Secure Compute-and-Forward in a Bidirectional Relay

Shashank Vatedka, *Student Member, IEEE,*
Navin Kashyap, *Senior Member, IEEE,* Andrew Thangaraj, *Senior Member, IEEE*

Abstract

We consider the basic bidirectional relaying problem, in which two users in a wireless network wish to exchange messages through an intermediate relay node. In the compute-and-forward strategy, the relay computes a function of the two messages using the naturally-occurring sum of symbols simultaneously transmitted by user nodes in a Gaussian multiple access (MAC) channel, and the computed function value is forwarded to the user nodes in an ensuing broadcast phase. In this paper, we study the problem under an additional security constraint, which requires that each user's message be kept secure from the relay. We consider two types of security constraints: perfect secrecy, in which the MAC channel output seen by the relay is independent of each user's message; and strong secrecy, which is a form of asymptotic independence. We propose a coding scheme based on nested lattices, the main feature of which is that given a pair of nested lattices that satisfy certain "goodness" properties, we can explicitly specify probability distributions for randomization at the encoders to achieve the desired security criteria. In particular, our coding scheme guarantees perfect or strong secrecy even in the absence of channel noise. The noise in the channel only affects reliability of computation at the relay, and for Gaussian noise, we derive achievable rates for reliable and secure computation. We also present an application of our methods to the multi-hop line network in which a source needs to transmit messages to a destination through a series of intermediate relays.

I. Introduction

Consider a network having three nodes, denoted by `A`, `B` and `R`, as shown in Fig. I. 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

S. Vatedka and N. Kashyap ({shashank,nkashyap}@ece.iisc.ernet.in) are with the Department of Electrical Communication Engineering, Indian Institute of Science, Bangalore, India.

A. Thangaraj (andrew@ee.iitm.ac.in) is with the Department of Electrical Engineering, Indian Institute of Technology, Madras, India.

This work was presented in part at ISIT 2012, Cambridge, Mass., USA, and at ISIT 2013, Istanbul, Turkey.

white Gaussian noise (AWGN) channels. Bidirectional relaying in such settings has been studied extensively in the recent literature [2], [25], [29], [36], [39].

We use the compute-and-forward framework proposed in [25], [36] 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. 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

$$W = U + V + Z, \tag{1}$$

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 $X \oplus Y$.

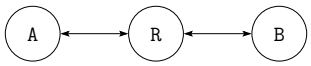

Fig. 1. Bidirectional relay.

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 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.

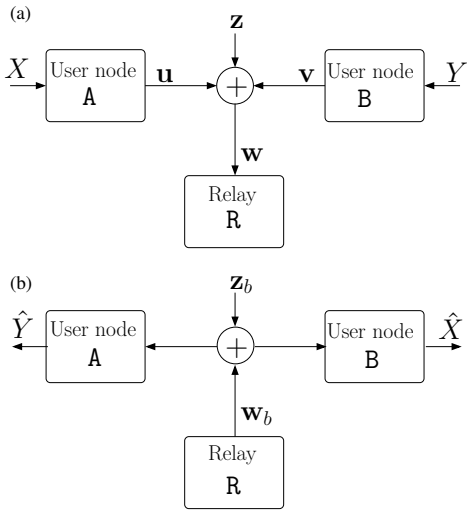

Fig. 2. Bidirectional relaying: (a) MAC phase, (b) Broadcast phase.

We propose a coding scheme for secure bidirectional relaying that uses a pair of nested lattices $(\Lambda^{(d)}, \Lambda_0^{(d)})$, with $\Lambda_0^{(d)} \subset \Lambda^{(d)}$. In our scheme, the messages are mapped to the cosets of the *coarse lattice* $\Lambda_0^{(d)}$ in the *fine lattice* $\Lambda^{(d)}$. 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 obtained by sampling and normalizing over Λ_j, a well-chosen density function f on $\mathbb{R}^d$. We will show that if the characteristic function of f is supported within the fundamental Voronoi region of the Fourier dual of $\Lambda_0^{(d)}$, 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* [24]. 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 [22] 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 [22] 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 [18], 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 [19], 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 [19]. 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., [4], [22], [28]). They have also been used in interference networks [1], and for secret key generation using correlated Gaussian sources [27].

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.

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

Organization of the paper

We establish some basic notation and recall some definitions related to lattices in Section II. We describe the secure bidirectional relaying problem in Section III, and then proceed to design coding schemes under the perfect secrecy constraint in Section IV. 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 V, then study the effect of additive noise and find achievable transmission rates in Section VI. Thereafter, we study the same problem under a strong secrecy constraint, design coding schemes, and evaluate the performance in Section VII, with the main result summarized in Theorem 16. In Section VIII, we show that these schemes can be extended to the multi-hop line network [18] and find achievable transmission rates under the two secrecy constraints. We make some concluding remarks in Section IX. Most of the technical proofs are given in appendices.

II. 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 $\mathbb{R}^+$ 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.

A. Lattices in $\mathbb{R}^d$

We briefly recall some definitions of lattices and their properties. For a more detailed treatment, see e.g., [3], [6].

Let k, d be positive integers with $k \le d$. 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, $\Lambda = \{\sum_{i=1}^k a_i \mathbf{u}_i : a_i \in \mathbb{Z}, 1 \le i \le k\}$, is called a k-dimensional *lattice* in $\mathbb{R}^d$. It is easy to verify that Λ forms an Abelian group under componentwise

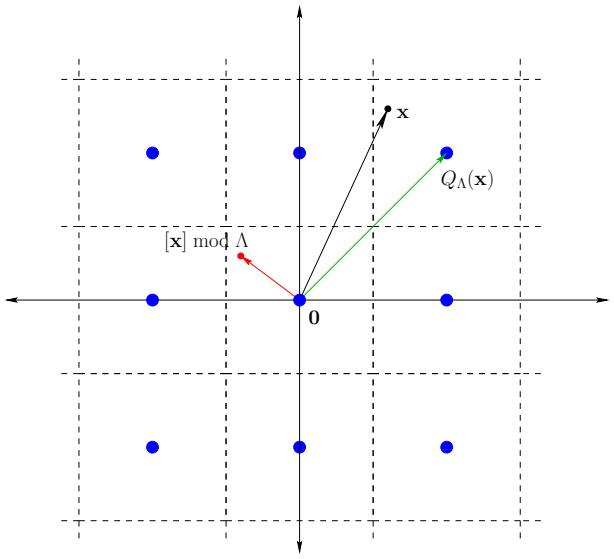

Fig. 3. Illustrating the $Q_\Lambda(.)$ and the $[.] \bmod \Lambda$ operation for the $\mathbb{Z}^2$ lattice.

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 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 $|\Lambda/\Lambda_0| = \det\Lambda_0/\det\Lambda$ [3, 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 $\{\mathbf{x} : \sum_{i=1}^d x_i y_i \in \mathbb{Z} \text{ for every } \mathbf{y} \in \Lambda\}$[3]. 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 $\det\Lambda$ [3], [6].

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 $Q_\Lambda(\cdot)$. Figure 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 $\mathbb{R}^d$

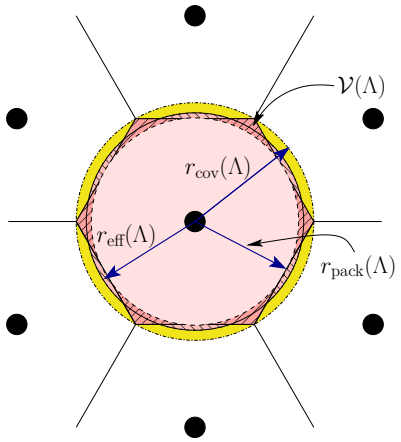

Fig. 4. Illustrating the covering, packing and effective radii of the hexagonal lattice.

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 $\mathbb{R}^d$ 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 $\mathbf{0}$ which 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. 4.

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

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

III. 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

$$\mathbf{W} = \mathbf{U} + \mathbf{V} + \mathbf{Z}, \tag{3}$$

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 $\mathbf{W}=\mathbf{U}+\mathbf{V}$ 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, $X\rightarrow\mathbf{U}+\mathbf{V}\rightarrow\mathbf{U}+\mathbf{V}+\mathbf{Z}$ 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) ***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*

$$R^{(d)}=\frac{\log_2 M^{(d)}}{d}. \tag{4}$$

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

$$P^{(d)}=\frac{1}{d}\mathbb{E}\|\mathbf{U}\|^2=\frac{1}{d}\mathbb{E}\|\mathbf{V}\|^2. \tag{5}$$

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

4) ***Decoder:*** *The relay* $\mathtt{R}$ *receives a vector* $\mathbf{W} \in \mathbb{R}^{(d)}$ *as given in (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*).*

IV. 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 $\mathtt{A}$ (resp. $\mathtt{B}$) 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.

V. 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 [17]), 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. ∎

A. 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).

1) 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

$$p_{UVXY}(k, l, a, b) = (1/2)(1/2)p_{U|a}(k)p_{V|b}(l) \tag{6}$$

for $k, l \in \mathbb{Z}$ and $a, b \in \{0, 1\}$. 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 $V = Y$. Property (S3) is satisfied by any U, V such that

$$\begin{aligned} p_{U|0}(k) = p_{V|0}(k) = 0 \quad &\text{for all odd } k \in \mathbb{Z}, \\ p_{U|1}(k) = p_{V|1}(k) = 0 \quad &\text{for all even } k \in \mathbb{Z}. \end{aligned} \tag{7}$$

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

$$p_U * p_V = p_{U|a} * p_V = p_U * p_{V|a} \quad \text{for } a \in \{0, 1\}. \tag{8}$$

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, (8) is equivalent to

$$\varphi_U\varphi_V = \varphi_{U|a}\varphi_V = \varphi_U\varphi_{V|a} \quad \text{for } a \in \{0, 1\}. \tag{9}$$

Note that $\varphi_U = (1/2)(\varphi_{U|0} + \varphi_{U|1})$ and $\varphi_V = (1/2)(\varphi_{V|0} + \varphi_{V|1})$. Hence, (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 (7) and (9). Then, the rvs U, V, X, Y with joint pmf given by (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

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

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 (9) must hold. The equality $\varphi_U\varphi_V = \varphi_U\varphi_{V|a}$ in (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, which we briefly recall here. Our description is based largely on Section XIX.5 in [17].

B. The Poisson Summation Formula

Fix a positive integer d, and let Λ be a full-rank lattice in $\mathbb{R}^d$. Recall from Section II-A 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 $\boldsymbol{\zeta} \in \mathbb{R}^d$,

$$\sum_{\mathbf{n}\in\hat{\Lambda}} \psi(\boldsymbol{\zeta} + \mathbf{n})\, e^{-i\langle \mathbf{n}, \mathbf{s}\rangle} = (\det \Lambda) \sum_{\mathbf{k}\in\Lambda} f(\mathbf{k} + \mathbf{s})\, e^{i\langle \mathbf{k}+\mathbf{s}, \boldsymbol{\zeta}\rangle}, \tag{10}$$

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., [34, Chapter VII, Corollary 2.6] or [3, Eq. (17.1.2)] — for (10) to hold. However, as argued by Feller in proving the formula in the one-dimensional setting [17, 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 (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 (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 (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 (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 rv supported within $\Lambda + \mathbf{s}$. In fact, by plugging in $\boldsymbol{\zeta} = \mathbf{0}$ in (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 variable supported within the set $\Lambda + \mathbf{s} := \{\mathbf{k} + \mathbf{s} : \mathbf{k} \in \Lambda\}$. The probability mass at the point $\mathbf{k} + \mathbf{s}$ 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., [17, Section XV.2, Table 1], [12], [31]. We also give an explicit construction in Example 1 in Section V-C.

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*

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

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 $Tf(kT + s)$.

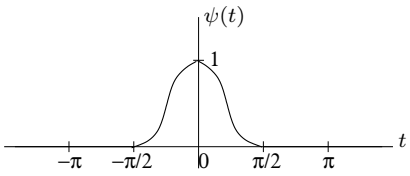

Fig. 5. A generic characteristic function supported on $[-\pi/2, \pi/2]$.

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

C. 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 Figure 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 Figure 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

$$p(k) = f(k) \text{ for all } k \in \mathbb{Z}. \tag{11}$$

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 Figure 7 for a generic ψ shown in Figure 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 Figure 7.

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

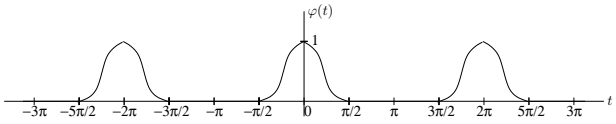

Fig. 6. Period-2π extension of generic ψ from Figure 5.

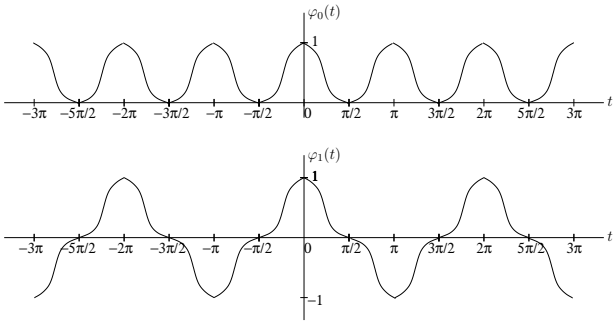

Fig. 7. The periodic functions φ_0 and φ_1 derived from ψ.

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

$$p_0(k) = \begin{cases} 2f(k) & \text{if } k \text{ is an even integer} \\ 0 & \text{otherwise.} \end{cases} \tag{12}$$

and that corresponding to φ_1 is

$$p_1(k) = \begin{cases} 2f(k) & \text{if } k \text{ is an odd integer} \\ 0 & \text{otherwise.} \end{cases} \tag{13}$$

From (11)–(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

$$\varphi^2 = \varphi\varphi_0 = \varphi\varphi_1. \tag{14}$$

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 (11). Clearly, (7) holds. To verify (9), note that, by virtue of (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 (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 [17, 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 (12) and 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 (12)) if the bit is 0, or with probability $p_1(k)$ (from (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, 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 [17, Section XV.2, Table 1])*

$$f(x) = \begin{cases} \frac{1}{2\pi} & \text{if } x = 0 \\ \frac{1-\cos x}{\pi x^2} & \text{if } x \neq 0 \end{cases} \tag{15}$$

which has characteristic function

$$\hat{f}(t) = \max\{0, 1 - |t|\} \tag{16}$$

The function $\hat{f}$ is plotted in Figure 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

$$g(t) = (\hat{f} * \hat{f})(t) = \begin{cases} \frac{1}{2}|t|^3 - t^2 + \frac{2}{3} & \text{if } |t| \leq 1 \\ \frac{1}{6}(2 - |t|)^3 & \text{if } 1 \leq |t| \leq 2 \\ 0 & \text{otherwise} \end{cases} \tag{17}$$

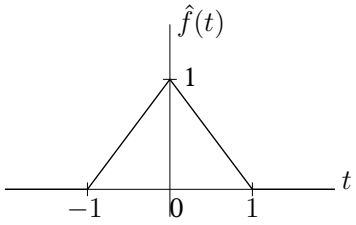

Fig. 8. $\hat{f}(t) = \max\{0, 1 - |t|\}$.

Now, define $h(x) := (3\pi^2/4)\,[f(\pi x/4)]^2$*, with* f *as in (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 (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 [31]. In fact, [31] 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 [17, 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 (9), we readily obtain

$$\varphi_{U|0}^2 = \varphi_{U|1}^2 \quad \text{and} \quad \varphi_{V|0}^2 = \varphi_{V|1}^2. \tag{18}$$

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

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 $t \in \mathbb{R}$ [23, 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 (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). ■

D. 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 $N_1, N_2, \ldots, N_k$ [20, 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 rvs U, V taking values in Λ, having

the properties (S1)–(S3). The following theorem shows that this is possible. Here, $\mathbb{R}^+$ denotes the set of all non-negative real numbers.

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:*

$$p_j(\mathbf{k}) = \begin{cases} |\det \Lambda_0| f(\mathbf{k}) & \text{if } \mathbf{k} \in \Lambda_j \\ 0 & \text{otherwise.} \end{cases} \tag{19}$$

Finally, define a random variable U (resp. V) jointly distributed with X (resp. Y) as follows: if $X = \Lambda_j$ (resp. $Y = \Lambda_j$), U (resp. V) is a random point from Λ_j picked according to the distribution p_j. Then, the resulting Λ-valued rvs U, V satisfy properties (S1)–(S3). Additionally, $\mathbb{E}\|U\|^2$ and $\mathbb{E}\|V\|^2$ are finite iff ψ is twice differentiable at $\mathbf{0}$, in which case $\mathbb{E}\|U\|^2 = \mathbb{E}\|V\|^2 = -\Delta\psi(\mathbf{0})$, 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., [12], [31] and Theorem 11 below. A smooth, compactly supported characteristic function in $\mathbb{R}^2$ is depicted in Figure 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 variables U and 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 variables $\mathbf{U}' = \alpha\mathbf{U}$ and $\mathbf{V}' = \alpha\mathbf{V}$, distributed over

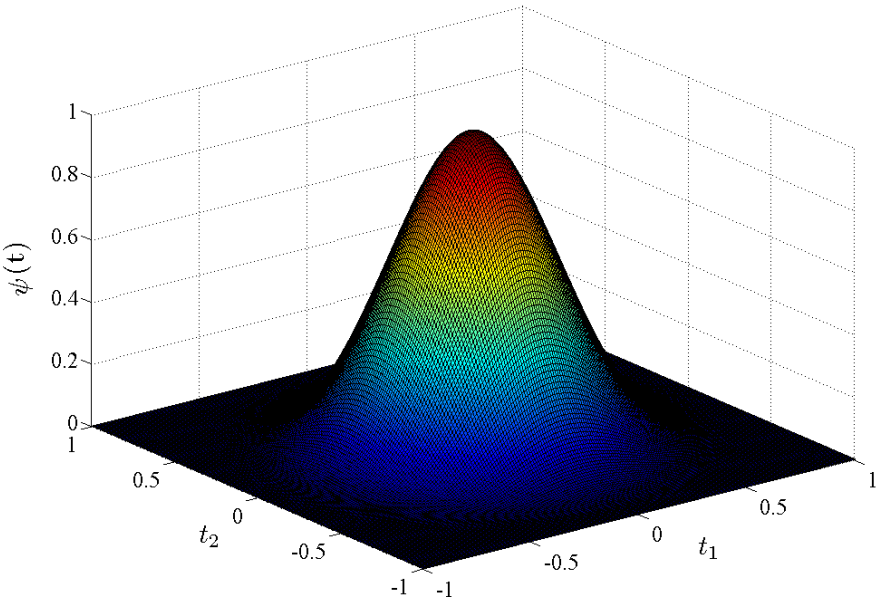

Fig. 9. Example of a characteristic function supported within $\mathcal{V}(2\mathbb{Z}^2)$.

$\alpha\Lambda := \{\alpha\mathbf{z} : \mathbf{z} \in \Lambda\}$ 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 [12], 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 ([12], Theorem 5.1)**.** *Fix a $\rho > 0$. If ψ is a characteristic function of a random variable distributed over $\mathbb{R}^d$ such that $\psi(\mathbf{t}) = 0$ for $\|\mathbf{t}\| \geq \rho$, then*

$$-\Delta\psi(\mathbf{0}) \geq \frac{4}{\rho^2}\, j^2_{\frac{d-2}{2}}, \tag{20}$$

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

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

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*

$$\tilde{f}(\mathbf{x}) = c_d \left(\frac{\Omega_d(\|\mathbf{x}\|/2)}{j_{\frac{d-2}{2}}^2 - (\|\mathbf{x}\|/2)^2} \right)^2, \tag{21}$$

where

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

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 [18], [19], [22] 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., [10], [32], [38]) that have low decoding complexity.*

VI. 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 [13], [25], and is described below.

A. Coding Scheme for Perfect Secrecy

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

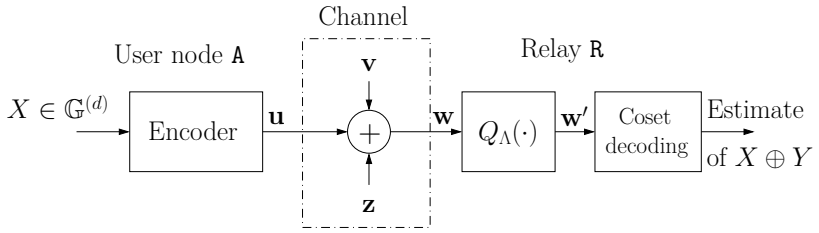

Fig. 10. The operations performed by the user nodes and the relay.

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 $\Lambda_0, \Lambda_1, \ldots, \Lambda_{M^{(d)}-1}$.

Encoding: We have messages X, Y at nodes A, 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 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 B transmits a random vector $\mathbf{V} \in \Lambda_k$ picked according to the distribution p_k. The rate of transmission from A or 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

$$P^{(d)} = \frac{4j_{\frac{d-2}{2}}^2}{d\left(r_{\text{pack}}(\hat{\Lambda}_0^{(d)})\right)^2}, \tag{22}$$

is achievable by a suitable choice of ψ. It was shown in [35] (see also [16]) 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,

$$P^{(d)} = \frac{d}{{r_{\text{pack}}}^2(\hat{\Lambda}_0^{(d)})}(1 + o_d(1)), \tag{23}$$

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

Decoding: The relay 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 $\Lambda_0^{(d)}$ represented by $Q_{\Lambda^{(d)}}(\mathbf{W})$, the closest vector to $\mathbf{W}$ in the lattice $\Lambda^{(d)}$. 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 $r_{\text{pack}}(\hat{\Lambda}_0^{(d)})$), 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 $(\Lambda^{(d)},\Lambda_0^{(d)})$ to those which satisfy the following "goodness" properties[6]:

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

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

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

Unlike prior work on nested lattices [1], [13], [25], [27] which only required $\{\Lambda_0^{(d)}\}$ and $\{\Lambda^{(d)}\}$ to satisfy properties (G_1) and (G_3) above, we have the additional requirement that the sequence of Fourier duals, $\{\hat{\Lambda}_0^{(d)}\}$ must be good for packing. While it is well established that there exist nested lattices satisfying (G_1) and (G_3) [13], [14], [25], it turns out that the duals of most of these lattices also satisfy the goodness properties. In the next section, we will formally describe an ensemble of lattices, also studied in [13], [25], and show that most of the lattices in this ensemble satisfy all the above properties.

B. Good Ensembles of Nested Lattices with Good Duals

Our description of the construction of the $(\Lambda^{(d)},\Lambda_0^{(d)})$ nested lattice codes is based on [13], [25]. Let d and k be positive integers with $k\le d$, 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 [14]) 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 ∞ [14]. 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: $\mathcal{C}'=(1/q)\mathcal{C}(\mathsf{G})=\{(1/q)\mathbf{x}:\mathbf{x}\in\mathcal{C}(\mathsf{G})\}$.

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

(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 [6]. 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 [14]

$$r_{\text{eff}}(\Lambda) = \left(\frac{\Gamma\left(\frac{d}{2}+1\right)}{\pi^{d/2} q^k} \right)^{1/d}. \tag{24}$$

Choose a sequence of coarse lattices $\{\Lambda_0^{(d)}\}$, each $\Lambda_0^{(d)}$ 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 $\Lambda_0^{(d)}$. For this choice of $\{\Lambda_0^{(d)}\}$, we construct another ensemble of lattices from which we pick the sequence of fine lattices $\{\Lambda^{(d)}\}$. 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 $\Lambda^{(d)} = (\mathsf{A}^{(d)})^T \tilde{\Lambda}_f^{(d)} := \{(\mathsf{A}^{(d)})^T \mathbf{y} : \mathbf{y} \in \tilde{\Lambda}_f^{(d)}\}$.

We will call this ensemble of $(\Lambda^{(d)}, \Lambda_0^{(d)})$ 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 $M^{(d)} = |\Lambda^{(d)}/\Lambda_0^{(d)}| = q_1^{k_1}$ with probability tending to 1 as $d - k$ tends to ∞ [25]. Hence, the rate of the $(\Lambda^{(d)}, \Lambda_0^{(d)})$ code will be

$$R^{(d)} = \frac{k_1}{d} \log_2(q_1). \tag{25}$$

We choose

$$k = \beta_0 d, \quad \text{and} \quad k_1 = \beta_1 d, \tag{26}$$

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

$$\lim_{d\to\infty} \frac{d}{q_1} = 0, \quad \text{and} \quad r_{\min}^{(0)} < r_{\text{eff}}(\Lambda_0^{(d)}) < 2 r_{\min}^{(0)}, \tag{27}$$

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 [14] for the details. We then have the following lemma, which is proved in Appendix D.

Lemma 13. *Let $(\Lambda^{(d)}, \Lambda_0^{(d)})$ 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 (26) and (27). Then, the probability that $(\Lambda^{(d)}, \Lambda_0^{(d)})$ satisfies (G_1)–(G_3) tends to one as d approaches infinity.*

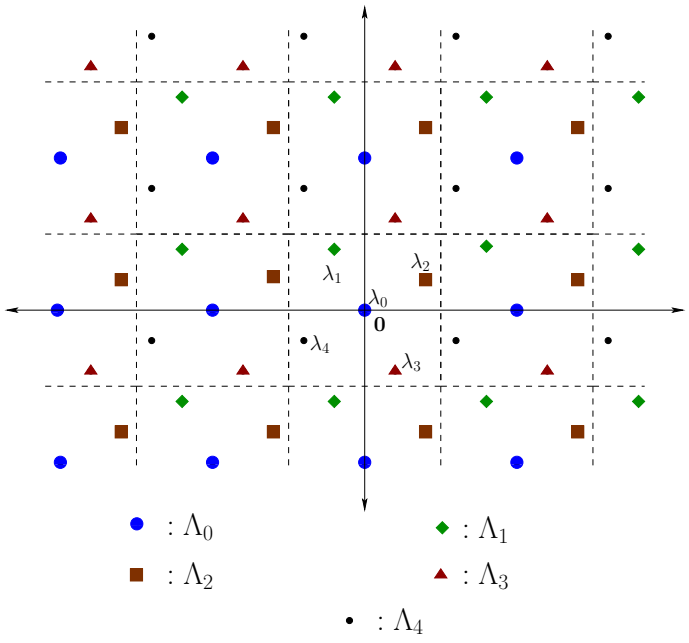

Fig. 11. Different cosets of Λ_0 in Λ. The coset representative of Λ_j within $\mathcal{V}(\Lambda_0)$ is λ_j.

C. 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 [13], [25], [26]. As defined in Section VI-A, 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 $\Lambda_0^{(d)}$ to which $Q_{\Lambda^{(d)}}(\mathbf{W})$ belongs. This is the same as the coset represented by $Q_{\Lambda^{(d)}}([\mathbf{W}] \bmod \Lambda_0^{(d)})$.

Each lattice point in $\Lambda^{(d)} \cap \mathcal{V}(\Lambda_0^{(d)})$ is a coset representative for a coset of $\Lambda_0^{(d)}$ in $\Lambda^{(d)}$. This is illustrated in Fig. 11. Suppose that Λ_j and Λ_k are the cosets which represent the messages X and Y, respectively. Let $\mathbf{X} = [\mathbf{U}] \bmod \Lambda_0^{(d)}$ and $\mathbf{Y} = [\mathbf{V}] \bmod \Lambda_0^{(d)}$ be the coset representatives of Λ_j and Λ_k, respectively. Then, $\Lambda_j \oplus \Lambda_k$ has $[\mathbf{X} + \mathbf{Y}] \bmod \Lambda_0^{(d)}$ as its representative. Therefore, the estimate $\mathcal{D}(\mathbf{W})$ has $\widehat{\mathbf{W}} = [Q_{\Lambda^{(d)}}(\mathbf{W})] \bmod \Lambda_0^{(d)}$ as its coset representative. This is equal to $\widehat{\mathbf{W}} = [Q_{\Lambda^{(d)}}([\mathbf{W}] \bmod \Lambda_0^{(d)})] \bmod \Lambda_0^{(d)}$. Let $\widetilde{\mathbf{W}} = [\mathbf{W}] \bmod \Lambda_0^{(d)}$. Then, $\widehat{\mathbf{W}} = [Q_{\Lambda^{(d)}}(\widetilde{\mathbf{W}})] \bmod \Lambda_0^{(d)}$. 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 [25]:

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

A channel of the form $\mathbf{W} = [\mathbf{X} + \mathbf{N}] \bmod \Lambda_0^{(d)}$, where $\mathbf{N}$ denotes the noise vector, is called a $\Lambda_0^{(d)}$-modulo lattice additive noise ($\Lambda_0^{(d)}$-MLAN) channel [13]. The random variable $\widetilde{\mathbf{W}}$ behaves like the output of a point-

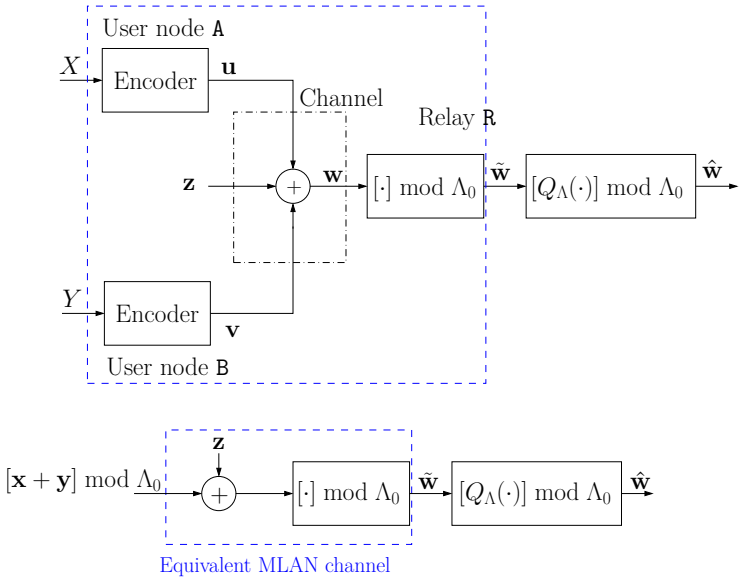

Fig. 12. MAC phase of the bidirectional relay and equivalent MLAN channel representation.

to-point transmission over a $\Lambda_0^{(d)}$-MLAN channel, with the transmitted vector being $[\mathbf{X}+\mathbf{Y}] \bmod \Lambda_0^{(d)}$. Looking from $\widetilde{\mathbf{W}}$, the "effective" channel is a $\Lambda_0^{(d)}$-MLAN channel, and the relay has to decode $[\mathbf{X}+\mathbf{Y}] \bmod \Lambda_0^{(d)}$ reliably from $\widetilde{\mathbf{W}}$. This is illustrated in Fig. 12. We will use the properties of the $\Lambda_0^{(d)}$-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 (26) and (27). Using the coding scheme of Section VI-A, 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 $\{\Lambda^{(d)}, \Lambda_0^{(d)}\}$ be a sequence of nested lattice pairs that satisfy (G_1)–(G_3), and scaled so as to satisfy $r_{\text{eff}}(\Lambda_0^{(d)}) = \sqrt{dM}$. Then, using the coding scheme of Section VI-A 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 [13, Theorem 4]; we omit the details.

D. Relating Achievable Rates to Transmit Power

From (23), we know that as long as the average transmit power per dimension is less than $\left(d/{r_{\text{pack}}}^2(\hat{\Lambda}_0^{(d)})\right)(1+o_d(1))$, we can guarantee perfect secrecy at the relay. From Proposition 14, we see that as long as the transmission rate is less than $\frac{1}{2}\log_2({r_{\text{eff}}}^2(\Lambda_0^{(d)})/(d\sigma^2))$, the relay can reliably compute $X \oplus Y$ from $\mathbf{W}$. In order to achieve positive rates, we need $r_{\text{eff}}(\Lambda_0^{(d)})$ to grow at least as fast as $\sqrt{d}$, i.e., $r_{\text{eff}}(\Lambda_0^{(d)}) = \Omega(\sqrt{d})$. Furthermore, to satisfy an average power constraint, we require $r_{\text{pack}}(\hat{\Lambda}_0^{(d)}) = \Omega(\sqrt{d})$. The rate is an

increasing function of $r_{\text{eff}}(\Lambda_0^{(d)})$, and the average transmit power per dimension is a decreasing function of $r_{\text{pack}}(\hat{\Lambda}_0^{(d)})$. Since we want to maximize the rate for a given power constraint, we would like both $r_{\text{eff}}(\Lambda_0^{(d)})$ and $r_{\text{pack}}(\hat{\Lambda}_0^{(d)})$ to be as large as possible. However, for any lattice $\Lambda_0^{(d)}$, we have $r_{\text{cov}}(\Lambda_0^{(d)})r_{\text{pack}}(\hat{\Lambda}_0^{(d)}) \leq \pi d$ [3, Theorem 18.3], and since $r_{\text{eff}}(\Lambda_0^{(d)}) \leq r_{\text{cov}}(\Lambda_0^{(d)})$, we get $r_{\text{eff}}(\Lambda_0^{(d)})r_{\text{pack}}(\hat{\Lambda}_0^{(d)}) \leq \pi d$. Hence, to obtain positive rates and at the same time satisfy the power constraint, both $r_{\text{eff}}(\Lambda_0^{(d)})$ and $r_{\text{pack}}(\hat{\Lambda}_0^{(d)})$ must grow roughly as $\sqrt{d}$. Therefore, we seek lattices satisfying properties (G_1)–(G_3), for which the product $r_{\text{eff}}(\Lambda_0^{(d)})r_{\text{pack}}(\hat{\Lambda}_0^{(d)})$ 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 $(1/d)r_{\text{eff}}(\Lambda_0^{(d)})r_{\text{pack}}(\hat{\Lambda}_0^{(d)})$,[7] as the following theorem shows.

Lemma 15. *Let $\{\Lambda_0^{(d)}\}$ be a sequence of coarse lattices, with each $\Lambda_0^{(d)}$ chosen from a (d,k,q) ensemble and k,q satisfying (26) and (27). If $\{\Lambda_0^{(d)}\}$ satisfies conditions (G_1)–(G_2), then,*

$$\lim_{d\to\infty} \frac{r_{\text{eff}}(\Lambda_0^{(d)})r_{\text{pack}}(\hat{\Lambda}_0^{(d)})}{d} \geq \frac{1}{2e}. \tag{28}$$

Proof: See Appendix E. ∎

E. Proof of Theorem 1

Let us choose $r_{\text{eff}}(\Lambda_0^{(d)}) = \frac{1}{2e}\sqrt{d\mathcal{P}}$, for a constant $\mathcal{P} > 4e^2\sigma^2$. Fix a $\delta > 0$. Using Lemma 15, we see that

$$r_{\text{pack}}(\hat{\Lambda}_0^{(d)}) \geq \frac{d}{2er_{\text{eff}}(\Lambda_0^{(d)})}(1 - o_d(1)) \geq \frac{\sqrt{d}}{\sqrt{\mathcal{P}}}(1 - o_d(1)). \tag{29}$$

From (23), we see that perfect secrecy can be achieved with an average power constraint as low as $P^{(d)} = \left(d/{r_{\text{pack}}}^2(\hat{\Lambda}_0^{(d)})\right)(1 + o_d(1))$. Combining this and (29), perfect secrecy can be achieved with an average transmission power,

$$P^{(d)} < \mathcal{P} + \delta \tag{30}$$

for all sufficiently large d. From Proposition 14, we have seen that the average probability of error can be made to go down to zero as long as

$$R^{(d)} < \mathcal{R} := \frac{1}{2}\log_2 \frac{\mathcal{P}}{(2e)^2\sigma^2}. \tag{31}$$

Therefore, for every $\delta > 0$, we can choose a sequence of nested lattice codes such that for all sufficiently large d, we have $R^{(d)} > \mathcal{R} - \delta$, $P^{(d)} < \mathcal{P} + \delta$ and $\eta^{(d)} < \delta$. 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 1. □

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

VII. Strong Secrecy

A natural question that arises is what happens if we replace f in Theorem 10 by a density function for which the support of the characteristic function goes beyond $\mathcal{V}(\hat{\Lambda}_0^{(d)})$. 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 $\rho > r_{\text{pack}}(\hat{\Lambda}_0^{(d)})$, and choose the characteristic function $\phi_{U|X=\mathbf{x}}(\mathbf{t}) = \sum_{\mathbf{n}\in\hat{\Lambda}_0^{(d)}} \psi(\mathbf{t}+\mathbf{n})e^{-i\langle \mathbf{n},\mathbf{x}\rangle}$. 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 (21), with the support of ψ chosen to be a ball of radius $\rho = \min\{r_{\text{eff}}(\hat{\Lambda}_0^{(d)}), 2r_{\text{pack}}(\hat{\Lambda}_0^{(d)})\}$.[8] Doing so would give us an improved rate of $\left[\frac{1}{2}\log_2 \frac{\mathcal{P}}{\sigma^2} - \log_2 e\right]^+$. However, for such a coding scheme, we are only able to show that the ℓ^2 norm of the difference between $p_{U+V,X}$ and $p_{U+V}p_X$ goes to zero as $d \to \infty$. Knowing only that the ℓ^2 norm of the difference between $p_{U+V,X}$ and $p_{U+V}p_X$ goes to zero as $d \to \infty$, 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 [22]. 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.

A. The Gaussian Density

We now introduce some notation that will be used in the sequel. Let Λ be a lattice in $\mathbb{R}^d$. For any $\mathbf{x} \in \mathbb{R}^d$, and any $\kappa > 0$, we define $g_{\kappa,\mathbf{x}}(\cdot)$ to be the Gaussian density with mean $\mathbf{x}$ and covariance matrix $\kappa^2\mathbf{I}_d$, i.e., $\forall \mathbf{z} \in \mathbb{R}^d$,

$$g_{\kappa,\mathbf{x}}(\mathbf{z}) := \frac{1}{(2\pi\kappa^2)^{d/2}} e^{-\frac{\|\mathbf{z}-\mathbf{x}\|^2}{2\kappa^2}}. \tag{32}$$

We also define

$$g_{\kappa,\mathbf{x}}(\Lambda) := \sum_{\lambda\in\Lambda} g_{\kappa,\mathbf{x}}(\lambda). \tag{33}$$

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.

B. Coding Scheme for Strong Secrecy

Code: Following Section VI-A, we use a $(\Lambda^{(d)}, \Lambda_0^{(d)})$ nested lattice code, with $\Lambda_0^{(d)} \subseteq \Lambda^{(d)}$. As before, the messages are chosen from $\mathbb{G}^{(d)} := \Lambda^{(d)}/\Lambda_0^{(d)}$, and $\oplus$ is the addition operation on $\mathbb{G}^{(d)}$. The $M^{(d)} := |\mathbb{G}^{(d)}|$ cosets of $\Lambda_0^{(d)}$ in $\Lambda^{(d)}$ are denoted by $\Lambda_0, \ldots, \Lambda_{M^{(d)}-1}$.

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

Encoding: For a coset Λ_j of $\Lambda_0^{(d)}$ in $\Lambda^{(d)}$, let λ_j denote its representative within $\mathcal{V}(\Lambda_0^{(d)})$ (see Fig. 11 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

$$p_j(\mathbf{u}) = \begin{cases} \frac{g_\kappa(\mathbf{u})}{g_{\kappa,-\lambda_j}(\Lambda_0^{(d)})} & \text{if } \mathbf{u} \in \Lambda_j, \\ \mathbf{0} & \text{otherwise.} \end{cases} \tag{34}$$

Decoding: The relay computes the closest point in $\Lambda^{(d)}$ to the linear minimum mean-squared error (MMSE) estimate of the received vector, as in [22], [13], [25], 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 $\widetilde{\mathbf{W}} = [\alpha^* \mathbf{W}] \bmod \Lambda_0^{(d)}$. The estimate of $X \oplus Y$, denoted by $\mathcal{D}(\mathbf{W})$, is then the coset to which $Q_{\Lambda^{(d)}}(\widetilde{\mathbf{W}})$ belongs.

Achievable power-rate pair: A power-rate pair of $(\mathcal{P}, \mathcal{R})$ is achievable if for every $\delta > 0$, there exists a sequence of $(\Lambda^{(d)}, \Lambda_0^{(d)})$ nested lattice codes such that for all sufficiently large d,

- the average transmit power per dimension is less than $\mathcal{P} + \delta$:
$$P^{(d)} := \frac{1}{d}\mathbb{E}\|\mathbf{U}\|^2 = \frac{1}{d}\mathbb{E}\|\mathbf{V}\|^2 < \mathcal{P} + \delta;$$
- the transmission rate is greater than $\mathcal{R} - \delta$:
$$R^{(d)} := \frac{1}{d}\log_2 M^{(d)} > \mathcal{R} - \delta;$$
- 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:
$$\mathcal{I}(X; \mathbf{U} + \mathbf{V}) = \mathcal{I}(Y; \mathbf{U} + \mathbf{V}) < \delta.$$

In the next two subsections, we will prove that

Theorem 16. *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 VII-B.

C. Strong Secrecy in the Absence of Noise

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

variation distance) between p_{U+V} and $p_{U+V|\mathbf{x}}(\cdot)$, defined as[9]

$$\mathbb{V}(p_{U+V}, p_{U+V|\mathbf{x}}) := \sum_{\mathbf{w}\in\Lambda^{(d)}} |p_{U+V}(\mathbf{w}) - p_{U+V|\mathbf{x}}(\mathbf{w})|, \tag{35}$$

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

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

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

Lemma 17 ([9], Lemma 1)**.** *For $|\mathbb{G}^{(d)}| \geq 4$, we have*

$$\mathcal{I}(X; \mathbf{U}+\mathbf{V}) \leq \overline{\mathbb{V}}\left(\log_2|\mathbb{G}^{(d)}| - \log_2(\overline{\mathbb{V}})\right). \tag{36}$$

Since $|\mathbb{G}^{(d)}|$ grows exponentially in d, it is sufficient to have $\overline{\mathbb{V}}$ going to zero as $o(1/d)$ for $\mathcal{I}(X; \mathbf{U}+\mathbf{V})$ to go to zero. We will in fact show that $\overline{\mathbb{V}}$ can be made to go to zero exponentially in d, which will guarantee that the mutual information also decays exponentially in d. In order to have $\overline{\mathbb{V}}$ going to zero exponentially in d, we will require the coarse and fine lattices to satisfy certain properties.

For any lattice Λ in $\mathbb{R}^d$, and any $\theta > 0$, the flatness factor $\epsilon_\Lambda(\theta)$ is defined as [22], [5]

$$\epsilon_\Lambda(\theta) := \frac{\max_{\mathbf{x}\in\mathcal{V}(\Lambda)} \left|\left(\sum_{\lambda\in\Lambda} g_{\theta,\lambda}(\mathbf{x})\right) - (1/\det\Lambda)\right|}{1/\det\Lambda}. \tag{37}$$

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 $\epsilon_\Lambda(a) \leq \epsilon_\Lambda(b)$ [22, Remark 2]. Following [22], we define a sequence of lattices $\{\Lambda^{(d)}\}$ to be *secrecy-good* if

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

It was shown in [22] that there exist lattices that are secrecy-good and also satisfy all the goodness properties described in Appendix C.

Let us choose κ in (34) 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 18. *If the sequence of nested lattice pairs $\{\Lambda^{(d)}, \Lambda_0^{(d)}\}$ satisfies $\epsilon^{(d)} := \epsilon_{\Lambda_0^{(d)}}(\sqrt{\mathcal{P}/2}) < 1/2$, then for every $\mathbf{x} \in \Lambda^{(d)} \cap \mathcal{V}(\Lambda_0^{(d)})$, we have*

$$\mathbb{V}(p_{U+V}, p_{U+V|\mathbf{x}}) \leq 216\epsilon^{(d)}. \tag{38}$$

[9]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 $\mathbb{V}(P_1, P_2) := \sup_{A\subseteq\mathcal{X}} |P_1(A) - P_2(A)|$. 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., [7, Section 11.6]). We have dropped the $\frac{1}{2}$ factor for simplicity.

A proof of the above theorem is given in Appendix F. 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 [22, Section V-B] tells us that if the flatness factor of the coarse lattice goes to zero as $d \to \infty$, then the average transmit power converges to $\mathcal{P}$.

Lemma 19. *If the flatness factor* $\epsilon_1 := \epsilon_{\Lambda_0^{(d)}}\left(\mathcal{P}\sqrt{1-1/(e\pi)}\right) < 1/2$*, then,*

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

Since $\sqrt{1-1/(e\pi)} > 1/\sqrt{2}$, it is sufficient to have (by monotonicity of the flatness factor) $\epsilon_{\Lambda_0^{(d)}}(\sqrt{\mathcal{P}/2}) \to 0$ to satisfy the power constraint for all sufficiently large d. From Theorem 18 and Lemma 17, we see that strong secrecy can be obtained in the noiseless scenario.

D. Strong Secrecy and Reliability of Decoding in the Presence of AWGN

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 $X \to (\mathbf{U}+\mathbf{V}) \to (\mathbf{U}+\mathbf{V}+\mathbf{Z})$ 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.

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

(L1) The sequence of coarse lattices, $\{\Lambda_0^{(d)}\}$, is good for covering, MSE quantization, and AWGN channel coding[10].

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

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

Using (44) in [22, Appendix II] and [22, Proposition 2], we can show that if Λ_0 is a lattice sampled uniformly at random from a (d,k,q) ensemble, where d,k,q satisfy (26) and (27), then for all sufficiently large d, we have $\mathbb{E}[\epsilon_{\Lambda_0}(\theta)] \leq 2\left(\frac{(\det(\Lambda_0))^{2/d}}{2\pi\theta^2}\right)^{d/2}$, which goes to zero exponentially in d as long as $\frac{(\det(\Lambda_0))^{2/d}}{2\pi\theta^2} < 1$. 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 d. From Lemma 13, we know that a randomly chosen nested lattice pair satisfies (L1) and (L3) with probability tending to 1 as $d \to \infty$. We can then use the union bound to conclude that a randomly chosen pair of nested lattices from the (d,k,q,k_1,q_1) ensemble satisfies (L1)–(L3) with probability at least $1/2$ as $d \to \infty$.

[10]For the definitions of lattices good for covering, MSE quantization, and AWGN channel coding, see Appendix C.

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

$$\begin{aligned}
\widetilde{\mathbf{W}} &= [\alpha^*(\mathbf{U}+\mathbf{V}+\mathbf{Z})] \bmod \Lambda_0^{(d)} \\
&= [\mathbf{U}+\mathbf{V}-(1-\alpha^*)(\mathbf{U}+\mathbf{V})+\alpha^*\mathbf{Z}] \bmod \Lambda_0^{(d)} \\
&= \Big[[\mathbf{X}+\mathbf{Y}] \bmod \Lambda_0^{(d)} + \mathbf{Z}'\Big] \bmod \Lambda_0^{(d)}, \qquad (39)
\end{aligned}$$

where $\mathbf{Z}' = (\alpha^*-1)(\mathbf{U}+\mathbf{V})+\alpha^*\mathbf{Z}$ is the effective noise of the MLAN channel. Unlike in Section VI-C, $\mathbf{Z}'$ is not statistically independent of $[\mathbf{X}+\mathbf{Y}] \bmod \Lambda_0^{(d)}$. However, as shown by the following lemma, 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 $\big(2(1-\alpha^*)^2\mathcal{P}+(\alpha^*)^2\sigma^2\big)I_d$. Given two density functions f_1 and f_2 over $\mathbb{R}^d$, the variational distance between f_1 and f_2, denoted by $\mathbb{V}(f_1,f_2)$, is defined as

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

Then, we have the following lemma proved in Appendix G.

Lemma 20. *If $\epsilon_{\Lambda_0^{(d)}}(\sqrt{\alpha^*\mathcal{P}}) < 1/2$, then for every $\mathbf{x}$ and $\mathbf{y}$ in $\mathbb{G}^{(d)}$,*

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

1) Proof of Theorem 16: If $\mathbb{P}_1$ and $\mathbb{P}_2$ are probability measures on $\mathbb{R}^d$ having densities f_1 and f_2 respectively, then $\sup_{A\subset\mathbb{R}^d} |\mathbb{P}_1(A)-\mathbb{P}_2(A)| = \frac{1}{2}\mathbb{V}(f_1,f_2)$, where the supremum is taken over all measurable subsets of $\mathbb{R}^d$ (assuming that both P_1 and P_2 are defined on a common event space) [11, Section 7.7]. Using this and Lemma 20, the probability of error of the decoder can be bounded by

$$\begin{aligned}
\eta^{(d)} &\le \Pr\Big[\mathbf{Z}' \notin \mathcal{V}(\Lambda^{(d)})\Big] \\
&\le \Pr\Big[\mathbf{N} \notin \mathcal{V}(\Lambda^{(d)})\Big] + 4\epsilon_{\Lambda_0^{(d)}}(\sqrt{\alpha^*\mathcal{P}}). \qquad (40)
\end{aligned}$$

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 $\epsilon_{\Lambda_0^{(d)}}(\sqrt{\alpha^*\mathcal{P}}) \to 0$ as $d\to\infty$, and the fine lattices are good for AWGN channel coding, then the probability of error at the relay goes to zero as long as $\frac{(\det(\Lambda^{(d)}))^{2/d}}{2\pi e\sigma_N^2} > 1$, or equivalently,

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

In other words,

$$R^{(d)} = \frac{1}{d}\log_2|\mathbb{G}^{(d)}| < \frac{1}{2}\log_2\left(\frac{(\det(\Lambda_0^{(d)}))^{2/d}}{2\pi e\sigma_N^2}\right). \qquad (41)$$

If we have $\alpha^* \geq 1/2$, then by monotonicity of the flatness factor, $\epsilon_{\Lambda_0^{(d)}}(\sqrt{\alpha^*\mathcal{P}}) \leq \epsilon_{\Lambda_0^{(d)}}(\sqrt{\mathcal{P}/2})$. This requires $\frac{2\mathcal{P}}{2\mathcal{P}+\sigma^2} \geq 1/2$, or $\mathcal{P} \geq \sigma^2/2$. Observe that having $\epsilon_{\Lambda_0^{(d)}}(\sqrt{\mathcal{P}/2}) \to 0$ has three important consequences: (a) strong secrecy, even in the absence of noise (Theorem 18); (b) the average transmit power converges to $\mathcal{P}$ (Lemma 19); and (c) the effective noise vector is "almost" independent of the message (Lemma 20).

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

$$\frac{\left(\det(\Lambda_0^{(d)})\right)^{2/d}}{2\pi(\mathcal{P}/2)} < 1. \tag{42}$$

Let us choose $\left(\det(\Lambda_0^{(d)})\right)^{2/d} = \pi\mathcal{P} - \delta$, for some arbitrary $\delta > 0$, so as to satisfy (42). Substituting this in (41), we get that for $\mathcal{P} \geq \sigma^2/2$, as long as

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

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 $d \to \infty$. Substituting for σ_N^2, we complete the proof of Theorem 16.

□

Remark 21. *In the perfect secrecy setting, we were not able to show that the technique of MMSE scaling can be used to obtain an additional* $1/2$ *in the rate expression. As in the strong-secrecy case, suppose that the relay computes* $\widetilde{\mathbf{W}} := [\alpha^*\mathbf{W}] \bmod \Lambda_0^{(d)}$*, where* $\alpha^* := (2\mathcal{P})/(2\mathcal{P}+\sigma^2)$*. The effective noise vector,* $\mathbf{Z}_{\mathit{eff}} = -(1-\alpha^*)(\mathbf{U}+\mathbf{V})+\alpha^*\mathbf{Z}$ *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* $\mathbf{Z}_{\mathit{eff}} \notin \mathcal{V}(\Lambda^{(d)})$*, 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.*

E. Prior Work on Strong Secrecy

The strongly secure scheme proposed by He and Yener in [19] 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 d. Unlike [19], 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 d. But unlike [19], 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 [19] was coupled with

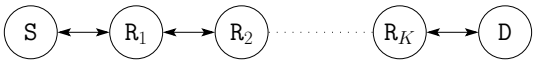

Fig. 13. Multi-hop line network with $K+1$ hops.

an Algebraic Manipulation Detection (AMD) code [8] 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 d. 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 [19].

VIII. 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 [18]. The structure of a multi-hop line network with $K+1$ hops is shown in Fig. 13. 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 [18], 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.

He and Yener showed that their scheme [18] 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.[11]

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 d channel uses. Let us choose a sequence of $(\Lambda^{(d)}, \Lambda_0^{(d)})$ nested lattice pairs that satisfy properties (L1)–(L3). Each node in the network employs the encoding and decoding scheme described in Section VII-B. Let $\mathcal{D} : \mathbb{R}^d \to \mathbb{G}^{(d)}$ denote

[11] 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.

the decoder map of Section VII-B. Also, for any $X \in \mathbb{G}^{(d)}$, let $\mathcal{E}(X)$ denote the encoded form of X as in Section VII-B.

- Each relay node i ($i = 1, \ldots, K$) generates a *jamming signal,* J_i, which is chosen uniformly at random from $\mathbb{G}^{(d)}$, 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 d-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: $\frac{1}{d}\mathbb{E}\|\mathbf{V}_i[n]\|^2 \leq P^{(d)}$ for $i = 0, 1, \ldots, K+1$ and $n = 1, 2, \ldots, K + 2N$.

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

$$R_N^{(d)} := \frac{N}{d(K+2N)} \log_2 |\mathbb{G}^{(d)}|. \tag{43}$$

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 for every $\delta > 0$, there exists a sequence of $(\Lambda^{(d)}, \Lambda_0^{(d)})$ nested lattice codes such that for all sufficiently large d, we have

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

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

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 22. *A power-rate pair of*

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

is achievable with strong secrecy[12]*, 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.

[12]If the scheme in [19] is used at each node, then the achievable rate with strong secrecy can be improved to $\left[\frac{1}{4}\log_2\left(\frac{1}{2} + \frac{\mathcal{P}}{\sigma^2}\right) - \frac{1}{2}\right]^+$.

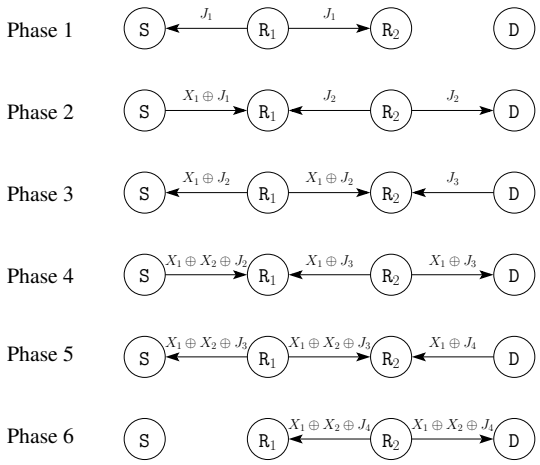

Fig. 14. Secure relaying of two messages in a 3-hop relay network.

Phase	Messages available at node at the end of phase			
	S	R_1	R_2	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 I

MESSAGES AVAILABLE AT VARIOUS NODES AT THE END OF EACH PHASE FOR THE PROTOCOL IN FIG. 14.

2) 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 [18]. 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. 14. Only the messages (elements of $\mathbb{G}^{(d)}$) 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 I. 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 $t = 0, 1, \ldots, N$.
- In the $(2t+i)$th phase ($t = 0, 1, 2, \ldots, N$), the ith node sends

$$\mathbf{V}_i[2t+i] = \mathcal{E}\big((\oplus_{p=1}^{t} X_p) \oplus J_{i+t}\big). \tag{44}$$

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.

3) Secrecy: Let us assume that all links are noiseless. As argued at the end of Section VII-C, 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

$$\mathbf{W}_k[2t+k-1] = \mathbf{V}_{k-1}[2t+k-1] + \mathbf{V}_{k+1}[2t+k-1] \tag{45}$$

$$= \mathcal{E}\Big((\oplus_{p=1}^{t} X_p) \oplus J_{k+t-1}\Big) + \mathcal{E}\Big((\oplus_{p=1}^{t-1} X_p) \oplus J_{k+t}\Big), \tag{46}$$

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

$$\Theta_{k,t} := \{J_k, J_{k-1}, \mathbf{W}_k[2m+k-1] : 1 \leq m \leq t\} \tag{47}$$

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 $\Theta_{k,0} := \{J_k, J_{k-1}\}$. 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 23. *Let $\epsilon^{(d)} := \epsilon_{\Lambda_0^{(d)}}(\sqrt{\mathcal{P}/2}) < 1/2$. Then, the total information available at the kth relay node at the end of all relaying phases can be bounded from above as follows:*

$$\mathcal{I}(X_1, \ldots, X_N; \Theta_{k,N}) \leq N\epsilon^{(d)} \left(\log_2 |\mathbb{G}^{(d)}| - \log_2 \epsilon^{(d)}\right). \tag{48}$$

Proof: See Appendix H. ∎

Since for our choice of nested lattices, $\epsilon^{(d)} \to 0$ exponentially in d, the mutual information $\mathcal{I}(X_1, \ldots, X_N; \Theta_{k,N})$ also goes to zero exponentially in d, thereby guaranteeing strong secrecy.

4) Achievable Rate and Proof of Theorem 22: 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 16, we can say that a power-rate pair of $\left(\mathcal{P}, \frac{N}{2(K+2N+1)}\left[\log_2\left(\frac{1}{2} + \frac{\mathcal{P}}{\sigma^2}\right) - \log_2 2e\right]^+\right)$ 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 22. The second part of the theorem can be proved in a similar manner.

IX. Conclusion

We have described two coding schemes for secure bidirectional relaying in presence of an honest-but-curious relay. We saw that using pmfs generated from density functions having compactly supported characteristic

functions, one can obtain perfect secrecy. We showed that reliable and perfectly secure computation at the relay is possible at transmission rates below $\left[\frac{1}{2}\log_2\frac{\mathcal{P}}{\sigma^2}-\log_2 2e\right]^+$. This is the first such result for perfect secrecy in the context of the bidirectional relay. In order to achieve higher transmission rates, we relaxed the secrecy constraint, and only required that the mutual information between $\mathbf{U}+\mathbf{V}$ and each individual message goes to zero for large block lengths. Using pmfs obtained from sampled Gaussian functions, we could achieve a rate of $\left[\frac{1}{2}\log_2\left(\frac{1}{2}+\frac{\mathcal{P}}{\sigma^2}\right)-\frac{1}{2}\log_2 2e\right]^+$. Prior work by He and Yener 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 with strong secrecy. These rates are within a constant gap of the best known achievable rate of $\left[\frac{1}{2}\log_2\left(\frac{1}{2}+\frac{\mathcal{P}}{\sigma^2}\right)\right]^+$ without secrecy constraints [25], [36].

The main theme of this paper was the use of nested lattice codes, and explicit pmfs having infinite support to obtain security. An inherent disadvantage of our scheme is that it is not possible to satisfy a maximum power constraint. One could study the scenario where the support of the distributions we described are truncated, and find the performance of such a scheme; we are yet to carry out this study.

All our results were derived under the assumptions that the messages are uniformly distributed, the channel gains from the user nodes to the relay are equal, and transmissions from both users are synchronized to arrive at the relay at the same time. Of course, in practice, these assumptions need not hold. Unfortunately, perfect secrecy does not appear to be robust to deviations from these assumptions. Indeed, if X and Y are not uniformly distributed, then we no longer have $(X\oplus Y)\perp\!\!\!\perp X$ and $(X\oplus Y)\perp\!\!\!\perp Y$. In general, if the channel gains are not equal and unknown at the user nodes, it is hard to get perfect secrecy. It can be shown that if $\mathbf{u},\mathbf{v}\in\Lambda^{(d)}$, and h_1,h_2 are real numbers such that h_1/h_2 is irrational, then it is possible to exactly recover $(\mathbf{u},\mathbf{v})$ from $h_1\mathbf{u}+h_2\mathbf{v}$. However, it may be possible to obtain strong secrecy even when some of these assumptions do not hold, but this is left as future work. But it is worth noting that our scheme guarantees perfect (strong) security even in the absence of noise, and hence it achieves perfect (strong) secrecy even when the distribution of the additive noise is arbitrary and unknown, as long as it is independent of the transmitted codewords.

The nested lattice coding schemes analyzed in this paper rely upon closest lattice point decoding, which is known to be computationally hard in general. However, recall that our randomization scheme for perfect secrecy works with *any* pair of nested lattices. In particular, it would work with nested lattice pairs on which practical coding schemes can be based, where by "practical coding schemes" we mean explicitly constructed nested lattice codes that admit reliable decoding with low computational complexity. Lattice coding schemes with low-complexity decoders have been studied in the literature, e.g., [10], [15], [32], [33], [38]. Our scheme for strong secrecy, on the other hand, requires that the nested lattices satisfy various goodness properties. Further investigation is needed to determine whether all these goodness properties can be found in lattices that admit low-complexity decoding.

Finally, in this paper, we only found achievable rates for secure and reliable computation at the relay. As remarked in [19], finding a converse result is much harder. Even without any secrecy constraints, a nontrivial outer bound on the capacity of a bidirectional relay is not known.

Appendix A: Technical Details of Example 1

We show here that the function $h(x) = (3\pi^2/4)\,[f(\pi x/4)]^2$, with f as in (15), is a density function whose characteristic function is given by

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

where g is as in (17).

Note first that $\hat{f}$ defined in (16) 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, $g = \hat{f} * \hat{f}$ is a density with characteristic function $4\pi^2 f^2$.

Now, f^2 is integrable since $(\hat{f})^2$ is integrable (see corollary to Theorem 3 of Section XV.3 of [17]). 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:

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

the last equality following from (17). 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 $\psi(t)$.

Appendix B: Proof of Theorem 10

We are given an index-M sublattice Λ_0 of the lattice Λ. Recall from Section II-A 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 variable U (resp. V) jointly distributed with X (resp. Y) as follows: if $X = \Lambda_j$ (resp. $Y = \Lambda_j$), U (resp. V) is a random point from Λ_j picked according to the distribution p_j. Then, U and V are identically distributed with $p_U = p_V = \frac{1}{M}\sum_{i=0}^{M-1} p_i$. Let φ_U, φ_V and φ_j, $j = 0, 1, \ldots, M-1$, be the characteristic functions corresponding to p_U, p_V and p_j, $j = 0, 1, \ldots, M-1$, respectively. We have the following straightforward generalization of Lemma 3.

Lemma 24. *Suppose that $\varphi_U \varphi_V = \varphi_j \varphi_V = \varphi_U \varphi_j$ for $j = 0, 1, \ldots, M-1$. Then, the random variables (U, V, X, Y) with joint pmf given by*

$$p_{UVXY}(\mathbf{k}, \mathbf{l}, \Lambda_i, \Lambda_j) = (1/M)(1/M) p_i(\mathbf{k}) p_j(\mathbf{l}) \quad \text{for } \mathbf{k}, \mathbf{l} \in \Lambda \text{ and } \Lambda_i, \Lambda_j \in \mathbb{G} \tag{49}$$

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 10. The function f can be retrieved from ψ by Fourier inversion:

$$\begin{aligned} f(\mathbf{x}) &= \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} \psi(\mathbf{t}) e^{-i\langle \mathbf{t}, \mathbf{x}\rangle}\, d\mathbf{t} \\ &= \frac{1}{(2\pi)^d} \int_{\mathcal{V}(\hat{\Lambda}_0)} \psi(\mathbf{t}) e^{-i\langle \mathbf{t}, \mathbf{x}\rangle}\, d\mathbf{t}. \end{aligned} \tag{50}$$

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

$$\varphi_j(\boldsymbol{\zeta}) = \sum_{\mathbf{n} \in \hat{\Lambda}_0} \psi(\boldsymbol{\zeta} + \mathbf{n})\, e^{-i\langle \mathbf{n},\, \mathbf{u}_j\rangle} \tag{51}$$

for all $\boldsymbol{\zeta} \in \mathbb{R}^d$. Then, by Proposition 5, we have that p_j is supported within Λ_j, and

$$p_j(\mathbf{k}) = (\det \Lambda_0)\, f(\mathbf{k}) \text{ for all } \mathbf{k} \in \Lambda_j. \tag{52}$$

Finally, define

$$\varphi(\boldsymbol{\zeta}) = \sum_{\mathbf{n} \in \hat{\Lambda}} \psi(\boldsymbol{\zeta} + \mathbf{n}) \tag{53}$$

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

We make two claims:

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

(ii) $\varphi = \varphi_U = \varphi_V$.

Given these claims, by Lemma 24, the random variables U, V satisfy the properties (S1)–(S3).

Both claims follow from the fact that $\hat{\Lambda}$ is a sublattice of $\hat{\Lambda}_0$. (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 (53) as

$$\varphi(\boldsymbol{\zeta}) = \sum_{\mathbf{n} \in \hat{\Lambda}} \psi(\boldsymbol{\zeta} + \mathbf{n})\, e^{-i\langle \mathbf{n},\, \mathbf{u}_j\rangle}. \tag{54}$$

This is possible because, for $\mathbf{n} \in \hat{\Lambda} = 2\pi\Lambda^*$ and $\mathbf{u}_j \in \Lambda$, we have $e^{-i\langle \mathbf{n},\, \mathbf{u}_j\rangle} = 1$. Comparing (51) and (54), and noting that ψ is supported within $\mathcal{V}(\hat{\Lambda}_0)$, it is evident that $\mathrm{supp}(\varphi) := \{\boldsymbol{\zeta} : \varphi(\boldsymbol{\zeta}) \neq 0\}$ is contained in $\mathrm{supp}(\varphi_j) := \{\boldsymbol{\zeta} : \varphi_j(\boldsymbol{\zeta}) \neq 0\}$. Furthermore, for all $\boldsymbol{\zeta} \in \mathrm{supp}(\varphi)$, we have $\varphi(\boldsymbol{\zeta}) = \varphi_j(\boldsymbol{\zeta})$. Claim (i) directly follows from this.

For Claim (ii), we note that $\mathcal{V}(\hat{\Lambda}_0) \subseteq \mathcal{V}(\hat{\Lambda})$, since $\hat{\Lambda}$ is a sublattice of $\hat{\Lambda}_0$. Hence, we can apply Proposition 5 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 (52) 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, $p = p_U = p_V$, which proves Claim (ii).

It remains to prove the statements concerning finiteness of $\mathbb{E}\|U\|^2$ and $\mathbb{E}\|V\|^2$. Theorem 1 in [37] 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 (53), 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 $\mathbb{E}\|U\|^2 = \mathbb{E}\|V\|^2 = -\Delta\psi(\mathbf{0})$. 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 $\mathbf{U} = (U_1, \ldots, U_d)$, so that $\|\mathbf{U}\|^2 = U_1^2 + \cdots + U_d^2$. We want to show that $\mathbb{E}[U_j^2] = -\frac{\partial^2}{\partial t_j^2}\psi(\mathbf{0})$, for $j = 1, \ldots, d$. For notational simplicity, we show this for $j = 1$. Note that the characteristic function of U_1 is given by $\varphi_{U_1}(t_1) = \varphi_U(t_1, 0, \ldots, 0)$. As argued prior to the statement of Theorem 7 in Section V-C, $\mathbb{E}[U_1^2] = -\varphi''_{U_1}(0)$. Now, $\varphi''_{U_1}(0) = \frac{\partial^2}{\partial t_1^2}\varphi_U(0, 0, \ldots, 0)$. From (53), we have that $\varphi_U = \psi$ in a small neighbourhood around $\mathbf{0} = (0, 0, \ldots, 0)$. Therefore, $\frac{\partial^2}{\partial t_1^2}\varphi_U(\mathbf{0}) = \frac{\partial^2}{\partial t_1^2}\psi(\mathbf{0})$, 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 10. □

Appendix C: "Good" Lattice Properties

In this appendix, we briefly review certain "good" lattice properties, and some results in the literature. This is almost entirely based on [14]. Let $\{\Lambda^{(d)}\}$ be a sequence of lattices, with each $\Lambda^{(d)}$ chosen uniformly at random from a (d, k, q) ensemble described in Section VI-B.

We say that that the sequence of lattices $\{\Lambda^{(d)}\}$ is *good for covering* if

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

We say that $\{\Lambda^{(d)}\}$ is *good for packing* if

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

Let $\mathcal{G}_{\Lambda^{(d)}}$ denote the normalized second moment per dimension of $\Lambda^{(d)}$, as defined in Section II-A. A sequence of lattices $\{\Lambda^{(d)}\}$ is said to be *good for MSE quantization* if $\mathcal{G}_{\Lambda^{(d)}} \to \frac{1}{2\pi e}$ as $d \to \infty$.

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

$$\mu := \frac{\text{vol}\big(\mathcal{V}(\Lambda^{(d)})\big)^{2/d}}{\sigma^2}.$$

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

$$\Pr[\mathbf{Z} \notin \mathcal{V}(\Lambda^{(d)})] \leq e^{-d\left(E_U(\mu) - o_d(1)\right)}$$

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

$$E_U(\mu) = \begin{cases} \frac{\mu}{16\pi e} & \text{if } 8\pi e \leq \mu \\ \frac{1}{2}\ln\frac{\mu}{8\pi} & \text{if } 4\pi e \leq \mu \leq 8\pi e \\ \frac{\mu}{4\pi e} - \frac{1}{2}\ln\frac{\mu}{2\pi} & \text{if } 2\pi e \leq \mu \leq 4\pi e. \end{cases} \tag{55}$$

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

$$\frac{\mathrm{vol}\big(\mathcal{V}(\Lambda^{(d)})\big)^{2/d}}{\sigma^2} \geq 2\pi e,$$

the probability that a lattice decoder decodes to a lattice point other than the one that was transmitted, decays exponentially in the dimension d, with the exponent given by (55).

It is worth noting that the above "goodness" properties are invariant to scaling. If $\{\Lambda^{(d)}\}$ is a sequence of lattices that is good for covering, packing, and AWGN channel coding, then for any $\alpha > 0$, $\{\alpha\Lambda^{(d)}\}$ is also good for covering, packing and AWGN channel coding. This is because of the fact that $r_{\mathrm{pack}}(\alpha\Lambda^{(d)}) = \alpha r_{\mathrm{pack}}(\Lambda^{(d)})$, $r_{\mathrm{cov}}(\alpha\Lambda^{(d)}) = \alpha r_{\mathrm{cov}}(\Lambda^{(d)})$, and $r_{\mathrm{eff}}(\alpha\Lambda^{(d)}) = \alpha r_{\mathrm{eff}}(\Lambda^{(d)})$.

Appendix D: Proof of Lemma 13

In proving Lemma 13, we use the following theorem from [14], which says that if the parameters k and q are selected appropriately, then almost all lattices in a (d, k, q) ensemble satisfy the "goodness" properties described in Appendix C.

Theorem 25 ([14], Theorem 5)**.** *Let $0 < r_{min} < \frac{1}{4}$ be chosen arbitrarily. Let $\Lambda^{(d)}$ be a sequence of lattices selected uniformly at random from a (d, k, q) ensemble, such that*

- *$k \leq \beta_1 d$ for some $0 < \beta_1 < 1$, but k grows faster than $\log^2 d$, and*
- *q is chosen so that $r_{\mathrm{eff}}(\Lambda^{(d)})$, as given by (24), satisfies $r_{min} < r_{\mathrm{eff}}(\Lambda^{(d)}) < 2r_{min}$.*

Then, the sequence of lattices $\Lambda^{(d)}$ is simultaneously good for covering, packing and MSE quantization, with probability approaching 1 as d tends to infinity. If, in addition, we have $\beta_1 < 1/2$, then the sequence of lattices is also simultaneously good for AWGN channel coding with probability tending to 1 as $d \to \infty$.

Therefore, if we choose k and q that satisfy the hypotheses of Lemma 13, then from the above theorem, the probability that a uniformly chosen $\Lambda_0^{(d)}$ satisfies condition (G_1) tends to 1 as $d \to \infty$.

Recall from Section II-A 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 A^{-1}. It turns out that the dual of a Construction-A lattice is also a Construction-A lattice, as seen from the following.

Proposition 26. *Suppose that G is the $k \times d$ systematic generator matrix of a (d, k) linear code $\mathcal{C}$ over $\mathbb{Z}_q$, q being prime, i.e., G has 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

$$\mathsf{A} = \frac{1}{q} \begin{bmatrix} \mathsf{I}_k & \mathsf{B} \\ \mathsf{0} & q\mathsf{I}_{(d-k)} \end{bmatrix} \tag{56}$$

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

Proof: We want to show that $\mathsf{A}^T\mathbb{Z}^d := \{\mathsf{A}^T\mathbf{y} : \mathbf{y} \in \mathbb{Z}^d\} = \Lambda(\mathcal{C})$. By definition, $\Lambda(\mathcal{C}) = \{\mathbf{x} \in \mathbb{R}^d : (q\mathbf{x}) \bmod q \in \mathcal{C}\}$. Fix any $\mathbf{z} \in \mathbb{Z}^d$. Then, it can be verified that $(q\mathsf{A}^T\mathbf{z}) \bmod q = (\mathsf{G}^T\hat{\mathbf{z}}) \bmod q$ (which is a codeword in $\mathcal{C}$) for some $\hat{\mathbf{z}} \in \{0, 1, \ldots, q-1\}^k$. Therefore, $(q\mathsf{A}^T\mathbf{z}) \bmod q \in \mathcal{C}$, and hence, $\mathsf{A}^T\mathbb{Z}^d \subseteq \Lambda(\mathcal{C})$. For the converse, define $\mathcal{C}' = \{\frac{1}{q}\mathbf{c} : \mathbf{c} \in \mathcal{C}\}$. Then, $\Lambda(\mathcal{C}) = \mathcal{C}' + \mathbb{Z}^d := \{\mathbf{c}+\mathbf{z} : \mathbf{c} \in \mathcal{C}', \mathbf{z} \in \mathbb{Z}^d\}$. The set $\mathsf{A}^T\mathbb{Z}^d$ forms a group under (componentwise) addition. Hence, it is sufficient to show that $\mathcal{C}' \subseteq \mathsf{A}^T\mathbb{Z}^d$, and $\mathbb{Z}^d \subseteq \mathsf{A}^T\mathbb{Z}^d$. Fix an arbitrary $\mathbf{c} \in \mathcal{C}$. Let $\mathbf{c}' = \frac{1}{q}\mathbf{c}$. By definition, there exists an $\mathbf{x} \in \mathbb{Z}_q^k$ such that

$$\begin{aligned}\mathbf{c} &= \left(\begin{bmatrix} \mathsf{I}_k & \mathsf{B} \end{bmatrix}^T \mathbf{x}\right) \bmod q \\ &= \begin{bmatrix} \mathbf{x} \\ \mathsf{B}^T\mathbf{x} \end{bmatrix} - q\begin{bmatrix} \mathbf{0} \\ \mathbf{z}' \end{bmatrix} \end{aligned} \tag{57}$$

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

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

Hence, there exists

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

so that $\mathbf{c}' = \mathsf{A}^T\mathbf{z}$. Therefore, we can say that $\mathcal{C}' \subseteq \mathsf{A}^T\mathbb{Z}^d$. Next, consider $\mathbf{z} \in \mathbb{Z}^d$. Let A^* be defined as

$$\mathsf{A}^* = \begin{bmatrix} q\mathsf{I}_k & \mathbf{0} \\ -\mathsf{B}^T & \mathsf{I}_{(d-k)} \end{bmatrix}, \tag{58}$$

and note that $\mathsf{A}^T\mathsf{A}^* = \mathsf{I}_d$, the $d \times d$ identity matrix. Let $\mathbf{z}' = \mathsf{A}^*\mathbf{z} \in \mathbb{Z}^d$. 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 $\mathbf{z} \in \mathbb{Z}^d$, there exists a $\mathbf{z}' \in \mathbb{Z}^d$ so that $\mathbf{z} = \mathsf{A}^T\mathbf{z}'$, and hence $\mathbb{Z}^d \subseteq \mathsf{A}^T\mathbb{Z}^d$, 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,

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

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

It is easy to verify that if A is full rank, then A^* defined in (58) is the inverse of A, and A^* is a generator matrix of $\Lambda^*(\mathcal{C})$. Since a permutation of the rows of a generator matrix of a lattice also yields a valid generator matrix for the same lattice,

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

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 [30]

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

We thus have the following result.

Lemma 27. *Let $\mathcal{C}$, G, $\Lambda(\mathcal{C})$ be as in Proposition 26. Then, the dual of $\Lambda(\mathcal{C})$, denoted by $\Lambda^*(\mathcal{C})$, has generator matrix*

$$\mathsf{A}^* = \begin{bmatrix} q\mathsf{I}_k & \mathsf{0} \\ -\mathsf{B}^T & \mathsf{I}_{(d-k)} \end{bmatrix}. \tag{59}$$

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

Since $\Lambda^*(\mathcal{C}) = q\Lambda(\mathcal{C}^\perp)$, if the generator matrix is full-rank, then $\Lambda(\mathcal{C}^\perp)$ belongs to a $(d, d-k, q)$ ensemble. Therefore, from [14], we can say that a randomly picked $\Lambda(\mathcal{C}^\perp)$ is good for packing and covering with probability tending to 1 as $d \to \infty$, as long as $d-k \leq \beta_1 d$ for some $0 < \beta_1 < 1$, and $d-k$ grows faster than $\log^2 d$. From the definitions, we see that the properties of covering and packing goodness are invariant to any scaling of the lattices. Therefore, if $\Lambda(\mathcal{C}^\perp)$ is good for packing and covering, then $q\Lambda(\mathcal{C}^\perp)$, and hence $\Lambda^*(\mathcal{C})$ is also good for packing and covering. We have seen that the probability of $\{\Lambda_0^{(d)}\}$ being simultaneously good for covering, packing and AWGN channel coding tends to 1 as d tends to ∞. If we choose $k = \beta_1 d$ for some $\beta_1 < 1/2$, then the sequence of dual lattices is good for packing with probability tending to 1 as $d \to \infty$. Using the union bound, we can argue that a randomly picked sequence of coarse lattices satisfies (G_1) and (G_2) with probability going to 1 as $d \to \infty$. .

It was also shown in [13] that if the coarse lattices are good for covering and AWGN channel coding, then as long as $d/q_1 \to 0$ as $d \to \infty$, the probability that a uniformly chosen sequence of fine lattices is good for AWGN channel coding tends to 1 as $d \to \infty$. This completes the proof of Lemma 13.

Appendix E: Proof of Lemma 15

For ease of notation, denote by r_{eff}, the effective radius of $\Lambda_0^{(d)}$. The index, d, in r_{eff} has been dropped but it must be understood that this is a function of d. Let $\mathcal{C}^{(d)}$ denote the (d,k) code over $\mathbb{Z}_q$ that is used to generate the coarse lattice. Using (24),

$$\begin{aligned} q^k &= \frac{\Gamma(d/2+1)}{\pi^{d/2} {r_{\text{eff}}}^d} \\ &= \sqrt{d\pi}\left(\frac{d}{2\pi e {r_{\text{eff}}}^2}\right)^{d/2}(1+o_d(1)), \end{aligned} \tag{60}$$

where the second step uses Stirling's approximation, and $o_d(1)$ is a term that approaches 0 as $d \to \infty$. From (26), $k = \beta_0 d$ for some $0 < \beta_0 < 1/2$. Substituting this in the above, and raising both sides to the power $1/d$, we get

$$q^{\beta_0} = (d\pi)^{\frac{1}{2d}}\left(\frac{d}{2\pi e {r_{\text{eff}}}^2}\right)^{1/2}(1+o_d(1))^{1/d} = (d\pi)^{\frac{1}{2d}}\frac{\sqrt{d}}{\sqrt{2\pi e} r_{\text{eff}}}(1+o_d(1)). \tag{61}$$

Let $\Lambda_0^{(d)*}$ denote the dual of $\Lambda_0^{(d)}$, and ${r_{\text{eff}}}^*$ denote the effective radius of $\Lambda_0^{(d)*}$. Let $\Lambda_0(\mathcal{C}^{(d)\perp})$ be the lattice obtained by applying Construction-A on the dual of $\mathcal{C}^{(d)}$, i.e., on $\mathcal{C}^{(d)\perp}$. As remarked in Appendix D, $\Lambda_0(\mathcal{C}^{(d)\perp})$ comes from a $(d, d-k, q)$ ensemble. From Lemma 27, $\Lambda_0^{(d)*} = q\Lambda_0(\mathcal{C}^{(d)\perp})$. Therefore, $(1/q)\Lambda_0^{(d)*} = \Lambda_0(\mathcal{C}^{(d)\perp})$ will satisfy

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

where $o_d(1) \to 0$ as $d \to \infty$. But $r_{\text{eff}}(\frac{1}{q}\Lambda_0^{(d)*}) = \frac{1}{q}{r_{\text{eff}}}^*$, and hence, analogous to (61), we have

$$q^{d(1-\beta_0)} = \sqrt{d\pi}\left(\frac{d}{2\pi e(1/q)^2({r_{\text{eff}}}^*)^2}\right)^{d/2}(1+o_d(1)). \tag{62}$$

Rearranging,

$${r_{\text{eff}}}^* = (d\pi)^{\frac{1}{2d}}\frac{\sqrt{d}q^{\beta_0}}{\sqrt{2\pi e}}(1+o_d(1))^{1/d}. \tag{63}$$

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

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

Substituting for q^{β_0} from (61) in the above equation, we get

$$\frac{r_{\text{eff}}(\Lambda_0^{(d)})r_{\text{pack}}(\Lambda_0^{(d)*})}{d} = \gamma(d)(d\pi)^{(1/d)}\frac{1}{2\pi e}(1+o_d(1)). \tag{64}$$

Therefore, as $d \to \infty$, the above expression converges to a value greater than or equal to $1/4\pi e$. Using $r_{\text{pack}}(\hat{\Lambda}_0^{(d)}) = 2\pi r_{\text{pack}}(\Lambda_0^{(d)*})$, we get Lemma 15. □

Appendix F: Proof of Theorem 18

The following lemma from [22] will be used in the proof.

Lemma 28 ([22], Lemma 4)**.** *Let Λ be a lattice in $\mathbb{R}^d$. Then, for all $\mathbf{z} \in \mathbb{R}^d$, and $\kappa > 0$,*

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

For ease of notation, we will suppress the index d in $\epsilon^{(d)}$, $\Lambda_0^{(d)}$ and $\Lambda^{(d)}$. We will find upper and lower bounds for $p_{U+V}(\mathbf{u})$ and $p_{U+V|\mathbf{x}}(\mathbf{u})$, 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 $S \subseteq \mathbb{R}^d$, 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 (34), with $\kappa = \sqrt{\mathcal{P}}$, we have

$$p_{U|\mathbf{x}}(\mathbf{u}) = \frac{g_{\sqrt{\mathcal{P}}}(\mathbf{u})}{g_{\sqrt{\mathcal{P}},-\mathbf{x}}(\Lambda_0)}\mathbf{1}_{\Lambda_0+\mathbf{x}}(\mathbf{u}). \tag{65}$$

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

$$p_U(\mathbf{u}) = \sum_{\mathbf{x}\in\mathbb{G}_X} \frac{g_{\sqrt{\mathcal{P}}}(\mathbf{u})}{g_{\sqrt{\mathcal{P}},-\mathbf{x}}(\Lambda_0)} \frac{\mathbf{1}_{(\Lambda_0+\mathbf{x})}(\mathbf{u})}{M}. \tag{66}$$

By monotonicity of the flatness factor, $\epsilon_{\Lambda_0}(\sqrt{\mathcal{P}}) < \epsilon_{\Lambda_0}(\sqrt{\mathcal{P}/2}) = \epsilon$, and using Lemma 28,

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

Using this in (66), we get for $\mathbf{u} \in \Lambda$,

$$\frac{g_{\sqrt{\mathcal{P}}}(\mathbf{u})}{Mg_{\sqrt{\mathcal{P}}}(\Lambda_0)} \le p_U(\mathbf{u}) \le \frac{g_{\sqrt{\mathcal{P}}}(\mathbf{u})}{Mg_{\sqrt{\mathcal{P}}}(\Lambda_0)} \frac{1+\epsilon}{1-\epsilon}. \tag{67}$$

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

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

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

$$\left(\frac{1-\epsilon}{1+\epsilon}\right) M g_{\sqrt{\mathcal{P}}}(\Lambda_0) \le g_{\sqrt{\mathcal{P}}}(\Lambda) \le M g_{\sqrt{\mathcal{P}}}(\Lambda_0). \tag{68}$$

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

$$\left(\frac{1-\epsilon}{1+\epsilon}\right) M g_{\sqrt{\frac{\mathcal{P}}{2}},\mathbf{a}}(\Lambda_0) \le g_{\sqrt{\frac{\mathcal{P}}{2}},\mathbf{a}}(\Lambda) \le M g_{\sqrt{\frac{\mathcal{P}}{2}},\mathbf{a}}(\Lambda_0). \tag{69}$$

We establish some more notation for convenience. Let

$$\alpha(\mathbf{w}) := \frac{g_{\sqrt{2\mathcal{P}}}(\mathbf{w})}{Mg_{\sqrt{\mathcal{P}}}(\Lambda_0)} \frac{g_{\sqrt{\frac{\mathcal{P}}{2}}}(\Lambda_0)}{g_{\sqrt{\mathcal{P}}}(\Lambda_0)}, \tag{70}$$

$$\beta(\mathbf{x},\mathbf{w}) := \left(\frac{g_{\sqrt{\frac{\mathcal{P}}{2}},\frac{\mathbf{w}}{2}-\mathbf{x}}(\Lambda_0)}{g_{\sqrt{\frac{\mathcal{P}}{2}}}(\Lambda_0)}\right) \left(\frac{g_{\sqrt{\mathcal{P}},-\mathbf{x}}(\Lambda_0)}{g_{\sqrt{\mathcal{P}}}(\Lambda_0)}\right)^{-1}. \tag{71}$$

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

Lemma 29. *For any lattice point $\mathbf{w} \in \Lambda$, and any $\mathbf{x} \in \mathbb{G}_X$, we have*

$$\left(\frac{1-\epsilon}{1+\epsilon}\right) \alpha(\mathbf{w}) \le p_{U+V}(\mathbf{w}) \le \left(\frac{1+\epsilon}{1-\epsilon}\right)^2 \alpha(\mathbf{w}) \tag{72}$$

$$\beta(\mathbf{x},\mathbf{w})\alpha(\mathbf{w}) \le p_{U+V|\mathbf{x}}(\mathbf{w}) \le \left(\frac{1+\epsilon}{1-\epsilon}\right) \beta(\mathbf{x},\mathbf{w})\alpha(\mathbf{w}). \tag{73}$$

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

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

Using (65) and (67) in the above equation, we obtain

$$\sum_{\mathbf{t}\in\Lambda_0+\mathbf{x}} \frac{g_{\sqrt{\mathcal{P}}}(\mathbf{t})}{g_{\sqrt{\mathcal{P}},-\mathbf{x}}(\Lambda_0)} \frac{g_{\sqrt{\mathcal{P}}}(\mathbf{w}-\mathbf{t})}{Mg_{\sqrt{\mathcal{P}}}(\Lambda_0)} \le p_{U+V|\mathbf{x}}(\mathbf{w}) \le \sum_{\mathbf{t}\in\Lambda_0+\mathbf{x}} \frac{g_{\sqrt{\mathcal{P}}}(\mathbf{t})}{g_{\sqrt{\mathcal{P}},-\mathbf{x}}(\Lambda_0)} \frac{g_{\sqrt{\mathcal{P}}}(\mathbf{w}-\mathbf{t})}{Mg_{\sqrt{\mathcal{P}}}(\Lambda_0)} \left(\frac{1+\epsilon}{1-\epsilon}\right). \tag{74}$$

Consider the term

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

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

$$= \frac{g_{\sqrt{2\mathcal{P}}}(\mathbf{w})}{Mg_{\sqrt{\mathcal{P}}}(\Lambda_0)g_{\sqrt{\mathcal{P}},-\mathbf{x}}(\Lambda_0)} \sum_{\mathbf{t}\in\Lambda_0+\mathbf{x}} g_{\sqrt{\frac{\mathcal{P}}{2}},\frac{\mathbf{w}}{2}}(\mathbf{t})$$

$$= \frac{g_{\sqrt{2\mathcal{P}}}(\mathbf{w})}{Mg_{\sqrt{\mathcal{P}}}(\Lambda_0)} \frac{g_{\sqrt{\frac{\mathcal{P}}{2}},\frac{\mathbf{w}}{2}-\mathbf{x}}(\Lambda_0)}{g_{\sqrt{\mathcal{P}},-\mathbf{x}}(\Lambda_0)}. \tag{75}$$

Substituting this in (74), and writing this in terms of α and β, we obtain (73). Similarly, bounding both p_U and p_V from above and below using (67), proceeding as above, and finally using (69) to bound $g_{\sqrt{\frac{\mathcal{P}}{2}},\frac{\mathbf{w}}{2}}(\Lambda)$, we get (72). ■

Observe that $\beta(\mathbf{x},\mathbf{w})$ in (71) is a ratio of two terms, both of which can be bounded using Lemma 28 to get

$$\left(\frac{1-\epsilon}{1+\epsilon}\right) \le \beta(\mathbf{x},\mathbf{w}) \le \left(\frac{1+\epsilon}{1-\epsilon}\right). \tag{76}$$

Let $\overline{p}_{U+V}$ and $\underline{p}_{U+V}$ respectively denote the upper and lower bounds for p_{U+V} in (72), and let $\overline{p}_{U+V|\mathbf{x}}$ and $\underline{p}_{U+V|\mathbf{x}}$ respectively denote the upper and lower bounds for $p_{U+V|\mathbf{x}}$ in (73). Then, we can say that $|p_{U+V|\mathbf{x}}(\mathbf{w}) - p_{U+V}(\mathbf{w})|$ is less than or equal to the maximum of $|\overline{p}_{U+V|\mathbf{x}}(\mathbf{w}) - \underline{p}_{U+V}(\mathbf{w})|$ and $|\underline{p}_{U+V|\mathbf{x}}(\mathbf{w}) - \overline{p}_{U+V}(\mathbf{w})|$.

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

$$|\overline{p}_{U+V|\mathbf{x}}(\mathbf{w}) - \underline{p}_{U+V}(\mathbf{w})| = \alpha(\mathbf{w}) \left(\frac{1-\epsilon}{1+\epsilon}\right) \left| \left(\frac{1+\epsilon}{1-\epsilon}\right)^2 \beta(\mathbf{x},\mathbf{w}) - 1 \right|. \tag{77}$$

However, from (76), 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,

$$|\overline{p}_{U+V|\mathbf{x}}(\mathbf{w}) - \underline{p}_{U+V}(\mathbf{w})| \le \alpha(\mathbf{w}) \left(\frac{1-\epsilon}{1+\epsilon}\right) 64\epsilon. \tag{78}$$

Similarly, expressing $|\underline{p}_{U+V|\mathbf{x}}(\mathbf{w}) - \overline{p}_{U+V}(\mathbf{w})|$ 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

$$|\underline{p}_{U+V|\mathbf{x}}(\mathbf{w}) - \overline{p}_{U+V}(\mathbf{w})| \le \alpha(\mathbf{w}) \left(\frac{1+\epsilon}{1-\epsilon}\right)^2 8\epsilon. \tag{79}$$

Rearranging (72), and observing that $\sum_{\mathbf{w}\in\Lambda} p_{U+V}(\mathbf{w}) = 1$, we have

$$\left(\frac{1-\epsilon}{1+\epsilon}\right)^2 \le \sum_{\mathbf{w}\in\Lambda} \alpha(\mathbf{w}) \le \left(\frac{1+\epsilon}{1-\epsilon}\right). \tag{80}$$

Combining (78) and (79), and summing over $\mathbf{w}$, we get

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

and using (80) to bound $\sum_{\mathbf{w}\in\Lambda} \alpha(\mathbf{w})$ from above, we get

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

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

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

thereby completing the proof. □

Appendix G: Proof of Lemma 20

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 $\mathbf{U}' := (\alpha^*-1)\mathbf{U}+\mathbf{N}'$ conditioned on $\mathbf{X}=\mathbf{x}$. Let $f_{\mathbf{V}'|\mathbf{y}}$ denote the density function of $\mathbf{V}' := (\alpha^*-1)\mathbf{V}+\alpha^*\mathbf{Z}$ 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

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

Therefore,

$$\mathbb{V}(f_{\mathbf{Z}'|\mathbf{x},\mathbf{y}}, f_{\mathbf{N}}) \leq \mathbb{V}(f_{\mathbf{V}'|\mathbf{y}}, f_{\mathbf{N}'}) + \mathbb{V}(f_{\mathbf{U}'|\mathbf{x}}, f_{\mathbf{N}}). \tag{81}$$

Lemma 30 ([22], Lemma 8)**.** *Let Λ be a lattice in $\mathbb{R}^d$, $\mathbf{x}\in\mathbb{R}^d$, 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*

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

Using Lemma 30 and the fact that for any constant $a>0$, $\epsilon_{a\Lambda_0^{(d)}}(a\theta) = \epsilon_{\Lambda_0^{(d)}}(\theta)$ [22, Remark 4], we get

$$\begin{aligned}
\mathbb{V}(f_{\mathbf{V}'|\mathbf{y}}, f_{\mathbf{N}'}) &\leq 4\epsilon_{\Lambda_0^{(d)}}\left(\frac{\alpha^*\sqrt{\mathcal{P}\sigma^2}}{\sqrt{(1-\alpha^*)^2\mathcal{P}+(\alpha^*)^2\sigma^2}}\right) &(82)\\
&\leq 4\epsilon_{\Lambda_0^{(d)}}\left(\frac{\alpha^*\sqrt{\mathcal{P}\sigma^2}}{\sqrt{(1-\alpha^*)^2 2\mathcal{P}+(\alpha^*)^2\sigma^2}}\right) &(83)\\
&= 4\epsilon_{\Lambda_0^{(d)}}\left(\sqrt{\alpha^*\mathcal{P}}\right), &(84)
\end{aligned}$$

where (83) is by the monotonicity of the flatness factor. Equation (84) is then obtained by substituting $\alpha^* = 2\mathcal{P}/(2\mathcal{P}+\sigma^2)$ and simplifying. By similar arguments, we can show that

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

Substituting in (81) completes the proof. □

Appendix H: Proof of Lemma 23

We want to show that $\mathcal{I}(X_1,\ldots,X_N;\Theta_{k,N})$ is arbitrarily small for all sufficiently large d. 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 VII to conclude that each of these terms go to zero as $d\to\infty$. 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 [19].

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

$$\begin{aligned}
\mathcal{I}(X_1,\ldots,X_N;\Theta_{k,N}) &= \sum_{t=1}^{N}\mathcal{I}(X_t;\Theta_{k,N}\big|X_1,\ldots,X_{t-1})\\
&= \sum_{t=1}^{N}\left[\mathcal{I}(X_t;J_k,J_{k-1}\big|X_1,\ldots,X_{t-1}) + \sum_{n=1}^{N}\mathcal{I}(X_t;\mathbf{W}_k[2n+k-1]\big|X_1,\ldots,X_{t-1},\Theta_{k,n-1})\right]\\
&= \sum_{t=1}^{N}\sum_{n=1}^{N}\mathcal{I}(X_t;\mathbf{W}_k[2n+k-1]\big|X_1,\ldots,X_{t-1},\Theta_{k,n-1}), &(85)
\end{aligned}$$

where the last step follows from the fact that $\mathcal{I}(X_t;J_k,J_{k-1}\big|X_1,\ldots,X_{t-1})=0$ 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 $n \in \{1, 2, \ldots, N\}$, $\Theta_{k,n-1}$ consists of random variables which are all functions of $X_1, \ldots, X_{n-1}$ and $J_{k-1}, \ldots, J_{k+n-1}$, which are all independent of X_t for $n \leq t$ (even when conditioned on the first $l-1 < t$ messages). Therefore,

Proposition 31. *Let $1 \leq t \leq N$, and $n, l \in \{1, 2, \ldots, t\}$. Then, the message X_t is conditionally independent of $\Theta_{k,n-1}$ given $X_1, X_2, \ldots, X_{l-1}$.*

Using a similar argument, we obtain

Proposition 32. *Let $1 \leq t < n \leq N$. The vector $\mathbf{W}_k[2n+k-1]$ received by the kth relay in the $(2n+k-1)$st phase is independent of $X_1, \ldots, X_t$ and $\Theta_{k,n-1}$.*

We now evaluate the terms in (85). Using Proposition 31, we get

$$\mathcal{I}(X_t; \mathbf{W}_k[2n+k-1] \big| X_1, \ldots, X_{t-1}, \Theta_{k,n-1}) = 0 \tag{86}$$

for all $1 \leq n < t \leq N$. Similarly, using Proposition 32,

$$\mathcal{I}(X_t; \mathbf{W}_k[2n+k-1] \big| X_1, \ldots, X_{t-1}, \Theta_{k,n-1}) = 0 \tag{87}$$

for all $1 \leq t < n \leq N$. Therefore, (85) reduces to

$$\mathcal{I}(X_1, \ldots, X_N; \Theta_{k,N}) = \sum_{t=1}^{N} \mathcal{I}(X_t; \mathbf{W}_k[2t+k-1] \big| X_1, \ldots, X_{t-1}, \Theta_{k,t-1}). \tag{88}$$

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

$$\begin{aligned}\mathcal{I}(X_t; \mathbf{W}_k[2t+k-1] \big| X_1, \ldots, X_{t-1}, \Theta_{k,t-1}) &= \mathcal{H}(\mathbf{W}_k[2t+k+1] \big| X_1, \ldots, X_{t-1}, \Theta_{k,t-1}) \\ &\quad - \mathcal{H}(\mathbf{W}_k[2t+k+1] \big| X_1, \ldots, X_t, \Theta_{k,t-1}). \end{aligned} \tag{89}$$

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

$$\begin{aligned}\mathcal{H}(\mathbf{W}_k[2t+k+1] \big| X_1, \ldots, X_t, \Theta_{k,t-1}) &\geq \mathcal{H}(\mathbf{W}_k[2t+k+1] \big| X_1, \ldots, X_t, \oplus_{p=1}^{t} X_p \oplus J_{k+t-1}, \Theta_{k,t-1}) \\ &= \mathcal{H}(\mathbf{W}_k[2t+k+1] \big| \oplus_{p=1}^{t} X_p \oplus J_{k+t-1}). \end{aligned} \tag{90}$$

The first step is true because conditioning reduces entropy. The second step requires more justification. Given $\oplus_{p=1}^{t} X_p \oplus J_{k+t-1}$, 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 $\oplus_{p=1}^{t} X_p \oplus J_{k+t-1}$. 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 $\oplus_{p=1}^{t} X_p \oplus J_{k+t-1}$, thus justifying (90). Now, define $X := \oplus_{p=1}^{t} X_p \oplus J_{k+t-1}$, and $Y := \oplus_{p=1}^{t-1} X_p \oplus J_{k+t}$. 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 32, the first term of (89), $\mathcal{H}(\mathbf{W}_k[2t+k+1]\big|X_1,\ldots,X_{t-1},\Theta_{k,t-1}) = \mathcal{H}\left(\mathcal{E}(X)+\mathcal{E}(Y)\right)$. 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 Y are independent and uniformly distributed over $\mathbb{G}^{(d)}$. The lemma now follows by using Theorem 18 and Lemma 17 to bound this quantity. □

Acknowledgement

We would like to thank the anonymous reviewers for carefully reading our manuscript and suggesting several improvements to the presentation. In particular, we thank the reviewer who suggested a means of avoiding the use of a random dither to achieve the rate in Section VII. We are also grateful to Manjunath Krishnapur for providing Proposition 9 and its proof.

References

[1] 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.

[2] I.-J. Baik and S.-Y. Chung, "Network coding for two-way relay channels using lattices," in *Proc. IEEE Int. Conf. Communications*, Beijing, China, 2008, pp. 3898–3902.

[3] A. Barvinok, *Math 669: Combinatorics, Geometry and Complexity of Integer Points*. [Online]. Available: http://www.math.lsa.umich.edu/∼barvinok/latticenotes669.pdf .

[4] J.-C. Belfiore and F. Oggier, "Secrecy gain: A wiretap lattice code design," in *Proc. 2010 Int. Symp. Information Theory and Its Applications*, Taichung, Taiwan, pp. 174–178.

[5] J.-C. Belfiore, "Lattice codes for the compute-and-forward protocol: The flatness factor," in *Proc. 2011 Information Theory Workshop*, Paraty, Brazil, pp. 1–4.

[6] J.H. Conway and N.J. Sloane, *Sphere Packings, Lattices and Groups*, New York: Springer-Verlag, 1988.

[7] T. Cover and J. Thomas, *Elements of Information Theory*, 2nd ed. New York: Wiley-Interscience, 1996.

[8] 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*, vol. 4965, pp. 471–488, 2008.

[9] I. Csiszár and P. Narayan, "Secrecy capacities for multiple terminals," *IEEE Trans. Inf. Theory*, vol. 50, no. 12, pp. 3047–3061, Dec. 2004.

[10] 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.

[11] A. Dasgupta, *Probability for Statistics and Machine Learning*, New York: Springer Texts in Statistics, 2011.

[12] 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.

[13] 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.

[14] 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.

[15] U. Erez, S. ten Brink, "A close-to-capacity dirty paper coding scheme," *IEEE Trans. Inf. Theory*, vol. 51, no. 10, pp. 3417–3432, Oct. 2005.

[16] 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.

[17] W. Feller, *An Introduction to Probability Theory and Its Applications, Vol. 2*, 2nd ed. New York: Wiley, 1971.

[18] 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.

[19] 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.

[20] I.N. Herstein, *Topics in Algebra*, 2nd ed. New York: Wiley, 1975.

[21] N. Kashyap, V. Shashank, and A. Thangaraj, "Secure computation in a bidirectional relay," in *Proc. 2012 IEEE Int. Symp. Information Theory*, Cambridge, MA, pp. 1162–1166.

[22] C. Ling, L. Luzzi, J.-C. Belfiore, and D. Stehlé, "Semantically secure lattice codes for the Gaussian wiretap channel," submitted for publication. [Online]. Available: http://arxiv.org/abs/1210.6673.

[23] E. Lukacs, *Characteristic Functions*, 2nd ed. London, U.K.: Griffin, 1970.

[24] U. Maurer and S. Wolf. "Information-theoretic key agreement: From weak to strong secrecy for free," *Proc. EUROCRYPT–2000 on Advances in Cryptology*, vol. 1807, pp. 351–368, Springer, 2000.

[25] 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.

[26] B. Nazer and M. Gastpar, "Reliable physical layer network coding," *Proc. of the IEEE*, vol. 99, no. 3, pp. 438–460, Mar. 2011.

[27] 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.

[28] F. Oggier, P. Solé, and J.-C. Belfiore, "Lattice codes for the wiretap Gaussian channel: construction and analysis," submitted for publication. [Online]. Available: http://arxiv.org/abs/1103.4086.

[29] 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.

[30] R.M. Roth, *Introduction to Coding Theory*, Cambridge, U.K.: Cambridge University Press, 2006.

[31] 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.

[32] N. Sommer, M. Feder, and O. Shalvi, "Low density lattice codes," *IEEE Trans. Inf. Theory*, vol. 54, no. 4, pp. 1561–1585, Apr. 2008.

[33] N. Sommer, M. Feder, and O. Shalvi, "Shaping methods for low-density lattice codes," in *Proc. 2009 Information Theory Workshop*, Taormina, Italy, pp. 238–242.

[34] E.M. Stein and G.L. Weiss, *Introduction to Fourier Analysis on Euclidean Spaces*, Princeton, NJ: Princeton Univ. Press, 1971.

[35] 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.

[36] 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.

[37] S.J. Wolfe, "On the finite series expansion of multivariate characteristic functions," *J. Multivariate Anal.*, vol. 3, pp. 328–335, 1973.

[38] Y. Yan, C. Ling, and X. Wu, "Polar lattices: Where Arikan meets Forney," in *Proc. 2013 IEEE Int. Symp. Information Theory*, Istanbul, Turkey, pp. 1292–1296.

[39] S. Zhang and S.-C. Liew, "Channel coding and decoding in a relay system operated with physical-layer network coding," *IEEE J. Sel. Areas Commun.*, vol. 27, no. 5, pp. 788–796, Jun. 2009.