Chapter 1

Introduction

Chapter 2

A Primer on Lattice Codes

Chapter 3

Perfectly Secure Bidirectional Relaying

3.1 Introduction

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, they are connected only to R, and not to each other directly. 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 [**?**, **?**, **?**, **?**, **?**].

We use the compute-and-forward framework proposed in [**?**, **?**] 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 $\mathtt{A} \to \mathtt{R}$ and $\mathtt{B} \to \mathtt{R}$ 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 $X \oplus Y$.

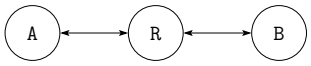

Figure 3.1: Bidirectional relay.

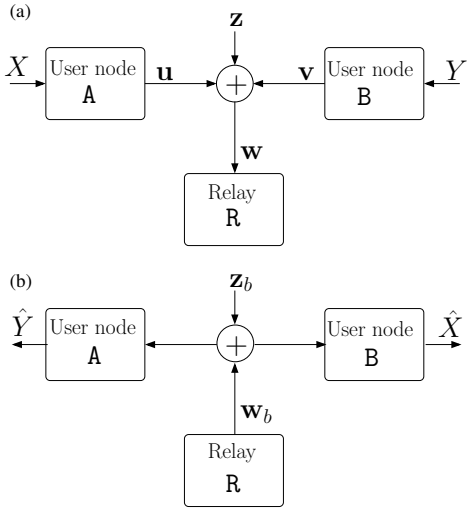

Figure 3.2: Bidirectional relaying: (a) MAC phase, (b) Broadcast phase.

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 d-dimensional real vectors $\mathbf{U}$ and $\mathbf{V}$ 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 paper.

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 $\mathbb{R}^d$. 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. 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 $\max\{x, 0\}$. 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 later 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* [**?**]. 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 [**?**] in the context of the Gaussian wiretap channel, and we will make use of the techniques developed there. Using this scheme, we 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 [**?**] 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.

In prior work, the problem of secure bidirectional relaying in the presence of an untrusted relay was studied by He and Yener in [**?**], who showed that the mutual information rate, defined to be $\frac{1}{d}\mathcal{I}(X;\mathbf{U}+\mathbf{V}) = \frac{1}{d}\mathcal{I}(Y;\mathbf{U}+\mathbf{V})$ goes to zero for large blocklengths d. They later studied the problem under a strong secrecy constraint in [**?**], 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 [**?**]. 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.

The idea of using nested lattice codes for secure communication is not new. They have been proposed for secure communication in other scenarios, particularly the Gaussian wiretap channel (see e.g., [**?**, **?**, **?**]). They have also been used in interference networks [**?**], and for secret key generation using correlated Gaussian sources [**?**].

Recall that the compute-and-forward protocol has two phases: a MAC phase and a broadcast phase. We will restrict our study exclusively to the MAC phase, since there is no security requirement in the broadcast phase and the relay can use a capacity-approaching code to broadcast $X \oplus Y$ to the users.

Organization of the paper

We establish some basic notation and recall some definitions related to lattices in Section 3.2. We describe the secure bidirectional relaying problem in Section 3.3, and then proceed to design coding schemes under the perfect secrecy constraint in Section 3.4. The main result under the perfect secrecy constraint is given in Theorem 1. We give a randomized encoding scheme for any arbitrary nested lattice code that achieves perfect secrecy in the absence of noise in Section 3.5, then study the effect of additive noise and find achievable transmission rates in Section 3.6. Thereafter, we study the same problem under a strong secrecy constraint, design coding schemes, and evaluate the performance in Section **??**, with the main result summarized in Theorem **??**. In Section **??**, we show that these schemes can be extended to the multi-hop line network [**?**] and find achievable transmission rates under the two secrecy constraints. We make some concluding remarks in Section **??**. Most of the technical proofs are given in appendices.

3.2 Definitions and Notation

We first describe the notation we will use throughout the paper. We denote the set of real numbers by $\mathbb{R}$, and integers by $\mathbb{Z}$. We use the notation $\mathbb{R}^+$

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

for the set of nonnegative real numbers. The number of elements in a finite set S is denoted by $|S|$. If x is a real number, then $[x]^+$ is defined as $\max\{x, 0\}$. Random vectors are denoted in boldface upper case, e.g., $\mathbf{U}$, and their instances in boldface lower case, as in $\mathbf{u}$. The components of the vectors are denoted in normal font, e.g., $\mathbf{x} = [x_1\ x_2]^T$. Matrices are represented in sans-serif, as in H. The Euclidean (ℓ^2) norm of a column vector $\mathbf{h}$ is denoted by $\|\mathbf{h}\|$. The identity matrix of size $M \times M$ is denoted by I_M.

The probability of an event A is denoted by $\Pr[A]$. If X is a random variable, then $\mathcal{H}(X)$ denotes the entropy of X. The symbol $\mathbb{E}[\cdot]$ denotes expectation. The characteristic function of a random variable X is the function $\psi(t) = \mathbb{E}[e^{iXt}]$, for $t \in \mathbb{R}$. For random variables X, Y, the notation $X \perp\!\!\!\perp Y$ means that X and Y are independent. The mutual information between X and Y is denoted by $\mathcal{I}(X;Y)$.

Let $f(n)$ and $g(n)$ be sequences of positive real numbers. We say that $g(n) = o(f(n))$ if $g(n)/f(n) \to 0$ as $n \to \infty$. Also, $g(n) = o_n(1)$ if $g(n) \to 0$ as $n \to \infty$. Furthermore, $g(n) = \Omega(f(n))$ if there exists a constant $K > 0$ such that $g(n) > Kf(n)$ for all sufficiently large n, and $g(n) = \mathcal{O}(f(n))$ if there exists a constant $K > 0$ such that $g(n) < Kf(n)$ for all sufficiently large n.

3.2.1 Lattices in $\mathbb{R}^d$

We briefly recall some definitions of lattices and their properties. For a more detailed treatment, see e.g., [**?**, **?**].

Let k, d be positive integers with $k \leq d$. Suppose $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k$ are linearly independent column vectors in $\mathbb{R}^d$. 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 $\mathbb{R}^d$. 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; clearly, the basis of a lattice is not unique, e.g., $\{-\mathbf{u}_1, -\mathbf{u}_2, \ldots, -\mathbf{u}_k\}$ is also a basis.

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

If Λ and Λ_0 are two lattices in $\mathbb{R}^d$ such that $\Lambda_0 \subset \Lambda$, then we say that Λ_0 is

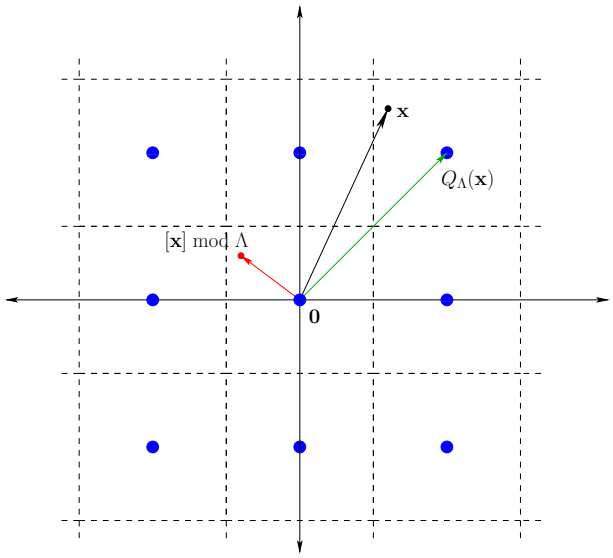

Figure 3.3: Illustrating the $Q_\Lambda(.)$ and the $[.] \bmod \Lambda$ operation for the $\mathbb{Z}^2$ lattice.

a *sublattice* of Λ, or 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* of Λ_0 in Λ, denoted by $|\Lambda/\Lambda_0|$. It is a standard fact that $|\Lambda/\Lambda_0| = \det\Lambda_0/\det\Lambda$ [**?**, Theorem 5.2].

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

For any $\mathbf{x} \in \mathbb{R}^d$, we define the nearest neighbour quantizer $Q_\Lambda(\mathbf{x}) := \arg\min_{\lambda\in\Lambda}\|\mathbf{x} - \lambda\|$ to be the function which maps $\mathbf{x}$ to the closest point in Λ. The *fundamental Voronoi region* of Λ is defined as $\mathcal{V}(\Lambda) := \{\mathbf{y} : Q_\Lambda(\mathbf{y}) = \mathbf{0}\}$. The volume of the fundamental Voronoi region, $\mathrm{vol}(\mathcal{V}(\Lambda))$ is equal to $\det\Lambda$ [**?**, **?**].

For any $\mathbf{x} \in \mathbb{R}^d$, we define the modulo-$\Lambda$ operation as $[\mathbf{x}] \bmod \Lambda := \mathbf{x} - Q_\Lambda(\mathbf{x})$. In other words, $[\mathbf{x}] \bmod \Lambda$ gives the quantization error of the nearest neighbour quantizer $Q_\Lambda(\cdot)$. Fig. 3.3 illustrates the $Q_\Lambda(\cdot)$ and the modulo-Λ operations.

The *covering radius* of Λ, denoted by $r_{\mathrm{cov}}(\Lambda)$, is defined as the radius of the smallest closed ball in $\mathbb{R}^d$ centered at $\mathbf{0}$ which contains $\mathcal{V}(\Lambda)$. The

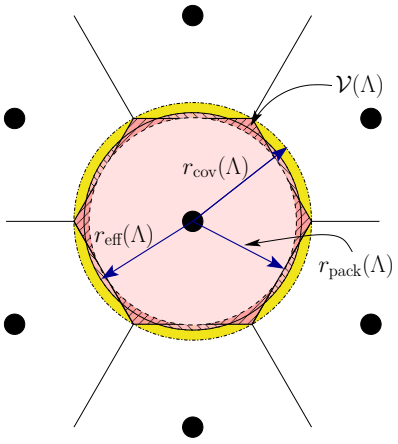

Figure 3.4: Illustrating the covering, packing and effective radii of the hexagonal lattice.

effective radius, $r_{\text{eff}}(\Lambda)$, is defined as the radius of a ball in $\mathbb{R}^d$ having the same volume as that of $\mathcal{V}(\Lambda)$. The *packing radius*, $r_{\text{pack}}(\Lambda)$, is the radius of the largest open ball centered at $\mathbf{0}$ that is contained 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. 3.4.

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

$$\mathcal{G}_\Lambda = \frac{1}{d\left(\det\Lambda\right)^{1+2/d}} \int_{\mathcal{V}(\Lambda)} \|\mathbf{y}\|^2 \, d\mathbf{y}. \tag{3.2}$$

3.3 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 $\mathbb{G}^{(d)}$, where the choice of $\mathbb{G}^{(d)}$ is left to the system designer. The mapping is such that the random variables X and Y remain uniformly distributed over $\mathbb{G}^{(d)}$, and we will see later that this distribution helps in achieving secrecy. The addition operation in the group $\mathbb{G}^{(d)}$ is denoted $\oplus$. The encoder at node `A` maps the given message X into a random d-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

$$\mathbf{W} = \mathbf{U} + \mathbf{V} + \mathbf{Z}, \tag{3.3}$$

where $\mathbf{Z}$ is a Gaussian random vector with zero mean and covariance matrix $\sigma^2 \mathsf{I}_d$, 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 one of the following:

- *Perfect secrecy*: The mutual information between $\mathbf{W}$ and each individual message is exactly zero[2], i.e., $\mathcal{I}(\mathbf{W}; X) = \mathcal{I}(\mathbf{W}; Y) = 0$.
- *Strong secrecy*: $\mathcal{I}(\mathbf{W}; X)$ and $\mathcal{I}(\mathbf{W}; Y)$ can be made arbitrarily small for all sufficiently large d.

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 either zero (perfect secrecy) or can be made arbitrarily small for all sufficiently large d (strong secrecy). 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, any scheme that achieves perfect (strong) secrecy in the absence of noise will also achieve perfect (strong) 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 $\mathbb{G}^{(d)}$, $\frac{1}{d}\log_2 |\mathbb{G}^{(d)}|$ 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 $\frac{1}{d}\log_2 |\mathbb{G}^{(d)}|$) for a given power constraint at the user nodes. To formalize these notions, we have the following definition:

Definition 1. *For a positive integer d, a $(d, M^{(d)})$ 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 $\mathbb{G}^{(d)}$ with $M^{(d)} = |\mathbb{G}^{(d)}|$ elements.*

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

2. ***Codebook:*** *The codebook, denoted by $\mathcal{C}$, is a discrete subset of $\mathbb{R}^d$, 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 $\mathbb{G}^{(d)}$ to $\mathbb{R}^d$, 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 $x, y \in \mathbb{G}^{(d)}$. At node* `A`*, given a message $x \in \mathbb{G}^{(d)}$ 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*

$$R^{(d)} = \frac{\log_2 M^{(d)}}{d}. \tag{3.4}$$

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

$$P^{(d)} = \frac{1}{d}\mathbb{E}\|\mathbf{U}\|^2 = \frac{1}{d}\mathbb{E}\|\mathbf{V}\|^2. \tag{3.5}$$

4. ***Decoder:*** *The relay* `R` *receives a vector $\mathbf{W} \in \mathbb{R}^{(d)}$ as given in (3.3). The decoder, $\mathcal{D}^{(d)} : \mathbb{R}^d \to \mathbb{G}^{(d)}$ maps the received vector to an element of the set of messages. The average probability of error of the decoder is defined as*

$$\eta^{(d)} := \mathbb{E}\big[Pr[\mathcal{D}^{(d)}(\mathbf{W}) \neq X \oplus Y]\big],$$

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

3.4 Perfect Secrecy

We first study the case where perfect statistical independence between $\mathbf{U} + \mathbf{V}$ and the individual messages is required, and the relay must be able to reliably compute $X \oplus Y$ (where $\oplus$ denotes addition within $\mathbb{G}^{(d)}$) from the received vector. To summarize, 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 $d = 1$, 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, for every $\delta > 0$, there exists a sequence of $(d, M^{(d)})$ codes such that*

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

and for all sufficiently large d,

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

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

Theorem 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.

3.5 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 d-dimensional setting.

We will show that there exist 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 2. *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 [**?**]), 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. □

3.5.1 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

$$p_{UVXY}(k, l, a, b) = (1/2)(1/2)p_{U|a}(k)p_{V|b}(l) \tag{3.6}$$

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

$$\begin{aligned} p_{U|0}(k) = p_{V|0}(k) = 0 \ \text{ for all odd } k \in \mathbb{Z}, \\ p_{U|1}(k) = p_{V|1}(k) = 0 \ \text{ for all even } k \in \mathbb{Z}. \end{aligned} \tag{3.7}$$

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

$$p_U * p_V = p_{U|a} * p_V = p_U * p_{V|a} \ \text{ for } a \in \{0, 1\}. \tag{3.8}$$

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, (3.8) is equivalent to

$$\varphi_U \varphi_V = \varphi_{U|a}\varphi_V = \varphi_U \varphi_{V|a} \ \text{ for } a \in \{0, 1\}. \tag{3.9}$$

Note that $\varphi_U = (1/2)(\varphi_{U|0} + \varphi_{U|1})$ and $\varphi_V = (1/2)(\varphi_{V|0} + \varphi_{V|1})$. Hence, (3.9) 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. *Suppose that the conditional pmfs $p_{U|a}$ and $p_{V|a}$, $a \in \{0,1\}$, satisfy (3.7) and (3.9). Then, the rvs U, V, X, Y with joint pmf given by (3.6) have properties (S1)–(S3).*

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

Proposition 4. *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 (3.9) must hold. The equality $\varphi_U \varphi_V = \varphi_U \varphi_{V|a}$ in (3.9) 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 4, 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.

We now give a construction of U, V that satisfy the hypotheses of Lemma 3. 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 summation formula,

[3] In fact, a stronger negative result can be shown — see Proposition 9.

which we briefly recall here. Our description is based largely on Section XIX.5 in [**?**].

3.5.2 The Poisson Summation Formula

Fix a positive integer d, and let Λ be a full-rank lattice in $\mathbb{R}^d$. Recall from Section 3.2.1 that $\hat{\Lambda}$ denotes the Fourier dual of Λ.

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

$$\sum_{\mathbf{n}\in\hat{\Lambda}} \psi(\boldsymbol{\zeta}+\mathbf{n})\, e^{-i\langle \mathbf{n},\, \mathbf{s}\rangle} = (\det\Lambda) \sum_{\mathbf{k}\in\Lambda} f(\mathbf{k}+\mathbf{s})\, e^{i\langle \mathbf{k}+\mathbf{s},\, \boldsymbol{\zeta}\rangle}, \tag{3.10}$$

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., [**?**, Chapter VII, Corollary 2.6] or [**?**, Eq. (17.1.2)] — for (3.10) to hold. However, as argued by Feller in proving the formula in the one-dimensional setting [**?**, 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 (3.10) 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 (3.10) 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 $\Lambda + \mathbf{s} := \{\mathbf{k}+\mathbf{s} : \mathbf{k} \in \Lambda\}$, 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 $\hat{\Lambda}$: $\psi(\mathbf{t}) = 0$ for all $\mathbf{t} \notin \mathcal{V}(\hat{\Lambda})$. In this case, we can readily show that the series on the LHS of (3.10) converges to a continuous function Ψ. Indeed, if we define $\widetilde{\psi}(\mathbf{t}) := \psi(\mathbf{t}) e^{-i\langle \mathbf{t},\mathbf{s}\rangle}$, then the series on the LHS of (3.10) may be written as $\Psi(\boldsymbol{\zeta}) := e^{i\langle \boldsymbol{\zeta},\mathbf{s}\rangle} \widetilde{\Psi}(\boldsymbol{\zeta})$, where

$$\widetilde{\Psi}(\boldsymbol{\zeta}) := \sum_{\mathbf{n}\in\hat{\Lambda}} \widetilde{\psi}(\boldsymbol{\zeta}+\mathbf{n}).$$

Now, recall that ψ, being a characteristic function, is continuous on $\mathbb{R}^d$; hence, so is $\widetilde{\psi}$. Also, by assumption, ψ is supported within $\mathcal{V}(\hat{\Lambda})$; hence, so is $\widetilde{\psi}$. In particular, by continuity, $\widetilde{\psi}$ must be 0 on the boundary of $\mathcal{V}(\hat{\Lambda})$; therefore, the supports of $\widetilde{\psi}(\cdot)$ and $\widetilde{\psi}(\cdot+\mathbf{n})$ do not intersect for any non-zero $\mathbf{n} \in \hat{\Lambda}$. 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 (3.10) 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 5. *Let Λ be a full-rank lattice in $\mathbb{R}^d$. Let $\psi : \mathbb{R}^d \to \mathbb{C}$ be a characteristic function such that $\psi(\mathbf{t}) = 0$ for all $\mathbf{t} \notin \mathcal{V}(\hat{\Lambda})$, and let $f : \mathbb{R}^d \to \mathbb{R}^+$ be the corresponding probability density function. Then, for any $\mathbf{s} \in \mathbb{R}^d$, the function $\Psi : \mathbb{R}^d \to \mathbb{C}$ defined by*

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

is the characteristic function of a random vector supported within the set $\Lambda + \mathbf{s} := \{\mathbf{k}+\mathbf{s} : \mathbf{k} \in \Lambda\}$. 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., [**?**, Section XV.2, Table 1], [**?**], [**?**]. We also give an explicit construction in Example 1 in Section 3.5.3.

Applying Proposition 5 to the one-dimensional lattice $T\mathbb{Z} = \{kT : k \in \mathbb{Z}\}$, with $T > 0$, we obtain the corollary below.

Corollary 6. *Let ψ be a characteristic function of a real-valued random variable 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*

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

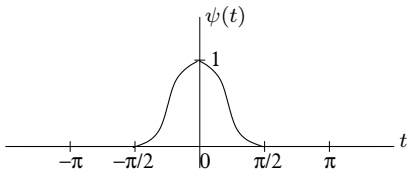

Figure 3.5: A generic characteristic function supported on $[-\pi/2, \pi/2]$.

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.

3.5.3 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. 3.5; 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 6 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. 3.6. Note that $\varphi(\zeta) = \sum_{n=-\infty}^{\infty} \psi(\zeta + 2\pi n)$. Thus, applying Corollary 6 with $T = 1$ and $s = 0$, we find that φ is the characteristic function of an integer-valued rv, with pmf given by

$$p(k) = f(k) \text{ for all } k \in \mathbb{Z}. \tag{3.11}$$

[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 $\psi\bar{\psi} = |\psi|^2$.

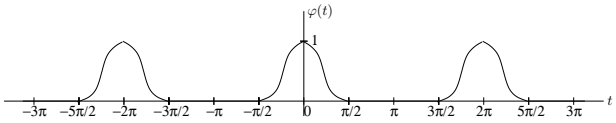

Figure 3.6: Period-2π extension of generic ψ from Fig. 3.5.

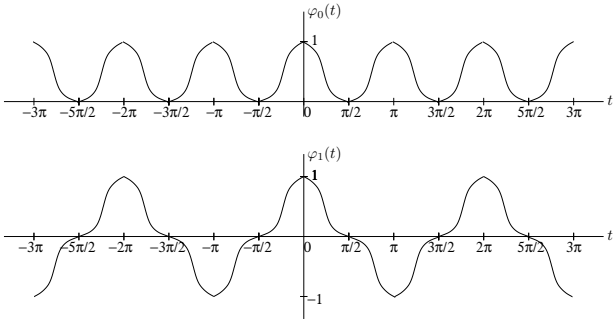

Figure 3.7: The periodic functions φ_0 and φ_1 derived from ψ.

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. 3.7 for a generic ψ shown in Fig. 3.5. 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. 3.7.

Applying Corollary 6 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

$$p_0(k) = \begin{cases} 2f(k) & \text{if } k \text{ is an even integer} \\ 0 & \text{otherwise.} \end{cases} \tag{3.12}$$

and that corresponding to φ_1 is

$$p_1(k) = \begin{cases} 2f(k) & \text{if } k \text{ is an odd integer} \\ 0 & \text{otherwise.} \end{cases} \tag{3.13}$$

From (3.11)–(3.13), 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

$$\varphi^2 = \varphi\varphi_0 = \varphi\varphi_1. \tag{3.14}$$

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 (3.11). Clearly, (3.7) holds. To verify (3.9), note that, by virtue of (3.14), 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, the rvs (U, V, X, Y) with joint pmf given by (3.6) have the properties (S1)–(S3).

Recall from the discussion following Proposition 4 that we need the rvs U and V to have finite variance. To ensure this, we use the fact [**?**, 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 7. *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 (3.12) and 3.13). 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 7, 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 (3.12)) if the bit is 0, or with probability $p_1(k)$ (from (3.13)) 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,

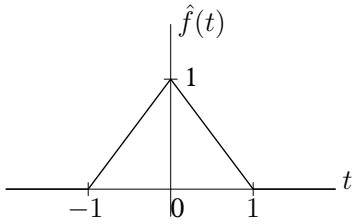

Figure 3.8: $\hat{f}(t) = \max\{0, 1 - |t|\}$.

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 [?, Section XV.2, Table 1])*

$$f(x) = \begin{cases} \frac{1}{2\pi} & \text{if } x = 0 \\ \frac{1-\cos x}{\pi x^2} & \text{if } x \neq 0 \end{cases} \tag{3.15}$$

which has characteristic function

$$\hat{f}(t) = \max\{0, 1 - |t|\} \tag{3.16}$$

The function $\hat{f}$ is plotted in Fig. 3.8. In particular, $\hat{f}(t) = 0$ for $|t| \geq 1$.

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

$$g(t) = (\hat{f} * \hat{f})(t) = \begin{cases} \frac{1}{2}|t|^3 - t^2 + \frac{2}{3} & \text{if } |t| \leq 1 \\ \frac{1}{6}(2 - |t|)^3 & \text{if } 1 \leq |t| \leq 2 \\ 0 & \text{otherwise} \end{cases} \tag{3.17}$$

Now, define $h(x) := (3\pi^2/4)\,[f(\pi x/4)]^2$, with f as in (3.15). We prove in Appendix A that 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 (3.17). It can be directly verified that ψ is non-negative with $\psi(t) = 0$ for $|t| \geq \pi/2$, and that ψ is twice differentiable, with $\psi''(0) = -48/\pi^2$.

Thus, rvs U and V can be constructed as in Theorem 7 with $var(U) = var(V) = 48/\pi^2$.

Remark 8. *It is even possible to construct compactly supported C^∞ characteristic functions. Constructions of such functions are given in [?]. In fact, [?] 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 [?, p. 512]). If such a characteristic function ψ is used in the construction described in Theorem 7, 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 $p(k) \leq C\lambda^{-|k|}$ for all sufficiently large $|k|$.

Proposition 9. *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 (3.9), we readily obtain

$$\varphi_{U|0}^2 = \varphi_{U|1}^2 \quad \text{and} \quad \varphi_{V|0}^2 = \varphi_{V|1}^2. \tag{3.18}$$

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

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

$t \in \mathbb{R}$ [?, 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 (3.18), 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).

3.5.4 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 7) 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 d dimensions, thereby obtaining a scheme that satisfies properties (S1)–(S3).

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 $N_1, N_2, \ldots, N_k$ [?, Theorem 2.14.1]. Here, $\mathbb{Z}_{N_j}$ denotes the group of integers modulo-N_j. Taking $\Lambda = \mathbb{Z}^d$ and $\Lambda_0 = \mathsf{A}^T\mathbb{Z}^d$, where A is the diagonal matrix $\mathrm{diag}(N_1, N_2, \ldots, N_k)$, 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 $\mathbb{R}^d$.

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

$$p_j(\mathbf{k}) = \begin{cases} |\det \Lambda_0| f(\mathbf{k}) & \text{if } \mathbf{k} \in \Lambda_j \\ 0 & \text{otherwise.} \end{cases} \tag{3.19}$$

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 $\Delta = \sum_{j=1}^d \partial_j^2$ is the Laplacian operator.

As with Theorem 7 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 5, in a manner completely analogous to Theorem 7. The interested reader is directed to Appendix B for the proof.

Constructing compactly supported twice-differentiable (or even C^∞) characteristic functions $\psi : \mathbb{R}^d \to \mathbb{R}^+$, $d \geq 1$, is straightforward, given our previous constructions of such functions from $\mathbb{R}$ to $\mathbb{R}^+$. Suppose that for $i = 1, 2, \ldots, d$, $\psi_i : \mathbb{R} \to \mathbb{R}^+$ 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 $X_1, X_2, \ldots, X_d$ are mutually independent. Then, $\psi(t_1, \ldots, t_d) = \prod_{i=1}^d \psi_i(t_i)$ is the characteristic function of the random vector $\mathbf{X} = (X_1, \ldots, X_d)$. Note that ψ is compactly supported: $\psi(\mathbf{t}) = 0$ for $\mathbf{t} \notin \prod_{i=1}^d (-\lambda_i, \lambda_i)$. 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., [**?**], [**?**] and Theo-

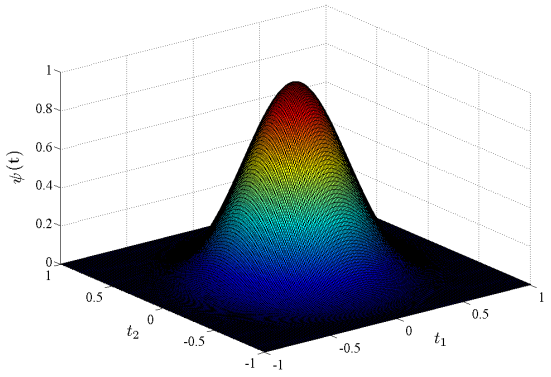

Figure 3.9: Example of a characteristic function supported within $\mathcal{V}(2\mathbb{Z}^2)$.

rem 11 below. A smooth, compactly supported characteristic function in $\mathbb{R}^2$ is depicted in Fig. 3.9.

Our objective is to design codes (as defined in Definition 1) for secure computation at the relay. With the construction described above, 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 $\mathcal{V}(\hat{\Lambda}_0)$ suffices for the construction of Theorem 10, 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 $\alpha\Lambda := \{\alpha\mathbf{z} : \mathbf{z} \in \Lambda\}$ 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 [**?**], 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 $\mathbb{R}^d$.

Theorem 11 ([**?**], Theorem 5.1)**.** *Fix a $\rho > 0$. If ψ is a characteristic function of a random vector distributed over $\mathbb{R}^d$ such that $\psi(\mathbf{t}) = 0$ for $\|\mathbf{t}\| \geq \rho$, then*

$$-\Delta\psi(\mathbf{0}) \geq \frac{4}{\rho^2}\, j^2_{\frac{d-2}{2}}, \tag{3.20}$$

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

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

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

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

and

$$\gamma_d^2 = \frac{4 j_{\frac{d-2}{2}}^{d-2}}{\pi^{d/2}\Gamma(d/2) J^2_{\frac{d}{2}}(j_{\frac{d-2}{2}})},$$

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

$$\tilde{f}(\mathbf{x}) = c_d \left(\frac{\Omega_d(\|\mathbf{x}\|/2)}{j^2_{\frac{d-2}{2}} - (\|\mathbf{x}\|/2)^2} \right)^2, \tag{3.21}$$

where

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

Remark 12. *Observe that Theorem 10 is true for any nested lattice pair* (Λ, Λ_0)*. As long as* $\psi(\mathbf{t})$ *is a characteristic function supported within* $\mathcal{V}(\hat{\Lambda}_0)$*, 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* $\mathcal{V}(\hat{\Lambda}_0)$*, then Theorem 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* $r_{\text{pack}}(\hat{\Lambda}_0)$*, where* $r_{\text{pack}}(\hat{\Lambda}_0)$ *denotes the packing radius of* $\hat{\Lambda}_0$*. 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 [***?***, ***?***, ***?****] 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., [***?***, ***?***, ***?****]) that have low decoding complexity.*

3.6 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 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 [**?**, **?**], and is described below.

3.6.1 Coding Scheme for Perfect Secrecy

We now describe the sequence of $(d, M^{(d)})$ (recall Definition 1) codes that achieve perfect secrecy.

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

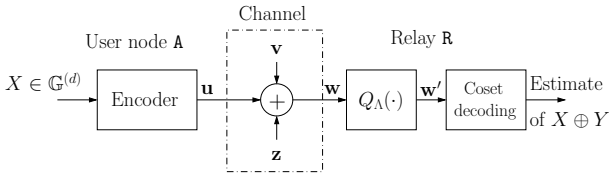

Figure 3.10: The operations performed by the user nodes and the relay.

Encoding: We have messages X, Y at nodes $\mathtt{A}, \mathtt{B}$ that are independent rvs, uniformly distributed over $\mathbb{G}^{(d)}$. We first pick a characteristic function ψ supported within $\mathcal{V}(\hat{\Lambda}_0^{(d)})$, as needed in Theorem 10. We impose the restriction that ψ be supported within a ball centered at $\mathbf{0}$ with radius equal to the packing radius, $r_{\text{pack}}(\hat{\Lambda}_0^{(d)})$, of the dual lattice $\hat{\Lambda}_0^{(d)}$. Recall that the packing radius is, by definition, the largest radius of a ball centered at $\mathbf{0}$ that is contained within $\mathcal{V}(\hat{\Lambda}_0^{(d)})$. So, if $\psi(\mathbf{t}) = 0$ for $\|\mathbf{t}\| \geq r_{\text{pack}}(\hat{\Lambda}_0)$, then $\psi(\mathbf{t})$ is certainly supported within $\mathcal{V}(\hat{\Lambda}_0)$. 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 10. 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 $R^{(d)} = \frac{1}{d}\log_2 M^{(d)}$. The average transmit power per dimension at each node is $P^{(d)} = \frac{-\Delta\psi(\mathbf{0})}{d}$, as in Theorem 10.

From Theorem 11, we see that an average transmit power per dimension as low as

$$P^{(d)} = \frac{4j_{\frac{d-2}{2}}^2}{d\left(r_{\text{pack}}(\hat{\Lambda}_0^{(n)})\right)^2}, \tag{3.22}$$

is achievable by a suitable choice of ψ. It was shown in [**?**] (see also [**?**]) 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 b is a constant independent of k. Therefore,

$$P^{(d)} = \frac{d}{{r_{\text{pack}}}^2(\hat{\Lambda}_0^{(d)})}(1 + o_d(1)), \tag{3.23}$$

where $o_d(1) \to 0$ as $d \to \infty$, is achievable by a suitable choice of ψ using Theorem 11.

Decoding: The relay $\mathtt{R}$ receives $\mathbf{W} = \mathbf{U} + \mathbf{V} + \mathbf{Z}$, where $\mathbf{Z}$ is a Gaussian noise vector with d 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 10 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 $\eta^{(d)}$ 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 $(\Lambda^{(d)}, \Lambda_0^{(d)})$ for which the following hold for sufficiently large d: $R^{(d)} > \mathcal{R} - \delta$, $P^{(d)} < \mathcal{P} + \delta$ and $\eta^{(d)} < \delta$.

For a given nested lattice pair, Theorem 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 $r_{\text{pack}}(\hat{\Lambda}_0^{(n)})$), 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[6]:

(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, $\{\hat{\Lambda}_0^{(n)}\}$, 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 [**?**, **?**, **?**, **?**] 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, $\{\hat{\Lambda}_0^{(n)}\}$ must be good for packing. While it is well established that there exist nested lattices satisfying (G_1) and (G_3) [**?**, **?**, **?**], it turns out that the duals of most of these lattices

[6]For definitions of lattices good for covering, packing, and AWGN channel coding, the reader is directed to Appendix C.

also satisfy the goodness properties. In the next section, we will formally describe an ensemble of lattices, also studied in [**?**, **?**], and show that most of the lattices in this ensemble satisfy all the above properties.

3.6.2 Good Ensembles of Nested Lattices with Good Duals

Our description of the construction of the $(\Lambda^{(n)}, \Lambda_0^{(n)})$ nested lattice codes is based on [**?**, **?**]. Let d and k be positive integers with $k \leq d$, and let q be a prime number. Let $\mathbb{Z}_q$ denote the field of integers modulo q. The (d, k, q) ensemble of lattices (in the terminology of [**?**]) is used in the construction. A lattice from the (d, k, q) ensemble is sampled as follows:

1. Choose a $k \times d$ matrix G with entries from $\mathbb{Z}_q$ uniformly at random. Note that G need not be full-rank. However, the probability that G is full-rank goes to 1 as $(d-k)$ tends to ∞ [**?**]. The linear code over $\mathbb{Z}_q$ generated by G is denoted by $\mathcal{C}(\mathsf{G}) = \{(\mathsf{G}^T\mathbf{y}) \bmod q : \mathbf{y} \in \mathbb{Z}_q^k\}$.

2. Apply Construction A on the code $\mathcal{C}(\mathsf{G})$. This is done as follows:

   (c_1) The codebook is scaled so that the scaled codewords lie within the d-dimensional unit cube: $\mathcal{C}' = (1/q)\mathcal{C}(\mathsf{G}) = \{(1/q)\mathbf{x} : \mathbf{x} \in \mathcal{C}(\mathsf{G})\}$.

   (c_2) The lattice is obtained by tessellating the entire space, $\mathbb{R}^d$, with copies of $\mathcal{C}'$, i.e., $\Lambda(\mathcal{C}) = \mathcal{C}' + \mathbb{Z}^d := \{\mathbf{c} + \mathbf{x} : \mathbf{c} \in \mathcal{C}', \mathbf{x} \in \mathbb{Z}^d\}$.

From the construction, it is clear that $\mathbb{Z}^d$ is a sublattice of $\Lambda(\mathcal{C})$. More detail regarding Construction-A lattices can be found in [**?**]. We would like to make note of one important property of these lattices: if the generator matrix of a Construction-A lattice Λ has rank d, then the effective radius of Λ is given by [**?**]

$$r_{\text{eff}}(\Lambda) = \left(\frac{\Gamma\left(\frac{d}{2} + 1\right)}{\pi^{d/2} q^k} \right)^{1/d}. \tag{3.24}$$

Choose a sequence of coarse lattices $\{\Lambda_0^{(n)}\}$, each $\Lambda_0^{(n)}$ selected uniformly at random from the (d, k, q) ensemble, where k and q may be functions of d chosen beforehand. For $d \in \{1, 2, 3, \ldots\}$, let $\mathsf{A}^{(d)}$ be the generator matrix of the coarse lattice $\Lambda_0^{(n)}$. 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, $\{\tilde{\Lambda}_f^{(d)}\}$, with each $\tilde{\Lambda}_f^{(d)}$ coming from the (d, k_1, q_1) ensemble of Construction-A lattices. As mentioned earlier, $\tilde{\Lambda}_f^{(d)}$ contains $\mathbb{Z}^d$ as a sublattice. If the generator matrix of $\tilde{\Lambda}_f^{(d)}$ has full rank, then the number of cosets of $\mathbb{Z}^d$ in $\tilde{\Lambda}_f^{(d)}$ is $q_1^{k_1}$.

(f_2) The lattice $\tilde{\Lambda}_f^{(d)}$ is subjected to a linear transformation by the matrix $(\mathsf{A}^{(d)})^T$, to get $\Lambda^{(n)} = (\mathsf{A}^{(d)})^T \tilde{\Lambda}_f^{(d)} := \{(\mathsf{A}^{(d)})^T \mathbf{y} : \mathbf{y} \in \tilde{\Lambda}_f^{(d)}\}$.

We will call this ensemble of $(\Lambda^{(n)}, \Lambda_0^{(n)})$ pairs as the (d, k, q, k_1, q_1) *ensemble*. The lattice pair can be scaled appropriately so as to satisfy the average power constraint. We have $M^{(d)} = |\Lambda^{(n)}/\Lambda_0^{(n)}| = q_1^{k_1}$ with probability tending to 1 as $d - k$ tends to ∞ [**?**]. Hence, the rate of the $(\Lambda^{(n)}, \Lambda_0^{(n)})$ code will be

$$R^{(d)} = \frac{k_1}{d} \log_2(q_1). \tag{3.25}$$

We choose

$$k = \beta_0 d, \quad \text{and} \quad k_1 = \beta_1 d, \tag{3.26}$$

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

$$\lim_{d\to\infty} \frac{d}{q_1} = 0, \quad \text{and} \quad r_{\min}^{(0)} < r_{\text{eff}}(\Lambda_0^{(n)}) < 2r_{\min}^{(0)}, \tag{3.27}$$

for some $0 < r_{\min}^{(0)} < 1/4$. It is possible to choose primes that satisfy the above conditions, and we direct the interested reader to [**?**] for the details. We then have the following lemma, which is proved in Appendix D.

Lemma 13. *Let $(\Lambda^{(n)}, \Lambda_0^{(n)})$ be a nested lattice pair chosen uniformly at random from the (d, k, q, k_1, q_1) ensemble, with the parameters k, q, k_1, q_1 chosen so as to satisfy (3.26) and (3.27). Then, the probability that $(\Lambda^{(n)}, \Lambda_0^{(n)})$ satisfies (G_1)–(G_3) tends to one as d approaches infinity.*

3.6.3 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 [**?**, **?**, **?**]. As defined in Section 3.6.1, 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 $Q_{\Lambda^{(n)}}([\mathbf{W}] \bmod \Lambda_0^{(n)})$.

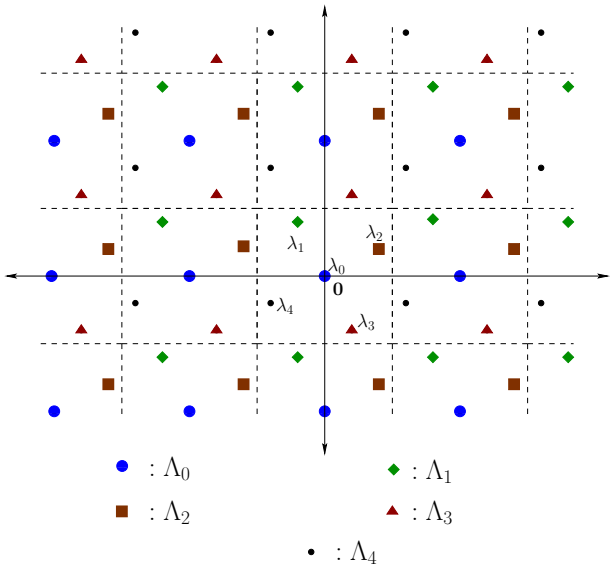

Figure 3.11: Different cosets of Λ_0 in Λ. The coset representative of Λ_j within $\mathcal{V}(\Lambda_0)$ is λ_j.

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. **??**. 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 $\widetilde{\mathbf{W}} = [\mathbf{W}] \bmod \Lambda_0^{(n)}$. 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 [**?**]:

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

A channel of the form $\mathbf{W} = [\mathbf{X} + \mathbf{N}] \bmod \Lambda_0^{(n)}$, where $\mathbf{N}$ denotes the

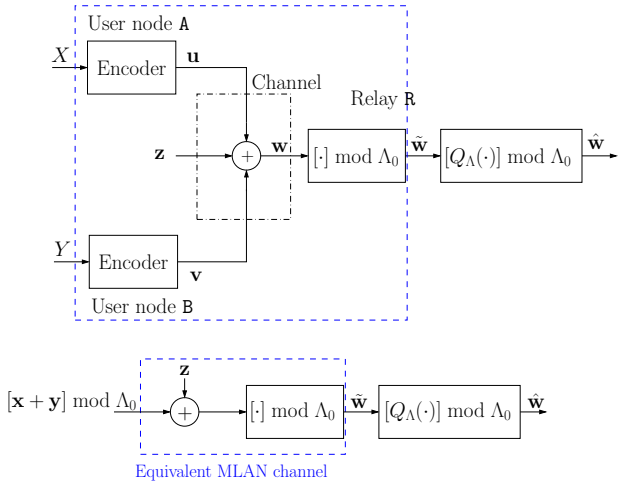

Figure 3.12: MAC phase of the bidirectional relay and equivalent MLAN channel representation.

noise vector, is called a $\Lambda_0^{(n)}$-modulo lattice additive noise ($\Lambda_0^{(n)}$-MLAN) channel [**?**]. 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. **??**. 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 (d, k, q, k_1, q_1) ensemble, where k, q, k_1 and q_1 satisfy (3.26) and (3.27). Using the coding scheme of Section 3.6.1, we can achieve perfect secrecy. The proposition below provides us with the means of determining the rates achievable with this coding scheme.

Proposition 14. *Let $M > 0$ 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 $r_{\text{eff}}(\Lambda_0^{(n)}) = \sqrt{dM}$. Then, using the coding scheme of Section 3.6.1 with this sequence of nested lattice pairs, any rate less than $\frac{1}{2}\log_2\left(\frac{M}{\sigma^2}\right)$ is achievable with perfect secrecy.*

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

3.6.4 Relating Achievable Rates to Transmit Power

From (3.23), we know that as long as the average transmit power per dimension is less than $\left(d/{r_{\text{pack}}}^2(\hat{\Lambda}_0^{(n)})\right)(1+o_d(1))$, we can guarantee perfect secrecy at the relay. From Proposition **??**, we see that as long as the transmission rate is less than $\frac{1}{2}\log_2({r_{\text{eff}}}^2(\Lambda_0^{(n)})/(d\sigma^2))$, 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 $\sqrt{d}$, i.e., $r_{\text{eff}}(\Lambda_0^{(n)}) = \Omega(\sqrt{d})$. Furthermore, to satisfy an average power constraint, we require $r_{\text{pack}}(\hat{\Lambda}_0^{(n)}) = \Omega(\sqrt{d})$. 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 $r_{\text{pack}}(\hat{\Lambda}_0^{(n)})$. Since we want to maximize the rate for a given power constraint, we would like both $r_{\text{eff}}(\Lambda_0^{(n)})$ and $r_{\text{pack}}(\hat{\Lambda}_0^{(n)})$ to be as large as possible. However, for any lattice $\Lambda_0^{(n)}$, we have $r_{\text{cov}}(\Lambda_0^{(n)})r_{\text{pack}}(\hat{\Lambda}_0^{(n)}) \leq \pi d$ [**?**, Theorem 18.3], and since $r_{\text{eff}}(\Lambda_0^{(n)}) \leq r_{\text{cov}}(\Lambda_0^{(n)})$, we get $r_{\text{eff}}(\Lambda_0^{(n)})r_{\text{pack}}(\hat{\Lambda}_0^{(n)}) \leq \pi d$. Hence, to obtain positive rates and at the same time satisfy the power constraint, both $r_{\text{eff}}(\Lambda_0^{(n)})$ and $r_{\text{pack}}(\hat{\Lambda}_0^{(n)})$ must grow roughly as $\sqrt{d}$. Therefore, we seek lattices satisfying properties (G_1)–(G_3), for which the product $r_{\text{eff}}(\Lambda_0^{(n)})r_{\text{pack}}(\hat{\Lambda}_0^{(n)})$ is close to the upper bound of πd.

For a sequence of Construction-A coarse lattices satisfying (G_1) and (G_2), we can find an asymptotic lower bound for $(1/d)r_{\text{eff}}(\Lambda_0^{(n)})r_{\text{pack}}(\hat{\Lambda}_0^{(n)})$,[7] as the following theorem shows.

Lemma 15. *Let $\{\Lambda_0^{(n)}\}$ be a sequence of coarse lattices, with each $\Lambda_0^{(n)}$ chosen from a (d,k,q) ensemble and k,q satisfying (3.26) and (3.27). If $\{\Lambda_0^{(n)}\}$ satisfies conditions (G_1)–(G_2), then,*

$$\lim_{d\to\infty} \frac{r_{\text{eff}}(\Lambda_0^{(n)})r_{\text{pack}}(\hat{\Lambda}_0^{(n)})}{d} \geq \frac{1}{2e}. \tag{3.28}$$

Proof. See Appendix E. □

[7]The product $r_{\text{eff}}(\Lambda_0^{(n)})r_{\text{pack}}(\hat{\Lambda}_0^{(n)})$ 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 $(1/\alpha)\hat{\Lambda}_0^{(n)}$.