Providing Secrecy With Structured Codes: Two-User Gaussian Channels

Xiang He, *Member, IEEE*, and Aylin Yener, *Senior Member, IEEE*

***Abstract*— Recent results have shown that structured codes can be used to construct good channel codes, source codes, and physical layer network codes for Gaussian channels. For Gaussian channels with secrecy constraints, however, efforts to date rely on Gaussian random codes. In this paper, we advocate that structure in random code generation is useful for providing secrecy as well. In particular, a Gaussian wiretap channel in the presence of a cooperative jammer is studied. Previously, the achievable secrecy rate for this channel was derived using Gaussian signaling, which saturated at high signal-to-noise ratio (SNR), owing to the fact that the cooperative jammer simultaneously helps by interfering with the eavesdropper, and hurts by interfering with the intended receiver. In this paper, a new achievable rate is derived through imposing a lattice structure on the signals transmitted by both the source and the cooperative jammer, which are aligned at the eavesdropper but remain separable at the intended receiver. We prove that the achieved secrecy rate does not saturate at high SNR for all values of channel gains except when the channel is degraded.**

***Index Terms*— Information theoretic secrecy, lattice codes, cooperative jamming, Gaussian wiretap channels.**

I. Introduction

THE notion of information theoretic secrecy was first proposed by Shannon [1] whereby a message transmitted to a receiver is guaranteed to be kept secret from an eavesdropper, irrespective of the computational power the eavesdropper possesses. In particular, it was shown that it is possible that the eavesdropper gains no information regarding the secret message having intercepted the cryptogram, albeit at the expense of very long keys [1]. Wyner, in [2], established that, if the signal received by the eavesdropper (Eve) is a degraded version of the signal observed by the receiver, the long secret keys needed to achieve secrecy per Shannon's notion are not necessary [2]. Csiszár and Körner [3] extended Wyner's setting to the general discrete memoryless wiretap channel and established its secrecy capacity.

Manuscript received July 28, 2009; revised December 31, 2011 and June 29, 2013; accepted November 22, 2013. Date of publication January 9, 2014; date of current version March 13, 2014. This work was supported in part by the National Science Foundation under Grants CNS- 0716325, CNS-0721445, and CCF-0964362, and in part by the DARPA ITMANET Program under Grant W911NF-07-1-0028. This paper was presented at the 2008 Allerton Conference on Communication, Control, and Computing, and the 2009 IEEE Globecom Conference.

X. He was with the Department of Electrical Engineering, Pennsylvania State University, University Park, PA 16802 USA. He is now with Microsoft, Redmond, WA 98052 USA (e-mail: xianghe@microsoft.com).

A. Yener is with the Department of Electrical Engineering, Pennsylvania State University, University Park, PA 16802 USA (e-mail: yener@ee.psu.edu).

Communicated by M. C. Gastpar, Associate Editor for Shannon Theory.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIT.2014.2298132

Numerous channel models have since been studied using the information theoretic secrecy framework. In this work, we are mainly interested in Gaussian channels, i.e., channels with additive Gaussian noise. The maximum reliable transmission rate with secrecy was identified for some of these models including the Gaussian wiretap channel [4], the MIMO wiretap channel [5], [6] and the MIMO Gaussian broadcast channel with confidential messages [7], [8]. On the other hand, secrecy capacity regions for models with multiple transmitters remain in general as open problems except for some special cases, e.g., sum secrecy capacity for a degraded Gaussian multiple access wiretap channel [9], [10]. Upper bounds, lower bounds and some asymptotic results on the secrecy capacity exist, see for example [11]–[16]. To prove achievability, Shannon's random coding argument is used in these works, in which the codewords are i.i.d. sequences sampled from a distribution defined over the channel inputs.

On the other hand, it is known that introducing a structure on the set from which the codewords are sampled can be helpful in proving certain information theoretical results [17]. This could be as simple as sampling codewords from a QAM constellation [18]. In [17], a lattice is used, which can be viewed as a constellation defined over N channel uses. This *structured random code* approach [17] is useful in multi-terminal problems: the structure of these codes makes it possible to align unwanted interference, for example, in Gaussian interference channels with more than two users [19]–[22]. Additionally, it renders the analysis of some network topologies feasible: for example, in [23], [24], using structured codes allows the relaying scheme to be equivalent to a modulo sum operation, making it easy to trace the signal over a multi-hop relay network.

A natural question therefore is whether this approach is useful for *secret* communication as well. In this work, we shall answer this question positively. In particular, we will consider the application of structured signaling in a two-user setting employing cooperative jamming.

Cooperative jamming is a frequently used strategy in secure communication, where the legitimate transmitters introduce judicious interference into the channel to confuse the eavesdropper while not causing excessive harm to the intended receiver [12], [25]. This strategy has been used in a number of channel models to improve secrecy rates; see [12], [13], [16], [26]–[28] for example.

In this work, we focus on the simplest Gaussian channel model where such a strategy is known to be useful. The model consists of a Gaussian wiretap channel and a

cooperative jammer. This model can also be viewed as a special case of a number of two-user Gaussian channel models considered in previous work [12], [16], [26], [29]. Hence improving the achievable secrecy rate for this model implies that the achievable rates for all these models can be improved as well. Previously, this model was studied in [27] with the optimal transmission power control strategy, where both the cooperative jammer and the sender of the message use codewords sampled from a Gaussian distribution. It was found that the secrecy rate saturates when the transmission power P increases, since the intended receiver is limited by the interference from the cooperative jammer. In this work, instead of sampling from a Gaussian distribution, we use codewords sampled from a nested lattice structure, i.e., each codeword is a sequence of fine lattice points in the Voronoi region of a coarse lattice. The transmission power is adjusted such that the lattice points sent by the transmitter and the cooperative jammer align at the eavesdropper. This reduces the information leaked to the eavesdropper regarding the lattice point sent by the transmitter for a given component of the codeword. The rate of this information leakage is quantified to be less than 1 bit per channel use and further eliminated through using a wiretap code as an outer code. Meanwhile, since these lattice points are not aligned at the intended receiver, by adjusting the power and nested ratio of the nested lattice structure properly, we ensure that the receiver can decode the lattice point sent from the transmitter with high probability and find out the codeword being sent. We prove that the achievable secrecy rate with this scheme increases with power P at the rate of $O(\log_2(P))$ when the channel is fully connected and not degraded, and consequently demonstrate that positive secure degrees of freedom (s.d.o.f.) are achievable for this channel improving the previous results [12], [27].

The rest of the paper is organized as follows. Section II describes the channel model, i.e., the wiretap channel model with a cooperative jammer. Section III describes the methodology for computing the secrecy rate with nested lattice codes in a cooperative jamming setting. Section IV applies this approach to the channel model introduced in Section III, derives the achievable secrecy rate, and shows that positive s.d.o.f. are achievable except for a set of channel gains of measure zero, for which the positivity of s.d.o.f. is further proved in Section V. In Section VII, we discuss the impact of imperfect channel state information. Section VIII concludes the paper. Some of the proofs are presented in Appendices to improve readability.

II. The Gaussian Wiretap Channel With a Cooperative Jammer

Consider the Gaussian wiretap channel with a cooperative jammer [12], [25], [27] shown in Fig. 1. In this model, node S_1 sends a message W_1 via X_1 to node D_1, which must be kept secret from node D_2. Node S_2, the cooperative jammer, transmits X_2. We assume the channel is fully connected, i.e., all channel gains are non-zero. After normalizing the channel gains of the two links to the eavesdroppers, the received signals

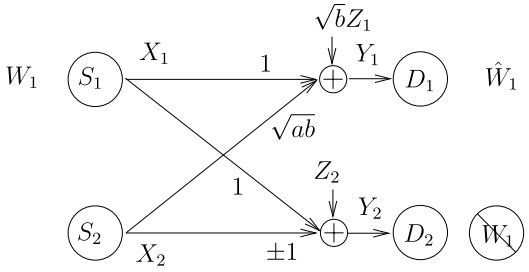

Fig. 1. The Gaussian Wiretap Channel with a Cooperative Jammer.

at the two receiving node D_1 and D_2 can be expressed as[1]

$$\begin{aligned} Y_1 &= X_1 + \sqrt{ab}X_2 + \sqrt{b}Z_1 \\ Y_2 &= X_1 \pm X_2 + Z_2. \end{aligned} \tag{1}$$

Let $\hat{W}_1$ be the estimate of W_1, estimated at node D_1. Let n be the total number of channel uses. D_1 recovers W_1 reliably if

$$\lim_{n\to\infty} \Pr\left(W_1 \neq \hat{W}_1\right) = 0. \tag{2}$$

In addition, since W_1 must be kept secret from D_2, we require[2]

$$\lim_{n\to\infty} \frac{1}{n} H\left(W_1\right) = \lim_{n\to\infty} \frac{1}{n} H\left(W_1 | Y_2^n\right). \tag{3}$$

There are two constraints on the input distribution to the channel model in (1): First, we assume there is no common randomness shared by the encoders of S_1 and S_2, i.e., the joint distribution of the input signals has the following form:

$$\Pr\left(X_1^n\right) \Pr\left(X_2^n\right). \tag{4}$$

Second, the average power of X_i is constrained to be $\bar{P}_i$. Define $X_{i,j}$ to be the jth component of X_i, we have:

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} E\left[X_{i,j}^2\right] \leq \bar{P}_i, i = 1, 2. \tag{5}$$

The secrecy rate R_s is defined as:

$$R_s = \lim_{n\to\infty} \frac{1}{n} H\left(W_1\right), \tag{6}$$

such that the conditions (2), (3) are satisfied simultaneously. The secrecy rate is the number of bits per channel use that can be reliably transmitted without leaking information to the eavesdropper at a positive rate.

In this work, we will mainly be concerned about the high SNR behavior of the secrecy rate. Namely,

Definition 1: The secure degrees of freedom is defined as:

$$\text{s.d.o.f.} = \limsup_{\bar{P}_1 = \bar{P}_2 = \rho \to \infty} \frac{R_s}{\frac{1}{2}\log_2(\rho)}. \tag{7}$$

[1] A general fully connected channel without loss of generality can be written as (1) by scaling and renaming the transmitted signals. For example, (1) can be written as $Y_1/\sqrt{b} = (X_1/\sqrt{b}) + \sqrt{a}X_2 + Z_1$, $Y_2 = \sqrt{b}(X_1/\sqrt{b}) \pm X_2 + Z_2$. Replacing $Y_1/\sqrt{b}$ with $\tilde{Y}_1$ and $X_1/\sqrt{b}$ with $\tilde{X}_1$ lead to the channel expression used in [27].

[2] The achievable rate in this work also holds for the secrecy constraint $\lim_{n\to\infty} H\left(W_1\right) = \lim_{n\to\infty} H\left(W_1|Y_2^n\right)$. See [30] for the additional steps to prove achievability for this stronger secrecy constraint.

Remark 1: In this paper, we focus our attention to real valued channel gains, where the best known achievable rates using Gaussian signaling do not yield positive secure degrees of freedom [27]. For treatment of complex channels, please see [31, Section 5.16][3]. □

III. Nested Lattice Codes and Cooperative Jamming

In this section, we derive the secrecy rate expression when nested lattice codes are used in a cooperative jamming setting. The derivation shall be carried out for the following simplified version of (1).

$$\begin{aligned} Y_1 &= X_1 + Z_1 \\ Y_2 &= X_1 \pm X_2 + Z_2, \end{aligned} \tag{8}$$

which corresponds to $a = 0, b = 1$ in (1). This simplification in the channel model allows us to explain all the necessary steps of computing the secrecy rate without going into the details on how nested lattice codes are decoded at the intended receiver, since the intended receiver is not affected by the cooperative jammer.

We begin by introducing a nested lattice code, which provides the set of codewords with structure that we will use. A nested lattice code is defined as an intersection of an N-dimensional "fine" lattice Λ and the fundamental Voronoi region of an N-dimensional "coarse" lattice Λ_c, denoted by $\mathcal{V}(\Lambda_c)$. $\Lambda, \Lambda_c \subset \mathbf{R}^N$. The term "nested" is due to the fact that $\Lambda_c \subset \Lambda$. The rate R_0 of the nested lattice code book $\Lambda \cap \mathcal{V}(\Lambda_c)$ is given by

$$R_0 = \frac{1}{N} \log_2 |\mathcal{V}(\Lambda_c) \cap \Lambda|. \tag{9}$$

where $|S|$ is the cardinality of a set S.

The modulo operation is defined as the quantization error of a point x with respect to the coarse lattice Λ_c:

$$x \bmod \Lambda_c = x - \arg\min_{u \in \Lambda_c} \|x - u\|_2, \tag{10}$$

where $\|x - y\|_2$ is the Euclidean distance between x and y in $\mathbf{R}^N$. It can be verified that $\Lambda \cap \mathcal{V}(\Lambda_c)$ is a finite Abelian group when the addition operation between two elements $x, y \in \Lambda \cap \mathcal{V}(\Lambda_c)$ is defined as

$$x + y \bmod \Lambda_c. \tag{11}$$

The signal X^N transmitted over N channel uses from a nested lattice codebook is given by

$$X^N = (u^N + d^N) \bmod \Lambda_c. \tag{12}$$

Here u^N is the lattice point chosen from $\Lambda \cap \mathcal{V}(\Lambda_c)$, and d^N is the dithering vector [32].

Remark 2: Conventionally, d^N is defined as a continuous random vector which is uniformly distributed over $\mathcal{V}(\Lambda_c)$ [32]. This so called dithering vector is used to facilitate the analysis of the probability of decoding errors for nested lattice codes, since it is easier to bound this probability averaged over the random dithering vectors than bounding it for a fixed dithering vector. The same technique is used here to prove that the lattice points can be decoded by the intended receiver in Appendices B and C. Note that, after proving the decodability, the random dithering vector can be replaced by a fixed dithering vector, since the performance under the former case is just the average performance under a fixed dithering vector, and thus must be attainable by at least one fixed dithering vector, see [31, Section 5.12] for details. Hence, in the main sections of the paper, we shall assume that d^N is a fixed vector and is perfectly known by all receiving nodes. □

[3]Although the real channel model can be viewed as a special case of the complex model, the scheme in [31, Section 5.16] does not yield a positive secure rate for the real channel model.

A. Coding Scheme

For our coding scheme, both node S_1 and S_2 shall use the same nested lattice code. Their transmitted signals are expressed in the form of (12):

$$X_i^N = (u_i^N + d_i^N) \bmod \Lambda_c, i = 1, 2. \tag{13}$$

The lattice point u_1^N corresponds to the effective channel input which conveys information. u_2^N is the judiciously introduced interference for cooperative jamming and is sampled in an i.i.d. fashion from the nested lattice codebook according to a uniform distribution.

The codeword used by node S_1 is a concatenation of m length-N vectors where m is a positive integer and each length-N vector is a shifted fine lattice point given by (13) for $i = 1$. The codebook contains 2^{nR_0} such codewords, where $n = mN$ and R_0 is given by (9). Each codeword is generated by (i) obtaining m independent copies of u_1^N by sampling uniformly and independently by m times from the set $\Lambda \cap \mathcal{V}(\Lambda_c)$, (ii) using (13) to translate each copy of u_1^N to X_1^N, (iii) concatenating the resulting m copies of X_1^N's to form a codeword.

The confidential message is mapped to the codeword using the wiretap coding scheme [2]: The codewords are randomly categorized into $2^{n[R_0-1]^+}$ bins. The encoder chooses a bin based on the realization of the confidential message, and randomly chooses a codeword from the bin to transmit. The overall coding scheme is a serial concatenation of a wiretap coding scheme and the nested lattice code. The nested lattice code is the inner code. The wiretap coding scheme is the outer code, which is used to eliminate the 1-bit information leak to achieve perfect secrecy mentioned in Section I. Although the wiretap codebook is randomly generated, it is not completely random since every N component of the codeword has to be a lattice point. This is consistent with the interesting terminology "structured random code" [17].

B. Secrecy Rate Computation

If each N group of channel uses are viewed as a single (meta) channel use with u_1^N being its input, we have in effect a *memoryless* wiretap channel at hand. This allows us to leverage the following result from reference [3]: Consider a memoryless wiretap channel described by $\Pr(Y, Z|X)$, where X is the channel input, Y is the observation of the legitimate receiver, Z is the observation of the eavesdropper. Then for a given input distribution $\Pr(X)$, any secrecy rate R_s such that

$$0 \le R_s < [I(X;Y) - I(X;Z)]^+ \tag{14}$$

is achievable where $[x]^+$ equals x if $x \geq 0$ and 0 otherwise.

In our case, this means that any R_s such that

$$0 \leq R_s < \tfrac{1}{N}[I(u_1^N; Y_1^N) - I(u_1^N; Y_2^N)]^+ \tag{15}$$

is achievable for the model in (8).

When evaluating (15), we expect that the first term $\frac{1}{N} I(u_1^N; Y_1^N)$ should be easy to compute. It should approach the rate of the lattice codebook R_0 as N increases, since this is the same AWGN setting considered by [32] and the receiver D_1 should be able to correctly decode the lattice points sent by S_1 if the lattice codebook is designed properly. Computing the exact value for *the second term* in (15) is challenging. However, as we shall show below, upper bounding its value is feasible. Since any value within the range given by (15) is an achievable secrecy rate, an upper bound on the $I(u_1^N; Y_2^N)$ leads to a computable achievable secrecy rate result.

We start with

$$\begin{aligned} & I(u_1^N; Y_2^N) && (16)\\ \leq\ & I(u_1^N; X_1^N \pm X_2^N, Z_2^N) && (17)\\ =\ & I(u_1^N; X_1^N \pm X_2^N). && (18)\end{aligned}$$

To bound (18), we introduce a new tool, which we will term the *representation theorem* from here on:

Theorem 1: Let $t_1^N, t_2^N, \ldots, t_K^N$ be K N-dimensional vectors taken from the fundamental Voronoi region of a given lattice Λ. There exists an integer T, such that $1 \leq T \leq K^N$, and $\sum_{k=1}^{K} t_k^N$ is uniquely determined by $\{T, \sum_{k=1}^{K} t_k^N \bmod \Lambda\}$. T is a function of $\sum_{k=1}^{K} t_k^N$.

Proof: The proof is given in Appendix A. ∎

Remark 3: As shown in its proof, Theorem 1 is a purely algebraic result and does not rely on the statistics of $t_1^N, t_2^N, \ldots t_K^N$. □

When $K = 2$, we have the following corollary:

Corollary 1: For $X_i^N, i = 1, 2$ computed according to (12), i.e., $X_i^N = (u_i^N + d_i^N) \bmod \Lambda_c$, there exists an integer T, such that $1 \leq T \leq 2^N$, and $X_1^N \pm X_2^N$ is uniquely determined by $\{T, X_1^N \pm X_2^N \bmod \Lambda_c\}$. T is a function of $X_1^N \pm X_2^N$.

Proof: Define $-\Lambda_c = \{-x^N : x^N \in \Lambda_c\}$. Since $0 \in \Lambda_c$ and the difference of any two lattice points is a lattice point, we have $-\Lambda_c = \Lambda_c$. This means that if $X_2^N \in \mathcal{V}(\Lambda_c)$, then $-X_2^N \in \mathcal{V}(-\Lambda_c)$. Since $-\Lambda_c = \Lambda_c$, this means that $-X_2^N \in \mathcal{V}(\Lambda_c)$. Hence the corollary follows from Theorem 1 by letting $t_1^N = X_1^N$ and $t_2^N = \pm X_2^N$. ∎

We next return to (18). Using Corollary 1, we find (18) can be written as:

$$\begin{aligned} & I(u_1^N; X_1^N \pm X_2^N) = I(u_1^N; X_1^N \pm X_2^N \bmod \Lambda_c, T) && (19)\\ =\ & I\left(u_1^N; X_1^N \pm X_2^N \bmod \Lambda_c\right) + I\left(u_1^N; T | X_1^N \pm X_2^N \bmod \Lambda_c\right) && (20)\\ \leq\ & I\left(u_1^N; X_1^N \pm X_2^N \bmod \Lambda_c\right) + H\left(T | X_1^N \pm X_2^N \bmod \Lambda_c\right) && (21)\\ \leq\ & I(u_1^N; X_1^N \pm X_2^N \bmod \Lambda_c) + H(T) && (22)\\ =\ & I(u_1^N; u_1^N \pm u_2^N \bmod \Lambda_c) + H(T), && (23)\end{aligned}$$

where T is the integer defined in Corollary 1. (23) follows from (22) since d_1^N and d_2^N are known by both the transmitters and the receivers.

Since $\Lambda \cap \mathcal{V}(\Lambda_c)$ is an Abelian group, when u_2^N is independent from u_1^N, and u_2^N is uniformly distributed over $\Lambda \cap \mathcal{V}(\Lambda_c)$, we have [33], [34]:

$$I(u_1^N; u_1^N \pm u_2^N \bmod \Lambda_c) = 0. \tag{24}$$

Applying it to (23), we find that (18) is upper bounded by

$$H(T) \leq N. \tag{25}$$

Equations (19)-(25) imply

$$\frac{1}{N} I\left(u_1^N; X_1^N \pm X_2^N\right) \leq 1. \tag{26}$$

Applying this result back to (15), we find that the secrecy rate approaches $[R_0 - 1]^+$ as N increases.

Remark 4: Note that X_i^N sent by Node S_i is always within the Voronoi region of Λ_c. As mentioned earlier, this restriction leads to a small amount of information being leaked which is eliminated by using the wiretap channel code as an outer code. Recently [35] proposed a lattice Gaussian signaling scheme for the Gaussian wiretap channel in which the transmitted signals could also be sampled outside of the Voronoi region of Λ_c and it was shown in [36] that this new scheme eliminates this information leakage without introducing a wiretap code as an outer code. □

IV. Achievable Secure Degrees of Freedom With Nested Lattice Codes

We next apply the procedure developed in Section III to the fully connected model (1).

We begin by reformulating the channel in (1). We notice that any $\sqrt{ab} \neq 0$, can be represented in the following form:

$$\sqrt{ab} = p/q + \gamma/q \tag{27}$$

where p, q are positive integers, and $-1 < \gamma < 1, \gamma \neq 0$. In this case, the channel model (1) can be expressed as:

$$\begin{aligned} qY_1 &= qX_1 + (p + \gamma) X_2 + q\sqrt{b} Z_1 && (28)\\ Y_2 &= X_1 \pm X_2 + Z_2. && (29)\end{aligned}$$

Using this notation, we have the following theorem regarding the achievable secrecy rate:

Theorem 2: For a given positive integer M, define P_{total} as

$$P_{total} = \frac{(\alpha\beta + 1)^M - 1}{\beta} q^2 b \tag{30}$$

where for $|\gamma| \leq 1/2$,

$$\alpha = \frac{1 - 2\gamma^2 + \sqrt{1 - 4\gamma^2}}{2\gamma^4} \tag{31}$$

and

$$\beta = q^2 + (p + \gamma)^2. \tag{32}$$

If P_{total} is available to node S_1 and S_2 as transmission power, i.e., $P_{total} \leq \min\{\bar{P}_1, \bar{P}_2\}$, then the following secrecy rate R_s is achievable for the channel model (1).

$$R_s = [(\frac{1}{4}\log_2(\alpha) - 1)M]^+. \tag{33}$$

Proof: Compared to the channel model in (8), the added complexity here is that node D_1 is interfered by node S_2. We use a layered coding scheme [37] to eliminate this interference. This scheme involves technical details related to how nested lattice codes are decoded at the receiver. For clarity, we keep these details in Appendix B while providing the essential steps on the computation of the secrecy rate as follows.

Let X_k^N be the signal sent by node S_k over N channel uses. In a layered coding scheme, X_k^N is the sum of codewords from M layers as shown below:

$$X_k^N = \sum_{i=1}^{M} X_{k,i}^N, \quad k = 1, 2. \tag{34}$$

$X_{k,i}^N$ is the signal sent by the node S_k in the ith layer.

For each layer, we use the nested lattice code described in Section III. The signal $X_{k,i}^N$ is computed as:

$$X_{k,i}^N = \left(u_{k,i}^N + d_{k,i}^N\right) \bmod \Lambda_{c,i} \quad k = 1, 2, i = 1, \ldots, M \tag{35}$$

where $d_{k,i}^N$ is the dithering vector. Let $u_{k,i}^N$ be the lattice point such that:

$$u_{k,i}^N \in \mathcal{V}\left(\Lambda_{c,i}\right) \cap \Lambda_i, \quad k = 1, 2. \tag{36}$$

Note that both node S_1 and S_2 use the same lattice codebook for each layer.

Define $\mathcal{M}$ as the set $\{1, \ldots, M\}$. We use the shorthand $A_{\mathcal{M}}$ to represent a set of vectors $A_i, i = 1, \ldots, M$. If the rate and the power are allocated properly among different layers, the intended receiver should be able to decode $u_{1,\mathcal{M}}^N$ with high probability. We denote the rate of the lattice codebook for the ith layer with R_i and its power by P_i. The value of R_i and P_i are established in Appendix B by (82)–(83) and (80) respectively.

As before, we then view each group of N channel uses as a single channel use and view $u_{1,\mathcal{M}}^N$ as the effective inputs of the equivalent channel. The distribution for $u_{2,\mathcal{M}}^N$ is chosen such that $u_{2,i}^N, i = 1, \ldots, M$ are independent and are also independent from $u_{1,\mathcal{M}}^N$. Each $u_{2,i}^N$ is uniformly distributed over $\mathcal{V}\left(\Lambda_{c,i}\right) \cap \Lambda_i$. Then, the signal transmitted by node S_2 is independent between every block of N-channel uses. This ensures the channel is memoryless over every N-channel uses and hence the result (14) from [3] can be used to derive the secrecy rate, i.e., secrecy rate R_s such that

$$0 \le R_s \le [\lim_{N\to\infty} \frac{1}{N} \left(I(u_{1,\mathcal{M}}^N; Y_1^N) - I(u_{1,\mathcal{M}}^N; Y_2^N)\right)]^+ \tag{37}$$

is achievable. To compute the secrecy rate, again we resort to deriving a lower bound to the right hand side of (37). We choose the distribution of $u_{1,\mathcal{M}}^N$ to be the same one as that of $u_{2,\mathcal{M}}^N$. This means that $u_{1,i}^N, i = 1, \ldots, M$ are independent and each of them is uniformly distributed over $\mathcal{V}\left(\Lambda_{c,i}\right) \cap \Lambda_i$. For this distribution,

$$\lim_{N\to\infty} \frac{1}{N} H(u_{1,\mathcal{M}}^N) = \sum_{i=1}^{M} R_i. \tag{38}$$

On the other hand, we know that, for a given M, $u_{1,\mathcal{M}}^N$ can be reliably decoded from Y_1^N. The decoding procedure is given in Appendix B. Let P_e denote the probability of decoding error which converges to 0 as N goes to ∞ [4]. Then, by Fano's inequality [38], we have

$$\frac{1}{N} H\left(u_{1,\mathcal{M}}^N | Y_1^N\right) \tag{39}$$

$$\le \frac{1}{N}\left(1 + P_e H(u_{1,\mathcal{M}}^N)\right) = \frac{1}{N} + P_e \sum_{i=1}^{M} R_i. \tag{40}$$

Therefore

$$\lim_{N\to\infty} \frac{1}{N} I(u_{1,\mathcal{M}}^N; Y_1^N) \tag{41}$$

$$= \lim_{N\to\infty} \frac{1}{N}(H(u_{1,\mathcal{M}}^N) - H\left(u_{1,\mathcal{M}}^N | Y_1^N\right)) \tag{42}$$

$$\ge \sum_{i=1}^{M} R_i - \frac{1}{N} - P_e \sum_{i=1}^{M} R_i. \tag{43}$$

By letting $N \to \infty$, (38) and (41)-(43) imply:

$$\lim_{N\to\infty} \frac{1}{N}(I(u_{1,\mathcal{M}}^N; Y_1^N) = \sum_{i=1}^{M} R_i, \tag{44}$$

where R_i is given by (83) in Appendix B.

The second term in (37) can be upper bounded as follows:

$$\frac{1}{N} I\left(u_{1,\mathcal{M}}^N; Y_2^N\right) \tag{45}$$

$$\le \frac{1}{N} I\left(u_{1,\mathcal{M}}^N; X_1^N \pm X_2^N\right) \tag{46}$$

$$\le \frac{1}{N} \sum_{i=1}^{M} I\left(u_{1,i}^N; X_{1,i}^N \pm X_{2,i}^N\right) \tag{47}$$

$$\le M. \tag{48}$$

(46) follows from the fact that the channel noise is independent from $u_{1,\mathcal{M}}^N$ and $X_1^N \pm X_2^N$. (47) is because the jamming signal $X_{2,i}^N$ of different layers are independent from each other. Finally, we apply (19)-(25) to each term inside the sum in (47) to obtain (48).

Substituting (44) and (48) into (37), we find that the following secrecy rate is achievable.

$$R_s = [\textstyle\sum_{i=1}^{M} R_i - M]^+. \tag{49}$$

We next apply the expression for R_i given in (82)-(83). This leads to the secrecy rate (33) claimed in the theorem.

The total power consumed by node $S_k, k = 1, 2$ can be computed by summing the expression for P_i in (80) from $i = 1$ to $i = M$, which leads to and is given by (30).

Hence we have completed the proof of the theorem. ■

Remark 5: Compared to Section III, the only difference is a layered coding scheme is used as the inner code. The outer code is still the stochastic wiretap coding scheme as described in Section III. Codewords are sampled from the M-fold Cartesian product of nested lattice code sets with a uniform distribution, with M being the number of layers and each component of the codeword is an M-tuple of lattice points. □

[4]It can be shown that this property is retained when the dithering vectors are fixed; see [31, Section 5.12] for details.

Remark 6: Note that P_i in (80) is computed with random dithering vectors. After fixing the dithering vector, the actual power used by the ith layer is not necessarily P_i. However, it can be shown that by choosing the fixed dithering vector properly, the difference in total power consumption caused by fixing the dithering vectors can be made negligible [31, Section 5.12]. □

From Theorem 2, we have the following corollary.

Corollary 2: The following number of secure degrees of freedom is achievable using nested lattice codes when $0 < |\gamma| < 0.5$:

$$\left[\frac{\frac{1}{4}\log_2(\alpha) - 1}{\frac{1}{2}\log_2(\alpha\beta + 1)}\right]^+, \tag{50}$$

where α and β are given by (31) and (32) respectively.

Proof: The corollary follows by substituting P_{total} and R_s into (7) and letting M go ∞. ■

Remark 7: A layered coding scheme with lattice codes was used in [37] for a K-user interference channel ($K \geq 3$) without secrecy constraints. The difference here from [37] is that in [37], a sphere shaped lattice code is used for each layer. Here, for each layer, a nested lattice code is used instead. As a result, the corresponding decoding algorithm and the error probability analysis are different. Reference [37] uses results from [39]. The rate derivation in our work uses results from reference [32]. □

A consequence of Corollary 2 is as follows:

Corollary 3: For $\sqrt{ab}$, such that $2\sqrt{ab}$ is not an integer and $1/\sqrt{ab}$ is not an integer, the number of secure degrees of freedom given by Corollary 2 is positive.

Proof: This corollary is proved in Appendix D. ■

V. Achieving Positive Secure Degrees of Freedom for Channel Gains not Covered by Corollary 3

Corollary 3 has shown that the secure degrees of freedom (s.d.o.f.) for the Gaussian wiretap channel with a cooperative jammer is positive with probability one if the channel gains are sampled from a continuous distribution. We next show how to achieve a positive s.d.o.f. for the cases not covered by Corollary 3. This requires a different decoding algorithm to be used at the intended receiver. For simplicity, the coding scheme shall be described for use with lattices of one-dimension only, as it is not clear how to prove the decodability for the general N-dimensional lattice used in the previous sections.

For parameter Q, a one-dimensional lattice is composed of points in the set $[0, Q) \cap \mathbf{Z}$ where $\mathbf{Z}$ is the set of all integers. The point can be scaled and shifted to obtain the actual transmitted signal.

Theorem 3: Using a one-dimensional lattice, we have achievable secure degrees of freedom results for the following four cases:

1) When $\sqrt{ab}$ is an algebraic irrational number, the secure degrees of freedom of $1/2$ is achievable.
2) When $\sqrt{ab} \in (0, 1/2] \cup [2, +\infty)$, we have the following result. Let $Q = \sqrt{ab}$ if $\sqrt{ab} \geq 2$. Otherwise, let $Q = 1/\sqrt{ab}$. Let $\lfloor Q \rfloor$ denote the largest integer smaller than or equal to Q. The following secure degrees of freedom is achievable:
$$\left[\frac{1}{2}\frac{\log_2 \lfloor Q \rfloor}{\log_2 Q} - \frac{f(\lfloor Q \rfloor)}{2\log_2 Q}\right]^+ \tag{51}$$
where $f(Q)$ is defined in (101). (51) is lower bounded by
$$\frac{1}{2}\frac{\log_2 \lfloor Q \rfloor}{\log_2 Q} - \frac{\log_2\left(2\pi e\left(\frac{1}{6}\right) - \frac{1}{12\lfloor Q \rfloor^2}\right)}{4\log_2(Q)}. \tag{52}$$
For $Q = 2$, (51) equals $1/4$.
3) When $\sqrt{ab} = 1$ and $Y_2 = X_1 - X_2 + Z_2$, the secure degrees of freedom of 0.0548 is achievable.
4) When $\sqrt{ab} = 1.5$, secure degrees of freedom of $1/6$ is achievable.

Proof: The proof for these four cases are provided in the subsections of Appendix E. ■

Remark 8: When $\sqrt{ab} = 1$ and all channel gains are positive, the channel is degraded and from the outer bound in [27], the number of secure degrees of freedom is 0. Since algebraic irrational numbers are dense on the real line, it follows that the number of secure degrees of freedom is discontinuous at $\sqrt{ab} = 1$. □

Theorem 3 and Corollary 3 together cover all possible $\sqrt{ab}$. Hence, we arrive at the following corollary:

Corollary 4: For the channel model in (1), a positive number of secure degrees of freedom is achievable except for the following case for which [10] proved this is not possible:

$$\begin{aligned} Y_1 &= X_1 + X_2 + \sqrt{b}Z_1 \\ Y_2 &= X_1 + X_2 + Z_2. \end{aligned} \tag{53}$$

VI. Numerical Results

We next provide numerical results for the rates in Theorem 2 and Theorem 3.

A. Secrecy Rate

In Fig. 2, we plot the secrecy rate versus power when $a = b = \sqrt{\frac{2}{3}}$, which makes $\sqrt{ab}$ an algebraic irrational number. The power P_{total} is the variance of X_1 or X_2 in (1).

For the integer lattice code, we use (114) to compute the secrecy rate; this provides the largest secure degrees of freedom, which is $1/2$.

From (104) and (105), the average power of this coding scheme is given by

$$P_{total} = P^{1/2+2\varepsilon}\frac{Q^2 - 1}{12}. \tag{54}$$

The first term in (54) is due to the scaling factor $P^{1/4+\varepsilon}$ in (105). We then choose different values for ε and plot the largest achievable power rate pair region in terms of $\{10\log_{10} P_{total}, R_s\}$ in Fig. 2.

For comparison, we also plot the largest possible secrecy rate offered by Gaussian random codebooks presented in [27]. This is done by removing the power constraints of the transmitters and computing the secrecy rate optimized over the

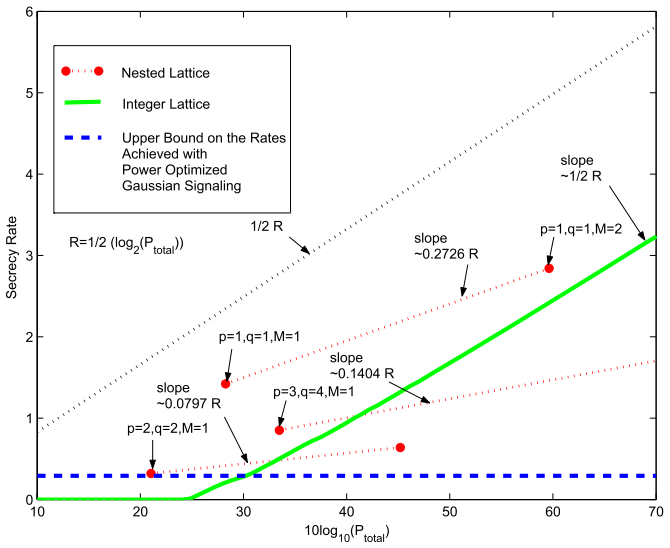

Fig. 2. Secrecy Rate with Finite Power, $a = b = \sqrt{\frac{2}{3}}$, P_{total} is the power of X_1 or X_2 in (1). $1/2R$ is shown as a reference line.

transmission powers. Clearly, this serves as an upper bound on the secrecy rate achievable in [27]. Since $a = b = \sqrt{2/3} < 1$, this upper bound can be computed following the derivation in [27, (11)][5]. The power unconstrained secrecy rate offered by the Gaussian random codebook is given by

$$\lim_{\bar{P}_k \to \infty, k=1,2} R_s = \frac{1}{2} \log_2 \frac{1}{ab}. \tag{55}$$

As shown by Fig. 2, the secrecy rate offered by the integer lattice is greater than the power unconstrained Gaussian signaling scheme when $10 \log_{10} P_{total} > 30 dB$.

Finally, we plot, in Fig. 2, the secrecy rate offered by the nested lattice code, by choosing different p, q values and the number of layers M. The secrecy rate is computed according to (33). P_{total} is given by (30). As expected, different choices of p, q result in different slopes for the secrecy rate curve, i.e., different secure degrees of freedom. When $10 \log_{10} P_{total} < 60 dB$, the nested lattice code can achieve a larger secrecy rate than integer lattice code due to its power efficiency. However, if P_{total} keeps increasing, the secrecy rate offered by the integer lattice is the largest, due to its higher achievable secure degrees of freedom.

To summarize, these results demonstrate that the coding scheme presented in this paper can provide a larger secrecy rate than random Gaussian signaling at medium to high SNR values.

B. Secure Degrees of Freedom

It is clear that the integer lattice provides a larger value of secure degrees of freedom for $\sqrt{ab}$ values in Case 1), 3), 4) in Theorem 3.

We next consider the remaining Case 2) in Theorem 3.

For clarity, we first plot in Fig. 3 the secure degrees of freedom achieved by this case. The actual performance

[5]Reference [27] only considers the case $Y_2 = X_1 + X_2 + Z_2$. However, since the Gaussian distribution is symmetric around zero, the case $Y_2 = X_1 - X_2 + Z_2$ has the same secrecy rate with Gaussian input distributions.

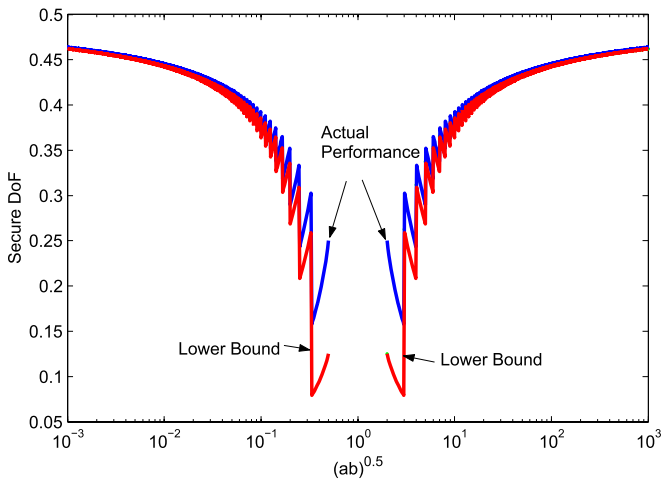

Fig. 3. Secure degrees of freedom achieved by integer lattice coding scheme in Theorem 3 Case 2.

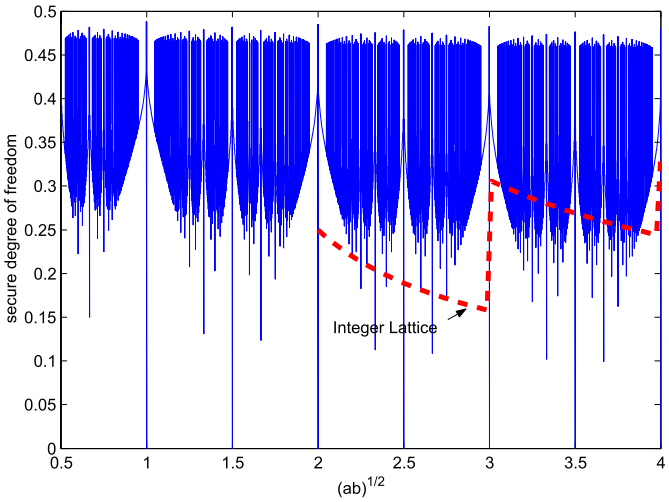

Fig. 4. Comparison of secure degrees of freedom achieved by nested lattice code (solid blue) and by integer lattice coding scheme in Theorem 3 Case 2 (dashed red).

curve in Fig. 3 corresponds to (51). The lower bound curve corresponds to (52). The zigzag shape of the curve is a consequence of the $\lfloor \ \rfloor$ operation on Q in (51). We notice as $\sqrt{ab}$ moves away from 1, the lower bound given by (52) becomes tighter, and the number of secure degrees of freedom converges to 0.5.

We next compare it with the secure degrees of freedom achieved by nested lattice code in Fig. 4. As shown by Fig. 4, neither scheme dominates the other in performance for all channel gains. The nested lattice code offers a larger number of secure degrees of freedom near the peak of spurs, while the integer lattice coding scheme described in this section is superior when $2\sqrt{ab}$ are integers and $\sqrt{ab} \geq 2$.

Remark 9: The smallest known upper bound on the number of secure degrees of freedom is found in [31] to be $2/3$, which still has a nonzero gap from $1/2$.

VII. Channel Gain Mismatch

The coding schemes presented in previous sections require aligning lattice points at the eavesdropper, which relies on

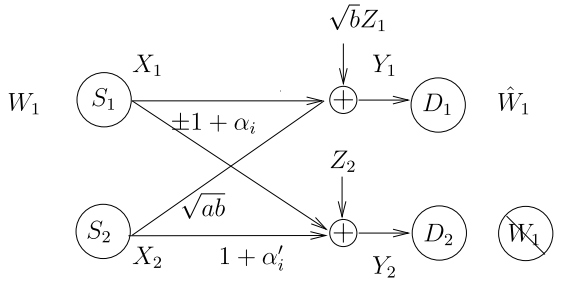

Fig. 5. Wiretap channel with a cooperative jammer S_2, where the channel gain between S_1 and D_1 has an estimation error $\alpha_i, \alpha_i' \in [-\alpha_{\max}, \alpha_{\max}]$ for the ith channel use. α_i and α_i' are known by D_2 but not known by S_1 or S_2.

accurate channel state information. It is conceivable that in practice such accurate channel state information may be difficult to obtain and it is reasonable to ask if the nested lattice coding scheme is still able to provide secrecy with imperfect channel state information. In this section, we explain how to compute the secrecy rate in this case.

The channel model is shown in Fig. 5. This is the same channel model we have been using except now, the channel gain between S_1 and D_2 has an estimation error α_i for the ith channel use and the channel gain between S_2 and D_2 has an estimation error α_i' for the ith channel use. For simplicity, we assume the remaining channels are estimated perfectly. The channel estimation errors, α_i, α_i', are independent from the signal transmitted by node S_1 and S_2, and $\alpha_i, \alpha_i' \in [-\alpha_{\max}, \alpha_{\max}]$. $\alpha_{\max} > 0$. We also assume the eavesdropper, node D_2, has perfect knowledge of α_i, α_i', but the other nodes only know $\alpha_{\max}$. In this setting, perfect channel state information at the eavesdropper is obviously a pessimistic assumption and is a measure of worst case performance.

With these assumptions, we have the following theorem:

Theorem 4: Let $\bar{P}_1 = \bar{P}_2 = \rho$ be the total power consumption for node S_1 and S_2. The *same* number of secure degrees of freedom given by Theorem 2 is achievable if for a constant $c > 0$,

$$\alpha_{\max}^2 \leq \frac{c}{\rho}. \tag{56}$$

Proof: The proof is given in Appendix F. ■

Remark 10: Imperfect channel state information for other links, for example, on the link between legitimate transmission pair S_1 and D_1 can be catered to by using the existing results for the model without secrecy constraints in [40, Section IV]. □

VIII. Conclusion

Structured codes were shown to be useful in a number of information theory problems. In this work, we have shown that structured codes are useful with cooperative jamming to prove higher achievable secrecy rates. For a Gaussian wiretap channel with a cooperative jammer, we have proved that a positive number of secure degrees of freedom is achievable as long as the channel is not degraded and is fully connected. Since the considered channel model is a special case of the multiple access wiretap channel, the two-user interference channel with confidential messages, and the two-user interference channel with an external eavesdropper, this result implies that the secure degrees of freedom for all these channels are positive as well. This is in contrast to earlier literature where Gaussian signaling was used and the achieved secure degrees of freedom was zero.

Appendix A
Proof of Theorem 1

Let $\mathcal{V}$ be the fundamental Voronoi region of the N-dimensional lattice Λ. For any set A, define αA as $\alpha A = \{\alpha x : x \in A\}$. Then we have:

$$\{\sum_{k=1}^{K} t_k^N : t_k^N \in \mathcal{V}, k = 1...K\} = K\mathcal{V}. \tag{57}$$

Theorem 1 follows if there are exactly K^N points in $K\mathcal{V}$ mapped to the same point in $\mathcal{V}$ by the modulus Λ operation. By (57), each point in $K\mathcal{V}$ can be expressed as $\sum_{k=1}^{K} t_k^N$, $t_k^N \in \mathcal{V}$. After the modulus Λ operation, it is mapped to $t'^N \triangleq \sum_{k=1}^{K} t_k^N \bmod \Lambda$. These two points are related by:

$$t'^N = \sum_{k=1}^{K} t_k^N + x^N, \quad \text{for some} \quad x^N \in \Lambda. \tag{58}$$

Hence we need to show there are exactly K^N possible lattice points for x^N for which we can find $t_k^N \in \mathcal{V}, k = 1...K$ such that (58) holds. This is equivalent to finding x^Ns that satisfy the following equation:

$$t'^N - x^N \in K\mathcal{V}, x^N \in \Lambda. \tag{59}$$

To show that there are exactly K^N x^N's in Λ that satisfy (59), we need to examine the structure of lattice Λ. Each point in the lattice, by definition, can be represented in the following form [41]: $x^N = \sum_{i=1}^{N} a_i v_i^N$, $v_i^N \in \mathbf{R}^N, a_i \in \mathbf{Z}$. $\{a_i\}$ is said to be the coordinates of the lattice point x under the basis described by vectors $\{v_i^N\}$. The set of basis vectors $\{v_i^N\}$ is a fundamental characterization of a lattice and does not change between lattice points. Based on this representation, we can define the following relationship $\sim$: Consider two points $x^N, y^N \in \Lambda$, with coordinates $\{a_i\}$ and $\{b_i\}$ respectively. Then $x^N \sim y^N$ if $a_i = b_i \bmod K$, $i = 1...N$. It is easy to see the relationship $\sim$ is an equivalence relationship. Therefore, it defines a partition over Λ. A partition defined in this way has the following properties:

1) There are K^N sets in this partition. This is because $a_i \bmod K$ can take values $0, 1, 2, \ldots, K-1$ and the coordinate $\{a_i\}$ has N components.
2) The sub-lattice $K\Lambda$ corresponds to one set in the partition for which $a_i \bmod K = 0, i = 1, \ldots, N$. The remaining $K^N - 1$ sets are its cosets.

Let C_i denote any one of these cosets or $K\Lambda$. Then C_i can expressed as $C_i = K\Lambda + y_i^N$, $y_i^N \in \Lambda$. It is easy to verify that $\{x^N + K\mathcal{V}, x^N \in C_i\}$ is a partition of $K\mathbf{R}^N + y_i^N$, which equals $\mathbf{R}^N$.

We now return to solve (59). Since $C_i, i = 1, ..., K^N$ is a partition of Λ, (59) can be solved by considering the following K^N equations:

$$t'^N - x^N \in K\mathcal{V}, \quad x^N \in C_i \tag{60}$$

This is equivalent to saying $t'^N \in x^N + K\mathcal{V}$ for some $x^N \in C_i$. Since $\{x^N + K\mathcal{V}, x^N \in C_i\}$ is a partition of $\mathbf{R}^N$, there is exactly one $x^N \in C_i$ that satisfies this requirement. This implies that for a given t'^N, and a given coset C_i, (60) has exactly one solution for x^N. Since there are K^N C_i's and hence K^N such equations given by (60) each yielding a distinct solution for x^N, equation (59) has exactly K^N solutions. By letting T to be the integer that indexes these K^N solutions, we obtain the theorem.

Appendix B
The Lattice Decoder for Theorem 2

In this section, we describe the decoding procedure at node D_1 and compute the rate and power for each layer in the layered coding scheme used in the proof of Theorem 2. We leave the technical details of decoding nested lattice codes to Appendix C for the interested reader and focus on the power and rate computation for the layered coding scheme in this section.

Let R_i be the rate of the codebook for the ith layer:

$$R_i = \frac{1}{N} \log_2 |\mathcal{V}(\Lambda_{c,i}) \cap \Lambda_i|. \tag{61}$$

We assume in this appendix that $d_{k,i}^N$ is uniformly distributed over $\mathcal{V}(\Lambda_{c,i})$, perfectly known by all receiving nodes and independently generated for each node, each layer and each block of N channel uses. As we explained in Remark 2, this technical detail is necessary to show the decodability of the lattice code, and after proving decodability, they can be replaced with fixed vectors.

Since $d_{k,i}^N$ is uniformly distributed over $\mathcal{V}(\Lambda_{c,i})$, the average power per dimension of the ith layer P_i can be computed as:

$$P_i = \frac{1}{N \ \mathbf{vol}(\mathcal{V}(\Lambda_{c,i}))} \int_{x \in \mathcal{V}(\Lambda_{c,i})} \|x\|_2^2 \, dx \tag{62}$$

where $\mathbf{vol}(\mathcal{V}(\Lambda_{c,i}))$ is the volume of the set $\mathcal{V}(\Lambda_{c,i})$.

We label layers with $1, \ldots, M$. The main idea is to decode the higher layer, subtract its contribution, and then continue to process the lower layers.

As shown in (28), Node D_1 receives qY_1, which, due to (34), can be written as:

$$qY_1 = \sum_{t=1}^{M} \left(qX_{1,t}^N + (p+\gamma) X_{2,t}^N \right) + q\sqrt{b} Z_1^N. \tag{63}$$

Assume that node D_1 has successfully decoded layers whose index is greater than i and now is ready to decode layer i. This means that the decoder at node D_1 can subtract the contribution from the former layers to obtain:

$$\sum_{t=1}^{i} \left(qX_{1,t}^N + (p+\gamma) X_{2,t}^N \right) + q\sqrt{b} Z_1^N. \tag{64}$$

Since p and q are both positive integers, $qu_{1,i}^N + pu_{2,i}^N \bmod \Lambda_{c,i} \in \Lambda_i \cap \mathcal{V}(\Lambda_{c,i})$. Node D_1 will first decode the integer part $qu_{1,i}^N + pu_{2,i}^N \bmod \Lambda_{c,i}$ and then decode the fractional part $\gamma u_{2,i}^N$.

To decode $qu_{1,i}^N + pu_{2,i}^N \bmod \Lambda_{c,i}$, D_1 computes

$$\hat{Y}_i = \tag{65}$$

$$\left[\begin{array}{l} \left(qu_{1,i}^N + pu_{2,i}^N \right) + \gamma X_{2,i}^N \\ + \sum_{t=1}^{i-1} \left(qX_{1,t}^N + (p+\gamma) X_{2,t}^N \right) + q\sqrt{b} Z_1^N \end{array} \right] \bmod \Lambda_{c,i} \tag{66}$$

from (64) since it knows $d_{1,i}^N$ and $d_{2,i}^N$.

Define A_i as

$$A_i = \sum_{t=1}^{i-1} \left(q^2 + (p+\gamma)^2 \right) P_t + q^2 b. \tag{67}$$

Then, from Lemma 1 in Appendix C, the probability that D_1 does not correctly decode $qu_{1,i}^N + pu_{2,i}^N \bmod \Lambda_{c,i}$ decreases exponentially fast with the lattice dimension N if

$$R_i \leq \frac{1}{2} \log_2 \left(\frac{P_i}{\gamma^2 P_i + A_i} \right). \tag{68}$$

where $\gamma^2 P_i + A_i$ is the variance of the term treated as noise which is $\gamma X_{2,i}^N + \sum_{t=1}^{i-1} \left(qX_{1,t}^N + (p+\gamma) X_{2,t}^N \right) + q\sqrt{b} Z_1^N$. It is independent from the signal term $qu_{1,i}^N + pu_{2,i}^N \bmod \Lambda_{c,i}$ thanks to the dithering vector $d_{2,i}^N$, hence the requirement of Lemma 1 is satisfied.

After decoding $qu_{1,i}^N + pu_{2,i}^N \bmod \Lambda_{c,i}$, node D_1 can recover:

$$[\gamma X_{2,i}^N + \sum_{t=1}^{i-1} \left(qX_{1,t}^N + (p+\gamma) X_{2,t}^N \right) + q\sqrt{b} Z_1^N] \bmod \Lambda_{c,i} \tag{69}$$

from (66). Then, we use the fact that as long as

$$P_i > \gamma^2 P_i + A_i, \tag{70}$$

(69) equals

$$\gamma X_{2,i}^N + \sum_{t=1}^{i-1} \left(qX_{1,t}^N + (p+\gamma) X_{2,t}^N \right) + q\sqrt{b} Z_1^N \tag{71}$$

with high probability. This fact is formally proved in Lemma 2 in Appendix C.

If (69) does not equal (71), a decoding error occurs.

Node D_1 then evaluates the following expression from (71):

$$\left[k \left(\begin{array}{l} \sum_{t=1}^{i-1} \left(qX_{1,t}^N + (p+\gamma) X_{2,t}^N \right) \\ + \gamma X_{2,i}^N + q\sqrt{b} Z_1^N \end{array} \right) - \gamma d_{2,i}^N \right] \bmod \gamma \Lambda_{c,i} \tag{72}$$

$$= \left[\begin{array}{l} \gamma u_{2,i}^N + (k-1)\gamma X_{2,i}^N + \\ k \left(\sum_{t=1}^{i-1} \left(qX_{1,t}^N + (p+\gamma) X_{2,t}^N \right) + q\sqrt{b} Z_1^N \right) \end{array} \right] \bmod \gamma \Lambda_{c,i} \tag{73}$$

where the scaling variable k is the Minimum Mean Squared Error (MMSE) coefficient introduced to minimize the variance per dimension of the term

$$(k-1)\,\gamma\,X_{2,i}^{N}+k\left(\sum_{t=1}^{i-1}\left(qX_{1,t}^{N}+(p+\gamma)\,X_{2,t}^{N}\right)+q\sqrt{b}Z_{1}^{N}\right). \tag{74}$$

(74) is called the effective noise in [32] and its minimum variance per dimension with the optimal k is given by:

$$\frac{\gamma^{2}P_{i}A_{i}}{\gamma^{2}P_{i}+A_{i}}. \tag{75}$$

The same scaling scheme was used in [32], see α in [32, (13)].

Node D_1 can correctly decode $u_{2,i}^{N}$ from (73) with high probability if

$$R_{i}\leq\frac{1}{2}\log_{2}\left(\frac{\gamma^{2}P_{i}}{\left(\frac{\gamma^{2}P_{i}A_{i}}{\gamma^{2}P_{i}+A_{i}}\right)}\right)=\frac{1}{2}\log_{2}\left(1+\frac{\gamma^{2}P_{i}}{A_{i}}\right). \tag{76}$$

Again this is a consequence of applying Lemma 1.

After decoding $u_{2,i}^{N}$, node D_1 can recover the following signal from (71):

$$\sum_{t=1}^{i-1}\left(qX_{1,t}^{N}+(p+\gamma)\,X_{2,t}^{N}\right)+q\sqrt{b}Z_{1}^{N} \tag{77}$$

which is identical to (64) with i replaced by $i-1$. This will be used by D_1 when decoding lower layers.

We next determine the power P_i and the rate R_i of each layer. As in [37], [42], we let the right hand side of (68) equal the right hand side of (76):

$$\frac{P_{i}}{\gamma^{2}P_{i}+A_{i}}=1+\frac{\gamma^{2}P_{i}}{A_{i}}. \tag{78}$$

It is easy to check that, with $\alpha=\frac{1-2\gamma^{2}+\sqrt{1-4\gamma^{2}}}{2\gamma^{4}}$, (78) has the following solution[6]:

$$P_{i}=\alpha A_{i}. \tag{79}$$

This leads to (31) in Theorem 2. For α to be a real number, we require $1-4\gamma^{2}\geq 0$, which means $|\gamma|\leq 0.5$.

By solving (79) and (67), we find that P_i is given by:

$$P_{i}=\alpha\,(\alpha\beta+1)^{i-1}\,q^{2}b \tag{80}$$

where $\beta=q^{2}+(p+\gamma)^{2}$. This leads to (32).

For this power allocation, A_i is given by

$$A_{i}=(\alpha\beta+1)^{i-1}\,q^{2}b. \tag{81}$$

Due to (78), R_i can be found by averaging (68) and (76) and substituting (80) and (81) into the result:

$$R_{i}=\frac{0.5}{2}\log_{2}\left(\frac{P_{i}}{\gamma^{2}P_{i}+A_{i}}\right)+\frac{0.5}{2}\log_{2}\left(1+\frac{\gamma^{2}P_{i}}{A_{i}}\right) \tag{82}$$

$$=\frac{0.5}{2}\log_{2}\left(\frac{P_{i}}{A_{i}}\right)=0.25\log_{2}(\alpha). \tag{83}$$

[6]The other solution is when $\alpha=\frac{1-2\gamma^{2}-\sqrt{1-4\gamma^{2}}}{2\gamma^{4}}$. It turns out this solution does not achieve positive secrecy rate.

It remains to check the requirement (70). To do that, we substitute (79) into (70) and get

$$\left(1-\gamma^{2}\right)\alpha>1 \tag{84}$$

where α can be expressed in terms of γ as shown in (31). It can be verified that the left hand side of (84) is always greater than 1 if $|\gamma|<0.5$. Hence (70) is satisfied.

Remark 11: We scaled the signal by k in (73) before performing the modulus operation. Doing so offers a slight gain in secure degrees of freedom than simply choosing $k=1$.

The scaling operation could also have been done in (66). However, the optimal scaling factor has a more complicated expression and it is difficult to derive an analytical expression for the secure degrees of freedom if this approach is followed. □

Appendix C
Some Useful Results on Decoding Nested Lattice Codes

In this section, we introduce the supporting results used in Appendix B which provide the rate constraint when decoding a lattice point from a signal under additive interference from other lattice points.

Consider a nested lattice pair $\{\Lambda,\Lambda_{p,0}\}$. $\Lambda_{p,0}\subset\Lambda$. Define t^N as a lattice point in $\Lambda\cap\mathcal{V}(\Lambda_{p,0})$, which is of interest to the decoder. The decoder observes Y^N given by

$$Y^{N}=\left(t^{N}+\sum_{i=1}^{K}U_{i}^{N}+Z^{N}\right)\bmod\Lambda_{p,0} \tag{85}$$

for $K\geq 0$, where Z^N is the additive channel noise vector composed of zero mean i.i.d. Gaussian random variables, each with variance σ^2. $\sum_{i=1}^{K}U_{i}^{N}$ represents the interference in the observation where U_i^N is a random variable uniformly distributed over the fundamental Voronoi region of a lattice denoted by $\Lambda_{p,i}$. By definition, $\sum_{i=1}^{0}U_{i}^{N}=0$. For convenience, we also define U_0^N which is uniformly distributed over $\mathcal{V}(\Lambda_{p,0})$. We assume $U_i^N, i=0,\ldots,K$ and Z^N are independent and the lattices $\Lambda_{p,i}, i=0,\ldots,K$ are Rogers-good for covering [43, Section I.B] and Poltyrev-good for channel coding [43, Section III.D].

The power of the interference is measured by $\sigma^2(U_i)$ which is the variance per dimension of U_i^N. When N increases, we scale Λ and $\Lambda_{p,i}, i=0,\ldots,K$ such that $\sigma^2(U_i)$ remains unchanged.

Define $\tilde{t}^N$ as the value for t^N decoded from Y^N using an Euclidean distance decoder. Then we have the following result:

Lemma 1: When $K\geq 1$, Y^N is given by (85), and

$$R_{0}<\frac{1}{2}\log_{2}\left(\frac{\sigma^{2}(U_{0})}{\sigma^{2}+\sum_{i=1}^{K}\sigma^{2}(U_{i})}\right), \tag{86}$$

for each N dimensions, there exist lattices $\Lambda,\Lambda_{p,i}$, $i=0,...,K$ such that $\Pr(t^N\neq\tilde{t}^N)$ decreases exponentially fast with N.

Proof: The proof is done by approximating the interference U_i^N with Gaussian noise. Define $O_i^N, i = 1, \ldots, K$, as zero mean Gaussian random variables such that

$$O_i^N \sim \mathcal{N}(\mathbf{0}, \sigma^2(O_i^N)\mathbf{I}) \tag{87}$$

where $\mathbf{I}$ is an $N \times N$ identity matrix. $\sigma^2(O_i^N)$ is a scalar chosen as the variance per dimension of a random variable uniformly distributed over the smallest ball covering $\mathcal{V}(\Lambda_{p,i})$. Then from [32, Lemma 6], there exists lattice $\Lambda_{p,i}$ such that this variance approaches the variance of the uniform variable over the Voronoi region:

$$\lim_{N\to\infty} \sigma^2\left(O_i^N\right) = \sigma^2\left(U_i\right). \tag{88}$$

Let $f_X(x)$ denote the probability density function of any continuous random variable X. We first use the following fact shown in [32, (200)]:

$$f_{U_i^N}(x) \le e^{N\varepsilon\left(\Lambda_{p,i}\right)} f_{O_i^N}(x) \tag{89}$$

which approximates U_i^N with the Gaussian random variable O_i^N. $\varepsilon\left(\Lambda_{p,i}\right)$ is defined [32, (67)], which was shown therein to have the following property:

$$\lim_{N\to\infty} \varepsilon\left(\Lambda_{p,i}\right) = 0. \tag{90}$$

Define $U_{sum}^N = \sum_{i=1}^K U_i^N$ and $O_{sum}^N = \sum_{i=1}^K O_i^N$. From (89), since $U_i^N, i = 1, \ldots, K, Z^N$ are independent, we have

$$f_{U_{sum}^N + Z^N}(x) \le e^{N\sum_{i=1}^K \varepsilon\left(\Lambda_{p,i}\right)} f_{O_{sum}^N + Z^N}(x) \tag{91}$$

which means:

$$\Pr\left(\sum_{i=1}^K U_i^N + Z^N \notin \mathcal{V}(\Lambda)\right) \tag{92}$$

$$\le e^{N\sum_{i=1}^K \varepsilon\left(\Lambda_{p,i}\right)} \Pr\left(\sum_{i=1}^K O_i^N + Z^N \notin \mathcal{V}(\Lambda)\right). \tag{93}$$

Since $\sum_{i=1}^K O_i^N + Z^N$ is an i.i.d. Gaussian vector, we can show that, for a given K,

$$\Pr\left(\sum_{i=1}^K O_i^N + Z^N \notin \mathcal{V}(\Lambda)\right) \tag{94}$$

decreases exponentially fast with N if

$$R_0 < \frac{1}{2}\log_2\left(\frac{\sigma^2(U_0)}{\sigma^2 + \lim_{N\to\infty}\sum_{i=1}^K \sigma^2\left(O_i^N\right)}\right), \tag{95}$$

by repeating the proof of [32, Theorem 5] when we choose the scalar α defined therein to be 1. The right hand side of (95) becomes (86) after applying (88). This and (90) together imply that (93) decreases exponentially fast with N. Therefore (92) decreases exponentially fast with N. Since (92) is an upper bound on $\Pr(t^N \neq \tilde{t}^N)$, we have proved Lemma 1. ∎

The following result is adapted from [32, (89)].

Lemma 2: Define μ as

$$\mu = \frac{\sigma^2(U_0)}{\sigma^2 + \sum_{i=1}^K \sigma^2(U_i)}. \tag{96}$$

Then if $\mu > 1$, the probability

$$\Pr\left(\sum_{i=1}^K U_i^N + Z^N \bmod \Lambda_{p,0} \neq \sum_{i=1}^K U_i^N + Z^N\right) \tag{97}$$

decreases exponentially fast with respect to N.

Proof: (97) is upper bounded by

$$\Pr\left(\sum_{i=1}^K U_i^N + Z^N \notin \mathcal{V}\left(\Lambda_{p,0}\right)\right) \tag{98}$$

which, in turn, by following similar steps which lead to (93), is upper bounded by

$$e^{N\sum_{i=1}^K \varepsilon\left(\Lambda_{p,i}\right)} \Pr\left(\sum_{i=1}^K O_i^N + Z^N \notin \mathcal{V}\left(\Lambda_{p,0}\right)\right). \tag{99}$$

Since $\Lambda_{p,0}$ is taken from the lattice code ensemble defined in [32, Section VII], according to [32, (78)], we have

$$\Pr\left(\sum_{i=1}^K O_i^N + Z^N \notin \mathcal{V}\left(\Lambda_{p,0}\right)\right) \le e^{-N(E_P(\mu) - o_N(1))} \tag{100}$$

where $E_P(\mu)$ is the Poltyrev exponent defined in [32, (56)]. $o_N(1)$ is any function of N such that $\lim_{N\to\infty} o_N(1) = 0$ [32, Section I]. Since $E_p(\mu)$ is positive for $\mu > 1$, we have proved Lemma 2. ∎

APPENDIX D
PROOF OF COROLLARY 3

We first verify Corollary 3 holds for interval $1 \le \sqrt{ab} \le 2$. The value of (50) is plotted in Fig. 6 for $1 < \sqrt{ab} < 2$. To prove that (50) is positive in this range, it suffices to choose $(p = 1, q = 1)$, $(p = 2, q = 1)$, $(p = 3, q = 2)$, and let $\sqrt{ab} = p/q + \gamma/q$. A higher secure number of degrees of freedom can be achieved by choosing other values for p, q, but for clarity, these curves are not plotted in Fig. 6.

Note that for a fixed pair of p, q, by changing γ, (50) takes the shape of a spur. Fig. 6 includes three such spurs. The value of (50) converges to 0.5 when γ converges to 0. However, since $\gamma \neq 0$, the peak of the spur is not included. Hence a positive number of secure degrees of freedom cannot be guaranteed by Corollary 2 only when $\sqrt{ab} = 1, 1.5$ or 2.

We next argue that Corollary 3 holds for interval $i \le \sqrt{ab} \le i+1$ for all integer i, $i \ge 1$. This follows from the fact that the denominator of (50) is always positive. Hence the positivity of (50) is only determined by its numerator, which is only a function of γ. When $i \le \sqrt{ab} \le i+1$, we can simply choose the following three pairs of (p, q): $(p = i, q = 1)$, $(p = i+1, q = 1)$, and $(p = 2i+1, q = 2)$. The positivity of (50) in this interval $[i, i+1]$ should be the same as the positivity of (50) in interval $[1, 2]$. Since (50) is verified to

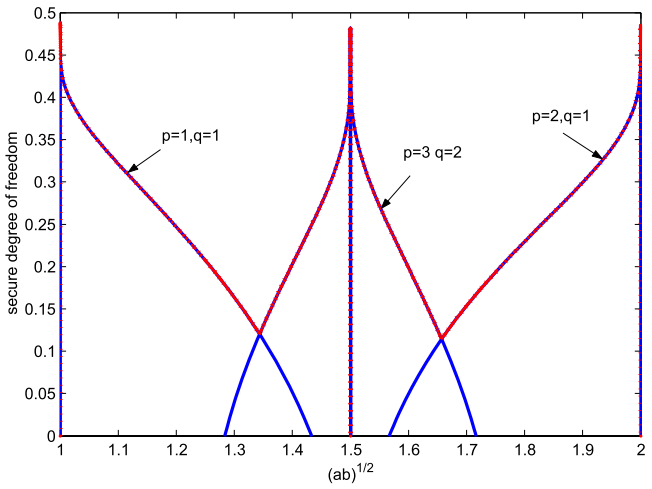

Fig. 6. The value of (50) when $1 < \sqrt{ab} < 2$. The red dashed line is the contour.

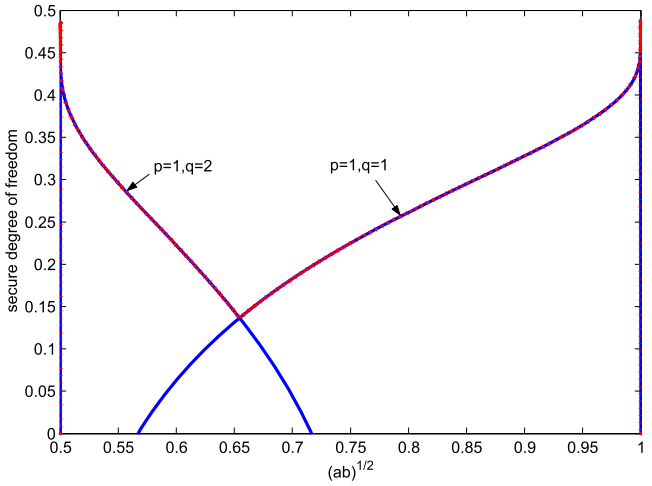

Fig. 7. The value of (50) when $1/2 < \sqrt{ab} < 1$. The red dashed line is the contour.

be positive in interval $[1, 2]$ except for the cases stated in Corollary 3, the same case holds for interval $[i, i+1]$.

We next verify Corollary 3 holds for the interval $1/2 \leq \sqrt{ab} \leq 1$. The value of (50) is plotted in Fig. 7. The two spurs in Fig. 7 follows from choosing $p = 1, q = 1$, and $p = 1, q = 2$.

Finally we consider the interval $1/(i+1) \leq \sqrt{ab} \leq 1/i$ for $i \geq 1$. For this interval, we can choose $p = 1, q = i+1$ and $p = 1, q = i$. Again, since the positivity of (50) is only determined by γ, Corollary 3 holds for this interval $[1/(i+1), 1/i]$ since it holds for interval $1/2 \leq \sqrt{ab} \leq 1$.

Hence we have completed the proof of Corollary 3.

Appendix E
Proof of Theorem 3

A. A Supporting Result

In this section, we present a supporting result. We consider the value of the following function:

$$f(Q) = I(X_1; X_1 \pm X_2) \tag{101}$$

where $X_i, i = 1, 2$ is uniformly distributed over $[0, Q) \cap \mathbf{Z}$. $f(Q)$ can be bounded by the following lemma:

Lemma 3:

$$f(Q) \leq \frac{1}{2}\log_2(2\pi e(\frac{1}{6} - \frac{1}{12Q^2})) < \frac{1}{2}\log_2(\frac{\pi e}{3}) < 0.8. \tag{102}$$

Proof: Equation (102) follows from [22, Lemma 12]. $H(X_1 + X_2)$ can be bounded as

$$H(X_1 + X_2) \leq \frac{1}{2}\log_2(2\pi e(2P' + \frac{1}{12})) \tag{103}$$

where P' is the variance of $X_k, k = 1, 2$, which is given by

$$P' = \frac{1}{Q}\left[\sum_{k=0}^{Q-1} k^2\right] - \left[\sum_{k=0}^{Q-1} k/Q\right]^2 = \frac{Q^2 - 1}{12}. \tag{104}$$

Substituting (104) into (103) we get (102). ■

As shown by (102), the leakage rate is smaller than the 1-bit bound for the nested lattice code, which suggests that integer lattice codes may be useful for cases where the secrecy rate offered by the nested lattice code is 0.

B. Proof of Theorem 3 Case 1

The proof for Theorem 3 Case 1 uses the Diophantine approximation theory [22]. We start with the same integer lattice as in [22, Theorem 1]. Let $\Lambda_{P,\varepsilon}$ be the scalar lattice defined as:

$$\Lambda_{P,\varepsilon} = \left\{x : x = P^{1/4+\varepsilon}z, z \in \mathbf{Z}\right\}. \tag{105}$$

The set of integer lattice points we use is denoted with $\mathcal{C}_{P,\varepsilon}$ and is given by:

$$\mathcal{C}_{P,\varepsilon} = \Lambda_{P,\varepsilon} \cap \left[-\sqrt{P}, \sqrt{P}\right] \tag{106}$$

where $P = \min\{\bar{P}_1, \bar{P}_2\}$. Hence $\lceil -P^{1/4-\varepsilon} \rceil \leq z \leq \lfloor P^{1/4-\varepsilon} \rfloor$, where $\lceil\ \rceil$ and $\lfloor\ \rfloor$ denote the ceiling and floor operation. $|\mathcal{C}_{P,\varepsilon}| \geq 2\left(P^{1/4-\varepsilon} - 1\right) + 1 = 2P^{1/4-\varepsilon} - 1$. This means that, for large enough P, we can write:

$$\log_2|\mathcal{C}_{P,\varepsilon}| \geq \log_2\left(2P^{1/4-\varepsilon} - 1\right) \geq \log_2\left(P^{1/4-\varepsilon}\right). \tag{107}$$

The same set of lattice points is used by both nodes S_1 and S_2 as the alphabet sets of input signals.

We use a uniform distribution over $\mathcal{C}_{P,\varepsilon}$ as the input distribution for S_2 and let the input X_2 from S_2 be an i.i.d. sequence. The channel then becomes a memoryless wiretap channel and the secrecy rate then follows by applying the result (14) from [3], i.e., any secrecy rate R_s such that

$$0 \leq R_s \leq [I(X_1; Y_1) - I(X_1; Y_2)]^+ \tag{108}$$

is achievable. Hence to compute the achievable secrecy rate, we need to find a lower bound to the right hand side of (108).

According to [22, Theorem 1], $\Pr(X_1)$ is chosen to be a uniform distribution over $\mathcal{C}_{P,\varepsilon}$. Therefore

$$H(X_1) = \log_2\left(|\mathcal{C}_{P,\varepsilon}|\right). \tag{109}$$

From [22, Theorem 1], when

$$P > \frac{1}{a^2 b^2}, \tag{110}$$

we have

$$H\left(X_1|Y_1\right) \leq 1 + 2\exp\left(-\tfrac{P^{2\varepsilon}}{8b}\right)\log_2\left(\left|\mathcal{C}_{P,\varepsilon}\right|\right). \tag{111}$$

The fact that $\sqrt{ab}$ is algebraic irrational is used in [22, Theorem 1], which uses a result from Diophantine approximation. (110) is due to [22, Lemma 2]. b in (111) is due to the fact that the variance of the Gaussian noise contained in Y_1 is b instead of unity.

From (109) and (111), we have:

$$I\left(X_1; Y_1\right) \geq \left(1 - 2\exp\left(-\frac{P^{2\varepsilon}}{8b}\right)\right)\log_2\left(\left|\mathcal{C}_{P,\varepsilon}\right|\right) - 1. \tag{112}$$

Since $\mathcal{C}_{P,\varepsilon}$ is simply a scaled and shifted version of the integer lattice code, we can use Lemma 3 to bound $I\left(X_1; Y_2\right)$:

$$I\left(X_1; Y_2\right) \leq I\left(X_1; Y_2, Z_2\right) = I\left(X_1; X_1 \pm X_2\right) = f(Q) \leq 0.8 \tag{113}$$

where $f(Q)$ is defined in (101). Q is determined by the range of z in (105). Hence $Q = 2\lfloor P^{1/4-\varepsilon}\rfloor + 1$. The last inequality in (113) follows from Lemma 3. Using (112) (113), and (107), we find (108) is lower bounded by

$$\left[\left(1 - 2\exp\left(-\frac{P^{2\varepsilon}}{8b}\right)\right)\left(\frac{1}{4} - \varepsilon\right)\log_2(P) - 1 - f(Q)\right]^+ \tag{114}$$

for sufficiently large P. Since $0 \leq f(Q) < 0.8$, and ε can be any value between $(0, 1/4)$, using the definition in (7), we find that the number of achieved secure degrees of freedom is $1/2$. Hence we have completed the proof.

C. Proof of Theorem 3 Case 2

In this appendix, we provide the proof for Theorem 3 Case 2. Here we use a Q-bit expansion scheme similar to the one in [19].

We begin by considering the case when $\sqrt{ab} \geq 2$. For this case, let $Q = \sqrt{ab}$ and

$$X_k = \sqrt{P_0}\sum_{i=0}^{M-1} a_{k,i}\, Q^{2i}, k = 1, 2, \tag{115}$$

where M is a constant positive integer, P_0 is a constant scaling factor. Both are related to the variance of X_k. $a_{k,i}$ is uniformly distributed over $[0, \lfloor Q\rfloor - 1] \cap \mathbf{Z}$. Due to the range limit imposed on $a_{k,i}$, we observe that $a_{k,i}$ is uniquely determined by X_k.

The signal received by node D_1 is given by

$$Y_1 = \sqrt{P_0}\left(\sum_{i=0}^{M-1} a_{1,i}\, Q^{2i} + \sum_{i=0}^{M-1} a_{2,i}\, Q^{2i+1}\right) + \sqrt{b}Z_1. \tag{116}$$

We then derive a lower bound to $[I(X_1; Y_1) - I(X_2; Y_2)]^+$ as we did for Theorem 1.

$I(X_1; Y_1)$ can be lower bounded by following a similar derivation from [22, Theorem 1]. Define $\mathcal{C}_k = \{\sqrt{P_0}\sum_{i=0}^{M-1} a_{k,i}\, Q^{2i} : a_{k,i} \in [0, \lfloor Q\rfloor - 1] \cap \mathbf{Z}\}$, $k = 1, 2$. We use the same maximum likelihood decoder used in [22, Theorem 1]:

$$\hat{Y}_1 = \arg \min_{X_1 + QX_2,\ \text{s.t. } X_k \in \mathcal{C}_k, k=1,2} |Y_1 - (X_1 + QX_2)|^2. \tag{117}$$

It is clear that given $\hat{Y}_1$, there is a unique pair of X_1, X_2 such that $X_1 + QX_2 = \hat{Y}_1$. Let this mapping from $\hat{Y}_1$ to X_1 be f. Define binary random variable A such that

$$A = \begin{cases} 0 & \text{if } |\sqrt{b}Z_1| < \sqrt{P_0}/2 \\ 1 & \text{otherwise} \end{cases} \tag{118}$$

Note that if $A = 0$, we have $X_1 = f(\hat{Y}_1)$. For this definition of A, we have

$$\begin{aligned} H\left(X_1|\hat{Y}_1\right) &\leq H\left(X_1, A|\hat{Y}_1\right) && (119)\\ &= H\left(A|\hat{Y}_1\right) + H\left(X_1|\hat{Y}_1, A\right) && (120)\\ &\leq 1 + \Pr(A=1)\, H\left(X_1|\hat{Y}_1, A=1\right) \\ &\quad + \Pr(A=0)\, H\left(X_1|\hat{Y}_1, A=0\right) && (121)\\ &= 1 + \Pr(A=1)\, H\left(X_1|\hat{Y}_1, A=1\right) && (122)\\ &\leq 1 + \Pr(A=1)\, H\left(X_1|A=1\right) && (123)\\ &= 1 + \Pr(A=1)\, H\left(X_1\right). && (124) \end{aligned}$$

(124) is because A is independent from X_1. $\Pr(A = 1)$ is bounded as follows:

$$\begin{aligned} \Pr(A=1) &= \int_{|t| \geq \sqrt{P_0}/2} \frac{1}{\sqrt{2\pi b}}\exp\left(-\frac{t^2}{2b}\right)dt && (125)\\ &\leq 2\exp\left(-\frac{P_0}{8b}\right). && (126) \end{aligned}$$

Substituting into (124), we get:

$$H\left(X_1|\hat{Y}_1\right) \leq 1 + 2\exp\left(-\tfrac{P_0}{8b}\right)H\left(X_1\right). \tag{127}$$

Therefore $I(X_1; Y_1)$ is lower bounded as:

$$\begin{aligned} I\left(X_1; Y_1\right) &\geq I\left(X_1; \hat{Y}_1\right) && (128)\\ &\geq \left(1 - 2\exp\left(-\frac{P_0}{8b}\right)\right)H\left(X_1\right) - 1. && (129) \end{aligned}$$

For $I(X_1; Y_2)$, we have:

$$\begin{aligned} I(X_1; Y_2) &\leq I(X_1; X_1 \pm X_2) \\ &\leq \sum_{i=0}^{M-1} I(a_{1,i}; a_{1,i} \pm a_{2,i}) = Mf(\lfloor Q\rfloor). && (130) \end{aligned}$$

For X_1 defined in (115), we have $H(X_1) = M\log_2\lfloor Q\rfloor$. Substituting it into (129) and combining it with (130), we find, from (108), that the following secrecy rate is achievable.

$$R_s = [M(1 - 2\exp(-\frac{P_0}{8b}))(\log_2\lfloor Q\rfloor) - 1 - Mf(\lfloor Q\rfloor)]^+. \tag{131}$$

The transmission power is given by:[7]

$$Var\left[X_k\right] = P_0\left(\frac{\lfloor Q\rfloor^2-1}{12}\right)\sum_{i=0}^{M-1} Q^{4i} \tag{132}$$

$$= P_0\left(\frac{\lfloor Q\rfloor^2-1}{12}\right)\frac{Q^{4M}-1}{Q^4-1}. \tag{133}$$

The secure degrees of freedom can then be computed by substituting the transmission power (133) and the secrecy rate (131) into (7), which yields:

$$\lim_{M\to\infty}\frac{[\left(\left(1-2\exp\left(-\frac{P_0}{8b}\right)\right)\log_2\lfloor Q\rfloor - f(\lfloor Q\rfloor)\right)M]^+}{\frac{1}{2}\log_2\left(Q^{4M}\right)} \tag{134}$$

$$= [\frac{1}{2}\left(1-2\exp\left(-\frac{P_0}{8b}\right)\right)\frac{\log_2\lfloor Q\rfloor}{\log_2 Q} - \frac{f(\lfloor Q\rfloor)}{2\log_2(Q)}]^+. \tag{135}$$

(135) can be made arbitrarily close to (51) by choosing a large enough P_0. (52) then follows from (51) via Lemma 3.

When $Q = 2$, it can be verified that $f(\lfloor Q\rfloor) = 0.5$, and (135) can be made to be arbitrarily close to $1/4$.

The case of $1/\sqrt{ab} \geq 2$ can be proved in a similar fashion. Let $1/\sqrt{ab} = Q$ and let $X_k = Q\sqrt{P_0}\sum_{i=0}^{M-1} a_{k,i}Q^{2i}$, $k = 1, 2$. Then all previous derivations apply. In particular, (116) becomes:

$$Y_1 = \sqrt{P_0}(\sum_{i=0}^{M-1} a_{1,i}Q^{2i+1} + \sum_{i=0}^{M-1} a_{2,i}Q^{2i}) + \sqrt{b}Z_1 \tag{136}$$

and (117) becomes:

$$\hat{Y}_1 = \arg\min_{X_1+X_2/Q,\ \text{s.t. } X_k\in\mathcal{C}_k,k=1,2} |Y_1-(X_1+\frac{1}{Q}X_2)|^2. \tag{137}$$

The achieved secrecy rate remains the same. The transmission power is scaled by Q^2:

$$Var\left[X_k\right] = Q^2P_0\left(\frac{\lfloor Q\rfloor^2-1}{12}\right)\frac{Q^{4M}-1}{Q^4-1}, \quad k = 1, 2. \tag{138}$$

Hence the number of secure degrees of freedom is still given by (52) when $Q > 2$ and $1/4$ when $Q = 2$.

D. Proof of Theorem 3 Case 3

Let X_k be given by (115) in Section E-C with $Q = 2$. However, unlike (115), here $a_{k,i}$ is not chosen to be uniformly distributed over $\{0, 1\}$. Instead, we choose its distribution to maximize

$$I(a_{1,i}; a_{1,i} + a_{2,i}) - I(a_{1,i}; a_{1,i} - a_{2,i}). \tag{139}$$

Let c_I denote the value of (139) when $\Pr(a_{1,i} = 1) = 0.1443$, $\Pr(a_{2,i} = 1) = 0.8557$. The value of c_I is close to and greater than 0.1095. We next derive the achievable secrecy rate by deriving a lower bound on $[I(X_1; Y_1) - I(X_2; Y_2)]^+$. Define $\mathcal{C}_k = \{\sqrt{P_0}\sum_{i=0}^{M-1} a_{k,i}Q^{2i} : a_{k,i} \in \{0, 1\}\}$, $k = 1, 2$. Define $f(Y_1)$ as

$$f(Y_1) = \arg\min_{X_1+X_2:X_k\in\mathcal{C}_k,k=1,2} |Y_1 - (X_1 + X_2)|^2. \tag{140}$$

Then:

$$I\left(X_1; X_1 + X_2\right) - I\left(X_1; f\left(Y_1\right)\right)$$

$$\leq I\left(X_1; X_1 + X_2 | f\left(Y_1\right)\right) \tag{141}$$

$$\leq H\left(X_1 + X_2 | f\left(Y_1\right)\right) \tag{142}$$

$$\leq 1 + 2\exp\left(-\frac{P_0}{8b}\right)H\left(X_1 + X_2\right). \tag{143}$$

Inequality (143) follows from (119)-(124) with $\hat{Y}_1$ replaced by $f(Y_1)$ defined in (140), and X_1 replaced with $X_1 + X_2$. Then we have

$$I\left(X_1; Y_1\right)$$

$$\geq I\left(X_1; f\left(Y_1\right)\right) \tag{144}$$

$$\geq \left(1 - \left(2\exp\left(-\frac{P_0}{8b}\right)\right)\right)H\left(X_1 + X_2\right) - H\left(X_2\right) - 1 \tag{145}$$

$$= I\left(X_1; X_1 + X_2\right) - 1 - 2\exp\left(-\frac{P_0}{8b}\right)H\left(X_1 + X_2\right) \tag{146}$$

$$= \sum_{i=0}^{M-1} I\left(a_{1,i}; a_{1,i} + a_{2,i}\right) - 1$$

$$-2\exp\left(-\frac{P_0}{8b}\right)\sum_{i=0}^{M-1} H\left(a_{1,i} + a_{2,i}\right). \tag{147}$$

In (145) we use the fact that X_1 is independent from X_2 and apply the result from (143). (147) is because there is a one-to-one mapping between $X_1 + X_2$ and $\{a_{1,i} + a_{2,i}, i = 0, \ldots, M - 1\}$.

For $I(X_1; Y_2)$, we have

$$I\left(X_1; Y_2\right) \leq I\left(X_1; X_1 - X_2\right) \tag{148}$$

$$\leq \sum_{i=1}^{M} I\left(a_{1,i}; a_{1,i} - a_{2,i}\right). \tag{149}$$

Therefore

$$I\left(X_1; Y_1\right) - I\left(X_1; Y_2\right)$$

$$\geq \sum_{i=1}^{M}\left(I\left(a_{1,i}; a_{1,i} + a_{2,i}\right) - I\left(a_{1,i}; a_{1,i} - a_{2,i}\right)\right) - 1$$

$$-2\exp\left(-\frac{P_0}{8b}\right)\sum_{i=0}^{M-1} H\left(a_{1,i} + a_{2,i}\right) \tag{150}$$

$$= Mc_I - 1 - 2\exp\left(-\frac{P_0}{8b}\right)MH\left(a_{1,i} + a_{2,i}\right). \tag{151}$$

Let $V_{a,k}$ be the variance of $a_{k,i}$. Then we have

$$Var\left[X_k\right] = P_0V_{a,k}\frac{Q^{4M}-1}{Q^2-1}, \quad k = 1, 2. \tag{152}$$

[7] In case it is desired for X_k to have zero mean, we can simply shift X_k by a constant, which will not alter the secrecy rate.

Hence the number of secure degrees of freedom is given by:

$$\lim_{M\to\infty} \frac{[I(X_1;Y_1) - I(X_1;Y_2)]^+}{\frac{1}{2}\log_2 Var[X_k]} \tag{153}$$

$$= [\frac{c_I}{2} - \exp\left(-\frac{P_0}{8b}\right) H(a_{1,i} + a_{2,i})]^+. \tag{154}$$

We can always choose a large enough P_0 to make the secure degrees of freedom to be arbitrarily close to $\frac{c_I}{2} > 0.0548$. This completes the proof of Theorem 3 Case 3).

E. Proof of Theorem 3 Case 4

Here we mimic the coding scheme in Section IV when $p = 1, q = 1$. Note that positive secure degrees of freedom can not be achieved by the nested lattice code in Section IV in this case. This is mainly a consequence of leaking 1 bit per layer in the layered coding scheme. Here, we replace the nested lattice code with Q-bit expansion similar to the one used in Appendix E-C and leverage the fact that integer lattice leaks less than 0.8 bit per layer, and show that positive secure degrees of freedom are achievable.

Let X_i be:

$$X_k = \sqrt{P_0} \sum_{i=0}^{3M-1} a_{k,i} Q^i, k = 1, 2 \tag{155}$$

where $Q = 2$. M is a positive integer and P_0 is a positive constant as defined in the proof of Theorem 2. Choose $a_{k,i}$ such that

$$a_{k,i} = 0, \text{ if } i \bmod 3 = 1 \text{ or } 2, \tag{156}$$

otherwise $a_{k,i}$ is uniformly distributed over $\{0, 1\}$. Here $a_{k,i}$ is forced to be zero at $i \bmod 3 = 1$ or 2 to make room for $0.5a_{2,j}, \forall j \bmod 3 = 0$ and the carry-over from $a_{1,j} + a_{2,j}, \forall j \bmod 3 = 0$.

Define $\mathcal{C}_k$ as the collection of points X_k defined by (155). Define $\hat{Y}_1$ as

$$\hat{Y}_1 = \arg \min_{X_1+1.5X_2,\ \text{s.t. } X_k\in\mathcal{C}_k, k=1,2} |Y_1 - (X_1 + 1.5X_2)|^2. \tag{157}$$

With this new definition of $\hat{Y}_1$, the same derivation that leads to (129) applies, where $H(X_1) = M$. On the other hand:

$$I(X_1;Y_2) \le I(X_1; X_1 \pm X_2) \tag{158}$$

$$\le \sum_{i=0}^{3M-1} I\left(a_{1,i}; a_{1,i} \pm a_{2,i}\right) = 0.5M. \tag{159}$$

From (129), (159) and (108), the following secrecy rate is achievable:

$$\left[\left(1 - 2\exp\left(-\frac{P_0}{8b}\right)\right) M - 1 - 0.5M\right]^+. \tag{160}$$

The transmission power is given by

$$Var[X_k] = P_0 \frac{Q^2 - 1}{12} \sum_{i=0}^{M-1} Q^{6i} = \frac{P_0}{4} \frac{Q^{6M} - 1}{Q^6 - 1} \tag{161}$$

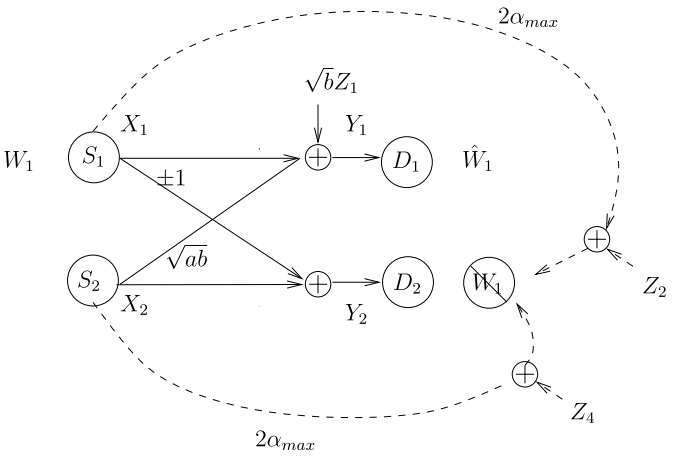

Fig. 8. Fig. 5 with an enhanced eavesdropper channel.

for $Q = 2$. The number of secure degrees of freedom is hence given by:

$$\lim_{M\to\infty} \frac{\left(1 - 2\exp\left(-\frac{P_0}{8b}\right)\right) M - 1 - 0.5M}{\frac{1}{2}\log_2\left(Q^{6M}\right)} \tag{162}$$

$$= \frac{1}{3}\left(1 - 2\exp\left(-\frac{P_0}{8b}\right) - 0.5\right) \tag{163}$$

which can be made arbitrarily close to $1/6$ by choosing a large enough P_0.

Remark 12: This scheme can be extended to the case when $\sqrt{ab} = 1 + 1/Q$ with Q being an integer greater than 2. In this case, we let $a_{k,i} = 0$ if $i \bmod 2 = 1$. Otherwise $a_{k,i}$ is taken from $[0, Q-1] \cap \mathbf{Z}$. However, to make room for the carryovers, the least significant bit of the binary representation of $a_{k,i}$ must be zero. Hence, 1 bit is lost per layer due to the carryovers. As a result, the achieved secure degrees of freedom turns out to be smaller than those achieved with nested lattice codes. Similar coding schemes can be designed for other values of $\sqrt{ab}$. However, it is difficult to find a uniform description of such codes that achieves a better performance than that of nested lattice code. □

APPENDIX F
PROOF OF THEOREM 4

We prove Theorem 4 by showing that the loss in secrecy rate due to the imperfectness of channel state information can be bounded by a constant. To prove this, we consider the channel in Fig. 8, where Z_4 has the same distribution as Z_2. Z_4, Z_1, Z_2 are independent. For this channel, we have the following lemma:

Lemma 4: Any secrecy rate achievable in Fig. 8 is also achievable in Fig. 5 with the *same* coding scheme.

Proof: Let n be the total number of channel uses. Let D and D' be two $n \times n$ diagonal matrices, whose diagonal elements at the ith row and ith column are α_i and α'_i respectively. Define a diagonal matrix $\bar{D}$ such that $\bar{D}$ is obtained from D by replacing the 0 elements on its diagonal line with $\alpha_{\max}$. $\bar{D}'$ is defined similarly from D'. Define $Z_i^n, i = 3, 4, 5$ as an independent random variable with the same Gaussian

distribution as Z_2^n. $Z_i^n, i = 2, 3, 4, 5$ are independent. Then the mutual information between W_1 and the knowledge of the eavesdropper in Fig. 5 can be upper bounded as follows:

$$0 \le I\left(W_1; Y_2^n, D, D'\right) \tag{164}$$
$$\le I\left(W_1; X_1^n \pm X_2^n + Z_2^n + DX_1^n + D'X_2^n, D, D'\right) \tag{165}$$
$$\le I\left(W_1; X_1^n \pm X_2^n, 0.5Z_2^n + DX_1^n, 0.5Z_4^n + D'X_2^n, D, D'\right). \tag{166}$$

We next divide the length-n vector $0.5Z_2^n + DX_1^n$ into two sets of variables depending on whether the diagonal term α_i is zero: $\{0.5Z_{2,i} + \alpha_i X_{1,i}\}_{i:\alpha_i \neq 0}$ and $\{0.5Z_{2,i}\}_{i:\alpha_i = 0}$. Similarly, we can rewrite $0.5Z_4^n + D'X_2^n$ as $\{0.5Z_{4,i} + \alpha_i' X_{2,i}\}_{i:\alpha_i' \neq 0}$ and $\{0.5Z_{4,i}\}_{i:\alpha_i' = 0}$. (166) then becomes

$$I(W_1; X_1^n \pm X_2^n, \{0.5Z_{2,i} + \alpha_i X_{1,i}\}_{i:\alpha_i \neq 0}, \{0.5Z_{2,i}\}_{i:\alpha_i = 0}, \{0.5Z_{4,i} + \alpha_i' X_{2,i}\}_{i:\alpha_i' \neq 0}, \{0.5Z_{4,i}\}_{i:\alpha_i' = 0}, D, D') \tag{167}$$
$$= I(W_1; X_1^n \pm X_2^n, \{0.5Z_{2,i} + \alpha_i X_{1,i}\}_{i:\alpha_i \neq 0}, \{0.5Z_{4,i} + \alpha_i' X_{2,i}\}_{i:\alpha_i' \neq 0}, D, D') \tag{168}$$
$$\le I\left(W_1; X_1^n \pm X_2^n, 0.5Z_2^n + \bar{D}X_1^n, 0.5Z_4^n + \bar{D}'X_2^n, \bar{D}, \bar{D}'\right) \tag{169}$$
$$= I\left(W_1; X_1^n \pm X_2^n, Z_2^n + 2\bar{D}X_1^n, Z_4^n + 2\bar{D}'X_2^n, \bar{D}, \bar{D}'\right) \tag{170}$$
$$= I(W_1; X_1^n \pm X_2^n, \frac{\bar{D}}{\alpha_{\max}} Z_2^n + \sqrt{1 - \frac{\bar{D}^2}{\alpha_{\max}^2}} Z_3^n + 2\bar{D}X_1^n, \frac{\bar{D}'}{\alpha_{\max}} Z_4^n + \sqrt{1 - \frac{\bar{D}'^2}{\alpha_{\max}^2}} Z_5^n + 2\bar{D}'X_2^n, \bar{D}, \bar{D}') \tag{171}$$
$$\le I(W_1; X_1^n \pm X_2^n, \frac{\bar{D}}{\alpha_{\max}} Z_2^n + 2\bar{D}X_1^n, \frac{\bar{D}'}{\alpha_{\max}} Z_4^n + 2\bar{D}'X_2^n, \bar{D}, \bar{D}') \tag{172}$$
$$= I(W_1; X_1^n \pm X_2^n, Z_2^n + 2\alpha_{\max}X_1^n, Z_4^n + 2\alpha_{\max}X_2^n, \bar{D}, \bar{D}'). \tag{173}$$

Note that (173) is exactly the same as the mutual information between W_1 and the eavesdropper's knowledge in Fig. 8. Hence if

$$\lim_{n\to\infty} \frac{1}{n} I(W_1; X_1^n \pm X_2^n, Z_2^n + 2\alpha_{\max}X_1^n, Z_4^n + 2\alpha_{\max}X_2^n, \bar{D}, \bar{D}') = 0. \tag{174}$$

Then we must have:

$$\lim_{n\to\infty} \frac{1}{n} I\left(W_1; Y_2^N, D, D'\right) = 0. \tag{175}$$

This means that to obtain the secrecy rate for the model in Fig. 5, we can as well compute the secrecy rate for the model in Fig. 8. ∎

We next compute the secrecy rate for Fig. 8. We use the same layered nested lattice coding scheme in Theorem 2 except now a different bin size must be chosen in the wiretap binning scheme.

Recall that the notation $A_{\mathcal{M}}$ represents $A_i, i = 1, \ldots, M$. We begin by applying the result (14) from [3]. It shows the following secrecy rate R_s is achievable:

$$0 \le R_s \le [\lim_{N\to\infty} \frac{1}{N}(I(u_{1,\mathcal{M}}^N; Y_1^N) - I(u_{1,\mathcal{M}}^N; X_1^N \pm X_2^N, 2\alpha_{\max}X_1^N + Z_2^N, 2\alpha_{\max}X_2^N + Z_4^N))]^+. \tag{176}$$

Note that we drop $\bar{D}$ and $\bar{D}'$ since we assume they are independent from the channel inputs and the additive channel noise.

The first term in (176) is still given by (44) since the signal received by D_1 remains the same. The second term in (176) can be upper bounded as follows:

$$\frac{1}{N} I\left(u_{1,\mathcal{M}}^N; X_1^N \pm X_2^N, Z_2^N + 2\alpha_{\max}X_1^N, 2\alpha_{\max}X_2^N + Z_4^N\right) \tag{177}$$
$$\le \frac{1}{N} I\left(u_{1,\mathcal{M}}^N; X_1^N \pm X_2^N\right) + \frac{1}{N} I\left(\begin{array}{l} u_{1,\mathcal{M}}^N, X_1^N \pm X_2^N; \\ Z_2^N + 2\alpha_{\max}X_1^N, 2\alpha_{\max}X_2^N + Z_4^N \end{array}\right) \tag{178}$$
$$\le \frac{1}{N} I\left(u_{1,\mathcal{M}}^N; X_1^N \pm X_2^N\right) + \frac{1}{N} I\left(\begin{array}{l} u_{1,\mathcal{M}}^N, u_{2,\mathcal{M}}^N, X_1^N \pm X_2^N; \\ Z_2^N + 2\alpha_{\max}X_1^N, 2\alpha_{\max}X_2^N + Z_4^N \end{array}\right) \tag{179}$$
$$= \frac{1}{N} I\left(u_{1,\mathcal{M}}^N; X_1^N \pm X_2^N\right) + \frac{1}{N} I\left(u_{1,\mathcal{M}}^N, u_{2,\mathcal{M}}^N; Z_2^N + 2\alpha_{\max}X_1^N, 2\alpha_{\max}X_2^N + Z_4^N\right) \tag{180}$$
$$\le \frac{1}{N} I\left(u_{1,\mathcal{M}}^N; X_1^N \pm X_2^N\right) + \frac{1}{N} I\left(u_{1,\mathcal{M}}^N; Z_2^N + 2\alpha_{\max}X_1^N\right) + \frac{1}{N} I\left(u_{2,\mathcal{M}}^N; Z_4^N + 2\alpha_{\max}X_2^N\right). \tag{181}$$

(180) is because $\{X_1^N \pm X_2^N\} - \{u_{k,\mathcal{M}}^N, k = 1, 2\} - \{Z_2^N + 2\alpha_{\max}X_1^N, Z_4^N + 2\alpha_{\max}X_2^N\}$ is a Markov chain, since $X_1^N \pm X_2^N$ is a deterministic function of $\{u_{k,\mathcal{M}}^N, k = 1, 2\}$. The first term in (181) is shown by (46)-(48) to be bounded by M. Hence we only need to bound the second and the third term. The second term is bounded by:

$$\frac{1}{N} I\left(u_{1,\mathcal{M}}^N; Z_2^N + 2\alpha_{\max}X_1^N\right) \tag{182}$$
$$\le \frac{1}{N} I\left(X_1^N; Z_2^N + 2\alpha_{\max}X_1^N\right) \tag{183}$$
$$\le C\left(4\alpha_{\max}^2 P_1\right) \tag{184}$$

where $C(x) = \frac{1}{2}\log_2(1 + x)$. P_1, the total transmission power consumed by S_1, is given by (30). The inequality (184) follows since the mutual information in (183) is maximized by a Gaussian input distribution. Similarly, the third term is bounded by:

$$\frac{1}{N} I\left(u_{2,\mathcal{M}}^N; Z_4^N + 2\alpha_{\max}X_2^N\right) \tag{185}$$
$$\le C\left(4\alpha_{\max}^2 P_2\right). \tag{186}$$

P_2 is the total transmission power consumed by S_2, which equals P_1 in our layered coding scheme.

Applying (184),(186), (46)-(48), and (44) in (176), we find the secrecy rate is given by

$$R_s = [(0.25\log_2(\alpha) - 1)M - 2C\left(4\alpha_{\max}^2 P_1\right)]^+. \quad (187)$$

Compared it to (33), the loss in secrecy rate is bounded by $2C\left(4\alpha_{\max}^2 P_1\right)$, which is bounded by a constant per the condition stated in Theorem 4. Hence the achieved secure degrees of freedom remain the same as in Theorem 2.

REFERENCES

[1] C. E. Shannon, "Communication theory of secrecy systems," *Bell Syst. Tech. J.*, vol. 28, no. 4, pp. 656–715, Sep. 1949.

[2] A. D. Wyner, "The wire-tap channel," *Bell Syst. Tech. J.*, vol. 54, no. 8, pp. 1355–1387, 1975.

[3] I. Csiszàr and J. Körner, "Broadcast channels with confidential messages," *IEEE Trans. Inf. Theory*, vol. 24, no. 3, pp. 339–348, May 1978.

[4] S. Leung-Yan-Cheong and M. Hellman, "The Gaussian wire-tap channel," *IEEE Trans. Inf. Theory*, vol. 24, no. 4, pp. 451–456, Jul. 1978.

[5] A. Khisti and G. Wornell, "Secure transmission with multiple antennas-I: The MISOME wiretap channel," *IEEE Trans. Inf. Theory*, vol. 56, no. 7, pp. 3088–3104, Jul. 2010.

[6] F. Oggier and B. Hassibi, "The secrecy capacity of the MIMO wiretap channel," in *Proc. IEEE Int. Symp. Inf. Theory*, Jul. 2008, pp. 524–528.

[7] R. Liu, T. Liu, and H. V. Poor, "Multiple-input multiple-output Gaussian broadcast channels with confidential messages," *IEEE Trans. Inf. Theory*, vol. 56, no. 9, pp. 4215–4227, Sep. 2010.

[8] E. Ekrem and S. Ulukus, "The secrecy capacity region of the Gaussian MIMO multi-receiver wiretap channel," *IEEE Trans. Inf. Theory*, vol. 57, no. 4, pp. 2083–2114, Apr. 2011.

[9] E. Tekin, S. Serbetli, and A. Yener, "On Secure signaling for the Gaussian multiple access wire-tap channel," in *Proc. 39th Annu. Asilomar Conf. Signals, Syst., Comput.*, Nov. 2005, pp. 1747–1751.

[10] E. Tekin and A. Yener, "The Gaussian multiple access wire-tap channel," *IEEE Trans. Inf. Theory*, vol. 54, no. 12, pp. 5747–5755, Dec. 2008.

[11] Y. Liang and H. V. Poor, "Multiple-access channels with confidential messages," *IEEE Trans. Inf. Theory*, vol. 54, no. 3, pp. 976–1002, Mar. 2008.

[12] E. Tekin and A. Yener, "The general Gaussian multiple access and two-way wire-tap channels: Achievable rates and cooperative jamming," *IEEE Trans. Inf. Theory*, vol. 54, no. 6, pp. 2735–2751, Jun. 2008.

[13] L. Lai and H. El Gamal, "Cooperation for secrecy: The relay-eavesdropper channel," *IEEE Trans. Inf. Theory*, vol. 54, no. 9, pp. 4005–4019, Sep. 2008.

[14] X. He and A. Yener, "Cooperation with an untrusted relay: A secrecy perspective," *IEEE Trans. Inf. Theory*, vol. 56, no. 8, pp. 3801–3827, Aug. 2010.

[15] X. He and A. Yener, "Two-hop secure communication using an untrusted relay," *Eurasip J. Wireless Commun. Netw.*, vol. 2009, pp. 305146-1–305146-13, Nov. 2009.

[16] O. Koyluoglu and H. El-Gamal, "Cooperative binning and channel prefixing for secrecy in interference channels," *submitted to IEEE Trans. Inf. Theory*, May 2009.

[17] B. Nazer and M. Gastpar, "The case for structured random codes in network capacity theorems," *Eur. Trans. Telecommun.*, vol. 19, no. 4, pp. 455–474, Jun. 2008.

[18] J. G. Proakis and M. Salehi, *Digital Communications*. New York, NY, USA: McGraw-Hill, 1995.

[19] V. R. Cadambe, S. A. Jafar, and S. Shamai, "Interference Alignment on the deterministic channel and application to fully connected AWGN interference networks," *IEEE Trans. Inf. Theory*, vol. 55, no. 1, pp. 269–274, Jan. 2009.

[20] G. Bresler, A. Parekh, and D. Tse, "The Approximate Capacity of the many-to-one and one-to-many Gaussian interference channels," in *Proc. 45th Allerton Conf. Commun., Control, Comput.*, Sep. 2007, pp. 1–27.

[21] S. Sridharan, A. Jafarian, S. Vishwanath, and S. A. Jafar, "Capacity of symmetric K-user Gaussian very strong interference channels," in *Proc. IEEE Global Telecommun. Conf.*, Nov. 2008, pp. 1–5.

[22] R. Etkin and E. Ordentlich, "The degrees-of-freedom of the K-user Gaussian interference channel is discontinuous at rational channel coefficients," *IEEE Trans. Inf. Theory*, vol. 55, no. 11, pp. 4932–4946, Nov. 2009.

[23] K. Narayanan, M. P. Wilson, and A. Sprintson, "Joint physical layer coding and network coding for Bi-directional relaying," in *Proc. 45th Allerton Conf. Commun., Control, Comput.*, Sep. 2007, pp. 1–14.

[24] W. Nam, S.-Y. Chung, and Y. H. Lee, "Capacity bounds for two-way relay channels," in *Proc. Int. Zurich Seminar Commun.*, Mar. 2008, pp. 144–147.

[25] E. Tekin and A. Yener, "The multiple access wire-tap channel: Wireless secrecy and cooperative jamming," in *Proc. Inf. Theory Appl. Workshop*, Jan. 2007, pp. 1–3.

[26] R. Liu, I. Maric, P. Spasojevic, and R. D. Yates, "Discrete memoryless interference and broadcast channels with confidential messages: Secrecy rate regions," *IEEE Trans. Inf. Theory*, vol. 54, no. 6, pp. 2493–2507, Jun. 2008.

[27] X. Tang, R. Liu, P. Spasojevic, and H. V. Poor, "The Gaussian wiretap channel with a helping interferer," in *Proc. IEEE Int. Symp. Inf. Theory*, Jul. 2008, pp. 389–393.

[28] S. Goel and R. Negi, "Guaranteeing secrecy using artificial noise," *IEEE Trans. Wireless Commun.*, vol. 7, no. 6, pp. 2180–2189, Jun. 2008.

[29] E. Ekrem and S. Ulukus, "On the secrecy of multiple access wiretap channel," in *Proc. 46th Allerton Conf. Commun., Control, Comput.*, Sep. 2008, pp. 1014–1021.

[30] X. He and A. Yener, "Interference channels with strong secrecy," in *Proc. 47th Allerton Conf. Commun., Control Comput.*, Sep. 2009, pp. 811–818.

[31] X. He, "Cooperation and information theoretic security in wireless networks," Ph.D. dissertation, Dept. Electr. Eng., Pennsylvania State Univ., Pennsylvania, PA, USA, Aug. 2010.

[32] U. Erez and R. Zamir, "Achieving 1/2 log (1+ SNR) on the AWGN channel with lattice encoding and decoding," *IEEE Trans. Inf. Theory*, vol. 50, no. 10, pp. 2293–2314, Oct. 2004.

[33] G. D. Forney, "On the role of MMSE estimation in approaching the information-theoretic limits of linear Gaussian channels: Shannon meets wiener," in *Proc. 41st Allerton Conf. Commun., Control, Comput.*, Sep. 2003, pp. 1–14.

[34] L. Lai, H. El Gamal, and H. V. Poor, "The wiretap channel with feedback: Encryption over the channel," *IEEE Trans. Inf. Theory*, vol. 54, no. 11, pp. 5059–5067, Nov. 2008.

[35] C. Ling, L. Luzzi, J.-C. Belfiore, and D. Stehlé, (2012, Oct.). *Semantically Secure Lattice Codes for the Gaussian Wiretap Channel* [Online]. Available: http://arxiv.org/abs/1210.6673v2

[36] S. Yang, P. Piantanida, M. Kobayashi, and S. Shamai, "Lattice coding for strongly secure compute-and-forward in a bidirectional relay," in *Proc. IEEE Int. Symp. Inf. Theory*, Jul. 2013, pp. 2775–2779.

[37] S. Sridharan, A. Jafarian, S. Vishwanath, S. A. Jafar, and S. Shamai, "A layered lattice coding scheme for a class of three user Gaussian interference channels," in *Proc. 46th Allerton Conf. Commun., Control, Comput.*, Sep. 2008, pp. 1–8.

[38] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York, NY, USA: Wiley, 2006.

[39] H. A. Loeliger, "Averaging bounds for lattices and linear codes," *IEEE Trans. Inf. Theory*, vol. 43, no. 6, pp. 1767–1773, Nov. 1997.

[40] B. Nazer and M. Gastpar, "Compute-and-forward: Harnessing interference with structured codes," in *Proc. Int. Symp. Inf. Theory*, Jul. 2008, pp. 6463–6486.

[41] J. H. Conway and N. J. A. Sloane, *Sphere Packings, Lattices and Groups*. New York, NY, USA: Springer-Verlag, 1999.

[42] X. He and A. Yener, "K-user interference channels: Achievable secrecy rate and degrees of freedom," in *Proc. IEEE Inf. Theory Workshop*, Jun. 2009, pp. 336–340.

[43] 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.

Xiang He (S'08–M'10) received B.S. and M.S. degrees in Electrical Engineering from Shanghai Jiao Tong University, Shanghai, China in 2003 and 2006 respectively. His master study was about high speed FPGA implementation of channel encoder, decoder and MIMO detectors. He received his Ph.D. degree in 2010 from the Department of Electrical Engineering at the Pennsylvania State University and joined Microsoft in that year. In 2010, he received the Melvin P. Bloom Memorial Outstanding Doctoral Research Award from the Department of Electrical Engineering at the Pennsylvania State University and the best paper award from the Communication Theory Symposium in IEEE International Conference on Communications (ICC). In 2011, he was named as one of the exemplary reviewers by IEEE COMMUNICATION LETTERS. His research interests include information theoretic secrecy, coding theory, queuing theory, optimization techniques, distributed detection and estimation.

Aylin Yener (S'91–M'00–SM'13) received the B.Sc. degree in electrical and electronics engineering, and the B.Sc. degree in physics, from Boğaziçi University, Istanbul, Turkey; and the M.S. and Ph.D. degrees in electrical and computer engineering from Wireless Information Network Laboratory (WINLAB), Rutgers University, New Brunswick, NJ. Commencing fall 2000, for three semesters, she was a P.C. Rossin Assistant Professor at the Electrical Engineering and Computer Science Department, Lehigh University, PA. In 2002, she joined the faculty of The Pennsylvania State University, University Park, PA, where she was an Assistant Professor, then Associate Professor, and has been a Professor of Electrical Engineering since 2010. During the academic year 2008-2009, she was a Visiting Associate Professor with the Department of Electrical Engineering, Stanford University, CA. Her research interests are in information theory, communication theory and network science, with recent emphasis on green communications and information security. She received the NSF CAREER award in 2003.

Dr. Yener previously served as a technical program chair or co-chair for various conferences for the IEEE Communications Society, as an associate editor for the IEEE TRANSACTIONS ON COMMUNICATIONS, as an associate editor and an editorial advisory board member for the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS. She served as the student committee chair for the IEEE Information Theory Society 2007-2011, and was the co-founder of the Annual School of Information Theory in North America co-organizing the school in 2008, 2009 and 2010. Dr. Yener currently serves on the board of governors of the IEEE Information Theory Society as its treasurer.