}$$

Note that the "skeleton matrix" $\widehat{H}$ is fixed beforehand, and the only randomness in H is in the coefficients. This matrix H is the parity check matrix of an n-length (Δ_V, Δ_C) regular LDPC code $\mathcal{C}$ over $\mathbb{F}_p$. The LDA lattice Λ is obtained by applying Construction A to the code $\mathcal{C}$, i.e., $\Lambda = \{\mathbf{x} \in \mathbb{Z}^n : \mathbf{x} \equiv \mathbf{c} \bmod p, \text{ for some } \mathbf{c} \in \mathcal{C}\}$. Equivalently, if Φ denotes the natural embedding of $\mathbb{F}_p^n$ into $\mathbb{Z}^n$, then $\Lambda = \Phi(\mathcal{C}) + p\mathbb{Z}^n$.

For a given $\mathbf{u} \in \mathbb{F}_p^{n(1-R)}$, let us define $\mathbb{S}(\mathbf{u})$ to be the set of all variable nodes that participate in the check equations i for which the ith entry of $\mathbf{u}$ (i.e., u_i) is nonzero. Formally, $\mathbb{S}(\mathbf{u}) \triangleq \bigcup_{i \in \text{Supp}(\mathbf{u})} \text{Supp}(\widehat{\mathbf{h}}_i)$. Equivalently, $i \in \mathbb{S}(\mathbf{u})$ iff there exists $1 \le j \le n(1-R)$ such that $u_j \neq 0$ and $\widehat{h}_{j,i} \neq 0$. This is illustrated in Fig. 6.1 and Fig. 6.2.

The rest of the chapter will be dedicated to proving the following theorem:



$$\hat{H} = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \end{bmatrix}$$

Figure 6.2: Illustration of $\mathbb{S}(\mathbf{u})$ for $\text{Supp}(\mathbf{u}) = \{2, 3\}$.

Theorem 6.2.2. *Let $A > 2(1+R)$, $B > 2(1+R)/(1-R)$,*

$$\epsilon = \frac{1-R}{A+1-R} \quad \textit{and} \quad \vartheta = \frac{1}{B(1-R)+1}.$$

Suppose that Δ_V satisfies the conditions of Lemma 6.2.1, and the corresponding $\mathcal{G}$ is (α, A, β, B)-good. Let

$$\lambda > \max\left\{ \frac{1}{R}, \frac{1}{1-R}, \frac{2}{A-2(1-R)}, \frac{2}{B(1-R)-2(1+R)}, 2\left(1 - \frac{1}{AB-1} - \frac{1}{A}\right)^{-1}, \right.$$
$$\left. \frac{1}{2(\alpha-1+R)}, \frac{2B+3/2}{B(1-R)-1} \right\}. \tag{6.5}$$

If we pick Λ at random from the $(\mathcal{G}, \lambda)$ LDA ensemble, then the probability that Λ is simultaneously good for packing, channel coding, and MSE quantization tends to 1 as $n \to \infty$. Moreover, the probability that Λ^* is also simultaneously good for packing, tends to 1 as $n \to \infty$.

We will prove each of the goodness properties in separate sections. The conditions on the parameters of the lattice to ensure goodness for channel coding are stated in Theorem 6.4.1. Goodness for packing is discussed in Theorem 6.5.1, and MSE quantization in Theorem 6.6.1. Sufficient conditions for the packing goodness of the duals of LDA lattices are given in Theorem 6.7.1. The above theorem can then be obtained by a simple application of the union bound.

Example 2. *Let us choose $R = 1/3$, $\alpha = 2.7$, $A = 3$, $\beta = 1.6$, and $B = 5$. Then, we get*

$$\epsilon = \frac{1-R}{A+1-R} \approx 0.182 \ \text{ and } \ \vartheta = \frac{1}{B(1-R)+1} \approx 0.231.$$

For these parameters, a choice of $\Delta_V = 21$ ensures that (6.3) is satisfied.

Substituting these values in (6.5), we get that choosing $\lambda > 6$ ensures that a random LDA lattice satisfies the guaranteed goodness properties.

Before we proceed to the main results, we will discuss some useful lemmas that we will need later on in the proofs.

6.3 Some Preliminary Lemmas

In this section, we record some basic results that will be used in the proofs. We start with the following elementary fact:

Lemma 6.3.1. *Let $\gamma > 0$, and s be a positive integer. Then,*

$$\sum_{j=s}^{\infty} n^{-j\gamma} = n^{-s\gamma}(1+o(1)).$$

Recall that V_n is the volume of a unit ball in n dimensions. We have the following upper bound on the number of integer points within a ball of radius r:

Lemma 6.3.2 (Corollary of [73], Lemma 1)**.** *Let $r > 0$, $\mathbf{y} \in \mathbb{R}^n$, and $\mathcal{B}$ denote the unit ball in n dimensions. Then,*

$$V_n\left(r - \frac{\sqrt{n}}{2}\right)^n \le |\mathbb{Z}^n \cap (\mathbf{y} + r\mathcal{B})| \le V_n\left(r + \frac{\sqrt{n}}{2}\right)^n.$$

Furthermore, if $m \le n$, then

$$|\{\mathbf{x} \in \mathbb{Z}^n \cap r\mathcal{B} : |\mathrm{Supp}(\mathbf{x})| \le m\}| \le \binom{n}{m} V_m \left(r + \frac{\sqrt{m}}{2}\right)^m.$$

Recall the randomized construction of the parity check matrix H from the (α, A, β, B)-good graph $\mathcal{G}$, described in the previous section. Also recall that for $\mathbf{u} \in \mathbb{F}_p^{n(1-R)}$, $\mathbb{S}(\mathbf{u})$ is the set of all variable nodes that participate in the check equations i for which $u_i \neq 0$. We have the following result which describes the distribution of $H^T\mathbf{u}$.

Lemma 6.3.3. *Let* $\mathbf{u} \in \mathbb{F}_p^{n(1-R)}$*, and* $\mathbf{x} \in \mathbb{F}_p^n$*. Then,*

$$\Pr[H^T\mathbf{u} = \mathbf{x}] = \begin{cases} \frac{1}{p^{|\mathbb{S}(\mathbf{u})|}} & \textit{if } \mathrm{Supp}(\mathbf{x}) \subset \mathbb{S}(\mathbf{u}) \\ 0 & \textit{else.} \end{cases} \tag{6.6}$$

Proof. Let $\mathbf{y} \triangleq H^T\mathbf{u}$. The jth entry of $\mathbf{y}$ is given by $y_j = \sum_{i=1}^{n(1-R)} h_{ji}u_i$. Consider any $j \in (\mathbb{S}(\mathbf{u}))^c$. From the definition of $\mathbb{S}(\mathbf{u})$, it is easy to see that the jth variable node does not participate in any of the parity check equations indexed by $\mathrm{Supp}(\mathbf{u})$. Hence, $h_{ij} = 0$ whenever $u_i \neq 0$ (see Fig. 6.2 to get a better picture). Therefore, $y_j = 0$. On the other hand, if $j \in \mathbb{S}(\mathbf{u})$, then there exists at least one i such that $h_{ij} \neq 0$. So, $y_j = \sum_{i \in \mathrm{Supp}(\mathbf{u})} h_{ji}u_i$, being a nontrivial linear combination of independent and uniformly distributed random variables, is also uniformly distributed over $\mathbb{F}_p$. Moreover, it is easy to see that the y_j's are independent. Therefore,

$$\Pr[y_j = a] = \begin{cases} 1/p & \text{if } j \in \mathbb{S}(\mathbf{u}) \\ 0 & \text{if } j \notin \mathbb{S}(\mathbf{u}) \text{ and } a \neq 0 \\ 1 & \text{if } j \notin \mathbb{S}(\mathbf{u}) \text{ and } a = 0. \end{cases}$$

This completes the proof. □

Recall that H defines a linear code over $\mathbb{F}_p$, where p is the smallest prime greater than n^λ. The following lemma gives a lower bound on the probability of a randomly chosen H not having full rank.

Lemma 6.3.4. *If $B > 2 + (1+\delta)/\lambda$ for some $\delta > 0$, then*

$$\Pr[H \textit{ is not full-rank }] \leq n^{-(2\lambda+\delta)}(1+o(1)).$$

Proof. We will prove that the probability that there is any nontrivial linear combination of the rows of H equal to zero tends to 0 as $n \to \infty$. Let $\mathbf{h}_i$ denote the ith row of H. For any $S \subseteq \{1, 2, \ldots, n(1-R)\}$, we define

$$\chi_S = \begin{cases} 1 & \text{if there is a nontrivial linear combination of } \{\mathbf{h}_i : i \in S\} \text{ that is zero,} \\ 0 & \text{otherwise.} \end{cases}$$

Let us also define

$$Y = \sum_{s=1}^{n(1-R)} \sum_{\substack{S \subset \{1,2,\ldots,n(1-R)\} \\ |S|=s}} \chi_S$$

Clearly, H is full rank if and only if $Y = 0$. Using Markov's inequality, we see that

$$\Pr[Y \geq 1] \leq \mathbb{E}[Y].$$

Therefore, it is enough to find an upper bound on the expectation of Y. Let

$$\eta(S) = |\bigcup_{i \in S} \mathrm{Supp}(\mathbf{h}_i)|.$$

In other words, $\eta(S)$ is the number of variable nodes that participate in the parity check equations indexed by S. This is also equal to the number of neighbours of S in $\mathcal{G}$, i.e., $|N(S)|$. Observe that there are at most p^{s-1} different linear combinations (not counting scalar multiples) of s rows of H. Using Lemma 6.3.3, the probability that a fixed linear combination of the S rows of H is zero is equal to $1/p^{\eta(S)}$. Using the union bound,

$$\Pr[\chi_S = 1] \leq \frac{p^{s-1}}{p^{\eta(S)}}.$$

Therefore, we have

$$\begin{aligned}
\mathbb{E}[Y] &\leq \sum_{s=1}^{n(1-R)} \sum_{\substack{S\subset\{1,2,\ldots,n(1-R)\}\\|S|=s}} \frac{p^{s-1}}{p^{\eta(S)}} \\
&= \sum_{s=1}^{\vartheta n(1-R)} \sum_{\substack{S\subset\{1,2,\ldots,n(1-R)\}\\|S|=s}} \frac{p^{s-1}}{p^{\eta(S)}} + \sum_{s=\vartheta n(1-R)}^{n(1-R)/2} \sum_{\substack{S\subset\{1,2,\ldots,n(1-R)\}\\|S|=s}} \frac{p^{s-1}}{p^{\eta(S)}} \\
&\qquad + \sum_{s=n(1-R)/2}^{n(1-R)} \sum_{\substack{S\subset\{1,2,\ldots,n(1-R)\}\\|S|=s}} \frac{p^{s-1}}{p^{\eta(S)}}
\end{aligned}$$

Since $\mathcal{G}$ is (α, A, β, B)-good, we have

$$\begin{aligned}
\mathbb{E}[Y] &\leq \sum_{s=1}^{\vartheta n(1-R)} \binom{n(1-R)}{s} \frac{p^{s-1}}{p^{Bs}} + \sum_{s=\vartheta n(1-R)}^{n(1-R)/2} \binom{n(1-R)}{s} \frac{p^{s-1}}{p^{\beta s}} \\
&\qquad + \sum_{s=n(1-R)/2}^{n(1-R)} \binom{n(1-R)}{s} \frac{p^{s-1}}{p^{s/(1-R)}}. \qquad (6.7)
\end{aligned}$$

We can further simplify this as follows,

$$\begin{aligned}
\mathbb{E}[Y] &\leq \sum_{s=1}^{\vartheta n(1-R)} n^s \frac{n^{\lambda(s-1)}}{n^{\lambda Bs}} + \sum_{s=\vartheta n(1-R)}^{n(1-R)/2} 2^n \frac{n^{\lambda(s-1)}}{n^{\lambda\beta s}} + \sum_{s=n(1-R)/2}^{n(1-R)} 2^n \frac{n^{\lambda(s-1)}}{n^{\lambda s/(1-R)}} \qquad (6.8) \\
&\leq \sum_{s=1}^{\vartheta n(1-R)} n^{s(1+\lambda(1-B))-\lambda} + n^{-c_1 n}(1+o(1)) \\
&= n^{(1+\lambda(1-B))-\lambda}(1+o(1)) + n^{-c_1 n}(1+o(1)), \qquad (6.9)
\end{aligned}$$

for some constant $c_1 > 0$, since β and $1/(1-R)$ are greater than 1, and $B > 1 + 1/\lambda$. Suppose that for some constant $\delta > 0$, we have $B > 2+(1+\delta)/\lambda$. Then, $(1+\lambda(1-B))-\lambda < -(2\lambda+\delta)$, and therefore,

$$\mathbb{E}[Y] \leq n^{-(2\lambda+\delta)}(1+o(1)).$$

Therefore, $\Pr[Y \geq 1]$, and hence the probability that H is not full rank, goes to zero as

$n \to \infty$.

□

Remark 6.3.5. *To prove that* $\mathbb{E}[Y] \to 0$ *in (6.9), it is sufficient to have* $B > 1 + 1/\lambda$*. The expected value of* Y*, and subsequently* $\Pr[H \textit{ is not full-rank}]$ *could then be bounded from above by* $n^{-(\lambda+\delta)}(1+o(1))$*. However, we need* $\Pr[H \textit{ is not full-rank}]$ *to be less than* $n^{-(2\lambda+\delta)}(1+o(1))$ *to prove that LDA lattices are good for MSE quantization (in particular, to show that the second term in (6.18) goes to zero), and hence we impose the stronger condition that* $B > 2 + 1/\lambda$*.*

We now proceed to prove the various goodness properties of LDA lattices.

6.4 Goodness for Channel Coding

Recall that a sequence of lattices $\{\Lambda^{(n)}\}$ is good for coding in presence of semi norm-ergodic noise if for any sequence of semi norm-ergodic noise vectors $\{\mathbf{z}^{(n)}\}$, with second moment per dimension equal to $\sigma^2 \triangleq \frac{1}{n}\mathbb{E}[\|\mathbf{z}^{(n)}\|^2]$,

$$\Pr[\mathbf{z}^{(n)} \notin \mathcal{V}(\Lambda^{(n)})] \to 0 \text{ as } n \to \infty$$

as long as $(\mathrm{vol}\Lambda^{(n)})^{2/n} > 2\pi e\sigma^2$ for all sufficiently large n. But we have

$$(\mathrm{vol}\Lambda^{(n)})^{2/n} = (r_{\text{eff}}(\Lambda^{(n)}))^2 V_n^{2/n} = (r_{\text{eff}}(\Lambda^{(n)}))^2 \frac{2\pi e}{n}(1+o(1))$$

using Stirling's approximation. Therefore, we can equivalently say that a sequence of lattices is good for coding in presence of semi norm-ergodic noise if $\Pr[\mathbf{z}^{(n)} \notin \mathcal{V}(\Lambda^{(n)})] \to 0$ as $n \to \infty$ as long as $r_{\text{eff}}(\Lambda^{(n)}) \geq \sqrt{n\sigma^2}(1-o(1))$. Note that if the noise is assumed to be iid Gaussian, then the above definition is weaker than the definition of AWGN (or Poltyrev) goodness defined in [29], since the probability $\Pr[\mathbf{z}^{(n)} \notin \mathcal{V}(\Lambda^{(n)})]$ is not required to go to zero exponentially in n. However, the above definition covers a much wider class of noise distributions. In particular, the "effective noise" that is present in the equivalent modulo-lattice additive noise channel in the compute-and-forward protocol [69] is semi norm-ergodic, as discussed in [73].

The following result was proved by di Pietro:

Theorem 6.4.1 ([21], Theorem 3.2). *Let Λ be a lattice chosen uniformly at random from a $(\mathcal{G}, \lambda)$ LDA ensemble, where $\mathcal{G}$ is (α, A, β, B)-good, and suppose that the hypotheses of Lemma 6.2.1 are satisfied. If*

$$\lambda > \max\left\{\frac{1}{2(\alpha - 1 + R)}, \frac{3}{2(A - 1 + R)}, \frac{1}{B(1 - R) - 1}\right\},$$

then the probability that Λ is good for coding in presence of semi norm-ergodic noise tends to 1 as $n \to \infty$.

For semi norm-ergodic noise $\{\mathbf{z}^{(n)}\}$, we have for every $\delta > 0$, $\Pr[\mathbf{z}^{(n)} \notin (\sqrt{(1+\delta)n\sigma^2})\mathcal{B}] \to 0$ as $n \to \infty$. To prove that $\{\Lambda^{(n)}\}$ is good for coding, it is then enough to show the absence of nonzero lattice points within a ball of radius $\sqrt{(1+\delta)n\sigma^2}$ around $\mathbf{z}$, for all $n\sigma^2 < (r_{\text{eff}}(\Lambda^{(n)}))^2$ and all sufficiently large n. In [21], di Pietro proved the following statement, thus establishing Theorem 6.4.1, and hence showing that LDA lattices are good for channel coding: For every $\mathbf{z} \in \sqrt{(1+\delta)n\sigma^2}\mathcal{B}$,

$$\sum_{\mathbf{x} \in \mathbb{Z}^n \cap (r_n\mathcal{B} + \mathbf{z}) \setminus p\mathbb{Z}^n} \Pr[\mathbf{x} \in \Lambda] \to 0 \text{ as } n \to \infty, \tag{6.10}$$

where $r_n = r_{\text{eff}}(\Lambda^{(n)})(1 + \delta_n)$, and $\delta_n \to 0$ as $n \to \infty$.

6.5 Goodness for Packing

Recall that $\{\Lambda^{(n)}\}$ is good for packing if

$$\limsup_{n\to\infty} \frac{r_{\text{pack}}(\Lambda^{(n)})}{r_{\text{eff}}(\Lambda^{(n)})} \geq \frac{1}{2}.$$

We want to prove the following result:

Theorem 6.5.1. *Let Λ be a lattice chosen uniformly at random from a $(\mathcal{G}, \lambda)$ LDA ensemble, where $\mathcal{G}$ is (α, A, β, B)-good, and the parameters satisfy the hypotheses of Lemma 6.2.1. Furthermore, let*

$$\lambda > \max\left\{\frac{1}{2(\alpha-1+R)}, \frac{3}{2(A-1+R)}, \frac{1}{B(1-R)-1}\right\}.$$

Then, the probability that Λ is good for packing tends to 1 as $n \to \infty$.

Let us choose $r_n = r_{\text{eff}}(\Lambda)(1-\delta_n)$, where δ_n is a quantity that goes to 0 as $n \to \infty$. We want to prove that

$$\Pr[r_{\text{pack}}(\Lambda) < r_n/2] \to 0 \text{ as } n \to \infty.$$

It is enough to show that the probability of any nonzero integer point within $r_n\mathcal{B}$ belonging to Λ goes to zero as $n \to \infty$, i.e.,

$$\sum_{\mathbf{x}\in\mathbb{Z}^n\cap r_n\mathcal{B}\setminus\{\mathbf{0}\}} \Pr[\mathbf{x} \in \Lambda] \to 0 \text{ as } n \to \infty$$

This requirement is similar to (6.10), and the proof of packing goodness of LDA lattices follows, *mutatis mutandis*, on similar lines as that for AWGN channel coding.

6.6 Goodness for MSE Quantization

In nested lattice coding for power-constrained transmission over Gaussian channels, the codebook is generally the set of all points of the fine lattice within the fundamental Voronoi region of the coarse lattice. Hence, the fine lattice determines the codeword points, while the coarse lattice defines the shaping region. In order to maximize the rate for a given power constraint, we want the shaping region to be approximately spherical. The loss in rate (penalty for not using a spherical shaping region) is captured by the normalized second moment, $G(\Lambda)$, of the coarse lattice Λ, and in order to minimize this loss, we want $G(\Lambda)$ to be as close to $1/(2\pi e)$ as possible. As defined in Section 6.1, $\{\Lambda^{(n)}\}$ is good for

MSE quantization if $G(\Lambda^{(n)}) \to \frac{1}{2\pi e}$ as $n \to \infty$. In this section, we will prove the following result:

Theorem 6.6.1. *Let $A > 2(1+R)$ and $B > 2(1+R)/(1-R)$. Fix*

$$\epsilon = \frac{1-R}{A+1-R} \quad \textit{and} \quad \vartheta = \frac{1}{B(1-R)+1}.$$

Suppose that Δ_V satisfies the conditions of Lemma 6.2.1, and $\mathcal{G}$ is (α, A, β, B)-good. Furthermore, let

$$\lambda > \max\left\{ \frac{1}{R}, \frac{1}{1-R}, \frac{2}{A-2(1+R)}, \frac{2}{B(1-R)-2(1+R)}, 2\left(1 - \frac{1}{AB-1} - \frac{1}{A}\right)^{-1} \right\}. \tag{6.11}$$

Let Λ be randomly chosen from a $(\mathcal{G}, \lambda)$ LDA ensemble. Then, the probability that Λ is good for MSE quantization tends to 1 as $n \to \infty$.

To prove the theorem, we will show that for every positive δ_1, δ_2, and all sufficiently large n,

$$\Pr\left[G(\Lambda) > \frac{1}{2\pi e} + \delta_1 \right] \le \delta_2. \tag{6.12}$$

Since $G(\Lambda) > 1/(2\pi e)$ for all Λ [29], the above statement guarantees the existence of a sequence of lattices, $\{\Lambda^{(n)}\}$, for which $G(\Lambda^{(n)}) \to 1/(2\pi e)$ as $n \to \infty$. Our proof of the above inequality is based on the techniques used in [73] and [21]. For a lattice Λ, and $\mathbf{x} \in \mathbb{R}^n$, we define $d(\mathbf{x}, \Lambda) \triangleq \min_{\mathbf{y}\in\Lambda} \|\mathbf{x} - \mathbf{y}\|$ to be the Euclidean distance between $\mathbf{x}$ and the closest point in Λ to $\mathbf{x}$. For ease of notation, let us define $r \triangleq r_{\text{eff}}(\Lambda)$. Our proof of inequality (6.12), and hence Theorem 6.6.1, will make use of the following lemmas.

Lemma 6.6.2. *Suppose that the hypotheses of Theorem 6.6.1 are satisfied. Let Λ be drawn uniformly at random from a $(\mathcal{G}, \lambda)$ LDA ensemble, and X be a random vector uniformly distributed over $\mathcal{V}(\Lambda)$. Then,*

$$\mathbb{E}_\Lambda[G(\Lambda)] \le \mathbb{E}_{\Lambda,X}\left[\frac{d^2(X,\Lambda)}{n(\text{vol}\Lambda)^{2/n}} \middle| H \textit{ is full rank} \right] + o(1). \tag{6.13}$$

Proof. Recall that V_n denotes the volume of an n-dimensional unit ball. Using Stirling's approximation, we get,

$$V_n^{1/n} = \left(\frac{\pi^{n/2}}{\Gamma(n/2+1)}\right)^{1/n} = \frac{\sqrt{2\pi e}}{n^{1/2}}(1+o(1)). \tag{6.14}$$

For any Construction-A lattice Λ, we have $p\mathbb{Z}^n \subset \Lambda$. If H is full-rank, then the number of points of Λ in $[0,p)^n$, (and therefore, within $\mathcal{V}(p\mathbb{Z}^n)$) is equal to p^{nR}, which is $|\Lambda/p\mathbb{Z}^n|$. Since $|\Lambda/p\mathbb{Z}^n| = \mathrm{vol}(p\mathbb{Z}^n)/\mathrm{vol}\Lambda$, we get $\mathrm{vol}\Lambda = p^{n(1-R)}$. Therefore,

$$r_{\text{eff}}(\Lambda) = \left(\frac{\mathrm{vol}\Lambda}{V_n}\right)^{1/n} = \frac{n^{\lambda(1-R)+1/2}}{\sqrt{2\pi e}}(1+o(1)). \tag{6.15}$$

For any $\mathbf{x} \in \mathbb{R}^n$, we have $d(\mathbf{x},\Lambda) = \min_{\mathbf{y}\in\Lambda} \|\mathbf{y}-\mathbf{x}\|$ to be the distance between $\mathbf{x}$ and the closest point in Λ to $\mathbf{x}$. Recall that X is a random vector uniformly distributed over the fundamental Voronoi region of Λ. The normalized second moment of Λ is then equal to

$$G(\Lambda) = \mathbb{E}_X\left[\frac{d^2(X,\Lambda)}{n(\mathrm{vol}\Lambda)^{2/n}}\right].$$

We can write

$$\begin{aligned}\mathbb{E}_\Lambda[G(\Lambda)] &= \mathbb{E}_\Lambda[G(\Lambda)|H \text{ is full rank}]\Pr[H \text{ is full rank}]\\ &\qquad + \mathbb{E}_\Lambda[G(\Lambda)|H \text{ is not full rank}]\Pr[H \text{ is not full rank}]\\ &= \mathbb{E}_{\Lambda,X}\left[\left.\frac{d^2(X,\Lambda)}{n(\mathrm{vol}\Lambda)^{2/n}}\right| H \text{ is full rank}\right]\Pr[H \text{ is full rank}]\\ &\qquad + \mathbb{E}_{\Lambda,X}\left[\left.\frac{d^2(X,\Lambda)}{n(\mathrm{vol}\Lambda)^{2/n}}\right| H \text{ is not full rank}\right]\Pr[H \text{ is not full rank}] \end{aligned} \tag{6.16}$$

Since $p\mathbb{Z}^n \subset \Lambda$, we have for every $\mathbf{x} \in \mathbb{R}^n$, $d(\mathbf{x},\Lambda) \le d(\mathbf{x},p\mathbb{Z}^n) \le p\sqrt{n}/2$. Additionally, since $\Lambda \subset \mathbb{Z}^n$, we have $\mathrm{vol}\Lambda \ge \mathrm{vol}(\mathbb{Z}^n) = 1$. Hence, we can say that for any Construction-A lattice,

$$\frac{d^2(X,\Lambda)}{n(\mathrm{vol}\Lambda)^{2/n}} \le \frac{p^2}{4} \tag{6.17}$$

with probability 1. Let δ be a positive constant that satisfies $\delta < \lambda(B-2)-1$. From the hypotheses of Theorem 6.6.1, we have $B > 2(1+R)/(1-R)$, and $\lambda > 1/R$. This guarantees that $\lambda(B-2)-1 > 4/(1-R)-1 > 0$, and hence, we can choose a $\delta > 0$. Using Lemma 6.3.4, we can bound $\Pr[H \text{ is not full rank}]$ from above by $n^{-2\lambda-\delta}$. Using this and (6.17) in (6.16), and the fact that $\Pr[H \text{ is full rank}] \leq 1$, we obtain

$$\begin{aligned} \mathbb{E}_{\Lambda}[G(\Lambda)] &\leq \mathbb{E}_{\Lambda,X}\left[\frac{d^2(X,\Lambda)}{n(\mathrm{vol}\Lambda)^{2/n}}\middle| H \text{ is full rank}\right] + \frac{p^2}{4}\frac{1}{n^{2\lambda+\delta}} \\ &= \mathbb{E}_{\Lambda,X}\left[\frac{d^2(X,\Lambda)}{n(\mathrm{vol}\Lambda)^{2/n}}\middle| H \text{ is full rank}\right] + o(1), \end{aligned} \tag{6.18}$$

thus completing the proof. □

The following lemma is formally proved in Section 6.6.1.

Lemma 6.6.3. *Suppose that the hypotheses of Theorem 6.6.1 are satisfied. Let* $0 < \omega < 1$. *There exists a* $\delta > 0$ *so that for every* $\mathbf{x} \in \mathbb{R}^n$,

$$\Pr\left[d(\mathbf{x},\Lambda) > r\left(1+\frac{1}{n^{\omega}}\right)\middle| H \textit{ is full rank}\right] \leq \frac{1}{n^{2\lambda R+\delta}}(1+o(1)). \tag{6.19}$$

Lemma 6.6.4. *Let* U *be a random vector uniformly distributed over* $[0,p)^n$, *and* X *be uniformly distributed over* $\mathcal{V}(\Lambda)$. *Then,*

$$\mathbb{E}_{\Lambda}\mathbb{E}_X[d^2(X,\Lambda)|H \textit{ is full rank}] = \mathbb{E}_U\mathbb{E}_{\Lambda}[d^2(U,\Lambda)|H \textit{ is full rank}]. \tag{6.20}$$

Proof. Recall that U is uniformly distributed over $[0,p)^n$, and X is uniformly distributed over $\mathcal{V}(\Lambda)$. We have,

$$\begin{aligned}
&\mathbb{E}_U\mathbb{E}_\Lambda[d^2(U,\Lambda)|H \text{ is full rank}] \\
&= \int_{\mathbf{u}\in[0,p)^n} \sum_{\Lambda_1} d^2(\mathbf{u},\Lambda_1)\frac{Pr[\Lambda=\Lambda_1|H \text{ is full rank}]}{p^n}d\mathbf{u} && (6.21)\\
&= \sum_{\Lambda_1}\int_{\mathbf{u}\in[0,p)^n} d^2(\mathbf{u},\Lambda_1)\frac{Pr[\Lambda=\Lambda_1|H \text{ is full rank}]}{p^n}d\mathbf{u} && (6.22)\\
&= \sum_{\Lambda_1}\sum_{\mathbf{z}\in\Lambda_1\cap[0,p)^n}\int_{\mathbf{x}\in\mathcal{V}(\Lambda_1)} d^2(\mathbf{x}+\mathbf{z},\Lambda_1)\frac{Pr[\Lambda=\Lambda_1|H \text{ is full rank}]}{p^n}d\mathbf{x}. && (6.23)
\end{aligned}$$

For all $\mathbf{z}\in\Lambda$, we have $d(\mathbf{x}+\mathbf{z},\Lambda)=d(\mathbf{x},\Lambda)$. Hence,

$$\begin{aligned}
\mathbb{E}_U\mathbb{E}_\Lambda[d^2(U,\Lambda)|H \text{ is full rank}] &= \sum_{\Lambda_1} p^{nR}\int_{\mathbf{x}\in\mathcal{V}(\Lambda_1)} d^2(\mathbf{x},\Lambda_1)\frac{Pr[\Lambda=\Lambda_1|H \text{ is full rank}]}{p^n}d\mathbf{x} && (6.24)\\
&= \sum_{\Lambda_1}\int_{\mathbf{x}\in\mathcal{V}(\Lambda_1)} d^2(\mathbf{x},\Lambda_1)\frac{Pr[\Lambda=\Lambda_1|H \text{ is full rank}]}{p^{n(1-R)}}d\mathbf{x} && (6.25)\\
&= \mathbb{E}_\Lambda\mathbb{E}_X[d^2(X,\Lambda)|H \text{ is full rank}]. && (6.26)
\end{aligned}$$

This completes the proof. □

6.6.1 Proof of Lemma 6.6.3

Recall that $r \triangleq r_{\text{eff}}(\Lambda)$. We want to show that for some $\delta>0$, the probability $\Pr[d(\mathbf{x},\Lambda) > r(1+n^{-\omega})|H \text{ is full rank}]$ goes to zero faster than $n^{-2\lambda R+\delta}$. The proof is along the same lines as di Pietro's proof of existence of lattices that achieve the capacity of the AWGN channel in [21]. The parameters chosen in [21] were not sufficient to show that the lattices are good for MSE quantization. We have adapted the proof to show that under stronger conditions (on the parameters of the lattice), we can obtain lattices which are good for

MSE quantization. For $\mathbf{y} \in \mathbb{Z}^n$, define

$$\xi_{\mathbf{y}} = \begin{cases} 1 & \text{if } H\mathbf{y} \equiv \mathbf{0} \bmod p, \\ 0 & \text{otherwise.} \end{cases}$$

Let $\rho = r(1 + n^{-\omega})$. Recall that $\mathbf{x} + \rho\mathcal{B}$ denotes an n-dimensional ball centered at $\mathbf{x}$ and having radius ρ. We define

$$X_\rho \triangleq \sum_{\mathbf{y} \in \mathbb{Z}^n \cap (\mathbf{x} + \rho\mathcal{B})} \xi_{\mathbf{y}},$$

which is simply the number of lattice points in $\mathbf{x} + \rho\mathcal{B}$. Let us define $\mathcal{E}(\rho) = |\mathbb{Z}^n \cap (\mathbf{x} + \rho\mathcal{B})|^2 \frac{1}{p^{2n(1-R)}}$. From [21, p. 119], we have

$$\mathbb{E}[X_\rho] \geq \sqrt{\mathcal{E}(\rho)}. \tag{6.27}$$

In [21, pp. 122–128], it was shown that the variance of X_ρ can be bounded from above as follows.[3]

$$\mathrm{Var}(X_\rho) \leq \sum_{s=1}^{\lfloor \frac{n(1-R)}{A+1-R} \rfloor} n^{s(2-\lambda(A-2))} \tag{6.28}$$

$$+ \sum_{\substack{i,j,t \\ j \leq n(1-R)/(B(1-R)+1) \\ i+j+t=n(1-R) \\ i+j>0}} \mathcal{E}(\rho) \left(1 + \frac{Bj}{n-Bj}\right)^{\frac{n-Bj+1}{2}} n^{j(1-\lambda(B(1-R)-2))}$$

$$\times \left(1 + \frac{Bi}{n-Bi}\right)^{\frac{n-Bi+1}{2}} n^{i(1-\lambda(B(1-R)-2))}(1+o(1)) \tag{6.29}$$

$$+ \sum_{\substack{i,j,t \\ j \leq n(1-R)/(B(1-R)+1) \\ i+j+t=n(1-R) \\ i+j>0}} \mathcal{E}(\rho) \left(1 + \frac{Bj}{n-Bj}\right)^{\frac{n-Bj+1}{2}} n^{j\lambda(2-B(1-R))}$$

$$\times n^{(j+t)\left(1+\lambda\left(\frac{1}{AB-1}+\frac{1}{A}-1\right)\right)} \frac{n^\lambda}{\sqrt{\mathcal{E}(\rho)}}(1+o(1)). \tag{6.30}$$

We show that (6.28), (6.29) and (6.30) are all bounded from above by $\mathcal{E}(\rho)n^{-2\lambda R-\delta}(1+o(1))$.

Let

$$\delta \triangleq \frac{1}{2}\min\{\lambda(A-2(1+R))-2,\ \lambda(B(1-R)-2(1+R))-1\}. \tag{6.31}$$

The hypotheses of Theorem 6.6.1 ensure that $\delta > 0$.

The First Term, (6.28)

We have

$$\sum_{s=1}^{\lfloor \frac{n(1-R)}{A+1-R} \rfloor} n^{s(2-\lambda(A-2))} = n^{2-\lambda(A-2)}(1+o(1)),$$

provided that the exponent is negative. As long as $2-\lambda(A-2) < -2\lambda R - \delta$, we have the first term bounded from above by $n^{-2\lambda R-\delta}(1+o(1))$. This condition is indeed satisfied, since by definition, $\delta < \lambda(A-2(1+R))-2$.

[3] The variance of X_ρ is upper bounded by a sum of three terms, (6.28), (6.29), and (6.30), which are equations (4.51), (4.56), and (4.60) respectively in [21]. We impose stronger constraints on B and λ so as to ensure that (6.19) goes to zero sufficiently fast as $n \to \infty$.

The Second Term, (6.29)

For all $x > 0$, we have $\ln(1+x) \le x$, and hence $(1+x)^{1/x} \le e$. With this, we get

$$\left(1 + \frac{Bj}{n - Bj}\right)^{\frac{n-Bj}{2}} \le e^{Bj/2}.$$

This implies that

$$\begin{aligned}\left(1 + \frac{Bj}{n - Bj}\right)^{\frac{n-Bj}{2}} n^{j(1-\lambda(B(1-R)-2))} &\le e^{Bj/2} n^{j(1-\lambda(B(1-R)-2))} \\ &= (c_1 n)^{j(1-\lambda(B(1-R)-2))},\end{aligned}$$

where $c_1 = e^{B/(2(1-\lambda(B(1-R)-2)))}$ is a positive constant. From (6.31), we have $\delta \le \frac{1}{2}(\lambda(B(1-R)-2(1+R))-1)$, and hence $1-\lambda(B(1-R)-2) \le -2\lambda R - 2\delta$. Moreover, $c_1^{-2\lambda R-2\delta} n^{-\delta} \le 1$ for sufficiently large n. Hence,

$$\left(1 + \frac{Bj}{n - Bj}\right)^{\frac{n-Bj}{2}} n^{j(1-\lambda(B(1-R)-2))} \le n^{j(-2\lambda R-\delta)} \tag{6.32}$$

for all sufficiently large n. Similarly,

$$\left(1 + \frac{Bi}{n - Bi}\right)^{\frac{n-Bi}{2}} n^{i(1-\lambda(B(1-R)-2))} \le n^{i(-2\lambda R-\delta)} \tag{6.33}$$

for all sufficiently large n. Hence, the second term is bounded from above by

$$\begin{aligned}&\sum_{\substack{i,j,t \\ j \le n(1-R)/(B(1-R)+1) \\ i+j+t=n(1-R) \\ i+j>0}} \mathcal{E}(\rho) n^{(i+j)(-2\lambda R-\delta)}(1+o(1)) \\ &\qquad = \sum_{\substack{i,j \\ j \le n(1-R)/(B(1-R)+1) \\ i+j \le n(1-R) \\ i+j>0}} \mathcal{E}(\rho) n^{(i+j)(-2\lambda R-\delta)}(1+o(1)) \\ &\qquad \le \mathcal{E}(\rho) n^{-2\lambda R-\delta}(1+o(1)),\end{aligned}$$

which follows from Lemma 6.3.1.

The Third Term, (6.30)

Since $B > 2/(1-R)$ and $\lambda > 2\left(1-\frac{1}{AB-1}-\frac{1}{A}\right)^{-1}$, we have for $j \neq 0$,

$$\left(1+\frac{Bj}{n-Bj}\right)^{\frac{n-Bj+1}{2}} n^{j\lambda(2-B(1-R))} = o(1), \text{ and} \tag{6.34}$$

$$n^{(j+t)\left(1+\lambda\left(\frac{1}{AB-1}+\frac{1}{A}-1\right)\right)} = o(1). \tag{6.35}$$

If $j = 0$, then the above terms are at most 1. Now,

$$\begin{aligned}\sqrt{\mathcal{E}(\rho)} &= |\mathbb{Z}^n \cap (\mathbf{x}+\rho\mathcal{B})|\frac{1}{p^{n(1-R)}} \\ &\geq V_n\left(\rho-\frac{\sqrt{n}}{2}\right)^n \frac{1}{p^{n(1-R)}} \qquad\qquad (6.36)\\ &= V_n r^n\left(1+\frac{1}{n^\omega}\right)^n\left(1-\frac{\sqrt{n}}{2\rho}\right)^n\frac{1}{p^{n(1-R)}},\end{aligned}$$

where (6.36) follows from Lemma 6.3.2. But $V_n r^n = p^{n(1-R)}$. Using this, and simplifying, we get

$$\begin{aligned}\sqrt{\mathcal{E}(\rho)} &\geq p^{n(1-R)}\exp\{n^{1-\omega}\}\exp\left\{\frac{\sqrt{2\pi e}}{2}n^{-\lambda(1-R)}n(1+n^{-\omega})^{-1}\right\}\frac{1}{p^{n(1-R)}}(1+o(1)) \\ &\geq \exp\{n^{1-\omega}-o(1)\}.\end{aligned} \tag{6.37}$$

Therefore, $1/\sqrt{\mathcal{E}(\rho)}$ goes to zero faster than any polynomial. Combining (6.34), (6.35), and (6.37), we can conclude that (6.30) is upper bounded by $\mathcal{E}(\rho)n^{-2\lambda R-\delta}(1+o(1))$. As a consequence, the variance of X_ρ is bounded from above by $3\mathcal{E}(\rho)n^{-2\lambda R-\delta}(1+o(1))$.

Proof of Lemma 6.6.3

We have already seen in (6.27) that $\mathbb{E}[X_\rho] \geq \sqrt{\mathcal{E}(\rho)}$ and in the previous subsections, we showed that $\mathrm{Var}(X_\rho) \leq \mathcal{E}(\rho) n^{-2\lambda R-\delta}(1+o(1))$. Therefore,

$$\begin{aligned}\Pr[d(\mathbf{x},\Lambda) > \rho] &= \Pr[X_\rho = 0] \leq \Pr[X_\rho \leq 0] \\ &= \Pr\big[X_\rho - \mathbb{E}[X_\rho] \leq -\mathbb{E}[X_\rho]\big] \\ &\leq \Pr\big[|X_\rho - \mathbb{E}[X_\rho]| \geq \mathbb{E}[X_\rho]\big]. \qquad (6.38)\end{aligned}$$

Using Chebyshev's inequality, we get

$$\Pr[d(\mathbf{x},\Lambda) > \rho] \leq \frac{\mathrm{Var}(X_\rho)}{(\mathbb{E}[X_\rho])^2} \leq \frac{3}{n^{2\lambda R+\delta}}(1+o(1)),$$

completing the proof of Lemma 6.6.3. □

6.6.2 Proof of Theorem 6.6.1

Recall that to prove the theorem, it is enough to prove inequality (6.12). To this end, we will show that the first term in (6.13) tends to $1/(2\pi e)$ as $n \to \infty$. We will use Lemma 6.6.3 to bound this term.

Recall that $r = r_{\text{eff}}(\Lambda)$. Since (6.19) holds for all $\mathbf{x} \in \mathbb{R}^n$, we can say that for any random vector U (having density function f) over $\mathbb{R}^n$, we have

$$\begin{aligned}\Pr[d(U,\Lambda) > r(1+n^{-\omega}) | H \text{ is full rank}] &= \int_{\mathbb{R}^n} \Pr[d(\mathbf{u},\Lambda) > r(1+n^{-\omega}) | H \text{ is full rank}] f(\mathbf{u}) d\mathbf{u} \\ &\leq n^{-(2\lambda R+\delta)}(1+o(1)). \qquad (6.39)\end{aligned}$$

Let us define $\rho = r(1+n^{-\omega})$. For any $\mathbf{u} \in \mathbb{R}^n$, and any Construction-A lattice Λ, we have

$d(\mathbf{u},\Lambda) \leq p\sqrt{n}/2$. Then, for any distribution on U,

$$\mathbb{E}_U\mathbb{E}_\Lambda[d^2(U,\Lambda)|H \text{ is full rank}] \leq \rho^2 \Pr[d(U,\Lambda) \leq \rho|H \text{ is full rank}] \\ + \frac{p^2n}{4}\Pr[d(U,\Lambda) > \rho|H \text{ is full rank}] \tag{6.40}$$

$$\leq \rho^2\left(1+\frac{p^2n}{4\rho^2}\frac{1}{n^{2\lambda R+\delta}}(1+o(1))\right). \tag{6.41}$$

Substituting $\rho = \frac{n^{\lambda(1-R)+1/2}}{\sqrt{2\pi e}}(1+o(1))$,

$$\mathbb{E}_U\mathbb{E}_\Lambda[d^2(U,\Lambda)|H \text{ is full rank}] \leq \rho^2\left(1+n^{2\lambda+1}\frac{2\pi e}{4n^{2\lambda(1-R)+1}}\frac{1}{n^{2\lambda R+\delta}}(1+o(1))\right) \tag{6.42}$$

$$= \rho^2\left(1+\frac{\pi e}{2n^\delta}(1+o(1))\right) \tag{6.43}$$

$$= r^2(1+o(1)). \tag{6.44}$$

From (6.44) and Lemma 6.6.4, we have

$$\mathbb{E}_\Lambda\mathbb{E}_X[d^2(U,\Lambda)|H \text{ is full rank}] \leq r^2(1+o(1)).$$

Recall that V_n denotes the volume of an n-dimensional unit ball. Using Stirling's approximation, we get,

$$V_n^{1/n} = \left(\frac{\pi^{n/2}}{\Gamma(n/2+1)}\right)^{1/n} = \frac{\sqrt{2\pi e}}{n^{1/2}}(1+o(1)).$$

Therefore,

$$n(\mathrm{vol}\Lambda)^{2/n} = (r_{\text{eff}}(\Lambda))^2 2\pi e(1+o(1)) = r^2 2\pi e(1+o(1))$$

and hence,

$$\mathbb{E}_\Lambda\mathbb{E}_X\left[\frac{d^2(U,\Lambda)}{n(\mathrm{vol}\Lambda)^{2/n}}\middle| H \text{ is full rank}\right] \leq \frac{1}{2\pi e}(1+o(1)).$$

Using this, and Lemma 6.6.2, we can write

$$\mathbb{E}[G(\Lambda)] \leq \frac{1}{2\pi e}(1+\delta(n)), \tag{6.45}$$

where $\delta(n)$ is a quantity that goes to 0 as $n \to \infty$. We also have $G(\Lambda) > 1/(2\pi e)$ for all Λ. For any $\gamma > 0$, we can write

$$\begin{aligned}\mathbb{E}[G(\Lambda)] &\geq \frac{1}{2\pi e}\Pr\left[\frac{1}{2\pi e} < G(\Lambda) \leq \frac{1}{2\pi e} + \gamma\right] + \left(\frac{1}{2\pi e} + \gamma\right)\Pr\left[G(\Lambda) > \frac{1}{2\pi e} + \gamma\right] \\ &= \frac{1}{2\pi e}\left(1 - \Pr\left[G(\Lambda) > \frac{1}{2\pi e} + \gamma\right]\right) + \left(\frac{1}{2\pi e} + \gamma\right)\Pr\left[G(\Lambda) > \frac{1}{2\pi e} + \gamma\right] \\ &= \frac{1}{2\pi e} + \gamma\Pr\left[G(\Lambda) > \frac{1}{2\pi e} + \gamma\right], \end{aligned} \tag{6.46}$$

and hence,

$$\Pr\left[G(\Lambda) > \frac{1}{2\pi e} + \gamma\right] \leq \frac{\mathbb{E}[G(\Lambda)] - 1/(2\pi e)}{\gamma} \tag{6.47}$$

Since the above inequality holds for every $\gamma > 0$, we can choose, for e.g., $\gamma = \sqrt{\delta(n)}$, and use (6.45) to obtain

$$\Pr\left[G(\Lambda) > \frac{1}{2\pi e} + \sqrt{\delta(n)}\right] \leq \sqrt{\delta(n)} \to 0 \text{ as } n \to \infty.$$

Therefore, we can conclude that the probability of choosing an LDA lattice which is good for MSE quantization tends to 1 as $n \to \infty$. □

6.7 Packing Goodness of the Duals of LDA Lattices

Recall that $r_{\text{pack}}(\Lambda)$ denotes the packing radius of Λ, and that a sequence of lattices $\{\Lambda^{(n)}\}$ is good for packing if

$$\frac{r_{\text{pack}}(\Lambda^{(n)})}{r_{\text{eff}}(\Lambda^{(n)})} \geq \frac{1}{2} \text{ as } n \to \infty.$$

Our motivation for studying the properties of the dual of a lattice comes from Chapter 3, where we studied perfectly secure bidirectional relaying. We showed that if the fine lattices are good for AWGN channel coding, the coarse lattices are good for MSE quantization, and the duals of the coarse lattices are good for packing, then a rate of $\frac{1}{2}\log_2\frac{P}{\sigma^2} - \log_2(2e)$ can be achieved with perfect secrecy. This motivates us to construct lattices whose duals are good for packing. In this section, we will prove the following result.

Theorem 6.7.1. *Let $\mathcal{G}$ be an (α, A, β, B)-good (Δ_V, Δ_C)-regular bipartite graph whose parameters satisfy the hypotheses of Lemma 6.2.1. If*

$$\lambda > \max\left\{\frac{1}{2(1-R)}, \frac{2B+3/2}{B(1-R)-1}\right\},$$

then the dual of a randomly chosen lattice from a $(\mathcal{G}, \lambda)$ LDA ensemble is good for packing with probability tending to 1 as $n \to \infty$.

6.7.1 Proof of Theorem 6.7.1

If Λ is a lattice obtained by applying Construction A to a linear code $\mathcal{C}$, and if Λ^* is the dual of Λ, then, $\frac{1}{p}\Lambda^*$ is obtained by applying Construction A to the dual code, $\mathcal{C}^\perp$ (Lemma 2.4.3). To show that the duals of LDA lattices are good for packing, it is enough to show that the Construction-A lattices generated by the duals of the nonbinary LDPC codes ($\mathcal{C}$) are good for packing.

Note that H (a parity check matrix for $\mathcal{C}$) is a generator matrix for $\mathcal{C}^\perp$. Let Λ' be the lattice obtained by applying Construction A on $\mathcal{C}^\perp$. We will prove that Λ' is good for packing. The lattice Λ' contains $p\mathbb{Z}^n$ as a sublattice, and the nesting ratio is $p^{n(1-R)}$ if H is full-rank. The volume of $\mathcal{V}(\Lambda')$ is equal to the ratio of the volume of $p\mathbb{Z}^n$ to the nesting ratio, and hence,

$$\mathrm{vol}(\Lambda') = \frac{p^n}{p^{n(1-R)}} = p^{nR}.$$

Recall that V_n is the volume of the unit ball in n dimensions. The effective radius of Λ' can therefore be written as,

$$r_{\text{eff}}(\Lambda') = \frac{p^R}{(V_n)^{1/n}}. \tag{6.48}$$

Let us define

$$r_n \triangleq \frac{p^R}{V_n^{1/n}}\zeta_n, \tag{6.49}$$

where ζ_n is a term that goes to 1 as $n \to \infty$, defined as follows:

$$\zeta_n = \frac{1}{n^{4/n}} \left(\frac{C_1}{e(1-R)\ln n} \right)^{\frac{4C_1}{(1-R)\ln n}} \left(1 - \frac{C_1}{(1-R)\ln n} \right)^2 . \tag{6.50}$$

Here,

$$C_1 \triangleq \frac{\ln\left(\frac{8}{1-(1-R)/(2\alpha)} \right)}{\lambda(1-(1-R)/\alpha)}. \tag{6.51}$$

We want to prove that the probability $\Pr[r_{\text{pack}}(\Lambda') < r_{\text{eff}}(\Lambda')/2] \to 0$ as $n \to \infty$. We will show that the probability of finding a nonzero lattice point within a ball of radius r_n centered at $\mathbf{0}$ goes to zero as $n \to \infty$.

Since $p\mathbb{Z}^n$ is always a sublattice of Λ', we must ensure that $r_{\text{eff}}(\Lambda') < p$. Substituting for $r_{\text{eff}}(\Lambda')$ from (6.48), we can see that $r_{\text{eff}}(\Lambda') < p$ is satisfied for all sufficiently large n as long as $\lambda > \frac{1}{2(1-R)}$, which is guaranteed by the hypothesis of Theorem 6.7.1.

We want

$$\Pr\left[\exists \mathbf{u} \in \mathbb{F}_p^{n(1-R)} \backslash \{\mathbf{0}\} : H^T\mathbf{u} \in (\mathbb{Z}^n \cap r_n\mathcal{B}) \bmod p\mathbb{Z}^n \right] \to 0 \text{ as } n \to \infty.$$

Instead, we will prove the following (stronger) statement.

$$\sum_{\mathbf{u} \in \mathbb{F}_p^{n(1-R)} \backslash \{\mathbf{0}\}} \Pr\left[H^T\mathbf{u} \in (\mathbb{Z}^n \cap r_n\mathcal{B}) \bmod p\mathbb{Z}^n \right] \to 0 \text{ as } n \to \infty.$$

The summation in the above statement can be expanded as follows:

$$\begin{aligned} &\sum_{\mathbf{u} \in \mathbb{F}_p^{n(1-R)} \backslash \{\mathbf{0}\}} \Pr\left[H^T\mathbf{u} \in (\mathbb{Z}^n \cap r_n\mathcal{B}) \bmod p\mathbb{Z}^n \right] \\ &= \sum_{t=1}^{n(1-R)} \sum_{\substack{\mathbf{u} \in \mathbb{F}_p^{n(1-R)} \\ |\text{Supp}(\mathbf{u})|=t}} \Pr\left[H^T\mathbf{u} \in (\mathbb{Z}^n \cap r_n\mathcal{B}) \bmod p\mathbb{Z}^n \right] \\ &= \sum_{t=1}^{n(1-R)} \sum_{\substack{\mathbf{u} \in \mathbb{F}_p^{n(1-R)} \\ |\text{Supp}(\mathbf{u})|=t}} \sum_{s=1}^{n} \sum_{\substack{\mathbf{x} \in \mathbb{Z}^n \cap r_n\mathcal{B} \\ |\text{Supp}(\mathbf{x})|=s}} \Pr[H^T\mathbf{u} \equiv \mathbf{x} \bmod p]. \end{aligned} \tag{6.52}$$

Fix $\mathbf{u} \in \mathbb{F}_p^{n(1-R)}$. Recall, from Section 6.2, that $\mathbb{S}(\mathbf{u})$ is the set of all variable nodes that participate in the check equations i for which $u_i \neq 0$. For $S \subset \{1, 2, \ldots, n\}$, define $\mathbf{1}_S(\mathbb{S}(\mathbf{u}))$ to be the function that takes the value 1 if $\mathbb{S}(\mathbf{u}) = S$, and zero otherwise. Note that this is a deterministic function of $\mathbf{u}$ since $\widehat{H}$ is fixed beforehand. Let us also define $\mathbb{1}_m(\mathbb{S}(\mathbf{u}))$ to be the function which takes the value 1 if $|\mathbb{S}(\mathbf{u})| = m$, and zero otherwise. Using Lemma 6.3.3, we have

$$\begin{aligned}\Pr[H^T\mathbf{u} \equiv \mathbf{x} \bmod p] &= \begin{cases} \frac{1}{p^{|\mathbb{S}(\mathbf{u})|}} & \text{if } \mathrm{Supp}(\mathbf{x}) \subset \mathbb{S}(\mathbf{u}) \\ 0 & \text{otherwise.} \end{cases} \\ &\leq \frac{1}{p^{|\mathbb{S}(\mathbf{u})|}} \\ &= \sum_{m=1}^{n} \mathbb{1}_m(\mathbb{S}(\mathbf{u}))\frac{1}{p^m}. \end{aligned} \tag{6.53}$$

We use this in (6.52) to obtain

$$\begin{aligned}&\sum_{\mathbf{u}\in\mathbb{F}_p^{n(1-R)}\setminus\{\mathbf{0}\}} \Pr\left[H^T\mathbf{u} \in (\mathbb{Z}^n \cap r_n\mathcal{B}) \bmod p\mathbb{Z}^n\right] \\ &\leq \sum_{t=1}^{n(1-R)} \sum_{\substack{\mathbf{u}\in\mathbb{F}_p^{n(1-R)} \\ |\mathrm{Supp}(\mathbf{u})|=t}} \sum_{m=1}^{n} \mathbb{1}_m(\mathbb{S}(\mathbf{u}))\frac{1}{p^m} \sum_{s=1}^{m} \sum_{\substack{\mathbf{x}\in\mathbb{Z}^n\cap r_n\mathcal{B} \\ |\mathrm{Supp}(\mathbf{x})|=s}} 1 \\ &= \sum_{t=1}^{n(1-R)} \sum_{\substack{\mathbf{u}\in\mathbb{F}_p^{n(1-R)} \\ |\mathrm{Supp}(\mathbf{u})|=t}} \sum_{m=1}^{n} \mathbb{1}_m(\mathbb{S}(\mathbf{u}))\frac{1}{p^m} |\{\mathbf{x} \in \mathbb{Z}^n \cap r_n\mathcal{B} : |\mathrm{Supp}(\mathbf{x})| \leq m\}|. \end{aligned} \tag{6.54}$$

All that remains now is to show that the above quantity goes to zero as $n \to \infty$. With that, we can say that the probability that the dual of a randomly chosen LDA lattice is good for packing goes to 1 as $n \to \infty$, completing the proof of Theorem 6.7.1.

We will now split the summation in (6.54) into four parts, and show that each quantity goes to zero as $n \to \infty$. The sum is divided into the following regimes:

1. $1 \leq t < \vartheta n(1-R)$,

2. $\vartheta n(1-R) \leq t < n(1-R)/2$,

3. $n(1-R)/2 \leq t < (1-R-C_1/\ln n)n - 1$,

4. $(1-R-C_1/\ln n)n - 1 \leq t \leq n$,

where C_1 is as defined in (6.51). In each case, we will use the appropriate expansion properties of the underlying Tanner graph to prove the desired result.

Case 1: $1 \leq t < \vartheta n(1-R)$

We will use property (R1) of the expander graph in this part of the proof. In this case, we have $t = |\mathrm{Supp}(\mathbf{u})| \leq \vartheta n(1-R)$. Therefore, $|N(\mathrm{Supp}(\mathbf{u}))| = |\mathbb{S}(\mathbf{u})| \geq Bt$, so that $\mathbb{1}_m(\mathbb{S}(\mathbf{u})) = 0$ for $m < Bt$. Consider

$$\begin{aligned}
\phi_1(n) &\triangleq \sum_{t=1}^{\vartheta n(1-R)} \sum_{\substack{\mathbf{u}\in\mathbb{F}_p^{n(1-R)} \\ |\mathrm{Supp}(\mathbf{u})|=t}} \sum_{m=1}^{n} \mathbb{1}_m(\mathbb{S}(\mathbf{u}))\frac{1}{p^m}|\{\mathbf{x}\in\mathbb{Z}^n\cap r_n\mathcal{B} : |\mathrm{Supp}(\mathbf{x})|\leq m\}| \\
&\leq \sum_{t=1}^{\vartheta n(1-R)} \sum_{\substack{\mathbf{u}\in\mathbb{F}_p^{n(1-R)} \\ |\mathrm{Supp}(\mathbf{u})|=t}} \sum_{m=Bt}^{n} \frac{1}{p^m}|\{\mathbf{x}\in\mathbb{Z}^n\cap r_n\mathcal{B} : |\mathrm{Supp}(\mathbf{x})|\leq m\}|.
\end{aligned}$$

Using Lemma 6.3.2, the above quantity can be bounded from above as

$$\begin{aligned}
\phi_1(n) &\leq \sum_{t=1}^{\vartheta n(1-R)} \sum_{\substack{\mathbf{u}\in\mathbb{F}_p^{n(1-R)} \\ |\mathrm{Supp}(\mathbf{u})|=t}} \sum_{m=Bt}^{n} \frac{1}{p^m}\binom{n}{m} V_m\left(r_n+\frac{\sqrt{m}}{2}\right)^m \\
&\leq \sum_{t=1}^{\vartheta n(1-R)} \sum_{\substack{\mathbf{u}\in\mathbb{F}_p^{n(1-R)} \\ |\mathrm{Supp}(\mathbf{u})|=t}} \sum_{m=Bt}^{n} \frac{1}{p^m}\binom{n}{m} V_m r_n^m\left(1+\frac{\sqrt{m}}{2r_n}\right)^m \\
&= \sum_{t=1}^{\vartheta n(1-R)} \sum_{\substack{\mathbf{u}\in\mathbb{F}_p^{n(1-R)} \\ |\mathrm{Supp}(\mathbf{u})|=t}} \sum_{m=Bt}^{n} \frac{1}{p^m}\binom{n}{m} V_m \frac{p^{mR}}{V_n^{m/n}}\zeta_n^m\left(1+\frac{\sqrt{m}}{2r_n}\right)^m. \qquad (6.55)
\end{aligned}$$

Using Stirling's approximation, we get

$$V_m = \frac{\pi^{m/2}}{\Gamma(1+m/2)} \leq \frac{\pi^{m/2}e^m}{(2\pi)^{1/2}m^{m+1/2}},$$

and

$$V_n \geq \frac{\pi^{n/2}e^n}{en^{n+1/2}}.$$

Therefore,

$$\frac{V_m}{V_n^{m/n}} \leq c' \left(\frac{n}{m}\right)^{m+1/2} (1+o(1)), \tag{6.56}$$

where c' is a positive constant. If $m > an$ for some $0 < a < 1$, then

$$\frac{V_m}{V_n^{m/n}} \leq c \left(\frac{n}{m}\right)^{m} (1+o(1)), \tag{6.57}$$

where c is a positive constant.

Observe that $\zeta_n < 1$ for all sufficiently large n, and $1+\frac{\sqrt{m}}{2r_n} \leq 2$. Using this, and (6.56), the inequality (6.55) reduces to

$$\begin{aligned}
\phi_1(n) &\leq c' \sum_{t=1}^{\vartheta n(1-R)} \sum_{\substack{\mathbf{u}\in\mathbb{F}_p^{n(1-R)} \\ |\text{Supp}(\mathbf{u})|=t}} \sum_{m=Bt}^{n} \frac{1}{p^{m(1-R)}} \begin{pmatrix} n \\ m \end{pmatrix} \left(2\frac{n}{m}\right)^m \left(\frac{n}{m}\right)^{1/2} (1+o(1)) \\
&\leq c' \sum_{t=1}^{\vartheta n(1-R)} \begin{pmatrix} n(1-R) \\ t \end{pmatrix} p^t \sum_{m=Bt}^{n} \frac{1}{p^{m(1-R)}} \begin{pmatrix} n \\ m \end{pmatrix} \left(2\frac{n}{m}\right)^m \left(\frac{n}{m}\right)^{1/2} (1+o(1)).
\end{aligned}$$

Using the inequalities $\begin{pmatrix} n \\ k \end{pmatrix} \leq n^k$ and $n/m \leq n$, we get

$$\begin{aligned}
\phi_1(n) &\leq c' \sum_{t=1}^{\vartheta n(1-R)} (n(1-R))^t p^t \sum_{m=Bt}^{n} \frac{(2n^2)^m}{p^{m(1-R)}} n^{1/2}(1+o(1)) \\
&= c' \sum_{t=1}^{\vartheta n(1-R)} (n(1-R))^t p^t \frac{(2n^2)^{Bt}}{p^{Bt(1-R)}} n^{1/2}(1+o(1)) \\
&= c' \sum_{t=1}^{\vartheta n(1-R)} (2^B(1-R))^t n^{t(1+\lambda+2B-\lambda B(1-R))} n^{1/2}(1+o(1)) \\
&\leq c' \sum_{t=1}^{\vartheta n(1-R)} (2^B(1-R))^t n^{t(3/2+\lambda+2B-\lambda B(1-R))} (1+o(1)).
\end{aligned} \tag{6.58}$$

But we have $3/2 + \lambda + 2B - \lambda B(1-R) < 0$, because the hypothesis of Theorem 6.7.1

guarantees that $\lambda > \frac{2B+3/2}{B(1-R)-1}$. Using Lemma 6.3.1, we can conclude that (6.58) is bounded from above by $(c''n)^{3/2+\lambda+2B-\lambda B(1-R)}(1+o(1))$ for some constant c'', and hence goes to zero as $n \to \infty$.

Case 2: $\vartheta n(1-R) \leq t < n(1-R)/2$

We will use property (R2) of the expander graph in this part of the proof. Since $|\mathrm{Supp}(\mathbf{u})| = t < n(1-R)/2$, we have $|N(\mathrm{Supp}(\mathbf{u}))| = |\mathbb{S}(\mathbf{u})| \geq \beta t$. Therefore, $\Pr[\mathbb{S}(\mathbf{u}) = m] = 0$ for $m < \beta t$. Proceeding along the same lines as in the previous subsection, we get

$$\begin{aligned}\phi_2(n) &\triangleq \sum_{t=\vartheta n(1-R)}^{n(1-R)/2} \sum_{\substack{\mathbf{u}\in\mathbb{F}_p^{n(1-R)} \\ |\mathrm{Supp}(\mathbf{u})|=t}} \sum_{m=1}^{n} \mathbb{1}_m(\mathbb{S}(\mathbf{u}))\frac{1}{p^m}|\{\mathbf{x}\in\mathbb{Z}^n \cap r_n\mathcal{B} : |\mathrm{Supp}(\mathbf{x})| \leq m\}| \\ &\leq \sum_{t=\vartheta n(1-R)}^{n(1-R)/2} \sum_{\substack{\mathbf{u}\in\mathbb{F}_p^{n(1-R)} \\ |\mathrm{Supp}(\mathbf{u})|=t}} \sum_{m=\beta t}^{n} \frac{1}{p^m}\binom{n}{m}\frac{V_m}{V_n^{m/n}}p^{mR}\left(1+\frac{\sqrt{m}}{2r_n}\right)^m(1+o(1)).\end{aligned}$$

Using (6.57), and the inequalities $\binom{n}{m} \leq 2^n$ and $1+\frac{\sqrt{m}}{2r_n} \leq 2$,

$$\begin{aligned}\phi_2(n) &\leq c\sum_{t=\vartheta n(1-R)}^{n(1-R)/2} \sum_{\substack{\mathbf{u}\in\mathbb{F}_p^{n(1-R)} \\ |\mathrm{Supp}(\mathbf{u})|=t}} \sum_{m=\beta t}^{n} \frac{1}{p^{m(1-R)}}\binom{n}{m}\left(\frac{n}{m}\right)^m\left(1+\frac{\sqrt{m}}{2r_n}\right)^m(1+o(1)) \\ &\leq c\sum_{t=\vartheta n(1-R)}^{n(1-R)/2} \sum_{\substack{\mathbf{u}\in\mathbb{F}_p^{n(1-R)} \\ |\mathrm{Supp}(\mathbf{u})|=t}} \sum_{m=\beta t}^{n} \frac{1}{p^{m(1-R)}}2^n\left(\frac{n}{m}\right)^m 2^m(1+o(1)).\end{aligned}$$

Since $n \geq m \geq \beta\vartheta n(1-R)$, we get

$$\begin{aligned}
\phi_2(n) &\leq c \sum_{t=\vartheta n(1-R)}^{n(1-R)/2} \sum_{\substack{\mathbf{u}\in\mathbb{F}_p^{n(1-R)} \\ |\mathrm{Supp}(\mathbf{u})|=t}} \sum_{m=\beta t}^{n} \frac{1}{p^{m(1-R)}} 2^n \left(\frac{1}{\beta\vartheta(1-R)}\right)^n 2^n(1+o(1)) \\
&\leq c \sum_{t=\vartheta n(1-R)}^{n(1-R)/2} \binom{n(1-R)}{t} p^t \sum_{m=\beta t}^{n} \frac{1}{p^{m(1-R)}} \left(\frac{4}{\beta\vartheta(1-R)}\right)^n (1+o(1)) \\
&\leq c \sum_{t=\vartheta n(1-R)}^{n(1-R)/2} 2^{n(1-R)} p^t \frac{1}{p^{\beta t(1-R)}} \left(\frac{4}{\beta\vartheta(1-R)}\right)^n (1+o(1)) \\
&\leq c 2^{n(1-R)} \frac{1}{p^{(\beta(1-R)-1)\vartheta n(1-R)}} \left(\frac{4}{\beta\vartheta(1-R)}\right)^n (1+o(1)),
\end{aligned} \tag{6.59}$$

which goes to zero as $n \to \infty$, since $\beta > 1/(1-R)$ from Definition 4.

Case 3: $n(1-R)/2 \leq t < (1-R-C_1/\ln n)n - 1$

We will use the following property of (α, A, β, B)-good expander graphs:

Lemma 6.7.2 ([21],Lemma 3.2)**.** *If $S \subset V$ is such that $|N(S)| < n(1-R)/2$, then $|S| \leq |N(S)|/\alpha$.*

Proof. Let us prove the contrapositive of the above statement. Suppose that $|S| > |N(S)|/\alpha$. Equivalently, $|N(S)| < \alpha|S|$. This implies that $|S| > n(1-R)/(2\alpha)$, otherwise we would be in violation of property (L2) of (α, A, β, B) graphs. But from (L2), we have $|N(S)| \geq \alpha n(1-R)/(2\alpha) = n(1-R)/2$, and this completes the proof. □

Since $T \triangleq \mathrm{Supp}(\mathbf{u})$ has at least $n(1-R)/2$ vertices, the set T^c has less than $n(1-R)/2$ vertices (see Fig. 6.3). If $S \triangleq \mathbb{S}(\mathbf{u}) = N(T)$, then, S^c has does not have any neighbours from T. Hence, $N(S^c) \subset T^c$. But $|T^c| < n(1-R)/2$ must imply that $|S^c| \leq |T^c|/\alpha$, from Lemma 6.7.2. Therefore, $n - |S| \leq (n(1-R) - |T|)/\alpha$, or $|S| \geq n(1-(1-R)/\alpha) + t/\alpha$. This means that $\Pr[\mathrm{Supp}(\mathbf{u}) = m] = 0$ for $m < n(1-(1-R)/\alpha) + t/\alpha$.

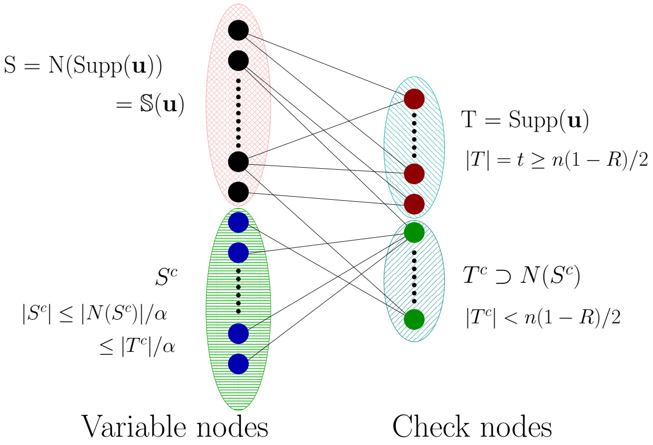

$\mathrm{S} = \mathrm{N}(\mathrm{Supp}(\mathbf{u}))$
$= \mathbb{S}(\mathbf{u})$
$\mathrm{T} = \mathrm{Supp}(\mathbf{u})$
$|T| = t \geq n(1-R)/2$
S^c
$|S^c| \leq |N(S^c)|/\alpha$
$\leq |T^c|/\alpha$
$T^c \supset N(S^c)$
$|T^c| < n(1-R)/2$
Variable nodes
Check nodes

Figure 6.3: Part 3 of proof.

Consider

$$\phi_3(n) \triangleq \sum_{t=n(1-R)/2}^{n(1-R-C_1/\ln n)} \sum_{\substack{\mathbf{u}\in\mathbb{F}_p^{n(1-R)} \\ |\mathrm{Supp}(\mathbf{u})|=t}} \sum_{m=1}^{n} \mathbb{1}_m(\mathbb{S}(\mathbf{u}))\frac{1}{p^m}|\{\mathbf{x}\in\mathbb{Z}^n\cap r_n\mathcal{B} : |\mathrm{Supp}(\mathbf{x})|\leq m\}|$$

$$\leq \sum_{t=n(1-R)/2}^{n(1-R-C_1/\ln n)} \sum_{\substack{\mathbf{u}\in\mathbb{F}_p^{n(1-R)} \\ |\mathrm{Supp}(\mathbf{u})|=t}} \sum_{m=n(1-(1-R)/\alpha)+t/\alpha}^{n} \frac{1}{p^m}|\{\mathbf{x}\in\mathbb{Z}^n\cap r_n\mathcal{B} : |\mathrm{Supp}(\mathbf{x})|\leq m\}|$$

Following the approach in the previous subsections, the above reduces to

$$\phi_3(n) \leq c \sum_{t=n(1-R)/2}^{n(1-R-C_1/\ln n)} \binom{n(1-R)}{t} p^t \sum_{m=n(1-(1-R)/\alpha)+t/\alpha}^{n} \frac{1}{p^{m(1-R)}} \binom{n}{m} \left(\frac{n}{m}\right)^m 2^m(1+o(1))$$

$$\leq c \sum_{t=n(1-R)/2}^{n(1-R-C_1/\ln n)} 8^n p^t \sum_{m=n(1-(1-R)/\alpha)+t/\alpha}^{n} \frac{1}{p^{m(1-R)}} \left(\frac{n}{m}\right)^n (1+o(1)),$$

where the last step uses the inequality $\binom{n}{k} \leq 2^n$. Since $m \geq n(1-(1-R)/\alpha)+t/\alpha \geq$

$n(1-(1-R)/\alpha+(1-R)/(2\alpha))$, we get

$$\begin{aligned}
\phi_3(n) &\le c \sum_{t=n(1-R)/2}^{n(1-R-C_1/\ln n)} 8^n p^t \sum_{m=n(1-(1-R)/\alpha)+t/\alpha}^{n} \frac{1}{p^{m(1-R)}} \left(\frac{1}{1-(1-R)/\alpha+(1-R)/(2\alpha)}\right)^n (1+o(1)) \\
&\le c \sum_{t=n(1-R)/2}^{n(1-R-C_1/\ln n)} \left(\frac{8}{1-(1-R)/(2\alpha)}\right)^n \frac{p^t}{p^{n(1-R)(1-(1-R)/\alpha)+t/\alpha}} (1+o(1)) \\
&= c \sum_{t=n(1-R)/2}^{n(1-R-C_1/\ln n)} n^{\frac{n\ln(8/(1-(1-R)/(2\alpha)))}{\ln(n)}} \frac{n^{\lambda t}}{n^{\lambda n(1-R)(1-(1-R)/\alpha)+\lambda t/\alpha}} (1+o(1)). \qquad (6.60)
\end{aligned}$$

If we have

$$\lambda n(1-R)\left(1-\frac{1-R}{\alpha}\right)+\lambda t\frac{(1-R)}{\alpha}-\lambda t-\frac{n}{\ln n}\ln\left(\frac{8}{1-(1-R)/(2\alpha)}\right) > 1+\delta$$

for some $\delta > 0$, then (6.60) is upper bounded by $cn \times n^{-1-\delta}(1+o(1))$, which goes to zero as $n \to \infty$. Simplifying the above quantity gives us the condition

$$t < n(1-R) - n\frac{C_1}{\ln n} - \frac{1+\delta}{\lambda(1-(1-R)/\alpha)},$$

which is satisfied in this regime, and hence, $\phi_3(n) \to 0$ as $n \to \infty$.

Case 4: $(1-R-C_1/\ln n)n - 1 \le t < n$

For any subset of parity check nodes, $T \subset C$, we have $|N(T)| \ge |T|/(1-R)$. This is because the number of edges between T and $N(T)$ is $|T|\Delta_V/(1-R)$, but the number of edges incident on each node in $N(T)$ from T is at most Δ_V. Therefore, we have

$$\begin{aligned}
\phi_4(n) &\triangleq \sum_{t=n(1-R-C_1/\ln n)}^{n} \sum_{\substack{\mathbf{u}\in\mathbb{F}_p^{n(1-R)}\\|\mathrm{Supp}(\mathbf{u})|=t}} \sum_{m=1}^{n} \mathbb{1}_m(\mathbb{S}(\mathbf{u}))\frac{1}{p^m}|\{\mathbf{x}\in\mathbb{Z}^n\cap r_n\mathcal{B} : |\mathrm{Supp}(\mathbf{x})|\leq m\}| \\
&\leq c\sum_{t=n(1-R-C_1/\ln n)}^{n}\binom{n(1-R)}{t}p^t\sum_{m=t/(1-R)}^{n}\frac{1}{p^{m(1-R)}}\binom{n}{m}\left(\frac{n}{m}\right)^m\zeta_n^m(1+o(1)) \\
&= c\sum_{t=n(1-R-C_1/\ln n)}^{n}\binom{n(1-R)}{n(1-R)-t}p^t\sum_{m=t/(1-R)}^{n}\frac{1}{p^{m(1-R)}}\binom{n}{n-m}\left(\frac{n}{m}\right)^m\zeta_n^m(1+o(1)).
\end{aligned}$$

Since $\binom{n}{n-k}$ is a decreasing function of k for $k>n/2$, we have

$$\begin{aligned}
\phi_4(n) \leq c &\sum_{t=n(1-R-C_1/\ln n)}^{n}\binom{n(1-R)}{nC_1/\ln n}p^t \\
&\times\sum_{m=t/(1-R)}^{n}\frac{1}{p^{m(1-R)}}\binom{n}{nC_1/((1-R)\ln n)}\left(\frac{n}{n-\frac{nC_1}{(1-R)\ln n}}\right)^m\zeta_n^m(1+o(1)).
\end{aligned}$$

Using the inequality $\binom{n}{m}\leq\left(\frac{ne}{m}\right)^m$ and simplifying, we get

$$\begin{aligned}
\phi_4(n) \leq c &\sum_{t=n(1-R-C_1/\ln n)}^{n}\left(\frac{e(1-R)\ln n}{C_1}\right)^{nC_1/\ln n}p^t\sum_{m=t/(1-R)}^{n}\frac{1}{p^{m(1-R)}} \\
&\times\left(\frac{e(1-R)\ln n}{C_1}\right)^{nC_1/((1-R)\ln n)}\left(\frac{1}{1-\frac{C_1}{(1-R)\ln n}}\right)^n\zeta_n^m(1+o(1)).
\end{aligned}$$

For all sufficiently large n, we have $m \geq n(1 - C_1/((1-R)\ln n)) > n/2$. Therefore, since $\zeta_n < 1$, we have

$$\begin{aligned}
\phi_4(n) &\leq c \sum_{t=n(1-R-C_1/\ln n)}^{n} \left(\frac{e(1-R)\ln n}{C_1}\right)^{nC_1/\ln n} p^t \sum_{m=t/(1-R)}^{n} \frac{1}{p^{m(1-R)}} \\
&\qquad \times \left(\frac{e(1-R)\ln n}{C_1}\right)^{nC_1/((1-R)\ln n)} \left(\frac{1}{1-\frac{C_1}{(1-R)\ln n}}\right)^n \zeta_n^{n/2}(1+o(1)) \\
&\leq c \sum_{t=n(1-R-C_1/\ln n)}^{n} \left(\frac{e(1-R)\ln n}{C_1}\right)^{2nC_1/((1-R)\ln n)} \left(\frac{1}{1-\frac{C_1}{(1-R)\ln n}}\right)^n \zeta_n^{n/2}(1+o(1)) \\
&\leq cn \left(\frac{e(1-R)\ln n}{C_1}\right)^{2nC_1/((1-R)\ln n)} \left(\frac{1}{1-\frac{C_1}{(1-R)\ln n}}\right)^n \zeta_n^{n/2}(1+o(1)),
\end{aligned}$$

which goes to zero as $n \to \infty$ because of our choice of ζ_n. This completes the proof of Theorem 6.7.1. □

6.8 Remarks

We now make some observations regarding our results and their applications to several problems. We first discuss the extension of our results to nested lattices, and then make some remarks regarding the choice of parameters and how it would affect the complexity of a BP decoder.

6.8.1 Construction of Nested Lattices

The main result of this chapter, namely Theorem 6.2.2, shows that a randomly chosen LDA lattice satisfies the desired "goodness" properties with high probability. In applications such as compute-and-forward, and coding for the AWGN channel, we need *nested lattices* which satisfy the necessary properties. Different nested lattice constructions have been proposed [28, 69, 73], and we briefly describe the construction by Ordentlich and Erez [73] here, since the results presented in this chapter can be easily extended to nested lattices using their construction.

Choose a $k_c \times n$ parity check matrix, H_c, over $\mathbb{F}_p$. Let $\mathcal{C}_c$ be the linear code that has parity check matrix H_c. Let H_f be the $k_f \times n$ parity check matrix ($k_f < k_c$) that consists of the first k_f rows of H_c, and $\mathcal{C}_f$ denote the corresponding linear code. Clearly, $\mathcal{C}_c$ is a subcode of $\mathcal{C}_f$. If Λ_c and Λ_f are lattices obtained by applying Construction A to $\mathcal{C}_c$ and $\mathcal{C}_f$ respectively, then $\Lambda_c \subset \Lambda_f$, with nesting ratio $p^{k_f - k_c}$ if the rows of H_c are linearly independent. The parity check matrix H_c can be chosen so that the Tanner graphs corresponding to both $\mathcal{C}_c$ and $\mathcal{C}_f$ have the required expansion properties [21, Section 4.3]. As long as λ and the parameters of the Tanner graph are chosen appropriately, the lattice Λ_c satisfies the goodness properties with probability tending to 1 as $n \to \infty$. Also, Λ_f satisfies the goodness properties with high probability. Using the union bound, we can argue that Λ_c and Λ_f simultaneously satisfy the goodness properties with probability tending to 1 as $n \to \infty$.

With this construction, we can use Theorem 6.2.2 to conclude that nested LDA lattices achieve the capacity of the AWGN channel, the capacity of the dirty paper channel, and the rates guaranteed by the compute-and-forward protocol [69]. Furthermore, they can also be used for secure bidirectional relaying, and achieve the rates guaranteed by Theorem 3.3.1. However, all of this is guaranteed under the assumption of a *closest lattice-point decoder* being used at the destination/relay. Although these lattices were empirically shown to give low error probability over the AWGN channel, their performance with belief propagation decoding still requires further study.

6.8.2 Choice of Parameters and Complexity of the BP Decoder

Theorem 6.2.2 gives sufficient conditions on the parameters required to obtain the structural goodness properties of a randomly chosen LDA lattice. In practice, one would want to optimize over the parameters in Theorem 6.2.2 to reduce the decoding complexity. At this point, we can only say that the achievability results for the various communication problems are valid with the assumption that a closest lattice-point decoder is used. However, in practice, we would want to use a belief propagation decoder instead. If this is done, then the decoding complexity would be roughly of the order of $np \log p$ (p messages need to be computed at each node, this having complexity $O(p \log p)$, and there are $O(n)$

nodes). Therefore, it is necessary to choose the smallest p for which the conditions of Theorem 6.2.2 are satisfied. Note that the condition $\lambda > \frac{2B+3/2}{B(1-R)-1}$ means that we should always have $\lambda > 2/(1-R)$. Choosing $R = 1/3$, we can make the lower bound on λ close to 3 by appropriately choosing A and B. This means that the decoding complexity would be roughly of the order of $n^4 \log n$. Although this means that we can decode in polynomial time, this complexity is still high when compared to the decoding complexity of the lattices presented in [86, 110]. However, it is still not known whether the lattices in [86, 110] have all the "goodness" properties that the LDA lattices satisfy.

Chapter 7

Concatenated Construction-A Lattices with Polynomial Encoding and Decoding Complexity

7.1 Introduction

As we discussed previously, a drawback with nested lattice codes is that their decoding complexity grows exponentially in the blocklength N. Unlike the case of linear codes for discrete memoryless channels, the encoding complexity of lattice codes also grows exponentially with N. While we now know that LDA lattices have all the "goodness" properties that we seek, studying the performance of these lattices under BP decoding seems to be a very difficult task.

As we noted in Chapter 1, there have been many efforts to design lattices with low-complexity decoding algorithms. Of particular interest is the polar lattice scheme proposed by Yan et al. [111] which can achieve the capacity of the AWGN channel with an encoding/decoding complexity of $O(N \log^2 N)$.[1] However, the probability of error is

[1] Yan et al. [111] also show that for a fixed error probability (as opposed to a probability of error that goes to zero as $N \to \infty$), the encoding/decoding complexity of polar lattices is $O(N \log N)$.

$e^{-O(\sqrt{N})}$. In this chapter, we give a concatenated coding scheme that achieves the capacity of the AWGN channel. The interesting feature of this scheme is that while the encoding and decoding complexities are polynomial in N, the probability of error decays exponentially in N.

Concatenated codes were introduced by Forney [34] as a technique for obtaining low-complexity codes that can achieve the capacity of discrete memoryless channels. Concatenating an inner random linear code with an outer Reed-Solomon code is a simple way of designing good codes. Using this idea, Joseph and Barron [45] proposed a capacity-achieving scheme for the AWGN channel with quadratic (in the blocklength N) encoding/decoding complexity. They used a concatenated coding scheme with an inner sparse superposition code and an outer Reed-Solomon code. The probability of decoding error goes to zero exponentially in $N/\log N$ [46].

In this chapter, we show that concatenating an inner nested lattice code with an outer Reed-Solomon code yields a capacity-achieving coding scheme whose encoding/decoding complexity is quadratic in the blocklength. Furthermore, the probability of error decays exponentially in N. Furthermore, we show that by replacing the Reed-Solomon code with an expander code [113] yields a capacity-achieving coding scheme with decoding complexity $O(N\log^2 N)$ and encoding complexity $O(N^2)$. To the best of our knowledge, this is the first capacity-achieving coding scheme for the AWGN channel whose encoding and decoding complexities are polynomial, and the probability of error decays exponentially in the blocklength. The techniques that we use are not new, and we use results from the works of Forney [34] and Erez and Zamir [28] to prove our results. An attractive feature of this technique is that it can also be used to reduce the complexity of nested lattice codes for several other Gaussian networks. It can be used as a tool to convert any nested lattice code having exponential decoding complexity to a code having polynomial decoding complexity. This comes at the expense of a minor reduction in performance (in terms of error probability) of the resulting code. As applications, we show how this can be used for the Gaussian wiretap channel and in reducing the decoding complexity of the compute-and-forward protocol for Gaussian networks. We will also use this technique to obtain polynomial-time lattice coding schemes for secret key generation from correlated sources in the next chapter.

Throughout this chapter, we measure complexity in terms of the number of binary operations required for decoding/encoding, and we are interested in how this complexity scales with the blocklength. We assume that arithmetic operations on real numbers are performed using floating-point arithmetic, and that each real number has a t-bit binary representation, with t being independent of the blocklength. The value of t would depend on the computer architecture used for computations (typically 32 or 64 bits). In essence, we assume that each floating-point operation has complexity $O(1)$.

The rest of the chapter is organized as follows. We describe the concatenated coding scheme for the AWGN channel in Section 7.2, with Theorem 7.2.2 summarizing the main result. In Section 7.3, we use an outer expander code to reduce the decoding complexity to $O(N\log^2 N)$. This is summarized by Theorem 7.3.2. Extension of the concatenated coding scheme to the Gaussian wiretap channel and the compute-and-forward protocol are outlined in Section 7.4.1 and Section 7.4.2 respectively.

7.2 Coding Scheme for the AWGN Channel

Let us consider the point-to-point AWGN channel where the source encodes its message M to $\mathbf{u}\in\mathbb{R}^n$ and transmits this to a destination that receives

$$\mathbf{w}=\mathbf{u}+\mathbf{z},$$

where $\mathbf{z}$ is the noise vector having independent and identically distributed (iid) Gaussian entries with mean zero and variance σ^2. Erez and Zamir [28] proposed a capacity-achieving nested lattice scheme for the AWGN channel, which we briefly describe here. The code is constructed using a pair of nested lattices $(\Lambda^{(n)},\Lambda_0^{(n)})$, where $\Lambda_0^{(n)}\subset\Lambda^{(n)}\subset\mathbb{R}^n$. The codebook consists of all the points of $\Lambda^{(n)}$ within the fundamental Voronoi region of $\Lambda_0^{(n)}$, i.e., the codebook is $\Lambda^{(n)}\cap\mathcal{V}(\Lambda_0^{(n)})$. The transmission rate is therefore $\frac{1}{n}\log_2|\Lambda^{(n)}\cap\mathcal{V}(\Lambda_0^{(n)})|$.

The source also generates a random dither vector $\mathbf{t}$, uniformly distributed over $\mathcal{V}(\Lambda_0^{(n)})$, which is assumed to be known to the decoder[2]. Each message M is mapped to a point $\mathbf{x}$ in the codebook $\Lambda^{(n)} \cap \mathcal{V}(\Lambda_0^{(n)})$. The encoder $\mathcal{E}^{(n)}$ takes the message M as input and outputs the vector $[\mathbf{x}-\mathbf{t}] \bmod \Lambda_0^{(n)}$, which is transmitted across the channel. This process of translating the message by $\mathbf{t}$ modulo $\Lambda_0^{(n)}$ prior to transmission is called *dithering*. The encoder satisfies a maximum transmit power constraint given by $\frac{1}{n}\max_{\mathbf{u}\in\mathcal{V}(\Lambda_0)}\|\mathbf{u}\|^2 = \frac{1}{n}{r_{\mathrm{cov}}}^2(\Lambda_0^{(n)}) < P$.

Upon receiving $\mathbf{w}$, the receiver uses a decoder $\mathcal{D}^{(n)}$ to estimate M, which does the following. It computes $\widetilde{\mathbf{w}} = [\alpha\mathbf{w} + \mathbf{t}] \bmod \Lambda_0^{(n)}$, where $\alpha = \frac{P}{P+\sigma^2}$. The estimate of M is the message that corresponds to $[Q_{\Lambda^{(n)}}(\widetilde{\mathbf{w}})] \bmod \Lambda_0^{(n)}$.

Let $C \triangleq \frac{1}{2}\log_2\left(1+\frac{P}{\sigma^2}\right)$. Erez and Zamir [28] showed that there exist nested lattices with which we can approach the capacity of the AWGN channel. Specifically,

Lemma 7.2.1 ([28], Theorem 5). *For every $\epsilon > 0$, there exists a sequence of nested Construction-A lattice pairs $(\Lambda^{(n)}, \Lambda_0^{(n)})$ such that for all sufficiently large n, the maximum transmit power,*

$$\frac{1}{n}{r_{\mathrm{cov}}}^2(\Lambda_0^{(n)}) \le P + \epsilon,$$

the transmission rate is[3]

$$R_{\mathrm{in}}^{(n)} \triangleq \frac{1}{n}\log_2|\Lambda^{(n)} \cap \mathcal{V}(\Lambda_0^{(n)})| \ge C - \epsilon,$$

and the probability of error decays exponentially in n for all $R_{\mathrm{in}}^{(n)} < C$, i.e., there exists a function $E : \mathbb{R}^+ \to \mathbb{R}^+$ so that for every $\mathbf{x} \in \Lambda^{(n)} \cap \mathcal{V}(\Lambda_0^{(n)})$ and all sufficiently large n, we have

$$\Pr[\mathbf{x} \neq \mathcal{D}^{(n)}(\mathcal{E}^{(n)}(\mathbf{x}) + \mathbf{z})] \le e^{-nE(R_{\mathrm{in}}^{(n)})}.$$

Furthermore, the quantity $E(R_{\mathrm{in}}^{(n)})$ is positive for all $R_{\mathrm{in}}^{(n)} < C$.

The decoding involves solving two closest lattice point problems, which are the $Q_{\Lambda^{(n)}}$

[2] In principle, the dither vector is not necessary. However, this technique of dithered transmission simplifies the analysis of the probability of error of the decoder.

[3] The subscript 'in' in $R_{\mathrm{in}}^{(n)}$ indicates that we intend to use the nested lattice code as an inner code in a concatenated coding scheme, which we describe in Section 7.2.1.

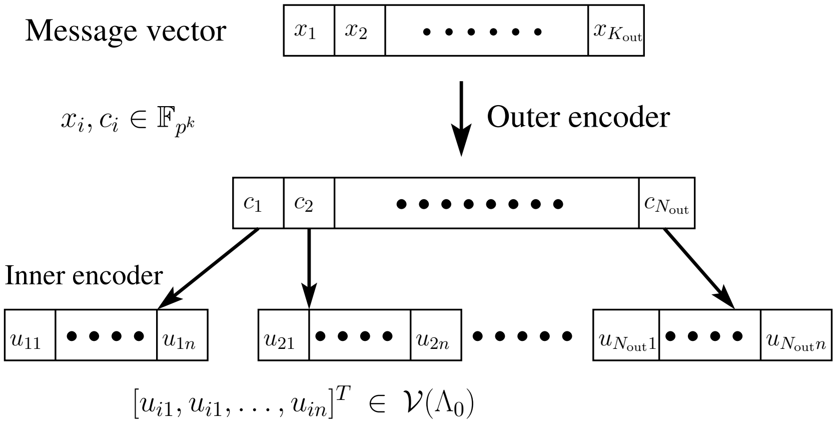


Figure 7.1: Illustration of the concatenated nested lattice coding scheme.

and mod$\Lambda_0^{(n)}$ operations. Therefore, the decoding complexity is $O(2^{nR_{\text{in}}})$. If the encoder uses a look-up table to map messages to codewords, the complexity would also be $O(2^{nR_{\text{in}}})$.

7.2.1 The Concatenated Coding Scheme for the AWGN Channel

Let us now give a brief description of the concatenated coding scheme. See [34] for a more detailed exposition and application to the discrete memoryless channel. The code has two components:

- Inner code: A nested Construction-A lattice code $(\Lambda^{(n)}, \Lambda_0^{(n)})$ with the fine lattice $\Lambda^{(n)}$ obtained from a (n, k) linear code over $\mathbb{F}_p$.
- Outer code: An $(N_{\text{out}}, K_{\text{out}}, d_{\text{out}})$ linear block code (where d_{out} is the minimum distance of the code) over $\mathbb{F}_{p^k}$.

The message set has size $p^{kK_{\text{out}}}$, and each message can be represented by a vector in $\mathbb{F}_{p^k}^{K_{\text{out}}}$. The outer code maps this vector to a codeword in $\mathbb{F}_{p^k}^{N_{\text{out}}}$ in a bijective manner. Let us call this $\mathbf{c}_{\text{out}} = [c_1 c_2 \cdots c_{N_{\text{out}}}]^T$, where each $c_i \in \mathbb{F}_{p^k}$. The inner code maps each $c_i \in \mathbb{F}_{p^k}$ to a point in $\Lambda^{(n)} \cap \mathcal{V}(\Lambda_0^{(n)})$. This results in a codeword of length nN_{out} having real-valued components. Each inner codeword is dithered by an independent dither vector prior to transmission. The encoding process is illustrated in Fig. 7.1. The receiver first uses the decoder for the inner code to estimate the components c_i, and finally uses the decoder for the outer code to recover the message. Since the outer code has minimum distance d_{out}, the message is guaranteed to be recovered correctly if not more than $(d_{\text{out}} - 1)/2$ inner codewords are in error. Furthermore, if all the inner codewords satisfy the (max) power constraint, then the concatenated code is also guaranteed to satisfy the same.

We now show that using this technique, we can achieve the capacity of the AWGN channel. Let us fix $\epsilon, \delta > 0$. Suppose we choose a sequence of nested lattice codes $(\Lambda^{(n)}, \Lambda_0^{(n)})$ that are guaranteed by Lemma 7.2.1. Recall that the number of cosets, $|\Lambda^{(n)} \cap \mathcal{V}(\Lambda_0^{(n)})|$, is equal to $p^k = 2^{nR_{\text{in}}}$. For every n, let us concatenate the nested lattice code with an outer $(N_{\text{out}}, K_{\text{out}}, N_{\text{out}} - K_{\text{out}} + 1)$ Reed-Solomon code over $\mathbb{F}_{p^k}$, where $N_{\text{out}} = p^k - 1$ and

$$K_{\text{out}} = N_{\text{out}}(1 - 2e^{-nE(R_{\text{in}})} - 2\delta). \tag{7.1}$$

Here and in the rest of this section, we drop the superscript in $R_{\text{in}}^{(n)}$ for convenience. The resulting code, which we denote $\mathcal{C}^{(n)}$, has blocklength $N = nN_{\text{out}} \approx n2^{nR_{\text{in}}}$, and rate

$$R^{(N)} = \frac{K_{\text{out}}}{nN_{\text{out}}} \log_2 p^k = (1 - 2e^{-nE(R_{\text{in}})} - 2\delta) R_{\text{in}}. \tag{7.2}$$

Theorem 7.2.2. *For every $\zeta > 0$, there exists a sequence of concatenated codes $\mathcal{C}^{(n)}$ with inner nested lattice codes and outer Reed-Solomon codes that satisfies the following for all sufficiently large n:*

- *the rate, $R^{(N)} \geq C - \zeta$,*
- *the maximum transmit power,*

$$\max_{\mathbf{x} \in \mathcal{C}^{(n)}} \frac{1}{N} \|\mathbf{x}\|^2 \leq P - \zeta,$$

- *the probability of error is at most $e^{-NE(R_{\text{in}})\zeta}$, and*
- *the encoding and decoding complexities grow as $O(N^2)$.*

Proof. The construction of the concatenated codes ensures that the power constraint can be satisfied. From Lemma 7.2.1, we are assured of nested lattice codes whose rate $R_{\text{in}} \geq C - \zeta/2$. Choosing a small enough δ and a large enough n in (7.2) guarantees that the effective rate $R^{(N)} \geq C - \zeta$ for all sufficiently large N.

Let us now proceed to analyze the probability of error. Clearly, the probability that an inner codeword is in error is upper bounded by $e^{-nE(R_{\text{in}})}$ by Lemma 7.2.1. Since the outer Reed-Solomon code has minimum distance $N_{\text{out}} - K_{\text{out}} + 1$, the decoder makes an error only if atleast $(N_{\text{out}} - K_{\text{out}} + 1)/2 = N_{\text{out}}(e^{-nE(R_{\text{in}})} + \delta + 1/(2N_{\text{out}}))$ inner codewords are in error. For all sufficiently large N, we can upper bound the probability of decoding error as follows:

$$\begin{aligned} P_e^{(N)} &\leq \binom{N_{\text{out}}}{N_{\text{out}}(e^{-nE(R_{\text{in}})} + \delta + 1/(2N_{\text{out}}))} \\ &\qquad \times \left(e^{-nE(R_{\text{in}})}\right)^{N_{\text{out}}(e^{-nE(R_{\text{in}})} + \delta + 1/(2N_{\text{out}}))} && (7.3) \\ &\leq \binom{N_{\text{out}}}{N_{\text{out}}(e^{-nE(R_{\text{in}})} + 2\delta)} \\ &\qquad \times \left(e^{-nE(R_{\text{in}})}\right)^{N_{\text{out}}(e^{-nE(R_{\text{in}})} + \delta)} && (7.4) \\ &\leq e^{N_{\text{out}} h(e^{-nE(R_{\text{in}})} + 2\delta)} \left(e^{-nE(R_{\text{in}})}\right)^{N_{\text{out}}(e^{-nE(R_{\text{in}})} + \delta)} && (7.5) \end{aligned}$$

where (7.3) is obtained using the union bound, and the last step from Stirling's formula. In (7.5), $h(\cdot)$ denotes the binary entropy function. For all sufficiently large n, we have $h(e^{-nE(R_{\text{in}})} + 2\delta) < h(3\delta)$. Using this in the above and simplifying, we get

$$P_e^{(N)} \leq \exp\Big(- nN_{\text{out}}\big(E(R_{\text{in}})(e^{-nE(R_{\text{in}})} + \delta) - h(3\delta)/n\big)\Big)$$

Let us define the error exponent as

$$E_{\text{conc}} \triangleq E(R_{\text{in}})(e^{-nE(R_{\text{in}})} + \delta) - h(3\delta)/n \tag{7.6}$$

It is clear that $E_{\text{conc}} > E(R_{\text{in}})\delta/2$ for all sufficiently large n. This proves that the probability of error decays exponentially in N.

Let us now inspect the encoding and decoding complexity. As remarked in the introduction, we assume that each floating-point operation requires a constant number of binary operations (i.e., independent of N) and has a complexity of $O(1)$. Encoding/decoding each inner (nested lattice) codeword requires $O(2^{nR_{\text{in}}})$ floating-point operations, and there are N_{out} many codewords, leading to a total complexity of $O(N^2)$. Furthermore, encoding/decoding a Reed-Solomon codeword requires $O(N_{\text{out}}^2)$ operations in $\mathbb{F}_{p^k}$ [79, Chapter 6]. Multiplication and inversion are the most computationally intensive operations in $\mathbb{F}_{p^k}$, and they can be performed using $O((k \log_2 p)^2) = O(n^2)$ binary operations [39, Chapter 2]. Therefore, the outer code has an encoding/decoding complexity of $O(N_{\text{out}}^2) \times O(n^2) = O(N^2)$. We can therefore conclude that encoding and decoding the concatenated code requires a complexity of $O(N^2)$. □

7.2.2 Complexity

Let us denote x to be the decoding complexity. From Theorem 7.2.2, we can conclude that for a fixed gap to capacity ($\gamma \triangleq C - R$), the probability of error for the concatenated coding scheme scales as $e^{-a\sqrt{x}}$ for some constant $a > 0$. As argued in [34, Section 5.1], this is a much stronger statement than saying that the decoding complexity is polynomial in the blocklength.

Previously, Joseph and Barron [45, 46] proposed a concatenated coding scheme with inner sparse superposition codes and outer Reed-Solomon codes. They showed that their scheme achieves the capacity of the AWGN channel with polynomial (in the blocklength N) time encoding/decoding. The decoding complexity is $x = O(N^2)$. However, the probability of error decays exponentially in $N/\log N$ for a fixed gap to capacity γ. Therefore, the probability of error is exponentially decaying in $\sqrt{x}/\log x$. More recently, Yan et al. [111] proposed a lattice-based scheme using polar codes that achieves capacity with an encoding/decoding complexity of $O(N \log N)$ for a fixed error probability, and $x = O(N \log^2 N)$ with a probability of error (for a fixed γ) which was exponential in

N^{β}, for some $0 < \beta < 0.5$. The probability of error is therefore $O(e^{-a(x/\log x)^{\beta}})$. The concatenated scheme we have studied here outperforms these works in the sense that the probability of error decays exponentially in the square root of x for a fixed γ. However, we have not been able to show that for a fixed probability of error, the decoding complexity is polynomial in the gap to capacity. The only such result that we are aware of is by Yan et al. [111], where they showed that polar lattices have a decoding complexity that is polynomial in the gap to the Poltyrev capacity (for the AWGN channel without restrictions/power constraint). Finding a capacity-achieving coding scheme for the power-constrained AWGN channel with a decoding complexity that scales polynomially in the gap to capacity for a fixed probability of error still remains an open problem.

7.3 Reduced Decoding Complexity using Expander Codes

In this section, we present a concatenation scheme that reduces the decoding complexity to $O(N \log^2 N)$. This is based on the parallel concatenation approach [4] of using outer expander-type codes to obtain “good” linear codes. It was shown in [4] that for binary channels, parallel concatenation yields an error performance similar to that of serial concatenation (concatenation using a Reed-Solomon code), but reduces the decoding complexity.

7.3.1 The Coding Scheme

Let us fix an $\epsilon > 0$ (where $\epsilon \ll 1$), and let our target rate be $R = C - \epsilon$, where C denotes the capacity of the AWGN channel. As in the previous section, let N_{out} and n denote the blocklengths of the outer and inner codes respectively. The overall blocklength is $N = nN_{\text{out}}$. Unlike the previous section, however, we fix n to be a sufficiently large constant, and we let N_{out} grow to infinity.

Let us fix the rate of the inner code to be

$$R_{\text{in}}^{(n)} = C - \frac{\epsilon}{2},$$

Then, Lemma 7.2.1 guarantees the existence of a sequence of nested Construction-A lattice codes $(\Lambda^{(n)}, \Lambda_0^{(n)})$ with rate $\frac{1}{n}\log_2|\Lambda^{(n)} \cap \mathcal{V}(\Lambda_0^{(n)})| \geq R_{\text{in}}^{(n)}$, for which the probability of error,

$$P_{e,\text{in}}^{(n)} \triangleq \Pr\Big[\mathbf{x} \neq \mathcal{D}^{(n)}(\mathcal{E}^{(n)}(\mathbf{x}) + \mathbf{z})\Big] \leq e^{-nE(R_{\text{in}}^{(n)})}$$

for all sufficiently large n. Furthermore, these lattices are obtained from linear codes over $\mathbb{F}_p$ for prime p (which is a function of n). Let us fix an n large enough so that

$$-\ln P_{e,\text{in}}^{(n)} = nE(R_{\text{in}}^{(n)}) \geq \frac{h\,(4\sqrt{\epsilon}) + \epsilon^2}{0.8\epsilon}, \tag{7.7}$$

where $h(\cdot)$ denotes the binary entropy function. For a fixed ϵ, the parameters n and p will remain constant, and we will let only N_{out} grow to infinity.

The outer code

We use an outer expander code whose construction is similar to the one in [113]. This has two components:

- A sequence of Δ-regular bipartite expander graphs $\mathcal{G}^{(N_{\text{out}})} = (\mathcal{A}^{(N_{\text{out}})}, \mathcal{B}^{(N_{\text{out}})}, \mathcal{E}^{(N_{\text{out}})})$ with vertex set $\mathcal{A}^{(N_{\text{out}})} \cup \mathcal{B}^{(N_{\text{out}})}$ and edge set $\mathcal{E}^{(N_{\text{out}})}$, with $|\mathcal{E}^{(N_{\text{out}})}| = N_{\text{out}}$. Here, $\mathcal{A}^{(N_{\text{out}})}$ denotes the set of left vertices and $\mathcal{B}^{(N_{\text{out}})}$ denotes the set of right vertices, with $|\mathcal{A}^{(N_{\text{out}})}| = |\mathcal{B}^{(N_{\text{out}})}| = \frac{N_{\text{out}}}{2\Delta}$. Here, Δ is a large constant independent of N_{out}. The second-largest eigenvalue of the adjacency matrix, $\lambda(\mathcal{G}^{(N_{\text{out}})}) \leq 2\sqrt{\Delta - 1}$. Explicit constructions of such graphs can be found in the literature [60] (see [63] for a stronger result). This graph is a normal factor graph for the outer expander code. We choose a sufficiently large Δ so that $\frac{2\sqrt{\Delta-1}}{\Delta} \leq \sqrt{\epsilon}$.

- A linear code $\mathcal{C}_0$ over $\mathbb{F}_{p^k}$ having blocklength Δ and dimension k_0. For convenience, let us choose $\mathcal{C}_0$ to be a (Δ, k_0) Reed Solomon code over $\mathbb{F}_{p^k}$ (assuming that $\Delta < p^k$)

with $k_0 = \Delta\left(1-4\sqrt{\epsilon}\right)+1$. The minimum Hamming distance of $\mathcal{C}_0$ is $d_0 = \Delta - k_0 + 1 = 4\sqrt{\epsilon}\Delta$. Let us define

$$\delta_0 \triangleq \frac{d_0}{\Delta} = 4\sqrt{\epsilon}. \tag{7.8}$$

Let us order the edges of $\mathcal{G}^{(N_{\text{out}})}$ in any arbitrary fashion, and for any $v \in \mathcal{A}^{(N_{\text{out}})} \cup \mathcal{B}^{(N_{\text{out}})}$, let $E_v \triangleq \{e_v(1), \ldots, e_v(\Delta)\}$ denote the set of edges incident on v, where $e_v(1) < e_v(2) < \cdots < e_v(\Delta)$ according to the order we have fixed. We define the expander code as follows: The codeword entries are indexed by the edges of $\mathcal{G}^{(N_{\text{out}})}$. A vector $\mathbf{x} \in \mathbb{F}_{p^k}$ is a codeword of the expander code iff for every $v \in \mathcal{A}^{(N_{\text{out}})} \cup \mathcal{B}^{(N_{\text{out}})}$, we have that $(x_{e_v(1)}x_{e_v(2)}\ldots x_{e_v(\Delta)})$ is a codeword in $\mathcal{C}_0$. Following [113], we will call this the $(\mathcal{G}^{(N_{\text{out}})}, \mathcal{C}_0)$ code. The $(\mathcal{G}^{(N_{\text{out}})}, \mathcal{C}_0)$ code has blocklength N_{out} and dimension at least $N_{\text{out}}\left(1-2\left(\frac{\Delta-k_0}{\Delta}\right)\right)$ [113].

Zémor [113] proposed an iterative algorithm to decode an expander code, which successively replaces $\mathbf{x}_v = (x_{e_v(1)}x_{e_v(2)}\ldots x_{e_v(\Delta)})$ by a nearest-neighbour codeword (to $\mathbf{x}_v$) in $\mathcal{C}_0$ for every $v \in \mathcal{A}^{(N_{\text{out}})}$, and then for every $v \in \mathcal{B}^{(N_{\text{out}})}$. Clearly, $\{E_v : v \in \mathcal{A}^{(N_{\text{out}})}\}$ forms a disjoint partition of the edge set $\mathcal{E}^{(N_{\text{out}})}$, and the nearest-neighbour decoding can be done in parallel for all the v's in $\mathcal{A}^{(N_{\text{out}})}$. The same holds for the vertices in $\mathcal{B}^{(N_{\text{out}})}$. We direct the interested reader to [113] for more details about the code and the iterative decoding algorithm.

Lemma 7.3.1 ([113])**.** *Let $\alpha < 1$ be fixed. The iterative decoding algorithm can be implemented in a circuit of size $O(N_{\text{out}} \log N_{\text{out}})$ and depth $O(\log N_{\text{out}})$ that always returns the correct codeword as long as the number of errors is less than $\frac{\alpha\delta_0 N_{\text{out}}}{2}\left(\frac{\delta_0}{2} - \frac{\lambda(\mathcal{G}^{(N_{\text{out}})})}{\Delta}\right)$.*

Since $\delta_0 = 4\sqrt{\epsilon}$ and $\frac{\lambda(\mathcal{G}^{(N_{\text{out}})})}{\Delta} \leq \frac{2\sqrt{\Delta-1}}{\Delta} \leq \sqrt{\epsilon}$, we see from Lemma 7.3.1 that the decoder can recover the transmitted outer codeword as long as the fraction of errors is less than $2\alpha\epsilon$. Although Lemma 7.3.1 was proved for binary expander codes, it can be verified that the result continues to hold in the case where the expander code is defined over $\mathbb{F}_{p^k}$, provided that p^k is a constant independent of N_{out}.

7.3.2 Performance of the Coding Scheme

We will show the following result.

Theorem 7.3.2. *For every $\epsilon > 0$, there exists a sequence of concatenated codes $\mathcal{C}^{(N)}$ with inner nested lattice codes and outer expander codes that satisfies the following for all sufficiently large N:*

- *the rate, $R^{(N)} \geq C - \epsilon$,*
- *the maximum transmit power,*

$$\max_{\mathbf{x}\in\mathcal{C}^{(n)}} \frac{1}{N}\|\mathbf{x}\|^2 \leq P - \epsilon,$$

- *the probability of error is at most $e^{-NE(R_{\text{in}})\epsilon}$,*
- *the encoding complexity grows as $O(N^2)$, and*
- *the decoding complexity grows as $O(N\log^2 N)$.*

Proof. Recall that the overall blocklength $N = nN_{\text{out}}$, where n is a sufficiently large constant. The probability that an inner (lattice) codeword is recovered incorrectly is at most $P_{e,\text{in}}^{(n)}$. Let us fix $\alpha = 0.9$ and define $\delta_{\text{out}} \triangleq \frac{\alpha\delta_0}{2}\left(\frac{\delta_0}{2} - \frac{\lambda(\mathcal{G}^{(N_{\text{out}})})}{\Delta}\right)$, the fraction of errors that the outer expander code is guaranteed to correct according to Lemma 7.3.1. From our choice of parameters, this quantity is at least 1.8ϵ. The probability of error of the concatenated code can be upper bounded as follows:

$$\begin{aligned}
P_{e,\text{concat}}^{(N)} &\leq \begin{pmatrix} N_{\text{out}} \\ \delta_{\text{out}}N_{\text{out}}+1\end{pmatrix}\left(P_{e,\text{in}}^{(n)}\right)^{\delta_{\text{out}}N_{\text{out}}+1} \\
&\leq e^{N_{\text{out}}\left(h\left(\delta_{\text{out}}+1/N_{\text{out}}\right)+o_{N_{\text{out}}}(1)\right)} \\
&\qquad \times e^{-nE(R_{\text{in}}^{(n)})(N_{\text{out}}\delta_{\text{out}}+1)} \\
&\leq e^{N_{\text{out}}(h(\delta_{\text{out}})+o_{N_{\text{out}}}(1))}e^{-nN_{\text{out}}\delta_{\text{out}}E(R_{\text{in}}^{(n)})} \qquad (7.9)
\end{aligned}$$

For all sufficiently large N_{out}, we can say that

$$\begin{aligned} P_{e,\text{concat}}^{(N)} &\leq \exp\left(-N\left(\delta_{\text{out}}E(R_{\text{in}}^{(n)}) - \frac{(h(\delta_{\text{out}})+\epsilon^2)}{n}\right)\right) \\ &= \exp\left(-NE_{\text{conc}}(R_{\text{in}}^{(n)}, \mathcal{G}^{(N_{\text{out}})}, \mathcal{C}_0)\right), \qquad (7.10) \end{aligned}$$

where the error exponent,

$$E_{\text{conc}}(R_{\text{in}}^{(n)}, \mathcal{G}^{(N_{\text{out}})}, \mathcal{C}_0) \triangleq \delta_{\text{out}}E(R_{\text{in}}^{(n)}) - \frac{(h(\delta_{\text{out}})+\epsilon^2)}{n}.$$

Since $1.8\epsilon \leq \delta_{\text{out}} < \delta_0 = 4\sqrt{\epsilon}$, we have

$$\begin{aligned} E_{\text{conc}}(R_{\text{in}}^{(n)}, \mathcal{G}^{(N_{\text{out}})}, \mathcal{C}_0) &\geq 1.8\epsilon E(R_{\text{in}}^{(n)}) - \frac{(h(4\sqrt{\epsilon})+\epsilon^2)}{n} \\ &= E(R_{\text{in}}^{(n)})\left(1.8\epsilon - \frac{(h(4\sqrt{\epsilon})+\epsilon^2)}{nE(R_{\text{in}}^{(n)})}\right) \\ &\geq E(R_{\text{in}}^{(n)})\epsilon \qquad (7.11) \end{aligned}$$

by our choice of n in (7.7).

Let us now inspect the encoding and decoding complexity. Recall that each floating-point operation has a complexity of $O(1)$. Since n is a constant, encoding/decoding each inner (nested lattice) codeword requires $O(1)$ floating-point operations, and there are N_{out} many codewords, leading to a total complexity of $O(N_{\text{out}})$. Since the outer code is linear, encoding requires $O(N_{\text{out}}^2)$ operations in $\mathbb{F}_{p^k}$. Since p^k is a constant, the outer code has an encoding complexity of $O(N_{\text{out}}^2) = O(N^2)$. Decoding the outer code requires $O(N_{\text{out}} \log^2 N_{\text{out}})$ operations in $\mathbb{F}_{p^k}$. We can therefore conclude that the decoding the concatenated code requires a complexity of $O(N \log^2 N)$, and encoding requires a complexity of $O(N^2)$. This completes the proof of Theorem 7.3.2. □

Scheme	**Decoding complexity (χ)**	**Encoding complexity**	**Error probability**	**Error probability as a function of χ**
Polar lattice [111]	$O(N\log^2 N)$	$O(N\log^2 N)$	$e^{-\Omega(\sqrt{N})}$	$e^{-\Omega(\sqrt{\chi/\log\chi})}$
Sparse superposition scheme [45]	$O(N^2)$	$O(N^2)$	$e^{-\Omega(N/\log N)}$	$e^{-\Omega(\sqrt{\chi}/\log\chi)}$
RS-concatenated lattice codes	$O(N^2)$	$O(N^2)$	$e^{-\Omega(N)}$	$e^{-\Omega(\sqrt{\chi})}$
Expander-concatenated lattice codes	$O(N\log^2 N)$	$O(N^2)$	$e^{-\Omega(N)}$	$e^{-\Omega(\sqrt{\chi})}$

Figure 7.2: A comparison of the performance of various polynomial-time capacity-achieving codes.

7.4 Discussion

The approach used in the previous sections can be used as a recipe for reducing the decoding complexity of optimal coding schemes for Gaussian channels. A nested lattice scheme that achieves a rate R over a Gaussian channel can be concatenated with a high-rate outer Reed-Solomon code or expander code to achieve any rate arbitrarily close to R. The only requirement is that the nested lattice code has a probability of error which decays exponentially in its blocklength. This procedure helps us bring down the decoding complexity to a polynomial function of the blocklength while ensuring that the probability of error continues to be an exponential function of the blocklength. As examples, we discuss two important applications: a scheme that achieves the secrecy capacity of the wiretap channel, and the compute-and-forward protocol for computing linear combinations of several messages over a Gaussian channel. Since our objective is only to suggest potential applications, we will keep the details to a minimum.

7.4.1 The Gaussian Wiretap Channel

The Gaussian wiretap channel [52] consists of three parties: a source, a destination and an eavesdropper. The source has a message M which is intended only for the destination but not the eavesdropper. The source encodes M to a real n-dimensional vector $\mathbf{u}$ and transmits it across to the destination. The destination receives $\mathbf{w}_{\mathrm{D}}$, which is modeled as

$$\mathbf{w}_{\mathrm{D}} = \mathbf{u} + \mathbf{z}_{\mathrm{D}},$$

where $\mathbf{z}_{\mathrm{D}}$ is AWGN with mean zero and variance σ^2. On the other hand, the eavesdropper observes

$$\mathbf{w}_{\mathrm{E}} = \mathbf{u} + \mathbf{z}_{\mathrm{E}},$$

where $\mathbf{z}_{\mathrm{E}}$ is AWGN with mean zero and variance $\sigma_{\mathrm{E}}^2 > \sigma^2$. We want to ensure that the destination recovers M with negligible probability of error, while the eavesdropper gets very little information about M. Specifically, we require $I(M; \mathbf{w}_{\mathrm{E}}) \to 0$ as $n \to \infty$. This is also called the *strong secrecy* constraint. A more detailed exposition of the wiretap and related channels studied in information-theoretic security can be found in [11]. We define the secrecy capacity of the wiretap channel as the supremum over all achievable rates while satisfying the strong secrecy constraint. If we define $C_{\mathrm{M}} \triangleq \frac{1}{2}\log_2(1 + P/\sigma^2)$ and $C_{\mathrm{E}} \triangleq \frac{1}{2}\log_2(1 + P/\sigma_{\mathrm{E}}^2)$ to be the capacities of the main and eavesdropper channels respectively, then the secrecy capacity of this wiretap channel is $C_{\mathrm{M}} - C_{\mathrm{E}}$.

Tyagi and Vardy [92] gave an explicit scheme using 2-universal hash functions that converts any coding scheme of rate R over the point-to-point AWGN (main) channel to a coding scheme that achieves a rate $R - C_{\mathrm{E}}$ over the wiretap channel while satisfying the strong secrecy constraint. This "conversion" adds an additional decoding complexity which is polynomial in the blocklength. We can therefore use this result with Theorem 7.2.2 or Theorem 7.3.2 to conclude that we can achieve the secrecy capacity of the Gaussian wiretap channel with polynomial time decoding/encoding.

7.4.2 Compute-and-forward

The compute-and-forward protocol was proposed by Nazer and Gastpar [69] for communication over Gaussian networks. Let us begin by describing the setup. We have L source nodes $\mathtt{S}_1, \mathtt{S}_2, \ldots, \mathtt{S}_L$, having independent messages $X_1, X_2, \ldots, X_L$ respectively. The messages are chosen from $\mathbb{F}_{p^k}^K$ for some prime number p and positive integers k, K. Let $\oplus$ denote the addition operator in $\mathbb{F}_{p^k}^K$. These messages are mapped to N-dimensional real vectors $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_L$ respectively and transmitted across a Gaussian channel to a destination $\mathtt{D}$ which observes

$$\mathbf{w} = \sum_{l=1}^{L} h_l \mathbf{u}_l + \mathbf{z}, \tag{7.12}$$

where $h_1, h_2, \ldots, h_L$ are real-valued channel coefficients and $\mathbf{z}$ is AWGN with mean zero and variance σ^2. The destination must compute $a_1X_1 \oplus a_2X_2 \oplus \cdots \oplus a_LX_L$, where $a_1, a_2, \ldots, a_L$ are integers. We assume that each source node must satisfy a maximum power constraint of P. We only consider symmetric rates here, i.e., all sources have identical message sets. The rate of the code is $\frac{kK}{N}\log_2 p$. This problem is relevant in many applications such as exchange of messages in bidirectional relay networks, decoding messages over the Gaussian multiple access channel [69], and designing good receivers for MIMO channels [114] to name a few. The basic idea is that instead of decoding the messages one at a time and using successive cancellation, it may be more efficient to decode multiple linear combinations of the messages. If we have L linearly independent such combinations, then we can recover all the individual messages.

We can extend the scheme of [69] to a concatenated coding scheme that achieves the rates guaranteed by [69], but now with encoders and decoders that operate in polynomial time. Recall that the messages are chosen from $\mathbb{F}_{p^k}^K$. We say that a rate $\mathcal{R}$ is achievable if for every $\epsilon > 0$, there exists a sequence of encoders and decoders so that for all sufficiently large blocklengths N, we have the transmission rate $R^{(N)} \triangleq \frac{kK}{N}\log_2 p > \mathcal{R} - \epsilon$, and the probability of error is less than ϵ. We can show the following:

Lemma 7.4.1. *Consider the problem of computing $a_1X_1 \oplus a_2X_2 \oplus \cdots \oplus a_LX_L$ from (7.12).*

Any rate

$$\mathcal{R} < \frac{1}{2}\log_2\left(\frac{P}{\alpha^2 + P\sum_{l=1}^{L}(\alpha h_l - a_l)^2}\right), \tag{7.13}$$

where

$$\alpha \triangleq \frac{P\sum_{l=1}^{L} h_l a_l}{\sigma^2 + P\sum_{l=1}^{L} h_l^2}, \tag{7.14}$$

is achievable with encoders and decoders whose complexities grow as $O(N^2)$ using an outer Reed-Solomon code, and a decoder whose complexity grows as $O(N\log^2 N)$ with an outer expander code. For transmission rates less than $\mathcal{R}$, the probability that the decoder makes an error goes to zero exponentially in N.

Proof. The technique used to justify this claim is a simple extension of the coding scheme of [69] using the methods described in Section 7.2. For completeness, we will briefly describe the scheme. For more details regarding the compute-and-forward protocol, see [69]. We use the concatenated coding scheme of Section 7.2.1. The inner code is obtained from nested Construction-A lattices $(\Lambda^{(n)}, \Lambda_0^{(n)})$. Suppose that $\Lambda^{(n)}$ is constructed using a (n,k) linear code over $\mathbb{F}_p$. The outer code is an $(N_{\text{out}}, K_{\text{out}}, N_{\text{out}} - K_{\text{out}} + 1)$ Reed-Solomon code, with $N_{\text{out}} = p^k - 1$ and K_{out} to be specified later. The transmission rate is $R^{(n)} = \frac{kK_{\text{out}}}{nN_{\text{out}}}\log_2 p$.

The messages are chosen from $\mathbb{F}_{p^k}^{K_{\text{out}}}$. Let the message at the lth user be $M_l = [m_1^{(l)}, m_2^{(l)}, \ldots, m_{K_{\text{out}}}^{(l)}]^T$, where $m_i^{(l)} \in \mathbb{F}_{p^k}$. The messages are mapped to an N_{out}-length codeword over $\mathbb{F}_{p^k}$ using the outer code. Let us denote the resulting codeword by $\mathbf{y}^{(l)} = [y_1^{(l)}, y_2^{(l)}, \ldots, y_{N_{\text{out}}}^{(l)}]^T$.

Each $y_i^{(l)}$ is then encoded to $\mathbf{u}_i^{(l)}$ using the inner code and then transmitted. Recall that there exists a group isomorphism from $\Lambda^{(n)}/\Lambda_0^{(n)}$ to $\mathbb{F}_{p^k}$. For $1 \le l \le L$ and $1 \le i \le N_{\text{out}}$, let $\mathbf{x}_i^{(l)}$ be the representative of $y_i^{(l)}$ in $\Lambda^{(n)} \cap \mathcal{V}(\Lambda_0^{(n)})$. Independent dither vectors $\mathbf{t}_1^{(l)}, \mathbf{t}_2^{(l)}, \ldots, \mathbf{t}_{N_{\text{out}}}^{(l)}$ are generated at the L sources. Transmitter l successively sends $\mathbf{u}_i^{(l)} = [\mathbf{x}_i^{(l)} - \mathbf{t}_i^{(l)}] \bmod \Lambda_0^{(n)}$ for $1 \le i \le N_{\text{out}}$ to the receiver.

The decoder, upon receiving $\mathbf{w}_i = \sum_{l=1}^{L}\mathbf{u}_i^{(l)} + \mathbf{z}$, computes $\widetilde{\mathbf{w}}_i = \left[\alpha\mathbf{w}_i + \sum_{l=1}^{L} a_l\mathbf{t}_i^{(l)}\right] \bmod \Lambda_0^{(n)}$. The estimate of $[\sum_{l=1}^{L} a_i\mathbf{x}_i^{(l)}] \bmod \Lambda_0^{(n)}$, is $[Q_{\Lambda^{(n)}}(\widetilde{\mathbf{w}}_i)] \bmod \Lambda_0^{(n)}$. Recall the definition of $\mathcal{R}$ in (7.13). Nazer and Gastpar showed in [69] that there exists a sequence of nested Construction-A lattices with $R_{\text{in}}^{(n)} = \frac{k}{n}\log_2 p$ for which the probability that the decoder makes an error in estimating the desired linear combination decays as $e^{-nE_c(R_{\text{in}}^{(n)})}$, where $E_c(\cdot)$ is some function which is positive for all $R_{\text{in}}^{(n)} < \mathcal{R}$. As we did before for the AWGN channel, we choose $K_{\text{out}} = N_{\text{out}}(1 - 2e^{-nE_c(R_{\text{in}}^{(n)})} - \epsilon)$. Assuming that fewer than $(N_{\text{out}} - K_{\text{out}})/2$ inner codewords are in error, the decoder can recover $\widehat{\mathbf{x}}_c = \Big[\big[\sum_l a_l \mathbf{x}_1^{(l)}\big] \bmod \Lambda_0^{(n)}, \ldots, \big[\sum_l a_l \mathbf{x}_{N_{\text{out}}}^{(l)}\big] \bmod \Lambda_0^{(n)}\Big]^T$ without error. Due to the existence of a group isomorphism between $\mathbb{F}_{p^k}$ and $\Lambda^{(n)}/\Lambda_0^{(n)}$, this implies that the decoder can recover $a_1\mathbf{y}^{(1)} \oplus \cdots \oplus a_L\mathbf{y}^{(L)}$, and hence, $a_1M_1 \oplus \cdots \oplus a_LM_L$. Arguing as in Section 7.2, the probability that the decoder makes an error goes to zero exponentially in N, and the decoding/encoding complexities grow as $O(N^2)$. The same arguments can be used to show that using an outer expander code, we can have the encoding complexity to be $O(N^2)$ and decoding complexity to be $O(N\log^2 N)$. □

7.4.3 Concluding Remarks

We have seen that concatenation can be a very powerful tool in reducing the asymptotic decoding complexity of nested lattice codes. However, it must be noted that achieving good performance using this scheme would require very large blocklengths. Although the probability of error decays exponentially in N, and the decoding/encoding complexities are polynomial in N, this is true only for very large values of N. The fact that N is at least exponential in the blocklength of the inner code is a major reason for this. Nevertheless, the concatenated coding approach shows that it is possible to obtain polynomial-time encoders and decoders for which the probability of error decays exponentially in the blocklength. The exponential decay is under the assumption that the gap between the transmission rate and capacity, $\gamma = C - R$, is kept fixed. For a fixed error probability P_e, the blocklength required by the concatenated coding scheme to achieve rate $R = C - \gamma$ and error probability P_e does not scale polynomially with $1/\gamma$. For a fixed error probability, we would like the complexity to not grow too fast as the rate approaches C. Ideally, we would want the minimum blocklength N required to achieve rate $R = C - \gamma$ and error probability P_e to be polynomial in $1/\gamma$ for a fixed P_e. It has been recently shown that polar codes have this property for binary memoryless symmetric channels [68]. Designing codes for the Gaussian channel whose decoding/encoding complexities are also polynomial in $1/\gamma$ for a fixed probability of error still remains an open problem.

Chapter 8

Secret Key Generation from Correlated Gaussian Markov Tree Sources

8.1 Introduction

In this chapter, we consider the application of lattice codes to the problem of secret key (SK) generation in the multiterminal source model, where m terminals possess correlated Gaussian sources. Each terminal observes N independent and identically distributed (iid) samples of its source. The terminals have access to a noiseless public channel of infinite capacity, and their objective is to agree upon a secret key by communicating across the public channel. The key must be such that any eavesdropper having access to the public communication must not be able to guess the key. In other words, the key must be independent (or almost independent) of the messages communicated across the channel. A measure of performance is the secret key rate that can be achieved, which is the number of bits of secret key generated per (source) sample. On the other hand, the probability that any terminal is unable to reconstruct the key correctly should be arbitrarily small.

The discrete setting — the case where the correlated sources take values in a finite alphabet — was studied by Csiszár and Narayan [19]. They gave a scheme for computing a secret key in this setting and found the secret key capacity, i.e., the maximum achievable secret key rate. This was later generalized by Nitinawarat and Narayan [71] to the case where the terminals possess correlated Gaussian sources.

In a practical setting, we can assume that the random sources are obtained by observing some natural parameters, e.g., temperature in a field. In other words, the underlying source is continuous. However, for the purposes of storage and computation, these sources must be quantized, and only the quantized source can be used for secret key generation. If each terminal uses a scalar quantizer, then we get the discrete source model studied in [19]. However, we could do better and instead use a vector quantizer to obtain a higher secret key rate. Nitinawarat and Narayan [71] found the secret key capacity for correlated Gaussian sources in such a setting. However, to approach capacity, the quantization rate and the rate of public communication at each terminal must approach infinity. In practice, it is reasonable to have a constraint on the quantization rate at each terminal. The terminals can only use the quantized source for secret key generation. Nitinawarat and Narayan [71] studied a two-terminal version of this problem, where a quantization rate constraint was imposed on only one of the terminals. They gave a nested lattice coding scheme and showed that it was optimal, i.e., no other scheme can give a higher secret key rate. In related work, Watanabe and Oohama [106] characterized the maximum secret key rate achievable under a constraint on the rate of public communication in the two-terminal setting. More recently, Ling et al. [57] gave a lattice coding scheme for the public communication-constrained problem and were able to achieve a secret key rate within $1/2$ nats of the maximum in [106].

We consider a multiterminal generalization of the two-terminal version studied by [71] where quantization rate constraints are imposed on each of the terminals. Terminal i has access to N iid copies of a Gaussian source $X_i(1), X_i(2), \ldots, X_i(N)$. The sources are correlated across the terminals. We assume that the joint distribution of the sources has a *Markov tree* structure [19, Example 7], which is a generalization of a Markov chain. Let us define this formally. Suppose that $T = (V, E)$ is a tree and $\{X_{\mathtt{v}} : \mathtt{v} \in V\}$ is a collection of random variables indexed by the vertices. Consider any two disjoint subsets

$\mathcal{I}$ and $\mathcal{J}$ of V. Let $\mathtt{v}$ be any vertex such that removal of $\mathtt{v}$ from T disconnects $\mathcal{I}$ from $\mathcal{J}$ (Equivalently, for every $\mathtt{i} \in \mathcal{I}$ and $\mathtt{j} \in \mathcal{J}$, the path connecting $\mathtt{i}$ and $\mathtt{j}$ passes through $\mathtt{v}$). For every such $\mathcal{I}, \mathcal{J}, \mathtt{v}$, if $\{X_{\mathtt{i}} : \mathtt{i} \in \mathcal{I}\}$ and $\{X_{\mathtt{j}} : \mathtt{j} \in \mathcal{J}\}$ are conditionally independent given X_v, then we say that $\{X_{\mathtt{i}} : \mathtt{i} \in T\}$ form a Markov chain on T. Alternatively, we say that $\{X_{\mathtt{i}} : \mathtt{i} \in V\}$ is a Markov tree source.

In this chapter, we study the problem of secret key generation in a Gaussian Markov tree source model with individual quantization rate constraints imposed at each terminal. We give a nested lattice-based scheme and find the achievable secret key rate. For certain classes of Markov trees, particularly homogeneous Markov trees[1], we show that our scheme achieves the secret key capacity as the quantization rates go to infinity. However, we also give examples where our scheme does not achieve the key capacity. A salient feature of our scheme is that the overall computational complexity required for quantization and key generation is polynomial in the number of samples N. To the best of our knowledge, this is the first polynomial-time scheme for secret key generation from Gaussian sources. It is also interesting to note that unlike the general schemes in [19, 71], we give a scheme where at least one terminal remains silent (does not participate in public communication), and omniscience is not attained.

8.2 Notation and Definitions

If $\mathcal{I}$ is an index set and $\{A_i : i \in \mathcal{I}\}$ is a class of sets indexed by $\mathcal{I}$, then their Cartesian product is denoted by $\bigtimes_{i\in\mathcal{I}} A_i$. Let $G = (V, E)$ be a graph. The distance between two vertices $\mathtt{u}$ and $\mathtt{v}$ in G is the length of the shortest path between $\mathtt{u}$ and $\mathtt{v}$. Given a rooted tree $T = (V, E)$ with root $\mathtt{r}(T)$ we say that a vertex $\mathtt{u}$ is the parent of $\mathtt{v} \neq \mathtt{r}(T)$, denoted $\mathtt{u} = \mathrm{par}(\mathtt{v})$, if $\mathtt{u}$ lies in the shortest path from $\mathtt{r}(T)$ to $\mathtt{v}$ and the distance between $\mathtt{u}$ and $\mathtt{v}$ is 1. Furthermore, for every $\mathtt{v} \in V$, we define $N_T(\mathtt{v})$ to be the set of all neighbours of $\mathtt{v}$ in T.

[1] We say that a Markov tree is homogeneous if $I(X_{\mathtt{u}}; X_{\mathtt{v}})$ is the same for all edges $(\mathtt{u}, \mathtt{v})$

8.3 Secret Key Generation from Correlated Gaussian Sources

8.3.1 The Problem

We now formally define the problem. We consider a multiterminal Gaussian source model [71], which is described as follows. There are m terminals, each having access to N independent and identically distributed (iid) copies of a correlated Gaussian source, i.e., the lth terminal observes $X_l(1), X_l(2), \dots, X_l(N)$ which are iid. Without loss of generality, we can assume that $X_l(i)$ has mean zero and variance 1. We can always subtract the mean and divide by the variance to ensure that this is indeed the case. The joint distribution of $\{X_l(i) : 1 \le l \le m\}$ can be described by their covariance matrix Φ.

Specifically, we assume that the sources form a Markov tree, defined in Sec. 8.1. Let $T = (V, E)$ be a tree having $|V| = m$ vertices, which defines the conditional independence structure of the sources. For $\mathtt{u}, \mathtt{v} \in V$, let us define $\rho_{\mathtt{uv}} \triangleq \mathbb{E}[X_{\mathtt{u}} X_{\mathtt{v}}]$.

We can therefore write

$$X_{\mathtt{u}} = \rho_{\mathtt{uv}} X_{\mathtt{v}} + \sqrt{1 - \rho_{\mathtt{uv}}^2}\, Z_{\mathtt{uv}}$$

where $Z_{\mathtt{uv}}$ is a zero-mean, unit-variance Gaussian random variable which is independent of $X_{\mathtt{v}}$. Similarly,

$$X_{\mathtt{v}} = \rho_{\mathtt{uv}} X_{\mathtt{u}} + \sqrt{1 - \rho_{\mathtt{uv}}^2}\, Z_{\mathtt{vu}}$$

where $Z_{\mathtt{vu}}$ is also a zero-mean, unit-variance Gaussian random variable which is independent of $X_{\mathtt{u}}$ (and different from $Z_{\mathtt{uv}}$).

Our objective is to generate a secret key using public communication. For $\mathtt{v} \in V$, let $\mathbf{X}_{\mathtt{v}}^N \triangleq (X_{\mathtt{v}}(1), X_{\mathtt{v}}(2), \dots, X_{\mathtt{v}}(N))$ denote the N iid copies of $X_{\mathtt{v}}$ available at terminal $\mathtt{v}$. We assume that the eavesdropper only has access to the public communication, and nothing else. Each terminal uses a vector quantizer $Q_{\mathtt{v}} : \mathbb{R}^N \to \mathcal{X}_{\mathtt{v}}$ of rate $R_{\mathrm{q}}^{(\mathtt{v})} \triangleq \frac{1}{N} \log_2 |\mathcal{X}_{\mathtt{v}}|$. Terminal $\mathtt{v}$ transmits $\mathbf{F}_{\mathtt{v}}^{(N)} \in \mathcal{F}_{\mathtt{v}}^{(N)}$ — which is a (possibly randomized) function of $Q_{\mathtt{v}}(\mathbf{X}_{\mathtt{v}}^N)$

— across a noiseless public channel that an eavesdropper may have access to[2]. Using the public communication and their respective observations of the quantized random variables, $Q_{\mathtt{v}}(\mathbf{X}_{\mathtt{v}}^N)$, the terminals must generate a secret key $\mathbf{K}^{(N)} \in \mathcal{K}^{(N)}$ which is concealed from the eavesdropper. Let $\mathcal{F}_G \triangleq \bigtimes_{\mathtt{v}\in V} \mathcal{F}_{\mathtt{v}}^{(N)}$.

Fix any $\epsilon > 0$. We say that $\mathbf{K}^{(N)}$ is an ϵ-secret key (ϵ-SK) if there exist functions $f_{\mathtt{v}} : (\mathcal{X}_{\mathtt{v}}, \mathcal{F}_G) \to \mathcal{K}^{(N)}$ such that:

$$\Pr\left[f_{\mathtt{v}}(Q_{\mathtt{v}}(\mathbf{X}_{\mathtt{v}}^N), \{\mathbf{F}_{\mathtt{u}}^{(N)} : \mathtt{u} \in V\}) \neq \mathbf{K}^{(N)}\right] < \epsilon,$$

$$\log_2 |\mathcal{K}^{(N)}| - H(\mathbf{K}^{(N)}) < \epsilon,$$

and

$$I\left(\{\mathbf{F}_{\mathtt{v}}^{(N)} : \mathtt{v} \in V\}; \mathbf{K}^{(N)}\right) < \epsilon.$$

We define $\frac{1}{N}\log_2|\mathcal{K}^{(N)}|$ to be the rate of the ϵ-secret key. We say that R_{key} is an achievable secret key rate if for every $\epsilon > 0$, there exist quantizers $\{Q_{\mathtt{v}}\}$, a scheme for public communication, $\{\mathbf{F}_{\mathtt{v}}^{(N)}\}$, and a secret key $\mathbf{K}^{(N)}$, such that for all sufficiently large N, $\mathbf{K}^{(N)}$ is an ϵ-SK, and $\frac{1}{N}\log_2|\mathcal{K}^{(N)}| \geq R_{\text{key}} - \epsilon$.

Consider the following procedure to obtain a class of rooted subtrees of T:

- Identify a vertex $\mathtt{v}$ in V as the root. The tree T with $\mathtt{v}$ as the root is a rooted tree. Call this $T'_{\mathtt{v}}$.
- Delete all the leaves of $T'_{\mathtt{v}}$. Call the resulting rooted subtree $T^*_{\mathtt{v}}$.

Let $\mathcal{T}^* \triangleq \{T^*_{\mathtt{v}} : \mathtt{v} \in V\}$ denote the set of all rooted subtrees of T obtained in the above manner. Fig. 8.1 illustrates this for a tree having four vertices. Note that there are $|V|$ trees in $\mathcal{T}^*$, one corresponding to each vertex. For any such rooted subtree $T^* = (V^*, E^*)$ in $\mathcal{T}^*$, let $\mathtt{r}(T^*)$ denote the root of T^*. We will see later that it is only the terminals that correspond to T^* that participate in the public communication while the other terminals remain silent. For any $\mathtt{v} \in V^*$, let $N_T(\mathtt{v})$ denote the set of all neighbours of $\mathtt{v}$ in T (and

[2] In this work, we only consider noninteractive communication, i.e., the public communication is only a function of the source and not of the prior communication.

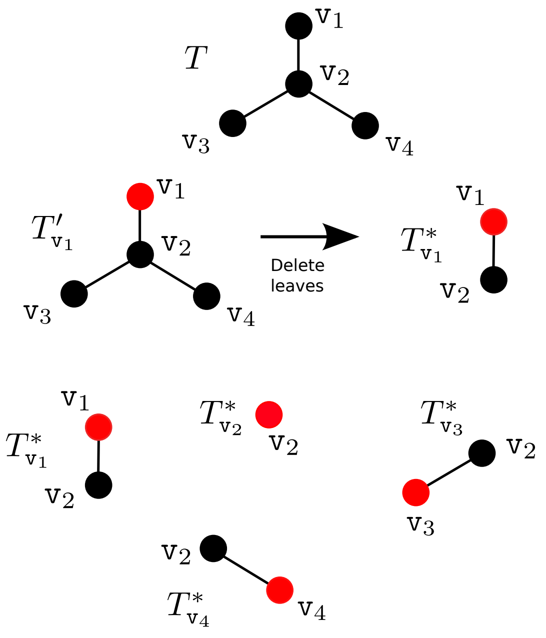


Figure 8.1: Illustration of $\mathcal{T}^*$ for a tree having four vertices.

not just those in T^*). Recall that each terminal $\mathtt{v}$ operates under a quantization rate constraint of $R_{\mathrm{q}}^{(\mathtt{v})}$. For every $T^* = (V^*, E^*)$, us define

$$\mathcal{R}_{\text{ent}} = R_q^{(\mathtt{r}(T^*))} + \sum_{\mathtt{u}\in V^*\setminus\mathtt{r}(T^*)} \frac{1}{2}\log_2\left((e^{2R_{\mathrm{q}}^{(\mathtt{u})}}-1)(1-\rho^2_{\mathtt{u},\mathrm{par}(\mathtt{u})})+1\right) \tag{8.1}$$

and

$$\mathcal{R}_{\text{com}} = \sum_{\mathtt{v}\in V^*} \max_{\mathtt{u}\in N_T(\mathtt{v})} \frac{1}{2}\log_2\left((e^{2R_{\mathrm{q}}^{(\mathtt{v})}}-1)(1-\rho^2_{\mathtt{uv}})+1+\frac{\rho^2_{\mathtt{uv}}e^{2R_{\mathrm{q}}^{(\mathtt{v})}}}{e^{2R_{\mathrm{q}}^{(\mathtt{u})}}-1}\right). \tag{8.2}$$

We will show that the joint entropy of the quantized sources is at least $\mathcal{R}_{\text{ent}}$ and the sum rate of public communication is at most $\mathcal{R}_{\text{com}}$ in our scheme. Also, the public communication that achieves $\mathcal{R}_{\text{com}}$ requires only the terminals in T^* to participate in the communication; the terminals in $V\setminus V^*$ are silent. Let us also define

$$\alpha \triangleq \frac{\max_{\mathtt{u}\in V^*} R_{\mathrm{q}}^{(\mathtt{u})}}{\min_{\mathtt{v}\in V^*} R_{\mathrm{q}}^{(\mathtt{v})}}. \tag{8.3}$$

Our aim is to prove the following result.

Theorem 8.3.1. *For a fixed quantization rate constraint* $\{R_{\mathrm{q}}^{(\mathtt{v})} : \mathtt{v} \in V\}$*, a secret key rate of*

$$R_{\mathrm{key}} = \max_{T^* \in \mathcal{T}^*} \left\{ \mathcal{R}_{\mathrm{ent}} - \mathcal{R}_{\mathrm{com}} \right\} \tag{8.4}$$

is achievable using a nested lattice coding scheme whose computational complexity grows as $O(N^{\alpha+1})$*.*

Note that if all terminals have identical quantization rate constraints, then the complexity is $O(N^2)$. Section 8.5 describes the scheme and contains the proof of the above theorem.

We now discuss some of the implications of the result. Letting the quantization rates $R_{\mathrm{q}}^{(\mathtt{u})}$ in (8.4) go to infinity, i.e., as $R_{\mathrm{q}}^{(\mathtt{v})} \to \infty$ for all $\mathtt{v}$, we get that

Corollary 8.3.2. *In the fine quantization limit, a secret key rate of*

$$R_{\mathrm{key}} = \max_{T^* \in \mathcal{T}^*} \left\{ \min_{\mathtt{v} \in N_T(\mathtt{r}(T^*))} \frac{1}{2} \log_2 \left(\frac{1}{1 - \rho^2_{\mathtt{r}(T^*)\mathtt{v}}} \right) + \sum_{\mathtt{u} \in V^* \setminus \mathtt{r}(T^*)} \min_{\mathtt{v} \in N_T(\mathtt{u})} \frac{1}{2} \log_2 \left(\frac{1 - \rho^2_{\mathtt{u},\mathrm{par}(\mathtt{u})}}{1 - \rho^2_{\mathtt{u},\mathtt{v}}} \right) \right\} \tag{8.5}$$

is achievable.

8.4 Remarks on the Achievable Secret Key Rate

8.4.1 The Two-User Case

Consider the two-user case with terminals $\mathtt{u}$ and $\mathtt{v}$. Let us define

$$\mathcal{R}(\mathtt{u}, \mathtt{v}) \triangleq \frac{1}{2} \log_2 \left(\frac{e^{2R_{\mathrm{q}}^{(\mathtt{u})}}}{(e^{2R_{\mathrm{q}}^{(\mathtt{u})}} - 1)(1 - \rho^2_{\mathtt{uv}}) + 1 + \frac{\rho^2_{\mathtt{uv}} e^{2R_{\mathrm{q}}^{(\mathtt{u})}}}{e^{2R_{\mathrm{q}}^{(\mathtt{v})}} - 1}} \right)$$

As we will see later, the above SK rate is achieved with $\mathtt{u}$ participating in the public communication and $\mathtt{v}$ remaining silent. The achievable secret key rate given by (8.4) is

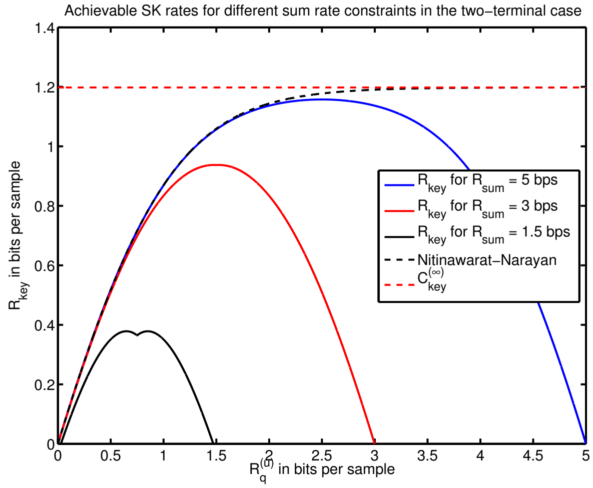


Figure 8.2: Plot of achievable secret key rates under a sum rate constraint for two terminals.

equal to $\max\{\mathcal{R}(\mathtt{u},\mathtt{v}), \mathcal{R}(\mathtt{v},\mathtt{u})\}$. A simple calculation reveals that

$$e^{-2\mathcal{R}(\mathtt{u},\mathtt{v})} - e^{-2\mathcal{R}(\mathtt{v},\mathtt{u})} = \rho_{\mathtt{uv}}^2 \left(\frac{1}{e^{2R_{\mathrm{q}}^{(\mathtt{v})}}(e^{2R_{\mathrm{q}}^{(\mathtt{v})}} - 1)} - \frac{1}{e^{2R_{\mathrm{q}}^{(\mathtt{u})}}(e^{2R_{\mathrm{q}}^{(\mathtt{u})}} - 1)} \right) \tag{8.6}$$

Hence, if $R_{\mathrm{q}}^{(\mathtt{v})} > R_{\mathrm{q}}^{(\mathtt{u})}$, then $\mathcal{R}(\mathtt{u},\mathtt{v}) > \mathcal{R}(\mathtt{v},\mathtt{u})$. This means that in order to obtain a higher secret key rate using our scheme, the terminal with the lower quantization rate must communicate, while the other must remain silent.

If we let $R_{\mathrm{q}}^{(\mathtt{v})}$ in $\mathcal{R}(\mathtt{u},\mathtt{v})$ go to infinity, then we get the rate R_{NN} achieved in [71], which was shown to be optimal when we only restrict the quantization rate of one terminal.

$$R_{\mathrm{NN}} = \frac{1}{2}\log_2 \left(\frac{e^{2R_{\mathrm{q}}^{(\mathtt{u})}}}{(e^{2R_{\mathrm{q}}^{(\mathtt{u})}} - 1)(1 - \rho_{\mathtt{uv}}^2) + 1} \right).$$

Fig. 8.2 illustrates the behaviour of the achievable rate for different sum-rate constraints ($R_{\mathrm{q}}^{(\mathtt{u})} + R_{\mathrm{q}}^{(\mathtt{v})} = R$). The rate achieved by the scheme of Nitinawarat and Narayan [71], R_{NN}, is also shown.

8.4.2 Optimality of R_{key} in the Fine Quantization Limit

If there are no constraints on the quantization rates, then from [71, Theorem 3.1] and [19, Example 7], we know that the maximum achievable secret key rate is

$$C_{\text{key}}^{(\infty)} = \min_{(\mathtt{u},\mathtt{v})\in E} \frac{1}{2}\log_2\left(\frac{1}{1-\rho_{\mathtt{uv}}^2}\right). \tag{8.7}$$

We present a class of examples where R_{key} is equal to the secret key capacity $C_{\text{key}}^{(\infty)}$ in the fine quantization limit. One such example is the class of *homogeneous* Markov trees, where $\rho_{\mathtt{uv}} = \rho$ for all edges $(\mathtt{u},\mathtt{v})$. In this case,

$$\min_{\mathtt{v}\in N_T(\mathtt{u})} \frac{1}{2}\log_2\left(\frac{1-\rho_{\mathtt{u},\text{par}(\mathtt{u})}^2}{1-\rho_{\mathtt{u},\mathtt{v}}^2}\right) = 0,$$

and hence, by Corollary 8.3.2,

$$R_{\text{key}} = \frac{1}{2}\log_2\left(\frac{1}{1-\rho^2}\right) = C_{\text{key}}^{(\infty)}.$$

This property holds for a wider class of examples. Consider the case where T has a rooted subtree T^* such that for every $\mathtt{u} \in V^*$, $\arg\min_{\mathtt{v}\in N_T(\mathtt{u})} \rho_{\mathtt{uv}} = \text{par}(\mathtt{u})$. Once again, we have

$$\min_{\mathtt{v}\in N_T(\mathtt{u})} \frac{1}{2}\log_2\left(\frac{1-\rho_{\mathtt{u},\text{par}(\mathtt{u})}^2}{1-\rho_{\mathtt{u},\mathtt{v}}^2}\right) = 0.$$

Moreover, the edge $(\mathtt{u},\mathtt{v}) \in E$ with the minimizing $\rho_{\mathtt{uv}}$ (and therefore, the minimizing mutual information) is incident on $\mathtt{r}(T^*)$. Hence, $R_{\text{key}} = C_{\text{key}}^{(\infty)}$.

8.4.3 Suboptimality of R_{key} in the Fine Quantization Limit

We can give several examples for which R_{key} in the fine quantization limit is strictly less than $C_{\text{key}}^{(\infty)}$. Note that

$$\min_{\mathtt{v}\in N_T(\mathtt{u})} \frac{1}{2}\log_2\left(\frac{1-\rho_{\mathtt{u},\text{par}(\mathtt{u})}^2}{1-\rho_{\mathtt{u},\mathtt{v}}^2}\right) \leq 0,$$

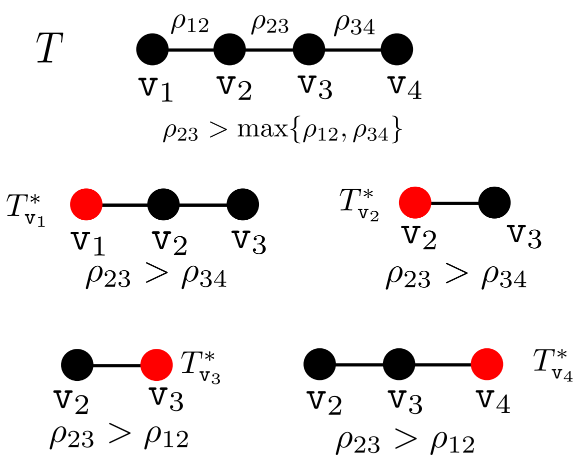


Figure 8.3: An example where our scheme is suboptimal.

and if these terms are nonzero for every $T^* \in \mathcal{T}^*$, then the scheme is suboptimal. As a specific example, consider the Markov chain of Fig. 8.3, where $\rho_{23} > \max\{\rho_{12}, \rho_{34}\}$. Let us further assume that $\rho_{12} = \rho_{34}$. The secret key capacity is

$$C_{\text{key}}^{(\infty)} = \frac{1}{2}\log_2 \frac{1}{1-\rho_{12}^2}.$$

Irrespective of which $T^* \in \mathcal{T}^*$ we choose (see Fig. 8.3), we have

$$\min_{\mathtt{v}\in N_T(\mathtt{r}(T^*))} \frac{1}{2}\log_2\left(\frac{1}{1-\rho_{\mathtt{r}(T^*)\mathtt{v}}^2}\right) = C_{\text{key}}^{(\infty)}.$$

Furthermore, the second term in (8.5) is negative for every T^*. This is because every T^* has some $\mathtt{u} \neq \mathtt{r}(T^*)$ for which $\arg\min_{\mathtt{v}\in V^*} \rho_{\mathtt{uv}} \neq \text{par}(\mathtt{v})$.

8.5 The Secret Key Generation Scheme

We now describe the lattice coding scheme that achieves the promised secret key rate. Our scheme is very similar to the scheme given by Nitinawarat and Narayan [71] for the two-terminal case.

We use a block encoding scheme just like the one in [71]. Recall that each terminal

$\mathtt{v}$ has a quantization rate constraint of $R_{\mathrm{q}}^{(\mathtt{v})}$. The total blocklength N is partitioned into N_{out} blocks of n samples each, i.e., $N = nN_{\mathrm{out}}$, where $N_{\mathrm{out}} = \min_{\mathtt{v}} 2^{nR_{\mathrm{q}}^{(\mathtt{v})}} - 1$. The secret key generation scheme comprises two phases: an information reconciliation phase, and a privacy amplification phase. The reconciliation phase is divided into two subphases: a lattice coding-based analog phase, which is followed by a Reed-Solomon coding-based digital phase. The privacy amplification phase employs a linear mapping to generate the secret key from the reconciled information. This uses the results of Nitinawarat and Narayan [71]. The digital phase is also inspired by the concatenated coding scheme used in Chapter 7 in the context of channel coding for Gaussian channels.

Let us briefly outline the protocol for secret key generation. Each terminal $\mathtt{v}$ uses a chain of nested lattices $(\Lambda_{\mathtt{v}}, \Lambda_{\mathtt{v}}^{(1)}, \Lambda_{\mathtt{v}}^{(2)})$ in $\mathbb{R}^n$, where $\Lambda_{\mathtt{v}}^{(2)} \subset \Lambda_{\mathtt{v}}^{(1)} \subset \Lambda_{\mathtt{v}}$. The Gaussian input $\mathbf{x}_{\mathtt{v}}$ at terminal $\mathtt{v}$ is processed blockwise, with n samples collected to form a block. Suppose $\mathbf{x}_{\mathtt{v}} = (\mathbf{x}_{\mathtt{v}}^{(1)}, \ldots, \mathbf{x}_{\mathtt{v}}^{(N_{\mathrm{out}})})$ where $\mathbf{x}_{\mathtt{v}}^{(i)}$ denotes the ith block of length n. Each terminal $\mathtt{v}$ also generates random dithers $\{\mathbf{d}_{\mathtt{v}}^{(i)} : 1 \le i \le N_{\mathrm{out}}\}$, which are all uniformly distributed over the fundamental Voronoi region of $\Lambda_{\mathtt{v}}$, and independent of each other. These are assumed to be known to all terminals.[3] The protocol for secret key generation is as follows.

- *Quantization:* Terminal $\mathtt{v} \in V$ computes

$$\mathbf{y}_{\mathtt{v}}^{(i)} = \left[Q_{\Lambda_{\mathtt{v}}}(\mathbf{x}_{\mathtt{v}}^{(i)} + \mathbf{d}_{\mathtt{v}}^{(i)}) - \mathbf{d}_{\mathtt{v}}^{(i)}\right] \bmod \Lambda_{\mathtt{v}}^{(2)}.$$

- *Information reconciliation: Analog phase:* Let $T^* = (V^*, E^*)$ be the rooted subtree which achieves the maximum in (8.4). Terminal $\mathtt{v} \in V^*$ broadcasts

$$\mathbf{w}_{\mathtt{v}}^{(i)} = [\mathbf{y}_{\mathtt{v}}^{(i)}] \bmod \Lambda_{\mathtt{v}}^{(1)}.$$

  across the public channel. Terminal $\mathtt{u}$ has access to $\mathbf{y}_{\mathtt{u}}^{(i)}$ and for every $\mathtt{v} \in N_{T^*}(\mathtt{u})$,

[3] In principle, the random dither is not required. Similar to [69], we can show that there exist fixed dithers for which all our results hold. One could avoid the use of dithers by employing the techniques in [57], but we do not take that approach here.

it estimates

$$\widehat{\mathbf{y}}_{\mathtt{v}}^{(i)} = \mathbf{w}_{\mathtt{v}}^{(i)} + Q_{\Lambda_{\mathtt{v}}^{(1)}}\left(\rho_{\mathtt{uv}}\mathbf{y}_{\mathtt{u}}^{(i)} - \mathbf{w}_{\mathtt{v}}^{(i)}\right).$$

Having estimated $\mathbf{y}_{\mathtt{v}}^{(i)}$ for all neighbours $\mathtt{v}$, it estimates $\mathbf{y}_{\mathtt{v}}^{(i)}$ for all $\mathtt{v}$ which are at a distance 2 from $\mathtt{u}$, and so on, till it has estimated $\{\mathbf{y}_{\mathtt{v}}^{(i)} : \mathtt{v} \in V^*, 1 \leq i \leq N_{\text{out}}\}$.

- *Information reconciliation: Digital phase:* To ensure that all N_{out} blocks can be recovered at all terminals with an arbitrarily low probability of error, we use a Slepian-Wolf scheme using Reed-Solomon codes. Each terminal uses an $(N_{\text{out}}, K_{\text{out}})$ Reed-Solomon code over $\mathbb{F}_{p^{k_{\mathtt{v}}}}$, where the parameters K_{out} and $p^{k_{\mathtt{v}}}$ will be specified later. The syndrome corresponding to[4] $(\mathbf{y}_{\mathtt{v}}^{(1)}, \ldots, \mathbf{y}_{\mathtt{v}}^{(N_{\text{out}})})$ in the code is publicly communicated by terminal $\mathtt{v}$. We show that this can be used by the other terminals to estimate all the $\mathbf{y}_{\mathtt{v}}^{(i)}$s with a probability of error that decays exponentially in N.

- *Key generation:* We use the result [71, Lemma 4.5] that there exists a linear transformation of the source symbols (viewed as elements of a certain finite field) that can act as the secret key. Since all terminals can estimate $\{\mathbf{y}_{\mathtt{v}}^{(i)} : \mathtt{v} \in V^*, 1 \leq i \leq N_{\text{out}}\}$ reliably, they can all compute the secret key with an arbitrarily low probability of error.

Before we go into the details of each step, we describe some specifics of the coding scheme. We want the nested lattices that form the main component of our protocol to satisfy certain "goodness" properties. We begin by describing the features that the lattices must possess.

8.5.1 Nested Lattices

Each terminal $\mathtt{v}$ uses a chain of n-dimensional nested lattices $(\Lambda_{\mathtt{v}}, \Lambda_{\mathtt{v}}^{(1)}, \Lambda_{\mathtt{v}}^{(2)})$, with $\Lambda_{\mathtt{v}}^{(2)} \subset \Lambda_{\mathtt{v}}^{(1)} \subset \Lambda_{\mathtt{v}}$. These are all Construction-A lattices [28, 29] obtained from linear codes of blocklength n over $\mathbb{F}_p$, with p chosen large enough to ensure that these lattices satisfy the required goodness properties. Furthermore, $\Lambda_{\mathtt{v}}^{(1)}$ and $\Lambda_{\mathtt{v}}^{(2)}$ are obtained from subcodes of

[4] We show that there is a bijection between $\Lambda_{\mathtt{v}}/\Lambda_{\mathtt{v}}^{(2)}$ and $\mathbb{F}_{p^{k_{\mathtt{v}}}}$.

linear codes that generate $\Lambda_{\mathtt{v}}$. Fix any $\delta > 0$. The lattices are chosen so that

$$\frac{1}{n}\log_2 |\Lambda_{\mathtt{v}} \cap \mathcal{V}(\Lambda_{\mathtt{v}}^{(2)})| = \frac{1}{n}\log_2 \frac{\mathrm{vol}(\Lambda_{\mathtt{v}}^{(2)})}{\mathrm{vol}(\Lambda_{\mathtt{v}})} = \frac{k_{\mathtt{v}}}{n}\log_2 p = R_{\mathrm{q}}^{(\mathtt{v})}, \tag{8.8}$$

$$\frac{\big(\mathrm{vol}(\Lambda_{\mathtt{v}}^{(2)})\big)^{2/n}}{2\pi e} = \big(1+\sigma^2(\Lambda_{\mathtt{v}})\big)(1+\delta), \tag{8.9}$$

and

$$\begin{aligned}\frac{\big(\mathrm{vol}(\Lambda_{\mathtt{v}}^{(1)})\big)^{2/n}}{2\pi e} = \max_{\mathtt{u}\in N_T(\mathtt{v})} &\Big(1-\rho_{\mathtt{uv}}^2+\sigma^2(\Lambda_{\mathtt{v}})+\rho_{\mathtt{uv}}^2\sigma^2(\Lambda_{\mathtt{u}})\Big)\\ &\times (1+\delta).\end{aligned} \tag{8.10}$$

Furthermore, these lattices satisfy the following “goodness” properties:

- $\Lambda_{\mathtt{v}}$ is good for covering.
- $\Lambda_{\mathtt{v}}^{(1)}$ and $\Lambda_{\mathtt{v}}^{(2)}$ are good for AWGN channel coding.

8.5.2 Quantization

Terminal $\mathtt{v}$ observes N samples $\mathbf{x}_{\mathtt{v}} = (x_{\mathtt{v}}(1), \ldots, x_{\mathtt{v}}(N))$. As mentioned earlier, the quantizer operates on blocks of n samples each, and there are N_{out} such blocks. We can write $\mathbf{x} = (\mathbf{x}_{\mathtt{u}}^{(1)}, \mathbf{x}_{\mathtt{u}}^{(2)}, \ldots, \mathbf{x}_{\mathtt{u}}^{(N_{\mathrm{out}})})$, where $\mathbf{x}_{\mathtt{u}}^{(j)} \in \mathbb{R}^n$ is given by $\mathbf{x}_{\mathtt{u}}^{(j)} = (x_{\mathtt{u}}((j-1)n+1), \ldots, x_{\mathtt{u}}(jn))$.

Terminal $\mathtt{v}$ also generates N_{out} dither vectors $\mathbf{d}_{\mathtt{v}}^{(1)}, \mathbf{d}_{\mathtt{v}}^{(2)}, \ldots, \mathbf{d}_{\mathtt{v}}^{(N_{\mathrm{out}})}$, which are all uniformly distributed over $\mathcal{V}(\Lambda_{\mathtt{v}})$, and independent of each other and of everything else. These dither vectors are assumed to be known to all the terminals, and to the eavesdropper.

For $1 \le i \le N_{\mathrm{out}}$ and $\mathtt{v} \in V$, let

$$\mathbf{y}_{\mathtt{v}}^{(i)} = [Q_{\Lambda_{\mathtt{v}}}(\mathbf{x}_{\mathtt{v}}^{(i)} + \mathbf{d}_{\mathtt{v}}^{(i)}) - \mathbf{d}_{\mathtt{v}}^{(i)}] \bmod \Lambda_{\mathtt{v}}^{(2)} \tag{8.11}$$

denote the output of the lattice quantizer at terminal $\mathtt{v}$. The terminals can only use $\mathbf{y}_{\mathtt{v}} \triangleq (\mathbf{y}_{\mathtt{v}}^{(1)}, \mathbf{y}_{\mathtt{v}}^{(2)}, \ldots, \mathbf{y}_{\mathtt{v}}^{(N_{\mathrm{out}})})$ for the secret key generation protocol. From (8.8) and (8.9), we can see that the quantization rates satisfy

$$R_{\mathrm{q}}^{(\mathtt{v})} = \frac{1}{2}\log_2\left(1+\frac{1}{\sigma^2(\Lambda_{\mathtt{v}})}\right) + \log_2(1+\delta) + o_n(1). \tag{8.12}$$

8.5.3 Information Reconciliation: The Analog Phase

Let $T^* = (V^*, E^*)$ denote the rooted tree in $\mathcal{T}^*$ which achieves the maximum in (8.4). The terminals in V^* are the only ones that communicate across the public channel. Terminal $\mathtt{v} \in V^*$ broadcasts

$$\begin{aligned}\mathbf{w}_{\mathtt{v}}^{(i)} &\triangleq [\mathbf{y}_{\mathtt{v}}^{(i)}] \bmod \Lambda_{\mathtt{v}}^{(1)} \\ &= [Q_{\Lambda_{\mathtt{v}}}(\mathbf{x}_{\mathtt{v}}^{(i)} + \mathbf{d}_{\mathtt{v}}^{(i)})] \bmod \Lambda_{\mathtt{v}}^{(1)} \end{aligned}\tag{8.13}$$

for $1 \le i \le N_{\mathrm{out}}$, across the public channel. Prior to the analog phase, terminal $\mathtt{u} \in V$ only has access to $\mathbf{y}_{\mathtt{u}} = (\mathbf{y}_{\mathtt{u}}^{(1)}, \mathbf{y}_{\mathtt{u}}^{(2)}, \ldots, \mathbf{y}_{\mathtt{u}}^{(N_{\mathrm{out}})})$. At the end of the information reconciliation phase, every terminal $\mathtt{u}$ will be able to recover $\{\mathbf{y}_{\mathtt{v}} : \mathtt{v} \in V^*\}$ with low probability of error. The analog phase ensures that every $\mathbf{y}_{\mathtt{v}}^{(i)}$ can be individually recovered with low probability of error. The digital phase guarantees that the entire block $\mathbf{y}_{\mathtt{v}}$ can be recovered reliably.

Now consider any $\mathtt{v} \in V^*$ and $\mathtt{u} \in N_T(\mathtt{v})$ (not necessarily in V^*). Suppose that some terminal $\mathtt{u}'$ (not necessarily $\mathtt{u}$) has a reliable estimate of $\mathbf{y}_{\mathtt{u}}^{(i)}$. From $\mathbf{y}_{\mathtt{u}}^{(i)}$ and $\mathbf{w}_{\mathtt{v}}^{(i)}$, terminal $\mathtt{u}'$ can estimate $\mathbf{y}_{\mathtt{v}}^{(i)}$ as follows:

$$\widehat{\mathbf{y}}_{\mathtt{v}}^{(i)} = \mathbf{w}_{\mathtt{v}}^{(i)} + Q_{\Lambda_{\mathtt{v}}^{(1)}}\left(\rho_{\mathtt{uv}}\mathbf{y}_{\mathtt{u}}^{(i)} - \mathbf{w}_{\mathtt{v}}^{(i)}\right). \tag{8.14}$$

We now have the following proposition.

Proposition 8.5.1. *Fix a $\delta > 0$. For every $1 \le i \le N_{\mathrm{out}}$, we have*

$$\mathbb{E}_{\mathbf{y}_{\mathtt{u}}^{(i)}} \Pr[\widehat{\mathbf{y}}_{\mathtt{v}}^{(i)} \ne \mathbf{y}_{\mathtt{v}}^{(i)}] \le e^{-nE_{\mathtt{uv}}(\delta)(1+o_n(1))} \tag{8.15}$$

where $E_{\mathtt{uv}}(\delta)$ is a quantity which is positive for all positive δ, as long as

$$\frac{\left(\mathrm{vol}(\Lambda_{\mathtt{v}}^{(1)})\right)^{2/n}}{2\pi e} > \max_{\mathtt{u}\in N_T(\mathtt{v})}\left(1-\rho_{\mathtt{uv}}^2+\sigma^2(\Lambda_{\mathtt{v}})+\rho_{\mathtt{uv}}^2\sigma^2(\Lambda_{\mathtt{u}})\right) \times (1+\delta), \tag{8.16}$$

$$\frac{\left(\mathrm{vol}(\Lambda_{\mathtt{v}}^{(2)})\right)^{2/n}}{2\pi e} > (1+\sigma^2(\Lambda_{\mathtt{v}}))(1+\delta), \tag{8.17}$$

and

$$\frac{\left(\mathrm{vol}(\Lambda_{\mathtt{u}}^{(2)})\right)^{2/n}}{2\pi e} > (1+\sigma^2(\Lambda_{\mathtt{u}}))(1+\delta). \tag{8.18}$$

Proof. Recall that

$$\begin{aligned}\mathbf{y}_{\mathtt{u}}^{(i)} &= [Q_{\Lambda_{\mathtt{u}}}(\mathbf{x}_{\mathtt{u}}^{(i)}+\mathbf{d}_{\mathtt{u}}^{(i)})-\mathbf{d}_{\mathtt{u}}^{(i)}] \bmod \Lambda_{\mathtt{u}}^{(2)} \\ &= \left[\mathbf{x}_{\mathtt{u}}^{(i)}-[\mathbf{x}_{\mathtt{u}}^{(i)}+\mathbf{d}_{\mathtt{u}}^{(i)}] \bmod \Lambda_{\mathtt{u}}\right] \bmod \Lambda_{\mathtt{u}}^{(2)} \\ &= [\mathbf{x}_{\mathtt{u}}^{(i)}+\widetilde{\mathbf{d}}_{\mathtt{u}}^{(i)}] \bmod \Lambda_{\mathtt{u}}^{(2)},\end{aligned} \tag{8.19}$$

where $\widetilde{\mathbf{d}}_{\mathtt{u}}^{(i)}$ is uniformly distributed over $\mathcal{V}(\Lambda_{\mathtt{u}})$ and is independent of $\mathbf{x}_{\mathtt{u}}^{(i)}$ [28, Lemma 1]. Since $\Lambda_{\mathtt{u}}$ is good for MSE quantization, $\Lambda_{\mathtt{u}}^{(2)}$ is good for AWGN and (8.18) is satisfied, we can use [29, Theorem 4] to assert that[5] the probability

$$\Pr[\mathbf{y}_{\mathtt{u}}^{(i)} \neq \mathbf{x}_{\mathtt{u}}^{(i)}+\widetilde{\mathbf{d}}_{\mathtt{u}}^{(i)}] \leq e^{-n(E_1(\delta)-o_n(1))} \tag{8.20}$$

where $E_1(\delta)>0$ for all $\delta>0$. Similarly, we can write

$$\mathbf{y}_{\mathtt{v}}^{(i)} = [\mathbf{x}_{\mathtt{v}}^{(i)}+\widetilde{\mathbf{d}}_{\mathtt{v}}^{(i)}] \bmod \Lambda_{\mathtt{v}}^{(2)},$$

[5] Note that there is a slight difference here since $\widetilde{\mathbf{d}}_{\mathtt{u}}^{(i)}$ is not Gaussian. However, the arguments in [28, Theorem 5] can be used to show that $\mathbf{x}_{\mathtt{u}}^{(i)}+\widetilde{\mathbf{d}}_{\mathtt{u}}^{(i)}$ can be approximated by a Gaussian since $\Lambda_{\mathtt{u}}$ is good for MSE quantization.

where $\widetilde{\mathbf{d}}_{\mathtt{v}}^{(i)}$ is independent of $\mathbf{x}_{\mathtt{v}}^{(i)}$, and

$$\Pr[\mathbf{y}_{\mathtt{v}}^{(i)} \neq \mathbf{x}_{\mathtt{v}}^{(i)} + \widetilde{\mathbf{d}}_{\mathtt{v}}^{(i)}] \leq e^{-n(E_2(\delta)-o_n(1))} \tag{8.21}$$

where $E_2(\delta) > 0$ for all $\delta > 0$.

Recall that $\mathbf{w}_{\mathtt{v}}^{(i)} = [\mathbf{y}_{\mathtt{v}}^{(i)}] \bmod \Lambda_{\mathtt{v}}^{(1)}$. We can write

$$\begin{aligned}
\widehat{\mathbf{y}}_{\mathtt{v}}^{(i)} &= \mathbf{w}_{\mathtt{v}}^{(i)} + Q_{\Lambda_{\mathtt{v}}^{(1)}}(\rho_{\mathtt{uv}}\mathbf{y}_{\mathtt{u}}^{(i)} - \mathbf{w}_{\mathtt{v}}^{(i)}) \\
&= \mathbf{w}_{\mathtt{v}}^{(i)} + Q_{\Lambda_{\mathtt{v}}^{(1)}}\left(\rho_{\mathtt{uv}}\mathbf{y}_{\mathtt{u}}^{(i)} - \mathbf{y}_{\mathtt{v}}^{(i)} + Q_{\Lambda_{\mathtt{v}}^{(1)}}(\mathbf{y}_{\mathtt{v}}^{(i)})\right) \\
&= \mathbf{w}_{\mathtt{v}}^{(i)} + Q_{\Lambda_{\mathtt{v}}^{(1)}}(\mathbf{y}_{\mathtt{v}}^{(i)}) + Q_{\Lambda_{\mathtt{v}}^{(1)}}\left(\rho_{\mathtt{uv}}\mathbf{y}_{\mathtt{u}}^{(i)} - \mathbf{y}_{\mathtt{v}}^{(i)}\right) \\
&= \mathbf{y}_{\mathtt{v}}^{(i)} + Q_{\Lambda_{\mathtt{v}}^{(1)}}\left(\rho_{\mathtt{uv}}\mathbf{y}_{\mathtt{u}}^{(i)} - \mathbf{y}_{\mathtt{v}}^{(i)}\right)
\end{aligned} \tag{8.22}$$

From (8.20) and (8.21), we know that $\mathbf{y}_{\mathtt{v}}^{(i)} = \mathbf{x}_{\mathtt{v}}^{(i)} + \widetilde{\mathbf{d}}_{\mathtt{v}}^{(i)}$ and $\mathbf{y}_{\mathtt{u}}^{(i)} = \mathbf{x}_{\mathtt{u}}^{(i)} + \widetilde{\mathbf{d}}_{\mathtt{u}}^{(i)}$ with high probability. Now,

$$\rho_{\mathtt{uv}}\mathbf{x}_{\mathtt{u}}^{(i)} - \mathbf{x}_{\mathtt{v}}^{(i)} = -\sqrt{1-\rho_{\mathtt{uv}}^2}\,\mathbf{z}_{\mathtt{vu}}^{(i)},$$

and again using the AWGN goodness property of $\Lambda_{\mathtt{v}}^{(1)}$ and (8.16), we have

$$\Pr\left[Q_{\Lambda_{\mathtt{v}}^{(1)}}(-\sqrt{1-\rho_{\mathtt{uv}}^2}\,\mathbf{z}_{\mathtt{vu}}^{(i)} + \widetilde{\mathbf{d}}_{\mathtt{u}}^{(i)} - \widetilde{\mathbf{d}}_{\mathtt{v}}^{(i)}) \neq \mathbf{0}\right] \leq e^{-n(E_3(\delta)-o_n(1))} \tag{8.23}$$

where $E_3(\delta) > 0$ for $\delta > 0$. Using (8.20), (8.21) and (8.23), we get that

$$\Pr[\widehat{\mathbf{y}}_{\mathtt{v}}^{(i)} \neq \mathbf{y}_{\mathtt{v}}^{(i)}] \leq \sum_{i=1}^{3} e^{-n(E_i(\delta)-o_n(1))}$$

which completes the proof of the proposition. □

Since terminal $\mathtt{u}$ has $\mathbf{y}_{\mathtt{u}}^{(i)}$, it can (with high probability) recover the corresponding quantized sources of its neighbours. Assuming that these have been recovered correctly, it can then estimate the quantized sources of all terminals at distance two from $\mathtt{u}$, and so on, till all $\mathbf{y}_{\mathtt{v}}^{(i)}$ for $\mathtt{v}$ in V^* have been recovered. Using the union bound, we can say that the probability that terminal $\mathtt{u}$ correctly recovers $\{\mathbf{y}_{\mathtt{v}}^{(i)} : \mathtt{v} \in V^*\}$ is at least

$1-\sum_{\mathtt{u}\in V^*}\max_{\mathtt{v}\in N_T(\mathtt{u})}e^{-nE_{\mathtt{uv}}(\delta)}$.

For all terminals to be able to agree upon the key, we must ensure that every terminal can recover all blocks $\{\mathbf{y}_{\mathtt{v}}:\mathtt{v}\in V^*\}$ with low probability of error. Since N_{out} is exponential in n, the analog phase does not immediately guarantee this. For that, we use the digital phase.

8.5.4 Information Reconciliation: The Digital Phase

Observe that $\mathbf{y}_{\mathtt{v}}^{(i)}\in\Lambda_{\mathtt{v}}\cap\mathcal{V}(\Lambda_{\mathtt{v}}^{(2)})$, where both $\Lambda_{\mathtt{v}}$ and $\Lambda_{\mathtt{v}}^{(2)}$ are Construction-A lattices obtained by linear codes over $\mathbb{F}_p$. As a result, $|\Lambda_{\mathtt{v}}\cap\mathcal{V}(\Lambda_{\mathtt{v}}^{(2)})|$ is always an integer power of p [29]. Let

$$|\Lambda_{\mathtt{v}}\cap\mathcal{V}(\Lambda_{\mathtt{v}}^{(2)})|=p^{k_{\mathtt{v}}}.$$

Then, there exists an (set) isomorphism $\varphi_{\mathtt{v}}$ from $\Lambda_{\mathtt{v}}\cap\mathcal{V}(\Lambda_{\mathtt{v}}^{(2)})$ to $\mathbb{F}_{p^{k_{\mathtt{v}}}}$. For every $\mathtt{v}\in V^*$ and $i\in\{1,2,\ldots,N_{\text{out}}\}$, let $y_{\mathtt{v}}^{(i)}=\varphi_{\mathtt{v}}(\mathbf{y}_{\mathtt{v}}^{(i)})$. Similarly, let $\widehat{y}_{\mathtt{v}}^{(i)}=\varphi_{\mathtt{v}}(\widehat{\mathbf{y}}_{\mathtt{v}}^{(i)})$.

The key component of the digital phase is a Reed-Solomon code over $\mathbb{F}_{p^{k_{\mathtt{v}}}}$. In [71], a Slepian-Wolf scheme with random linear codes was used for the digital phase. Using a Reed-Solomon code, we can ensure that the overall computational complexity (including all the phases of the protocol) is polynomial in N.

For every $\mathtt{v}$, let $\mathcal{C}_{\mathtt{v}}$ be a Reed-Solomon code of blocklength N_{out} and dimension

$$K_{\text{out}}=N_{\text{out}}(1-2\delta).\tag{8.24}$$

Let $\mathsf{y}_{\mathtt{v}}^{N_{\text{out}}}=(y_{\mathtt{v}}^{(1)},y_{\mathtt{v}}^{(2)},\ldots,y_{\mathtt{v}}^{(N_{\text{out}})})$ and $\widehat{\mathsf{y}}_{\mathtt{v}}^{N_{\text{out}}}=(\widehat{y}_{\mathtt{v}}^{(1)},\widehat{y}_{\mathtt{v}}^{(2)},\ldots,\widehat{y}_{\mathtt{v}}^{(N_{\text{out}})})$. We can write

$$\widehat{\mathsf{y}}_{\mathtt{v}}^{N_{\text{out}}}=\mathsf{y}_{\mathtt{v}}^{N_{\text{out}}}+\mathsf{e}_{\mathtt{v}}^{N_{\text{out}}},$$

where $\mathsf{e}_{\mathtt{v}}^{N_{\text{out}}}=(e_{\mathtt{v}}^{(1)},e_{\mathtt{v}}^{(2)},\ldots,e_{\mathtt{v}}^{(N_{\text{out}})})$ is the error vector, and from the previous section, we have

$$\Pr[e_{\mathtt{v}}^{(i)}\neq 0]\leq\sum_{\mathtt{v}\in V^*}\max_{\mathtt{u}\in N_T(\mathtt{v})}e^{-nE_{\mathtt{uv}}(\delta)}\leq\delta$$

for all sufficiently large n. Every $\mathsf{y}_{\mathsf{v}}^{N_{\text{out}}}$ can be written uniquely as

$$\mathsf{y}_{\mathsf{v}}^{N_{\text{out}}} = \mathsf{c}_{\mathsf{v}}^{N_{\text{out}}} + \mathsf{s}_{\mathsf{v}}^{N_{\text{out}}} \tag{8.25}$$

where $\mathsf{c}_{\mathsf{v}}^{N_{\text{out}}} \in \mathcal{C}_{\mathsf{v}}$, and $\mathsf{s}_{\mathsf{v}}^{N_{\text{out}}}$ is a minimum Hamming weight representative of the coset to which $\mathsf{y}_{\mathsf{v}}^{N_{\text{out}}}$ belongs in $\mathbb{F}_{p^{k_{\mathsf{v}}}}^{N_{\text{out}}}/\mathcal{C}_{\mathsf{v}}$. Terminal $\mathtt{v}$ broadcasts $\mathsf{s}_{\mathsf{v}}^{N_{\text{out}}}$ across the public channel. This requires a rate of public communication of at most

$$\frac{1}{N}\log_2|\mathbb{F}_{p^{k_{\mathsf{v}}}}^{N_{\text{out}}}/\mathcal{C}_{\mathsf{v}}| = \frac{2N_{\text{out}}\delta}{N}\log_2(p^{k_{\mathsf{v}}}) = 2\delta R_{\mathrm{q}}^{(\mathsf{v})}. \tag{8.26}$$

From $\mathsf{s}_{\mathsf{v}}^{N_{\text{out}}}$ and $\widehat{\mathsf{y}}_{\mathsf{v}}^{N_{\text{out}}}$, terminal $\mathtt{u}$ can compute

$$\widehat{\mathsf{c}}_{\mathsf{v}}^{N_{\text{out}}} = \widehat{\mathsf{y}}_{\mathsf{v}}^{N_{\text{out}}} - \mathsf{s}_{\mathsf{v}}^{N_{\text{out}}} = \mathsf{c}_{\mathsf{v}}^{N_{\text{out}}} + \mathsf{e}_{\mathsf{v}}^{N_{\text{out}}}.$$

We have from Theorem 7.2.2 that the probability of the Reed-Solomon decoder incorrectly decoding $\mathsf{c}_{\mathsf{v}}^{N_{\text{out}}}$ from $\widehat{\mathsf{c}}_{\mathsf{v}}^{N_{\text{out}}}$ decays exponentially in N. Therefore, for sufficiently large n, the probability that $\mathbf{y}_{\mathsf{v}}^{(i)}$ is estimated incorrectly is less than δ, and terminal $\mathtt{u}$ can recover $\mathsf{c}_{\mathsf{v}}^{N_{\text{out}}}$ with high probability using the decoder for the Reed-Solomon code. Having recovered $\mathsf{c}_{\mathsf{v}}^{N_{\text{out}}}$ reliably, the terminals can obtain $\mathsf{y}_{\mathsf{v}}^{N_{\text{out}}}$ using (8.25). Therefore, at the end of the digital phase, all terminals can recover $\{\mathbf{y}_{\mathsf{v}} : \mathsf{v} \in V^*\}$ with a probability of error that decays exponentially in N.

8.5.5 Secret Key Generation

Let $k \triangleq \sum_{\mathsf{v}\in V^*} k_{\mathsf{v}}$. There exists a (set) bijection ϕ from $\bigtimes_{\mathsf{v}\in V^*}\mathbb{F}_{p^{k_{\mathsf{v}}}}$ to $\mathbb{F}_{p^k}$. Let $y^{(i)} = \phi(\mathbf{y}_{\mathsf{v}}^{(i)} : \mathsf{v} \in V^*)$. We use the following result by Nitinawarat and Narayan [71], which says that there exists a linear function of the sources that can act as the secret key.

Lemma 8.5.2 (Lemma 4.5, [71]). *Let Y be a random variable in a Galois field $\mathbb{F}_q$ and D be an $\mathbb{R}^n$-valued random variable jointly distributed with Y. Consider N_{out} iid repetitions of (Y, D), namely $(Y^{N_{\text{out}}}, D^{N_{\text{out}}}) = ((Y_1, D_1), \ldots, (Y_{N_{\text{out}}}, D_{N_{\text{out}}}))$.*

Let $B = B^{(N_{\text{out}})} \in \mathcal{B}^{(N_{\text{out}})}$ be a finite-valued rv with a given joint distribution with

$(Y^{N_{\text{out}}}, D^{N_{\text{out}}})$.

Then, for every $\delta > 0$ and every

$$R < H(Y|D) - \frac{1}{N_{\text{out}}} \log |\mathcal{B}^{(N_{\text{out}})}| - 2\delta,$$

there exists a $\lfloor \frac{N_{\text{out}} R}{\log q} \rfloor \times N_{\text{out}}$ matrix L with $\mathbb{F}_q$ -valued entries such that

$$N_{\text{out}} R - H(LY^{N_{\text{out}}}) + I(LY^{N_{\text{out}}}; D^{N_{\text{out}}}, B)$$

vanishes exponentially in N_{out}.

In other words, $LY^{N_{\text{out}}}$ is an ϵ-SK for suitable ϵ. Let $q = p^k$ and $B = (\mathbf{w}_{\mathtt{v}}, s_{\mathtt{v}}^{N_{\text{out}}} : \mathtt{v} \in V^*)$. Then, the above lemma guarantees the existence of an $\mathbb{F}_{p^k}$-valued matrix L, so that $L(y^{(1)}, \ldots, y^{(N_{\text{out}})})^T$ is a secret key with a rate of

$$R_{\text{key}} = \frac{1}{N} H(\mathbf{y}_{\mathtt{v}} : \mathtt{v} \in V^* | \mathbf{d}_{\mathtt{v}} : \mathtt{v} \in V^*) - \sum_{\mathtt{v} \in V^*} R_{\text{com}}^{(\mathtt{v})}, \tag{8.27}$$

where $R_{\text{com}}^{(\mathtt{v})}$ denotes the total rate of communication of terminal $\mathtt{v}$. We give a lower bound on R_{key} by bounding $H(\mathbf{y}_{\mathtt{v}} : \mathtt{v} \in V^* | \mathbf{d}_{\mathtt{v}} : \mathtt{v} \in V^*)$ in the next section.

8.5.6 Joint Entropy of the Quantized Sources

We now give a lower bound on the joint entropy of $\{\mathbf{y}_{\mathtt{v}}^{(i)} : \mathtt{v} \in V^*\}$ conditioned on the dithers $D_i = \{\mathbf{d}_{\mathtt{v}}^{(i)} : \mathtt{v} \in V^*\}$.

Lemma 8.5.3. *Fix a $\delta > 0$, and let $D_i \triangleq (\mathbf{d}_{\mathtt{v}}^{(i)} : \mathtt{v} \in V^*)$. For all sufficiently large n, we have*

$$\frac{1}{n} H(\mathbf{y}_{\mathtt{v}}^{(i)} : \mathtt{v} \in V^* | D_i) \geq \frac{1}{2} \log_2 \left(1 + \frac{1}{\sigma^2(\Lambda_{\mathtt{r}(T^*)})} \right) + \sum_{\mathtt{u} \in V^*} \frac{1}{2} \log_2 \left(1 + \frac{1 - \rho_{\mathtt{u},\text{par}(\mathtt{u})}^2}{\sigma^2(\Lambda_{\mathtt{u}})} \right) - \delta \tag{8.28}$$

Proof. We prove the result by expanding the joint entropy using the chain rule, and then use a lower bound on the entropy of a quantized Gaussian. To do this, we will expand the joint entropy in a particular order. Let $\mathcal{S}$ be any (totally) ordered set containing the vertices of T^* and satisfying the following properties:

- $\max_{\mathtt{v}} \mathcal{S} = \mathtt{r}(T^*)$, i.e., $\mathtt{r}(T^*) \geq \mathtt{v}$ for all $\mathtt{v} \in \mathcal{S}$.
- $\mathtt{v} > \mathtt{u}$ if the distance between $\mathtt{v}$ and $\mathtt{r}(T^*)$ is less than that between $\mathtt{u}$ and $\mathtt{r}(T^*)$.

Essentially, $\mathtt{v} > \mathtt{u}$ if $\mathtt{v}$ is closer to $\mathtt{r}(T^*)$ than $\mathtt{u}$, and we do not care how the vertices at the same level (vertices at the same distance from $\mathtt{r}(T^*)$) are ordered. Let $D = (\mathbf{d}_{\mathtt{v}}^{(i)} : \mathtt{v} \in V^*)$. Then,

$$\begin{aligned} H(\mathbf{y}_{\mathtt{v}}^{(i)} : \mathtt{v} \in V^* | D) &= H(\mathbf{y}_{\mathtt{r}(T^*)}^{(i)} | D) + \sum_{\mathtt{v} \in V^* \backslash \mathtt{r}(T^*)} H(\mathbf{y}_{\mathtt{v}}^{(i)} | D, \mathbf{y}_{\mathtt{u}}^{(i)} : \mathtt{u} > \mathtt{v}) \\ &\geq H(\mathbf{y}_{\mathtt{r}(T^*)}^{(i)} | D) + \sum_{\mathtt{v} \in V^* \backslash \mathtt{r}(T^*)} H(\mathbf{y}_{\mathtt{v}}^{(i)} | D, \mathbf{x}_{\mathtt{u}}^{(i)} : \mathtt{u} > \mathtt{v}) \qquad (8.29) \\ &= H(\mathbf{y}_{\mathtt{r}(T^*)}^{(i)} | D) + \sum_{\mathtt{v} \in V^* \backslash \mathtt{r}(T^*)} H(\mathbf{y}_{\mathtt{v}}^{(i)} | D, \mathbf{x}_{\mathrm{par}(\mathtt{v})}^{(i)}) \qquad (8.30) \end{aligned}$$

where (8.29) follows from the data processing inequality. We would like to remark that (8.30) is the only place where we use Markov tree assumption. The rest of the proof closely follows [71, Lemma 4.3], and we give an outline. The idea is to find the average mean squared error distortion in representing $\mathbf{x}_{\mathtt{v}}^{(i)}$ by $\mathbf{y}_{\mathtt{v}}^{(i)}$ (with or without the side information $\mathbf{x}_{\mathrm{par}(\mathtt{v})}^{(i)}$), and then argue that the rate of such a quantizer must be greater than or equal to the rate-distortion function.

Claim 2.

$$\frac{1}{n} H(\mathbf{y}_{\mathtt{r}(T^*)}^{(i)} | D) \geq \frac{1}{2} \log_2 \left(1 + \frac{1}{\sigma^2(\Lambda_{\mathtt{r}(T^*)})} \right) - o_n(1). \qquad (8.31)$$

Making minor modifications to the proof of [71, Lemma 4.3], we can show that conditioned on D, the average MSE distortion (averaged over D) between $\mathbf{x}_{\mathtt{r}(T^*)}^{(i)}$ and

$$\widehat{\mathbf{x}}_{\mathtt{r}(T^*)}^{(i)} = \frac{1}{1+\sigma^2(\Lambda_{\mathtt{r}(T^*)})}\mathbf{y}_{\mathtt{r}(T^*)}^{(i)}$$

is at most $\frac{\sigma^2(\Lambda_{\mathtt{r}(T^*)})}{1+\sigma^2(\Lambda_{\mathtt{r}(T^*)})} + o_n(1)$. Since any rate-distortion code for quantizing $\mathbf{x}_{\mathtt{r}(T^*)}^{(i)}$ must have a rate at least as much as the rate-distortion function, we can show that (again following the proof of [71, Lemma 4.3]) Claim 2 is true.

Claim 3.

$$\frac{1}{n}H(\mathbf{y}_{\mathtt{v}}^{(i)}|D, \mathbf{x}_{\mathrm{par}(\mathtt{v})}^{(i)}) \geq \frac{1}{2}\log_2\left(1+\frac{1-\rho_{\mathtt{v},\mathrm{par}(\mathtt{v})}^2}{\sigma^2(\Lambda_{\mathtt{v}})}\right) - o_n(1) \tag{8.32}$$

The proof of the above claim also follows the same technique. We can show that conditioned on D and $\mathbf{x}_{\mathrm{par}(\mathtt{v})}^{(i)}$, the average MSE distortion between $\sqrt{1-\rho_{\mathtt{uv}}^2}\mathbf{z}_{\mathtt{vu}}^{(i)}$ and

$$\widehat{\mathbf{z}}_{\mathtt{vu}}^{(i)} = \frac{(1-\rho_{\mathtt{uv}}^2)}{1-\rho_{\mathtt{uv}}^2+\sigma^2(\Lambda_{\mathtt{v}})}\left[\mathbf{y}_{\mathtt{v}}^{(i)} - \rho_{\mathtt{uv}}\mathbf{x}_{\mathrm{par}(\mathtt{v})}^{(i)}\right] \bmod \Lambda_{\mathtt{v}}^{(2)}$$

is $\frac{(1-\rho_{\mathtt{uv}}^2)\sigma^2(\Lambda_{\mathtt{v}})}{1-\rho_{\mathtt{uv}}^2+\sigma^2(\Lambda_{\mathtt{v}})} + o_n(1)$. Arguing as before, the claim follows.

Finally, using (8.31) and (8.32) in (8.30) completes the proof of Lemma 8.5.3. □

For every $\mathtt{v}$, $\{\mathbf{y}_{\mathtt{v}}^{(i)} : 1 \leq i \leq N_{\mathrm{out}}\}$ are independent and identically distributed. If $D \triangleq \{D_i : 1 \leq i \leq N_{\mathrm{out}}\}$, then $H(\mathbf{y}_{\mathtt{v}} : \mathtt{v} \in V^*|D) = N_{\mathrm{out}}H(\mathbf{y}_{\mathtt{v}}^{(i)} : \mathtt{v} \in V^*|D_i)$. Substituting for $\sigma^2(\Lambda_{\mathtt{v}})$ from (8.12) in (8.28), we get

$$\frac{1}{N}H(\mathbf{y}_{\mathtt{v}} : \mathtt{v} \in V^*|D) \geq \mathcal{R}_{\mathrm{ent}} - g(\delta) - o_n(1),$$

where $\mathcal{R}_{\mathrm{ent}}$ is defined in (8.1), and $g(\delta)$ is a quantity that goes to 0 as $\delta \to 0$.

8.5.7 Achievable Secret Key Rate and Proof of Theorem 8.3.1

Lemma 8.5.2 guarantees the existence of a strong secret key which is a linear transformation of $(y^{(1)}, \ldots, y^{(N_{\mathrm{out}})})$. From Propositions 8.5.1 and Section 8.5.4, all terminals are able to recover $(y^{(1)}, \dots, y^{(N_{\text{out}})})$ with a probability of error that decays exponentially in $N = nN_{\text{out}}$.

During the analog phase, each terminal $\mathtt{v}$ in V^* publicly communicates

$$\begin{aligned} R_{\text{analog}}^{(\mathtt{v})} &= \frac{1}{n} \log_2 \frac{\text{vol}(\Lambda_{\mathtt{v}}^{(1)})}{\text{vol}(\Lambda_{\mathtt{v}})} \\ &\leq \max_{\mathtt{u} \in N_T(\mathtt{v})} \frac{1}{2} \log_2 \frac{(1 - \rho_{\mathtt{uv}}^2 + \sigma^2(\Lambda_{\mathtt{v}}) + \rho_{\mathtt{uv}}^2 \sigma^2(\Lambda_{\mathtt{u}}))}{\sigma^2(\Lambda_{\mathtt{v}})} + o_n(1) + \delta. \end{aligned} \tag{8.33}$$

bits per sample. Here, we have used the fact that an MSE quantization-good lattice $\Lambda_{\mathtt{v}}$ satisfies $\text{vol}(\Lambda_{\mathtt{v}}) \to 2\pi e \sigma^2(\Lambda)$ as $n \to \infty$. We know from (8.26) that during the digital phase, terminal $\mathtt{v}$ communicates $2\delta R_{\text{q}}^{(\mathtt{v})}$ bits per sample across the public channel. The total rate of communication by terminal $\mathtt{v}$ is therefore

$$R_{\text{com}}^{(\mathtt{v})} \leq \max_{\mathtt{u} \in N_T(\mathtt{v})} \frac{1}{2} \log_2 \frac{(1 - \rho_{\mathtt{uv}}^2 + \sigma^2(\Lambda_{\mathtt{v}}) + \rho_{\mathtt{uv}}^2 \sigma^2(\Lambda_{\mathtt{u}}))}{\sigma^2(\Lambda_{\mathtt{v}})} + \delta(1 + 2R_{\text{q}}^{(\mathtt{v})}) + o_n(1) \tag{8.34}$$

bits per sample. Using Lemma 8.5.3 and (8.34) in (8.27), and finally substituting (8.12), we obtain (8.4). All that remains now is to find an upper bound on the computational complexity of our scheme.

8.5.8 Computational Complexity

We now show that the computational complexity is polynomial in the number of samples N. The complexity is measured in terms of the number of binary operations required, and we make the assumption that each floating-point operation (i.e., operations in $\mathbb{R}$) requires $O(1)$ binary operations. In other words, the complexity of a floating-point operation is independent of N.

Recall that $N = nN_{\text{out}}$, where $N_{\text{out}} = \min_{\mathtt{v} \in V^*}(2^{nR_{\text{q}}^{(\mathtt{v})}} - 1)$. Also, $\alpha = \frac{\max_{\mathtt{v} \in V^*} R_{\text{q}}^{(\mathtt{v})}}{\min_{\mathtt{v} \in V^*} R_{\text{q}}^{(\mathtt{v})}}$.

- *Quantization*: Each lattice quantization operation has complexity at most $O(2^{nR_{\text{q}}^{(\mathtt{v})}}) = O(N_{\text{out}}^{\alpha})$. There are N_{out} such quantization operations to be performed at each terminal, and hence the total complexity is at most $O(N_{\text{out}}^{\alpha+1})$.

- *Analog Phase*: Terminal $\mathtt{v}$ performs N_{out} quantization and mod$\Lambda_{\mathtt{v}}^{(1)}$ operations to compute $\{\mathbf{w}_{\mathtt{v}}^{(i)} : 1 \le i \le N_{\text{out}}\}$, and this requires a total complexity of $O(N_{\text{out}}^{\alpha+1})$. Computation of $\{\widehat{\mathbf{y}}_{\mathtt{v}}^{(i)} : 1 \le i \le N_{\text{out}}, \mathtt{v} \in V^*\}$ requires at most $N_{\text{out}}(|V^*| - 1)$ quantization operations, which also has a total complexity of $O(N_{\text{out}}^{\alpha+1})$.

- *Digital Phase*: Each terminal has to compute the coset representative. This is followed by the decoding of the Reed-Solomon code. Both can be done using the Reed-Solomon decoder, and this requires $O(N_{\text{out}} \log_2 N_{\text{out}})$ operations in $\mathbb{F}_{p^{k_{\mathtt{v}}}}$. Each finite field operation on the other hand requires $O(\log_2^2 p^{k_{\mathtt{v}}}) = O(n^2)$ binary operations [39, Chapter 2]. The total complexity is therefore $O(N^2)$.

- *Secret Key Generation*: This involves multiplication of a $\lfloor \frac{N_{\text{out}} R}{\log_2 q} \rfloor \times N_{\text{out}}$ matrix with an N_{out}-length vector, which requires $O(N_{\text{out}}^2 / \log q)$ operations over $\mathbb{F}_q$. Hence, the complexity required is $O(N_{\text{out}}^2 \log q) = O(N^2)$.

From all of the above, we can conclude that the complexity required is at most $O(N^{\alpha+1})$. If the quantization rate constraints are the same, i.e., $R_{\text{q}}^{(\mathtt{u})} = R_{\text{q}}^{(\mathtt{v})}$, then the complexity is $O(N^2)$. This completes the proof of Theorem 8.3.1. □

Chapter 9

Summary and Extensions

9.1 Summary

Let us very briefly summarize the results of this thesis. In Chapter 3, we explored the problem of secure bidirectional relaying in presence of an "honest-but-curious" relay. For the noiseless setting, we gave a perfectly secure scheme for the user nodes to exchange a single bit each using a single time slot each for communication. We saw that by choosing the pmfs for randomization at the encoders appropriately, we could obtain perfect secrecy. We also showed that these pmfs could be chosen so as to satisfy an average power constraint. Moreover, we saw the impossibility of having a scheme that satisfied a maximum power constraint. We extended this technique to n dimensions, and gave a general lattice coding scheme for perfectly secure exchange of messages. We also established that our scheme guarantees perfect secrecy irrespective of the distribution of the additive noise, and it did not matter whether this distribution is known to the user nodes or not. Moreover, our scheme was explicit, in the sense that given any pair of nested lattices, we explicitly gave a pmf for randomization that guarantees perfect security. Under the assumption that the additive noise is Gaussian, we derived the achievable rates for our scheme under an average power constraint.

We then relaxed the secrecy constraint to strong secrecy to obtain higher rates in Chapter 4. By using a sampled Gaussian pmf for randomization, we could obtain strong secrecy. Once again, our scheme guaranteed secrecy in the absence of noise, which established that we can have secure communication irrespective of the noise distribution. Unlike the scheme we presented in Chapter 3, which guaranteed perfect secrecy for any pair of nested lattices, our strongly secure scheme required the lattices to be secrecy good. Under the assumption of additive Gaussian noise, we derived achievable rates under an average power constraint. We also extended our perfectly-secure and strongly-secure schemes to a multihop line network with honest-but-curious relay nodes, and derived achievable rates.

In Chapter 5, we studied the robustness of our secure bidirectional relaying schemes in the presence of channel imperfections. We assumed that the user nodes had no knowledge of the channel gains, whereas the eavesdropper knew these perfectly. We saw that secrecy is guaranteed when the ratio of the channel gains is rational, but also saw that in the noiseless scenario, no secrecy can be provided when the ratio of channel gains is irrational.

We then studied some low-complexity lattice codes. In Chapter 6, we showed that LDA lattices are good for packing and MSE quantization, and their duals are good for packing. This enabled us to conclude that the rates that we derived in 3 for perfectly secure bidirectional relaying can be achieved using LDA lattices. However, these results are only guaranteed assuming that we use exponential-time lattice decoders at the relay.

We proposed a concatenated lattice coding scheme in Chapter 7. Using inner nested lattice codes and outer Reed-Solomon codes, we could bring down the decoding/encoding complexity to $O(N^2)$ while having the probability of error go to zero as $e^{-\Omega(N)}$. Replacing the Reed-Solomon codes with expander codes further reduced the decoding complexity to $O(N \log^2 N)$. While most of Chapter 7 focused on proving that this scheme achieves the capacity of the AWGN channel, we also gave examples of the Gaussian wiretap channel and compute-and-forward to show that this approach of concatenation can be used to obtain optimal coding schemes in several other scenarios.

In Chapter 8, we studied the problem of secret key generation from correlated Gaussian sources. We used the techniques developed in Chapter 7 to design a secret key generation scheme whose overall computational complexity is polynomial in the number of samples. Assuming that the joint distribution of the Gaussian sources had a Markov tree structure, we derived the achievable secret key rate. We saw that for certain special cases, our scheme achieves the secret key capacity in the fine quantization limit. However, we also gave examples where our scheme cannot achieve the key capacity.

In essence, we looked at two problems in information-theoretic security and used lattice codes to obtain near-optimal solutions. We also made some progress in making lattice codes more "practical" by developing codes with low encoding/decoding complexity. In doing so, we only managed to raise more questions than we could answer. Some of these questions are outlined in the next section.

9.2 Problems to Ponder On

9.2.1 Secure Bidirectional Relaying

We showed that a rate of $\frac{1}{2}\log_2\frac{P}{\sigma^2}-\log_2 2e$ can be achieved with perfect secrecy, and $\frac{1}{2}\log_2\left(\frac{1}{2}+\frac{P}{\sigma^2}\right)-\frac{1}{2}\log_2 2e$ can be achieved with strong secrecy over the bidirectional relay. He and Yener showed that one can achieve a rate of $\frac{1}{2}\log_2\left(\frac{1}{2}+\frac{P}{\sigma^2}\right)-1$ with strong secrecy. On the other hand, the best known achievable rate without any secrecy constraints for the bidirectional relay is $\frac{1}{2}\log_2\left(\frac{1}{2}+\frac{P}{\sigma^2}\right)$ [69, 107]. An open problem is to obtain an upper bound better than the cut-set bound of $\frac{1}{2}\log_2\left(1+\frac{P}{\sigma^2}\right)$ for bidirectional relaying (with or without secrecy constraints).

Another interesting question is whether the rate $\frac{1}{2}\log_2\left(\frac{1}{2}+\frac{P}{\sigma^2}\right)-1$ or even a rate of $\frac{1}{2}\log_2\left(\frac{1}{2}+\frac{P}{\sigma^2}\right)-\frac{1}{2}\log_2 2e$ can be achieved with perfect secrecy. In several problems, a weakly secure scheme can be converted to a strongly secure scheme using privacy amplification [11]. Therefore, the rates achievable using strong secrecy are the same as those achievable using weak secrecy for these problems. This is true for the case of secure bidirectional relaying as well [41]. However, can we achieve the same rates if we demand perfect secrecy?

The schemes for secure compute-and-forward involve sampling from pmfs over high-dimensional lattices. Wang and Ling [104] proposed sampling schemes based on the Metropolis-Hastings algorithm to efficiently sample from a lattice Gaussian distribution.

They also studied the rate of convergence and showed that their algorithm is geometrically ergodic. An open problem is to develop efficient algorithms to sample from pmfs having compactly supported characteristic functions. Unlike sampled Gaussians, these are heavy-tailed pmfs, and it would be interesting to study MCMC algorithms to sample from such distributions.

Practical systems use a random number generator (that typically outputs a sequence of iid Bernoulli(1/2) rvs) for randomization at the encoder. In such a scenario, this randomness is itself a resource, and one can ask how many random bits are required to achieve a certain rate for perfectly/strongly secure communication. A more important question is: Given a constraint on the number of random bits available at the encoders, what is the maximum rate of secure communication that we can achieve?

In Chapter 5, we could only guarantee secrecy in the noiseless setting when the channel gains are rational. However, we also indicated that using a wiretap-type approach, it is indeed possible to obtain secrecy in all settings in the presence of noise. It would be interesting to explore if our strongly secure scheme could provide strong secrecy in presence of AWGN when the channel gains are unknown (or known with some error) to the user nodes. Some basic calculations led us to believe that the techniques used in Chapters 4 and 5 cannot be used directly to guarantee strong secrecy in this setting, as the conditions on the flatness factor for strong secrecy ensured that we cannot guarantee reliable decoding at the relay. This suggests that a different approach might be required, and this may require some new mathematical tools for analysis.

9.2.2 Secret Key Generation

Although we were only able to derive achievable secret key rates for the case where the sources form a Markov tree, the protocol that we proposed can be used to generate secret keys for arbitrarily correlated Gaussian sources. However, without the tree structure, we would be unable to use the conditional independence of the various random variables to derive bounds on the joint entropy of the quantized sources, which in turn was used to compute the secret key rate. We also saw that for certain classes of Markov tree sources, our scheme does not achieve key capacity in the fine quantization limit. One of the main reasons why this could be the case is because in order to estimate the quantized source of terminal $\mathtt{v}$, we only use the estimates of the sources of $N(\mathtt{v})$ (and the public communication of $\mathtt{v}$). This implies that there is scope to improve the protocol for secret key generation.

An open (and very difficult) problem is to characterize the secret key capacity for correlated Gaussian sources under a quantization rate constraint. An alternate problem is to limit the sum rate of public communication (or impose limits on the individual rates of public communication of each terminal) and ask for the maximum achievable secret key rate. This is a much harder problem, and is not well understood even when the sources are discrete.

9.2.3 LDA Lattices

The analysis of LDA lattices with belief propagation decoding is a challenging problem, and any progress in this area would be significant. The main difficulty in this problem is that studying the density evolution for nonbinary LDPC codes is extremely hard. In the binary case, the messages sent on each edge are log-likelihood ratios, and we only need to keep track of this variable in each iteration. However, in the nonbinary case, each edge carries a p-length vector (assuming that the LDPC code is designed over $\mathbb{F}_p$), and analyzing the evolution of its density after each update becomes very difficult.

In proving the various "goodness" properties, we used many simple bounds, which are loose in many cases. We saw that the minimum field-size (p) required to guarantee these "goodness" properties was quite large. This suggests that one could use tighter bounds, or use a different approach altogether to improve the constraints on p. However, this may also make the analysis more complicated (even with "simple" bounds, the proofs were quite lengthy).

In this thesis, we did not discuss two important "goodness" properties, namely covering goodness, and secrecy goodness [58] of LDA lattices. The property of secrecy goodness was crucially used in designing nested lattice codes for the wiretap channel in [58], and for strongly secure bidirectional relaying in Chapter 4. Covering goodness would help ensure that LDA lattices achieve the guaranteed rates in several Gaussian channels with a maximum power constraint (as opposed to an average power constraint). Whether LDA lattices satisfy these properties is still an open problem.

Another interesting problem to consider is the performance of spatially coupled LDA lattices [91], which are LDA lattices constructed from spatially coupled LDPC codes. There have been many new and exciting results on the optimality of spatially coupled binary LDPC codes [50, 51], and a similar analysis of spatially coupled nonbinary LDPC codes may give insights into solving this problem.

9.2.4 Concatenated Lattices

The technique of concatenating lattices and linear codes is a powerful tool to obtain low-complexity solutions to several problems in communication and secrecy. As we mentioned previously, the main disadvantage is that for a fixed error probability P_e, the minimum blocklength, N, required grows exponentially in $1/\gamma$ (where $\gamma = C - R$ is the gap to capacity). The next step would be to develop coding schemes where N grows polynomially in $1/\gamma$ for a fixed P_e. The most promising result so far is the polar lattice coding scheme [111], which satisfies this property for the channel coding problem without power constraints (or the 'coding without restrictions' problem) of Poltyrev [74].

References

[1] R. Ahlswede and I. Csiszár, "Common randomness in information theory and cryptography. Part I: Secret sharing," *IEEE Trans. Inf. Theory*, vol. 39, no. 7, pp. 1121–1132, Jul. 1993.

[2] S. Agrawal and S. Vishwanath, "On the secrecy rate of interference networks using structured codes," *Proc. 2009 IEEE Int. Symp. Information Theory*, Seoul, Korea, pp. 2091–2095.

[3] I.-J. Baik and S.-Y. Chung, "Network coding for two-way relay channels using lattices," *Proc. IEEE Int. Conf. Communications*, Beijing, China, 2008, pp. 3898–3902.

[4] A. Barg and G. Zémor, "Concatenated codes: Serial and parallel," *IEEE Trans. Inf. Theory,* vol. 51, no. 5, pp. 1625–1634, May 2005.

[5] A. Barvinok, *Math 669: Combinatorics, Geometry and Complexity of Integer Points*. [Online]. Available: `http://www.math.lsa.umich.edu/~barvinok/latticenotes669.pdf`.

[6] J.-C. Belfiore and F. Oggier, "Secrecy gain: A wiretap lattice code design," *Proc. 2010 Int. Symp. Information Theory and Its Applications*, Taichung, Taiwan, pp. 174–178.

[7] J.-C. Belfiore, "Lattice codes for the compute-and-forward protocol: The flatness factor," in *Proc. 2011 Information Theory Workshop*, Paraty, Brazil.

[8] M. Bellare, S. Tessaro, and A. Vardy, "Semantic security for the wiretap channel,"

Adv. Cryptology–CRYPTO 2012, Berlin, Germany: Springer Berlin-Heidelberg, pp. 294–311, 2012.

[9] A. L. Bertozzi, J. B. Garnett, and T. Laurent, "Characterization of radially symmetric finite time blowup in multidimensional aggregation equations," *SIAM J. Math. Anal.*, vol. 44, no. 2, pp. 651–681, 2012.

[10] M. Bloch, J. Barros, M.R. Rodrigues, and S.W. McLaughlin, "Wireless information-theoretic security," *IEEE Trans. Inf. Theory*, vol. 54, no. 6, pp. 2515–2534, Jun. 2008.

[11] M. Bloch and J. Barros, *Physical-Layer Security: From Information Theory to Security Engineering,* Cambridge, U.K.: Cambridge University Press, 2011.

[12] F. Boccardi, R.W. Heath, A. Lozano, T.L. Marzetta, and P. Popovski, "Five disruptive technology directions for 5G," *IEEE Comm. Mag.*, vol. 52, no. 2, pp. 74–80, Feb. 2014.

[13] G. Bresler, A. Parekh, and D.N.C. Tse, "The approximate capacity of the many-to-one and one-to-many Gaussian interference channels," *IEEE Trans. Inf. Theory*, vol. 56, no. 9, pp. 4566–4592, Sep. 2010.

[14] J.H. Conway and N.J. Sloane, *Sphere Packings, Lattices and Groups*, New York: Springer-Verlag, 1988.

[15] T. Cover and J. Thomas, *Elements of Information Theory*, 2nd ed. New York: Wiley-Interscience, 1996.

[16] R. Cramer, Y. Dodis, S. Fehr, C. Padro, and D. Wichs, "Detection of algebraic manipulation with applications to robust secret sharing and fuzzy extractors," *Adv. Cryptology–EUROCRYPT 2008*, Berlin, Germany: Springer Berlin-Heidelberg, pp. 471–488, 2008.

[17] I. Csiszár and J. Körner. "Broadcast channels with confidential messages," *IEEE Trans. Inf. Theory,* vol. 24, no. 3, pp. 339–348, May 1978.

[18] I. Csiszár and P. Narayan, “Common randomness and secret key generation with a helper,” *IEEE Trans. Inf. Theory,* vol. 46, no. 3, pp. 344–366, Mar. 2000.

[19] I. Csiszár and P. Narayan, “Secrecy capacities for multiple terminals,” *IEEE Trans. Inf. Theory*, vol. 50, no. 12, pp. 3047–3061, Dec. 2004.

[20] R. de Buda, “Some optimal codes have structure,” *IEEE J. Sel. Areas in Communications*, vol. 7, no. 6, pp. 893–899, 1989.

[21] N. di Pietro, “On infinite and finite lattice constellations for the additive white Gaussian noise channel,” Ph.D. dissertation, Math. Dept., Univ. Bordeaux, Bordeaux, France, 2014.

[22] N. di Pietro, J.J. Boutros, G. Zémor, and L. Brunel, “Integer low-density lattices based on Construction A,” *Proc. 2012 Information Theory Workshop*, Lausanne, Switzerland, 2012, pp. 422–426.

[23] N. di Pietro, J.J. Boutros, G. Zémor, and L. Brunel, “New results on low-density integer lattices,” *Proc. 2013 Information Theory and Applications Workshop*, San Diego, 2013, pp. 10–15.

[24] N. di Pietro, G. Zémor, and J. J. Boutros, “New results on Construction A lattices based on very sparse parity-check matrices,” in *Proc. 2013 IEEE Int. Symp. Information Theory*, Istanbul, Turkey, pp. 1675–1679.

[25] N. di Pietro, G. Zémor, and J. J. Boutros, “LDA Lattices Without Dithering Achieve Capacity on the Gaussian Channel,” arXiv preprint, 2016. [Online] Available: `http://arxiv.org/abs/1603.02863`.

[26] A. Dasgupta, *Probability for Statistics and Machine Learning*, New York: Springer Texts in Statistics, 2011.

[27] W. Ehm, T. Gneiting, and D. Richards, “Convolution roots of radial positive definite functions with compact support,” *Trans. AMS*, vol. 356, no. 11, pp. 4655–4685, May 2004.

[28] U. Erez and R. Zamir, “Achieving 1/2log(1+SNR) on the AWGN channel with lattice encoding and decoding,” *IEEE Trans. Inf. Theory*, vol. 50, no. 10, pp. 2293–2314, Oct. 2004.

[29] U. Erez, S. Litsyn, and R. Zamir, “Lattices which are good for (almost) everything,” *IEEE Trans. Inf. Theory*, vol. 51, no. 10, pp. 3401–3416, Oct. 2005.

[30] U. Erez, S. Shamai, and R. Zamir, “Capacity and lattice strategies for canceling known interference,” *IEEE Trans. Inf. Theory*, vol. 51, no. 11, pp. 3820–3833, Nov. 2005.

[31] U. Erez and S. ten Brink, “A close-to-capacity dirty paper coding scheme,” *IEEE Trans. Inf. Theory*, vol. 51, no. 10, pp. 3417–3432, Oct. 2005.

[32] A. Elbert and A. Laforgia, “An asymptotic relation for the zeros of Bessel functions,” *J. Math. Analysis and Applications*, vol. 98, no. 2, pp. 502–510, 1984.

[33] W. Feller, *An Introduction to Probability Theory and Its Applications, Vol. 2*, 2nd ed. New York: Wiley, 1971.

[34] G.D. Forney, *Concatenated Codes,* Cambridge, U.K.: MIT press, 1966.

[35] G.D. Forney and M.D. Trott, “The dynamics of group codes: Dual abelian group codes and systems,” *IEEE Trans. Inf. Theory*, vol. 50, no. 12, pp. 2935–2965, Dec. 2004.

[36] A. Gersho, “Asymptotically optimal block quantization,” *IEEE Trans. Inf. Theory,* vol. 25, no. 4, pp. 373–380, 1979.

[37] R.M. Gray and D.L. Neuhoff, “Quantization,” *IEEE Trans. Inf. Theory,* vol. 44, no. 6, pp. 2325–2383, Oct. 1998.

[38] T.C. Gulcu and A. Barg, “Achieving secrecy capacity of the wiretap channel and broadcast channel with a confidential component,” in *Proc. 2015 Inf. Theory Workshop,* Jerusalem, Israel.

[39] D. Hankerson, A.J. Menezes, and S. Vanstone, *Guide to Elliptic Curve Cryptography,* Springer Science & Business Media, 2006.

[40] X. He and A. Yener, "Providing secrecy with lattice codes," *Proc. 46th Annual Allerton Conf. on Communication, Control, and Computing*, Monticello, IL, 2008, pp. 1199–1206.

[41] X. He and A. Yener, "Strong secrecy and reliable Byzantine detection in the presence of an untrusted relay," *IEEE Trans. Inf. Theory*, vol. 59, no. 1, pp. 177–192, Jan. 2013.

[42] X. He and A. Yener, "End-to-end secure multi-hop communication with untrusted relays," *IEEE Trans. Wireless Comm.*, vol. 12, no. 1, pp. 1–11, Jan. 2013.

[43] X. He and A. Yener, "Providing secrecy with structured codes: Two-user Gaussian channels," *IEEE Trans. Inf. Theory*, vol. 60, no. 4, pp. 2121–2138, Apr. 2014.

[44] I.N. Herstein, *Topics in Algebra*, 2nd ed. New York: Wiley, 1975.

[45] A. Joseph and A.R. Barron, "Least squares superposition codes of moderate dictionary size are reliable at rates up to capacity," *IEEE Trans. Inf. Theory*, vol. 58, no. 5, pp. 2541–2557, May. 2012.

[46] A. Joseph and A.R. Barron, "Fast sparse superposition codes have near exponential error probability for $R < C$," *IEEE Trans. Inf. Theory*, vol. 60, no. 2, pp. 919–942, Feb. 2014.

[47] D.A. Karpuk and A. Chorti, "Perfect secrecy in physical layer network coding systems from structured interference," *arXiv preprint,* arXiv:1507.01098, 2015.

[48] N. Kashyap, V. Shashank, and A. Thangaraj, "Secure computation in a bidirectional relay," *Proc. 2012 IEEE Int. Symp. Information Theory*, Cambridge, MA, pp. 1162–1166.

[49] D. Krithivasan and S. Pradhan, "A proof of the existence of good nested lattices," Univ. Michigan, Jul. 2007 [Online]. Available: `http://www.eecs.umich.edu/techreports/systems/cspl/cspl-384.pdf`.

[50] S. Kudekar, T.J. Richardson, and R.L. Urbanke, "Threshold saturation via spatial coupling: why convolutional LDPC ensembles perform so well over the BEC," *IEEE Trans. Inf. Theory*, vol. 57, no. 2, pp. 803–834, Feb. 2011.

[51] S. Kudekar, T. Richardson, and R.L. Urbanke, "Spatially coupled ensembles universally achieve capacity under belief propagation," *IEEE Trans. Inf. Theory*, vol. 59, no. 12, pp. 7761–7813, Dec. 2013.

[52] S. Leung-Yan-Cheong and M.E. Hellman, "The Gaussian wire-tap channel," *IEEE Trans. Inf. Theory*, vol. 24, no. 4, pp. 451–456, Jul. 1978.

[53] Y. Liang and H.V. Poor, "Information theoretic security," *Found. Trends in Comm. Inf. Theory,* vol. 5, no. 4-5, pp. 355–580, 2009.

[54] E. Lieb, "Sharp constant in the Hardy-Littlewood-Sobolev and related inequalities," *Annals of Mathematics*, vol. 118, no. 2, pp. 349–374, Sep. 1983.

[55] S.C. Liew, S. Zhang, and L. Lu, "Physical-layer network coding: Tutorial, survey, and beyond", *Phys. Comm.*, vol. 6, pp. 4–42, Mar. 2013.

[56] T. Linder, C. Schlegel, and K. Zeger, "Corrected proof of de Buda's theorem [lattice channel codes]," *IEEE Trans. Inf. Theory,* vol. 39, no. 5, pp. 1735–1737, Sep. 1993.

[57] C. Ling, L. Luzzi, and M.R. Bloch, "Secret key generation from Gaussian sources using lattice hashing," *Proc. 2013 IEEE International Symposium on Information Theory (ISIT)*, Istanbul, Turkey, pp. 2621–2625.

[58] C. Ling, L. Luzzi, J.-C. Belfiore, and D. Stehlé, "Semantically secure lattice codes for the Gaussian wiretap channel," *IEEE Trans. Inf. Theory*, vol. 60, no. 10, pp. 6399–6416, Oct. 2014.

[59] H.A. Loeliger, “Averaging bounds for lattices and linear codes,” *IEEE Trans. Inf. Theory,* vol. 43, no. 6, pp. 1767–1773, Nov. 1997.

[60] A. Lubotzky, R. Phillips, and P. Sarnak, “Ramanujan graphs,” *Combinatorica,* vol. 8, no. 3, pp. 261–277, 1988.

[61] E. Lukacs, *Characteristic Functions*, 2nd ed. London, U.K.: Griffin, 1970.

[62] H. Mahdavifar and A. Vardy, “Achieving the secrecy capacity of wiretap channels using polar codes,” *IEEE Trans. Inf. Theory,* vol. 57, no. 10, Oct. 2011.

[63] A.W. Marcus, D.A. Spielman, and N. Srivastava, “Interlacing families I: bipartite Ramanujan graphs of all degrees,” *Annals of Mathematics,* vol. 182, no. 1, pp. 307–325, Jul. 2015.

[64] U.M. Maurer, “Secret key agreement by public discussion from common information,” *IEEE Trans. Inf. Theory*, vol. 39, no. 5, pp. 733–742, May 1993.

[65] U. Maurer and S. Wolf. “Information-theoretic key agreement: From weak to strong secrecy for free,” in *Adv. Cryptology–EUROCRYPT 2000*, Berlin, Germany: Springer Berlin-Heidelberg, pp. 351–368, 2000.

[66] D. Micciancio and O. Regev, “Lattice-based cryptography,” in *Post-Quantum Cryptography,* D.J. Bernstein, Eds. Berlin, Germany: Springer Berlin-Heidelberg, 2009.

[67] H. Minkowski, *Geometrie der Zahlen*, Teubner, Leipzig, 1896.

[68] M. Mondelli, S.H. Hassani, and R. Urbanke, “Unified scaling of polar codes: Error exponent, scaling exponent, moderate deviations, and error floors,” arXiv preprint, 2015. [Online] Available: `http://arxiv.org/abs/1501.02444`.

[69] B. Nazer and M. Gastpar, “Compute-and-forward: harnessing interference through structured codes,” *IEEE Trans. Inf. Theory*, vol. 57, no. 10, pp. 6463–6486, Oct. 2011.

[70] B. Nazer and M. Gastpar, "Reliable physical layer network coding," *Proc. of the IEEE*, vol. 99, no. 3, pp. 438–460, Mar. 2011.

[71] S. Nitinawarat and P. Narayan, "Secret key generation for correlated Gaussian sources," *IEEE Trans. Inf. Theory*, vol. 58, no. 6, pp. 3373–3391, Jun. 2012.

[72] F. Oggier, P. Solé, and J.-C. Belfiore, "Lattice codes for the wiretap Gaussian channel: construction and analysis," arXiv preprint, 2011. [Online] Available:: `http://arxiv.org/abs/1103.4086`.

[73] O. Ordentlich and U. Erez, "A simple proof for the existence of "good" pairs of nested lattices," in *Proc. 2012 IEEE 27th Conv. Electrical and Electronics Engineers in Israel,* Eilat, Israel, pp. 1–12.

[74] G. Poltyrev, "On coding without restrictions for the AWGN channel," *IEEE Trans. Inf. Theory*, vol. 40, no. 2, pp. 409–417, Mar. 1994.

[75] P. Popovski and H. Yomo, "Physical network coding in two-way wireless relay channels," in *Proc. IEEE Int. Conf. Communications*, Glasgow, Scotland, 2007, pp. 707–712.

[76] S. Ramanujan, "A proof of Bertrand's postulate," *J. Indian Math. Soc.*, vol. 11, pp. 181–182, 1919.

[77] T. Richardson and R. Urbanke, *Modern Coding Theory,* Cambridge, U.K.: Cambridge University Press, 2008.

[78] C.A. Rogers, *Packing and Covering,* Cambridge, U.K.: Cambridge University Press, 1964.

[79] R.M. Roth, *Introduction to Coding Theory*, Cambridge, U.K.: Cambridge University Press, 2006.

[80] H. Rubin and T.M. Sellke, "Zeroes of infinitely differentiable characteristic functions," in *A Festschrift for Herman Rubin*, Anirban DasGupta, ed., Institute of

Mathematical Statistics Lecture Notes – Monograph Series, vol. 45, pp. 164–170, 2004.

[81] M.R. Sadeghi, A.H. Banihashemi, and D. Panario, “Low-density parity-check lattices: construction and decoding analysis,” *IEEE Trans. Inf. Theory*, vol. 52, no. 10, pp. 4481–4495, Oct. 2006.

[82] A. Sakzad, M.R. Sadeghi, and D. Panario, “Construction of turbo lattices,” *Proc 2010 48th Ann. Allerton Conference Comm. Contr. Computing*, Allerton, IL, 2010, pp. 14-21.

[83] A. Sakzad, M.R. Sadeghi, and D. Panario, “Turbo lattices: Construction and error decoding performance,” arXiv preprint, 2015. [Online] Available: `http://arxiv.org/abs/1108.1873`.

[84] C.E. Shannon, “A mathematical theory of communication,” *Bell Syst. Tech. J.,* vol. 27, no. 3, pp. 379–423, Jul. 1948.

[85] C.E. Shannon, “Communication theory of secrecy systems,” *Bell Syst. Tech. J.,* vol. 28, no. 4, pp. 656–715, Oct. 1949.

[86] N. Sommer, M. Feder, and O. Shalvi, “Low density lattice codes,” *IEEE Trans. Inf. Theory*, vol. 54, no. 4, pp. 1561–1585, Apr. 2008.

[87] E.M. Stein and G.L. Weiss, *Introduction to Fourier Analysis on Euclidean Spaces*, Princeton, NJ: Princeton Univ. Press, 1971.

[88] A. Subramanian, A. Thangaraj, M. Bloch and S.W. McLaughlin, “Strong secrecy on the binary erasure wiretap channel using large-girth LDPC codes,” *IEEE Trans. Inf. Forensics Security,* vol. 6, no. 3, pp. 585–594, Sept. 2011.

[89] A. Thangaraj, S. Dihidar, A.R. Calderbank, S.W. McLaughlin, and J.M. Merolla, “Applications of LDPC codes to the wiretap channel,” *IEEE Trans. Inf. Theory,* vol. 53, no. 8, pp. 2933–2945, Aug. 2007.

[90] F.G. Tricomi, "Sulle funzioni di Bessel di ordine e argomento pressoché uguali," *Atti Accad. Sci. Torino Cl. Sci. Fis. Mat. Natur.*, vol. 83, pp. 3–20, 1949.

[91] N.E. Tunali, K.R. Narayanan, and H.D. Pfister, "Spatially-coupled low density lattices based on Construction A with applications to compute-and-forward" *Proc. 2013 Information Theory Workshop*, Sevilla, Spain, 2013, pp. 1–5.

[92] H. Tyagi and A. Vardy, "Explicit capacity-achieving coding scheme for the Gaussian wiretap channel," *Proc. 2014 IEEE Int. Symp. Information Theory (ISIT),* Honolulu, HI, 2014, pp.956-960.

[93] R. Urbanke and B. Rimoldi, "Lattice codes can achieve capacity on the AWGN channel," *IEEE Trans. Inf. Theory*, vol. 44, no. 1, pp. 273–278, Jan. 1998.

[94] S. Vatedka and N. Kashyap, "Lattice coding for strongly secure compute-and-forward in a bidirectional relay," *Proc. 2013 Int. Symp. Inf. Theory*, Istanbul, Turkey, 2013, pp. 2775–2779.

[95] S. Vatedka, N. Kashyap, and A. Thangaraj, "Secure compute-and-forward in a bidirectional relay," *IEEE Trans. Inf. Theory,* vol. 51, no. 5, pp. 2531–2556, May 2015.

[96] S. Vatedka and N. Kashyap, "Nested lattice codes for secure bidirectional relaying with asymmetric channel gains," in *Proc. 2015 IEEE Inf. Theory Workshop*, Jerusalem, Israel, 2015.

[97] S. Vatedka and N. Kashyap, "Nested lattice codes for secure bidirectional relaying with asymmetric channel gains," arXiv preprint, 2015. [Online] Available: `http://arxiv.org/abs/1506.02152`.

[98] S. Vatedka and N. Kashyap, "Some goodness properties of LDA lattices," in *Proc. 2015 IEEE Inf. Theory Workshop*, Jerusalem, Israel, 2015.

[99] S. Vatedka and N. Kashyap, "Some goodness properties of LDA lattices," submitted, *Prob. Inf. Transm.*, Dec. 2015.

[100] S. Vatedka and N. Kashyap, "A capacity-achieving coding scheme for the AWGN channel with polynomial encoding and decoding complexity," in *Proc. 2016 Nat. Conf. Comm.*, Guwahati, India, pp. 36–41.

[101] S. Vatedka and N. Kashyap, "A capacity-achieving coding scheme for the AWGN channel with polynomial encoding and decoding complexity," arXiv preprint, 2016. [Online] Available: `http://arxiv.org/abs/1603.08236`

[102] S. Vatedka and N. Kashyap, "A lattice coding scheme for secret key generation from Gaussian Markov tree sources," accepted, *2016 IEEE Int. Symp. Inf. Theory*, Barcelona, Spain, 2016. [Online]. Available: `http://arxiv.org/abs/1603.08236`

[103] S. Vishwanath and S.A. Jafar, "Generalized degrees of freedom of the symmetric Gaussian K-User interference channel," *IEEE Trans. Inf. Theory*, vol. 56, no. 7, pp. 3297–3303, Jul. 2010.

[104] Z. Wang and C. Ling, "On the geometric ergodicity of Metropolis-Hastings algorithms for lattice Gaussian sampling," submitted, *IEEE Trans. Inf. Theory,* 2015. [Online]. Available: `http://arxiv.org/abs/1501.05757`.

[105] C.X. Wang et al., "Cellular architecture and key technologies for 5G wireless communication networks," *IEEE Comm. Mag.*, vol. 52, no. 2, pp. 122–130, Feb. 2014.

[106] S. Watanabe and Y. Oohama, "Secret key agreement from vector Gaussian sources by rate limited communication," *IEEE Trans. Inf. Forensics Security,* vol. 6, no. 3, pp. 541–550, Sep. 2011.

[107] M. Wilson, K. Narayanan, H. Pfister, and A. Sprintson, "Joint physical layer coding and network coding for bidirectional relaying," *IEEE Trans. Inf. Theory*, vol. 56, no. 11, pp. 5641–5654, Nov. 2010.

[108] S.J. Wolfe, "On the finite series expansion of multivariate characteristic functions," *J. Multivariate Anal.*, vol. 3, pp. 328–335, 1973.

[109] A. D. Wyner, “The wire-tap channel,” *Bell Syst. Tech. J.*, vol. 54, no. 8, pp. 1355–1387, Oct. 1975

[110] Y. Yan, C. Ling, and X. Wu, “Polar lattices: Where Arikan meets Forney,” *Proc. 2013 IEEE Int. Symp. Information Theory*, Istanbul, Turkey, pp. 1292–1296.

[111] Y. Yan, L. Liu, C. Ling, and X. Wu, “Construction of capacity-achieving lattice codes: Polar lattices,” arXiv preprint, 2015. [Online] Available: `http://arxiv.org/abs/1411.0187`.

[112] R. Zamir, *Lattice Coding for Signals and Networks*, Cambridge, U.K.: Cambridge University Press, 2014.

[113] G. Zémor, “On expander codes,” *IEEE Trans. Inf. Theory,* vol. 47, no. 2, pp. 835–837, Feb. 2001.

[114] J. Zhan, B. Nazer, U. Erez, and M. Gastpar, “Integer-forcing linear receivers,” *IEEE Trans. Inf. Theory*, vol. 60, no. 12, pp. 7661–7685, Dec. 2014.

[115] S. Zhang, L. Fan, M. Peng, H.V. Poor, “Near-optimal modulo-and-forward scheme for the untrusted relay channel,” arXiv preprint, 2015. [Online] Available: `http://arxiv.org/abs/1503.08928`.

[116] S. Zhang and S.-C. Liew, “Channel coding and decoding in a relay system operated with physical-layer network coding,” *IEEE J. Sel. Areas Comm.*, vol. 27, no. 5, pp. 788–796, Jun. 2009.