

# Optimal rate-limited secret key generation from Gaussian sources using lattices

Laura Luzzi, Cong Ling and Matthieu R. Bloch

## Abstract

We propose a lattice-based scheme for secret key generation from Gaussian sources in the presence of an eavesdropper, and show that it achieves the strong secret key capacity in the case of degraded source models, as well as the optimal secret key / public communication rate trade-off. The key ingredients of our scheme are the use of the modulo lattice operation to extract the channel intrinsic randomness, based on the notion of flatness factor, together with a randomized lattice quantization technique to quantize the continuous source. Compared to previous works, we introduce two new notions of flatness factor based on  $L^1$  distance and KL divergence, respectively, which might be of independent interest. We prove the existence of secrecy-good lattices under  $L^1$  distance and KL divergence, whose  $L^1$  and KL flatness factors vanish for volume-to-noise ratios up to  $2\pi e$ . This improves upon the volume-to-noise ratio threshold  $2\pi$  of the  $L^\infty$  flatness factor.

## Index Terms

Secret key generation, strong secrecy, lattice coding, flatness factor.

## I. INTRODUCTION

Secret key generation (also known as key agreement) at the physical layer was first investigated by Maurer [3] and Ahlswede and Csiszár [4], who showed that correlated observations of noisy phenomena could be used to distill secret keys by exchanging information over a public channel. In recent years, this subject has received considerable attention in literature (see, e.g., [5–10]). The setup has been extended to the vector case [11, 12], the multi-terminal case [13–16], the quantum case [17] and the case with feedback [18]. Second-order asymptotics have been derived in [19, 20]. Code constructions for the discrete memoryless case have been proposed, e.g. [21, 22].

Most existing secret key generation schemes rely heavily on the assumption of discrete random sources over finite or countable alphabets. In order to apply these techniques to wireless communications, it is necessary to extend the key generation framework to the case of continuous sources, such as Gaussian sources [11, 23–25]<sup>1</sup>. In [25], the authors study a multi-terminal scenario for secret key generation in the special case for which the eavesdropper only has access to the public channel. Beside providing a characterization of the optimal strongly secret key rate, the authors show that this optimal rate can be achieved using lattice codes (for information reconciliation only).

We consider here the problem of secret key generation between two terminals, Alice and Bob, who observe correlated Gaussian sequences  $X^n$  and  $Y^n$ , in the presence of an eavesdropper, Eve, who also obtains a correlated sequence  $Z^n$ . For simplicity, we suppose that a single round of unidirectional public communication takes place in order to establish the key. Our main contribution is to show that, in the case of a degraded source model, the strong secret key capacity can be achieved by a complete lattice-coding scheme considerably different from and

The work of L. Luzzi was supported in part by the INEX CY Initiative AAP2017 Lattice Hashing and the INEX CY Initiative Ambition AAP2020 PHEBE. The work of C. Ling was supported in part by the Engineering and Physical Sciences Research Council (EPSRC) under Grant No. EP/S021043/1. The work of M. Bloch was supported in part by the National Science Foundation under award 1955401. This work was presented in part at the IEEE International Symposium on Information Theory (ISIT 2013), Istanbul, Turkey [1], and in part at the International Zurich Seminar on Communications (IZS 2018) [2].

L. Luzzi is with ETIS, UMR 8051 (CY Cergy Paris Université, ENSEA, CNRS), 95014 Cergy-Pontoise, France (e-mail: laura.luzzi@ensea.fr).

C. Ling is with the Department of Electrical and Electronic Engineering, Imperial College London, London SW7 2AZ, U.K. (e-mail: cling@ieee.org).

M. R. Bloch is with School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, Georgia (email: matthieu.bloch@ece.gatech.edu).

<sup>1</sup>An extension of the key distillation framework to quantum Gaussian states has also been considered [26, 27].

perhaps simpler than [25]<sup>2</sup>. This extends our previous work [1], in which it was shown that a secret key rate up to half a nat from the optimal was achievable.

Typically, secret key generation consists of two distinct procedures: *information reconciliation*, in which public messages are exchanged to ensure that Alice and Bob can construct the same data sequence with vanishing error probability, and *privacy amplification* to extract from this shared sequence a secret key that is statistically independent from Eve’s observation and from the public messages.

*Privacy amplification and randomness extraction:* Our privacy amplification strategy is based on the concept of *channel intrinsic randomness*, or the maximum bit rate that can be extracted from a channel output independently of its input [30–32]. We show that the reduction modulo a suitable lattice can be used to extract the intrinsic randomness. Although our main objective in this paper is to solve the problem of privacy amplification, our lattice extractor is also an intriguing result in its own right, which could have other applications.

*The flatness factor and its variants:* In our previous work [1], we provided a characterization of the class of lattices that are good for randomness extraction, which was based on a computable parameter, the *flatness factor*, measuring the  $L^\infty$  distance between the “folded” Gaussian distribution modulo the lattice and the uniform distribution on the corresponding fundamental region. The concept of flatness factor is related to the smoothing parameter used in lattice-based cryptography [33], and was first introduced in [34] in the context of physical-layer network coding. In [35], two of the authors also showed the relevance of the flatness factor for secrecy and introduced the notion of *secrecy-good lattices* for the wiretap channel. In this work, we consider two extended notions of flatness factor by which the  $L^\infty$  distance is replaced respectively by the  $L^1$  distance and the Kullback-Leibler (KL) divergence. These new flatness conditions are satisfied by a wider range of variance parameters, resulting in improved volume conditions for the chain of lattices under consideration, which allows us to achieve the secret key capacity. The existence of lattices with vanishing  $L^1$  and KL flatness factors follows by leveraging an existence result for resolvability codes for regular channels [36]. We note that the  $L^1$  smoothing parameter was already considered in [37, 38], while  $L^1$  and KL flatness factors were used implicitly earlier in [39, p. 1656]. An upper bound on the  $L^1$  flatness factor based on the Cauchy-Schwarz inequality was given in [40]. The independent work [41] studied  $L^1$  smoothing parameters both for lattices and for codes, also based on the Cauchy-Schwarz inequality. Our approach bypasses the Cauchy-Schwarz inequality, therefore leading to a tighter bound than [40]. We note however that [41] obtained a bound on the  $L^1$  smoothing parameter as tight as that in this paper, by decomposing the discrete Gaussian distribution into a convex combination of uniform ball distributions. The smoothing parameter is of fundamental importance in lattice and code-based cryptography [41], so our method for the  $L^1$  flatness factor may also be useful in these areas.

*Information reconciliation and Wyner-Ziv coding:* Our strategy for information reconciliation follows the outline of [23, 25]: first, the source  $X^n$  is vector quantized; then, a public message is generated in the manner of Wyner-Ziv coding, so that Bob can decode the quantized variable using the sequence  $Y^n$  as side information. The existence of good nested lattices for Wyner-Ziv coding has been established in [42] (see also [43, 44]). We show that this construction is compatible with the secrecy-goodness property to conclude our existence proof.

*Randomized quantization technique:* Unlike our previous work [1], the quantization performed at Alice’s side is not deterministic. We introduce a new *randomized quantization* step inspired by the randomized rounding technique in [45]. Essentially, this technique allows to round a continuous Gaussian into a *discrete Gaussian distribution* with slightly larger variance, provided that the  $L^\infty$  flatness factor of the lattice is small. We partially extend the result of [45] under an  $L^1$  flatness factor criterion. We show that randomized quantization with uniform dithering (where the dither is known by all parties, including the eavesdropper) achieves the optimal trade-off between public communication rate and secret key rate established in [23]. The dithering technique has been used to achieve capacity in literature [46, 47]. Besides, the discrete Gaussian distribution is widely used in lattice coding [35] and lattice-based cryptography [37, 45]. However, its application to quantization is new, to the best of our knowledge.

*Relation to fuzzy extractors:* Fuzzy extractors [48] allow to extract a secret key from a noisy measurement, which means that it is resilient to small measurement errors. Fuzzy extractors for continuous signals were proposed in [49, 50]. Our proposed lattice code is also robust to measurement errors, thanks to its channel coding component of Wyner-Ziv coding. A notable difference is that min-entropy is used to measure the available randomness in

<sup>2</sup>The scheme in [25, Section IV-B] requires the repetition of a dithered quantization and public communication step over  $N$  blocks, each of dimension  $n$ . This is needed to achieve strong secrecy from weak secrecy by using the technique in [28]. In contrast, our scheme achieves strong secrecy with a single block and bounds the mutual information using the variational distance, as in [29].

fuzzy extractors, while Shannon entropy is used in our key generation model. Moreover, for fuzzy extractors the measurement error is assumed to have bounded Hamming weight or Euclidean norm, while in our model it follows a Gaussian distribution.

*Organization:* This paper is organized as follows. In Section II we provide basic definitions about lattices and introduce the flatness factor and its  $L^1$  variant, which allows us to define the notion of  $L^1$  secrecy-good lattices. In Section III, we introduce the Gaussian source model and describe our lattice-based secret key generation scheme. Finally, in Section IV we offer some conclusions and perspectives. For ease of reading, the additional technical tools needed to prove the existence of good nested lattices are presented in the Appendix. More precisely, Appendix A summarizes some relevant results on the existence of resolvability codes for regular channels. Appendix B presents the KL flatness factor and its properties, and discusses how the modulo lattice operation allows to extract the intrinsic randomness of the additive Gaussian channel. The existence of lattices that are KL secrecy-good and, consequently, also  $L^1$  secrecy-good is proven in Appendix C. Finally, the existence of the sequences of nested lattices required in our key generation scheme is proven in Appendix D.

## II. LATTICES AND FLATNESS FACTOR

*Notation:* All logarithms in this paper are assumed to be natural logarithms, and information is measured in nats. Given a set  $A$ , the notation  $\mathcal{U}_A$  stands for the uniform distribution over  $A$ . The notation  $\mathbb{F}_p$  refers to the finite field of order  $p$ . We denote the variational distance between two (discrete or continuous) distributions  $p, q$  by  $\mathbb{V}(p, q)$ , and their KL divergence by  $\mathbb{D}(p||q)$ .

### A. Lattice definitions

In this section, we introduce the mathematical tools we use to describe and analyze our proposed scheme.

An  $n$ -dimensional lattice  $\Lambda$  in the Euclidean space  $\mathbb{R}^n$  is the discrete set defined by

$$\Lambda = \mathcal{L}(\mathbf{B}) = \{\mathbf{B}\mathbf{x} : \mathbf{x} \in \mathbb{Z}^n\}$$

where the columns of the basis matrix  $\mathbf{B} = [\mathbf{b}_1 \cdots \mathbf{b}_n]$  are linearly independent.

Given a lattice  $\Lambda$ , its dual lattice  $\Lambda^*$  is defined as the set of vectors  $\lambda^*$  in  $\mathbb{R}^n$  such that  $\langle \lambda^*, \lambda \rangle \in \mathbb{Z}$  for all  $\lambda \in \Lambda$ .

A measurable set  $\mathcal{R}(\Lambda) \subset \mathbb{R}^n$  is called a fundamental region of the lattice  $\Lambda$  if the disjoint union  $\cup_{\lambda \in \Lambda} (\mathcal{R}(\Lambda) + \lambda) = \mathbb{R}^n$ . Examples of fundamental regions include the fundamental parallelepiped  $\mathcal{P}(\Lambda)$  and the Voronoi region  $\mathcal{V}(\Lambda)$ . All the fundamental regions have equal volume  $V(\Lambda)$ .

Given a lattice  $\Lambda$  and a fundamental region  $\mathcal{R}(\Lambda)$ , any point  $\mathbf{x} \in \mathbb{R}^n$  can be written uniquely as a sum

$$\mathbf{x} = \lambda + \bar{\mathbf{x}},$$

where  $\lambda \in \Lambda$  and  $\bar{\mathbf{x}} \in \mathcal{R}(\Lambda)$ . The vector  $\lambda$  is the quantization of  $\mathbf{x}$  with respect to  $\mathcal{R}(\Lambda)$  and is denoted as  $Q_{\mathcal{R}(\Lambda)}(\mathbf{x})$ , where boundary points are decided systematically. Thus we define

$$[\mathbf{x}] \bmod \mathcal{R}(\Lambda) = \mathbf{x} - Q_{\mathcal{R}(\Lambda)}(\mathbf{x}) = \bar{\mathbf{x}}. \quad (1)$$

In particular, for any  $\mathbf{x} \in \mathbb{R}^n$ , the nearest-neighbor quantizer associated with  $\Lambda$  is given by

$$Q_{\Lambda}(\mathbf{x}) = Q_{\mathcal{V}(\Lambda)}(\mathbf{x}) = \arg \min_{\lambda \in \Lambda} \|\lambda - \mathbf{x}\|$$

where ties are broken systematically. Note that  $\mathbf{x} \bmod \mathcal{V}(\Lambda) = \mathbf{x} - Q_{\Lambda}(\mathbf{x})$ . The modulo lattice operation satisfies the distributive law [51, Proposition 2.3.1], i.e.,  $\forall \lambda \in \Lambda$

$$[\mathbf{x} + \lambda] \bmod \mathcal{R}(\Lambda) = [\mathbf{x}] \bmod \mathcal{R}(\Lambda). \quad (2)$$

The following property [52, equation (35)] will also be used in the paper: given two lattices  $\Lambda \subseteq \Lambda_1$ ,  $\mathbf{x} \in \mathbb{R}^n$ , and a fundamental region  $\mathcal{R}(\Lambda)$ ,

$$[Q_{\Lambda_1}(\mathbf{x})] \bmod \mathcal{R}(\Lambda) = [Q_{\Lambda_1}([\mathbf{x}] \bmod \mathcal{R}(\Lambda))] \bmod \mathcal{R}(\Lambda). \quad (3)$$

Given a sublattice  $\Lambda' \subset \Lambda$ , the quotient group  $\Lambda/\Lambda'$  is defined as the group of distinct cosets  $\lambda + \Lambda'$  for  $\lambda \in \Lambda$ . It can be identified by a set of coset representatives  $\Lambda \cap \mathcal{R}(\Lambda')$ , where  $\mathcal{R}(\Lambda')$  is any fundamental region of  $\Lambda'$ .

Furthermore,  $\mathcal{R}(\Lambda')$  can be written as a disjoint union of translates of any fundamental region  $\mathcal{R}(\Lambda)$  as follows [51, equation (8.33)]:

$$\mathcal{R}(\Lambda') = \bigcup_{\lambda \in \Lambda \cap \mathcal{R}(\Lambda')} ([\lambda + \mathcal{R}(\Lambda)] \bmod \mathcal{R}(\Lambda')). \quad (4)$$

### B. Gaussian distributions and the $L^\infty$ flatness factor

Suppose that  $\mathbf{X}^n$  is an  $n$ -dimensional i.i.d. Gaussian random variable of variance  $\sigma^2$  with distribution

$$f_\sigma(\mathbf{x}) = \frac{1}{(\sqrt{2\pi}\sigma)^n} e^{-\frac{\|\mathbf{x}\|^2}{2\sigma^2}},$$

for  $\mathbf{x} \in \mathbb{R}^n$ . The following useful property characterizing the product of Gaussian distributions was proven in [45, Fact 2.1]<sup>3</sup>:

*Lemma 1:* Given  $\sigma_1, \sigma_2 > 0$ , let  $\sigma$  and  $\bar{\sigma}$  be such that  $\sigma^2 = \sigma_1^2 + \sigma_2^2$ , and  $\frac{1}{\bar{\sigma}^2} = \frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}$ . Moreover, let  $\mathbf{c}_1, \mathbf{c}_2 \in \mathbb{R}^n$ , and  $\bar{\mathbf{c}} = \frac{\bar{\sigma}^2}{\sigma_1^2} \mathbf{c}_1 + \frac{\bar{\sigma}^2}{\sigma_2^2} \mathbf{c}_2$ . Then  $\forall \mathbf{x} \in \mathbb{R}^n$ ,

$$f_{\sigma_1}(\mathbf{x} - \mathbf{c}_1) f_{\sigma_2}(\mathbf{x} - \mathbf{c}_2) = f_\sigma(\mathbf{c}_1 - \mathbf{c}_2) f_{\bar{\sigma}}(\mathbf{x} - \bar{\mathbf{c}}).$$

Given a lattice  $\Lambda$ , we define the  $\Lambda$ -periodic function

$$f_{\sigma, \Lambda}(\mathbf{x}) = \frac{1}{(\sqrt{2\pi}\sigma)^n} \sum_{\lambda \in \Lambda} e^{-\frac{\|\mathbf{x} + \lambda\|^2}{2\sigma^2}}, \quad (5)$$

for all  $\mathbf{x} \in \mathbb{R}^n$ . We denote by  $f_{\sigma, \mathcal{R}(\Lambda)} = f_{\sigma, \Lambda}|_{\mathcal{R}(\Lambda)}$  its restriction to the fundamental region  $\mathcal{R}(\Lambda)$ . Note that  $f_{\sigma, \mathcal{R}(\Lambda)}$  is the probability density of  $\bar{\mathbf{X}}^n = [\mathbf{X}^n] \bmod \mathcal{R}(\Lambda)$ . Given  $\mathbf{c} \in \mathbb{R}^n$ , we will also use the notation

$$f_{\sigma, \Lambda, \mathbf{c}}(\mathbf{x}) = f_{\sigma, \Lambda}(\mathbf{x} - \mathbf{c})$$

to denote a shifted  $\Lambda$ -periodic function.

*Definition 1 ( $L^\infty$  Flatness factor [35]):* For a lattice  $\Lambda$  and for a parameter  $\sigma$ , the  $L^\infty$  flatness factor is defined by:

$$\epsilon_\Lambda(\sigma) \triangleq \max_{\mathbf{x} \in \mathcal{R}(\Lambda)} |V(\Lambda) f_{\sigma, \Lambda}(\mathbf{x}) - 1|.$$

In other words,  $\epsilon_\Lambda(\sigma)$  characterizes the  $L^\infty$  distance of  $f_{\sigma, \Lambda}(\mathbf{x})$  to the uniform distribution  $\mathcal{U}_{\mathcal{R}(\Lambda)}$  over  $\mathcal{R}(\Lambda)$ .

The  $L^\infty$  flatness factor is independent of the choice of the fundamental region  $\mathcal{R}(\Lambda)$  and can be computed from the theta series of the lattice

$$\Theta_\Lambda(\tau) = \sum_{\lambda \in \Lambda} e^{-\pi\tau\|\lambda\|^2} \quad (6)$$

using the identity [35, Proposition 2]

$$\epsilon_\Lambda(\sigma) = \left( \frac{\gamma_\Lambda(\sigma)}{2\pi} \right)^{\frac{n}{2}} \Theta_\Lambda \left( \frac{1}{2\pi\sigma^2} \right) - 1, \quad (7)$$

where  $\gamma_\Lambda(\sigma) = \frac{V(\Lambda)^{\frac{2}{n}}}{\sigma^2}$  is the volume-to-noise ratio (VNR). Moreover, the following relation holds between the flatness factor of  $\Lambda$  and the theta series of its dual lattice  $\Lambda^*$  [35, Corollary 1]:

$$\Theta_{\Lambda^*}(2\pi\sigma^2) = \epsilon_\Lambda(\sigma) + 1. \quad (8)$$

*Remark 1:* We have shown in [35] that  $\epsilon_\Lambda$  is a monotonically decreasing function, i.e., for  $\sigma < \sigma'$ , we have  $\epsilon_\Lambda(\sigma') \leq \epsilon_\Lambda(\sigma)$ .

The notion of secrecy-goodness characterizes lattice sequences whose  $L^\infty$  flatness factors vanish exponentially fast as  $n \rightarrow \infty$ .

*Definition 2 (Secrecy-good lattices under  $L^\infty$  flatness factor [35]):* A sequence of lattices  $\Lambda^{(n)}$  is *secrecy-good* under the  $L^\infty$  flatness factor if  $\epsilon_{\Lambda^{(n)}}(\sigma) = e^{-\Omega(n)}$  for all fixed  $\gamma_{\Lambda^{(n)}}(\sigma) < 2\pi$ .

<sup>3</sup>Note that although the statement in [45] refers to (unnormalized) Gaussian functions, one can check that it also holds for Gaussian distributions.

In [35] we have proven the existence of sequences of secrecy-good lattices under  $L^\infty$  flatness factor as long as

$$\gamma_\Lambda(\sigma) < 2\pi. \quad (9)$$

### C. The $L^1$ flatness factor

In this section, we introduce a weaker notion of flatness based on the  $L^1$  distance.

*Definition 3:* Given a lattice  $\Lambda$ , a fundamental region  $\mathcal{R}(\Lambda)$  and  $\sigma > 0$ , we define the  $L^1$  flatness factor as follows:

$$\epsilon_\Lambda^1(\sigma) = \int_{\mathcal{R}(\Lambda)} \left| f_{\sigma, \Lambda}(\mathbf{x}) - \frac{1}{V(\Lambda)} \right| d\mathbf{x} = \mathbb{V}(f_{\sigma, \mathcal{R}(\Lambda)}, \mathcal{U}_{\mathcal{R}(\Lambda)}). \quad (10)$$

Similarly to the  $L^\infty$  flatness factor, the  $L^1$  flatness factor does not depend on the choice of the fundamental region. Moreover, it is shift-invariant, i.e.  $\forall \mathbf{c} \in \mathbb{R}^n$ ,

$$\epsilon_\Lambda^1(\sigma) = \mathbb{V}(f_{\sigma, \Lambda, \mathbf{c}}|_{\mathcal{R}(\Lambda)}, \mathcal{U}_{\mathcal{R}(\Lambda)}). \quad (11)$$

*Remark 2:* For any lattice  $\Lambda$ ,  $\forall \sigma > 0$ , we have  $\epsilon_\Lambda^1(\sigma) \leq \epsilon_\Lambda(\sigma)$ .

The  $L^1$  flatness factor is related to the  $L^1$  smoothing parameter, which was discussed in [37, 38].

The following Lemma confirms the intuition that folded additive Gaussian noise with larger variance looks more uniform:

*Lemma 2:* The  $L^1$  flatness factor is monotonic, i.e. for any lattice  $\Lambda$ ,  $\forall \sigma' > \sigma$ ,

$$\epsilon_\Lambda^1(\sigma') \leq \epsilon_\Lambda^1(\sigma).$$

*Proof:* Let  $W^n \sim \mathcal{N}(0, \sigma^2 I_n)$  and  $X^n = W^n \bmod \mathcal{R}(\Lambda) \sim f_{\sigma, \mathcal{R}(\Lambda)}$ .

Given  $\sigma_0 > 0$ , let  $W_0^n \sim \mathcal{N}(0, \sigma_0^2 I_n)$  and consider

$$Y^n = [X^n + W_0^n] \bmod \mathcal{R}(\Lambda) = [[W^n] \bmod \mathcal{R}(\Lambda) + W_0^n] \bmod \mathcal{R}(\Lambda) \stackrel{(a)}{=} [W^n + W_0^n] \bmod \mathcal{R}(\Lambda) \sim f_{\sqrt{\sigma^2 + \sigma_0^2}, \mathcal{R}(\Lambda)},$$

where (a) follows from the distributive property (2). Now consider the random variable  $U^n \sim \mathcal{U}_{\mathcal{R}(\Lambda)}$ . By the Crypto Lemma [51, Lemma 4.1.1],

$$[U^n + W_0^n] \bmod \mathcal{R}(\Lambda) \sim \mathcal{U}_{\mathcal{R}(\Lambda)}.$$

Then using the data processing inequality for the variational distance,

$$\epsilon_\Lambda^1 \left( \sqrt{\sigma^2 + \sigma_0^2} \right) = \mathbb{V} \left( f_{\sqrt{\sigma^2 + \sigma_0^2}, \mathcal{R}(\Lambda)}, \mathcal{U}_{\mathcal{R}(\Lambda)} \right) = \mathbb{V}(Y^n, U^n) \leq \mathbb{V}(X^n, U^n) = \mathbb{V}(f_{\sigma, \mathcal{R}(\Lambda)}, \mathcal{U}_{\mathcal{R}(\Lambda)}) = \epsilon_\Lambda^1(\sigma).$$

Since this is true for any  $\sigma_0 > 0$ , the conclusion follows.  $\square$

We will next show that lattices that are good for secrecy in the  $L^1$  sense exist and that the corresponding volume condition is less stringent than the condition (9) for secrecy-goodness based on the  $L^\infty$  metric.

*Definition 4:* A sequence of lattices  $\{\Lambda^{(n)}\}$  is  $L^1$  secrecy-good if for all fixed  $\gamma_{\Lambda^{(n)}}(\sigma) < 2\pi e$ ,  $\forall c > 0$ ,  $\epsilon_{\Lambda^{(n)}}^1(\sigma) = o\left(\frac{1}{n^c}\right)$ , i.e., the  $L^1$  flatness factor vanishes super-polynomially.

The following theorem, which was presented in [2], is the first main result of this paper:

*Theorem 1:* If  $\gamma_\Lambda(\sigma) < 2\pi e$  is fixed, then there exists a sequence  $\{\Lambda^{(n)}\}$  of lattices which are  $L^1$ -secrecy good.

The proof of Theorem 1 is given in Appendix C. Our proof is information-theoretic and does not require the knowledge of the theta series, in contrast to the  $L^\infty$  flatness factor. We outline the key ideas here. In order to show the existence of a sequence of lattices  $\Lambda^{(n)}$  such that  $\epsilon_{\Lambda^{(n)}}^1(\sigma) = \mathbb{V}(f_{\sigma, \mathcal{R}(\Lambda^{(n)})}, \mathcal{U}_{\mathcal{R}(\Lambda^{(n)})}) \rightarrow 0$ , we actually prove a stronger result, namely that  $\mathbb{D}(f_{\sigma, \mathcal{R}(\Lambda^{(n)})} || \mathcal{U}_{\mathcal{R}(\Lambda^{(n)})}) \rightarrow 0$ . This requires some additional technical tools that are presented in Appendix B. We build the required lattices using Construction A, and their existence follows from the existence of linear resolvability codes in [36] (see Appendix A for more details).

*Remark 3:* It is worth mentioning that as soon as the VNR exceeds  $2\pi$ , the  $L^\infty$  flatness factor increases exponentially. In fact, it is easy to see that the bound  $\gamma_\Lambda(\sigma) < 2\pi$  is sharp: the  $L^\infty$  flatness factor of a lattice cannot vanish for any  $\gamma_\Lambda(\sigma) > 2\pi$ . This is simply because (7) implies that

$$\epsilon_\Lambda(\sigma) > \left( \frac{\gamma_\Lambda(\sigma)}{2\pi} \right)^{\frac{n}{2}} - 1$$

since  $\Theta_\Lambda(\tau) > 1$  for any  $\tau > 0$ . Thus, as the VNR approaches  $2\pi e$ , the  $L^\infty$  flatness factor  $\approx e^{n/2}$ , but the  $L^1$  flatness factor can still be brought under control. This demonstrates the advantage of the  $L^1$  flatness factor.

Also note that the VNR of an  $L^1$ -secrecy-good lattice approaches  $2\pi e$  from below, while that of an AWGN-good lattice approaches  $2\pi e$  from above. Recall that the normalized second moment of a quantization-good lattice approaches  $1/(2\pi e)$  [51], so all three types of lattices finally share the same VNR threshold  $2\pi e$ .

*Remark 4:* In the following, we discuss the implication of Theorem 1 on the smoothing parameter<sup>4</sup> that is commonly used in lattice-based cryptography.

*Definition 5 (Smoothing parameter):* For a lattice  $\Lambda$  and for  $\varepsilon > 0$ , the  $L^\infty$  and  $L^1$  smoothing parameters  $\eta_\varepsilon(\Lambda)$  and  $\eta_\varepsilon^1(\Lambda)$ , respectively, are the smallest  $\sigma > 0$  such that  $\epsilon_\Lambda(\sigma), \epsilon_\Lambda^1(\sigma) \leq \varepsilon$ .

Theorem 1 implies the existence of lattices whose smoothing parameters  $\eta_{\varepsilon_n}^1(\Lambda) \approx \frac{V(\Lambda)^{1/n}}{\sqrt{2\pi e}}$  for a suitable sequence  $\varepsilon_n \rightarrow 0$ . This improves upon the result  $\eta_{\varepsilon_n}(\Lambda) \approx \frac{V(\Lambda)^{1/n}}{\sqrt{2\pi}}$ . Using the Cauchy-Schwarz inequality, the following bound was proven in [40]<sup>5</sup>

$$\epsilon_\Lambda^1(\sigma) \leq \sqrt{\epsilon_\Lambda(\sqrt{2}\sigma)} \quad (12)$$

which implies the bound  $\eta_\varepsilon^1(\Lambda) \leq \frac{V(\Lambda)^{1/n}}{2\sqrt{\pi}}$ . However, this bound is not optimal.

#### D. Discrete Gaussians and randomized rounding

Given an  $n$ -dimensional lattice  $\Lambda$  in  $\mathbb{R}^n$  and a vector  $\mathbf{c} \in \mathbb{R}^n$ , we define the *discrete Gaussian distribution* over  $\Lambda$  centered at  $\mathbf{c}$  as the following discrete distribution taking values in  $\lambda \in \Lambda$ :

$$D_{\Lambda, \sigma, \mathbf{c}}(\lambda) = \frac{f_{\sigma, \mathbf{c}}(\lambda)}{f_{\sigma, \mathbf{c}}(\Lambda)} \quad \forall \lambda \in \Lambda,$$

where  $f_{\sigma, \mathbf{c}}(\Lambda) \triangleq \sum_{\lambda \in \Lambda} f_{\sigma, \mathbf{c}}(\lambda)$ . We write  $D_{\Lambda, \sigma} = D_{\Lambda, \sigma, \mathbf{0}}$ .

Following Peikert [45, Section 4.1], we introduce the notion of randomized rounding with respect to  $\Lambda$ :

*Definition 6 (Randomized rounding):* Given an input vector  $\mathbf{x} \in \mathbb{R}^n$ , we define the random variable

$$\lfloor \mathbf{x} \rfloor_{\Lambda, \sigma} \sim D_{\Lambda, \sigma, \mathbf{x}}.$$

Note that  $\lfloor \mathbf{x} \rfloor_{\Lambda, \sigma}$  is a discrete random variable taking values in  $\Lambda$ .

It was shown in [45] that when  $X^n$  is i.i.d. Gaussian with variance  $\sigma^2$ , the randomly rounded variable  $\lfloor X^n \rfloor_{\Lambda, \sigma_Q}$  is close in variational distance to the discrete Gaussian  $D_{\Lambda, \tilde{\sigma}}$ , where  $\tilde{\sigma}^2 = \sigma^2 + \sigma_Q^2$ , provided that the  $L^\infty$  flatness factor  $\epsilon_\Lambda(\sigma_Q)$  is small:

*Proposition 1 (Adapted from Theorem 3.1 of [45]):* Let  $X^n \sim \mathcal{N}(0, \sigma^2 I_n)$  and  $\boldsymbol{\mu} \in \mathbb{R}^n$ , and consider  $X_Q = \lfloor X^n + \boldsymbol{\mu} \rfloor_{\Lambda, \sigma_Q}$ . If  $\epsilon_\Lambda(\sigma_Q) < 1/2$ , then

$$\mathbb{V}(p_{X_Q}, D_{\Lambda, \tilde{\sigma}, \boldsymbol{\mu}}) \leq 4\epsilon_\Lambda(\sigma_Q),$$

where  $\tilde{\sigma}^2 = \sigma^2 + \sigma_Q^2$ .

In the following, we prove a partial generalization of this result under an  $L^1$  flatness factor condition, for randomized rounding with uniform dithering, which may be of independent interest.

*Lemma 3:* Given a Gaussian random vector  $X^n \sim \mathcal{N}(0, \sigma^2 I_n)$ , a dither  $\mathbf{U} \sim \mathcal{U}_{\mathcal{R}(\Lambda)}$  uniform over a fundamental region  $\mathcal{R}(\Lambda)$  and independent of  $X^n$ , and a constant  $\boldsymbol{\mu} \in \mathbb{R}^n$ , let  $X_Q = \lfloor X^n + \mathbf{U} + \boldsymbol{\mu} \rfloor_{\Lambda, \sigma_Q}$ . Then

$$\mathbb{E}_{\mathbf{U}} [\mathbb{V}(p_{X_Q|\mathbf{U}}, D_{\Lambda, \tilde{\sigma}, \mathbf{U} + \boldsymbol{\mu}})] \leq 2\epsilon_\Lambda^1(\sigma_Q).$$

In order to prove Lemma 3, we need the following intermediate Lemma, which will be used several times throughout the paper.

*Lemma 4:* Suppose that  $\tilde{\sigma}^2 = \sigma^2 + \sigma_Q^2$ , and let  $\mathcal{R}(\Lambda)$  be a fundamental region of  $\Lambda$ . Then the following inequality holds:

$$\sum_{\mathbf{x}_Q \in \Lambda} \int_{\mathcal{R}(\Lambda)} \left| \int_{\mathbb{R}^n} \frac{f_\sigma(\mathbf{x} - \boldsymbol{\mu}) f_{\sigma_Q}(\mathbf{x}_Q - \mathbf{x} - \mathbf{u})}{V(\Lambda) f_{\sigma_Q}(\Lambda - \mathbf{x} - \mathbf{u})} d\mathbf{x} - f_{\tilde{\sigma}}(\mathbf{x}_Q - \mathbf{u} - \boldsymbol{\mu}) \right| d\mathbf{u} \leq \epsilon_\Lambda^1(\sigma_Q).$$

<sup>4</sup>We remark that this definition differs slightly from the one in [33], where  $\sigma$  is scaled by a constant factor  $\sqrt{2\pi}$  (i.e.,  $s = \sqrt{2\pi}\sigma$ ).

<sup>5</sup>A similar bound was given in [41] using the statistical distance, which differs from the  $L^1$  distance by a factor  $\frac{1}{2}$ .

*Proof of Lemma 4:* By Lemma 1,

$$f_{\sigma_Q}(\mathbf{x}_Q - \mathbf{x} - \mathbf{u})f_{\sigma}(\mathbf{x} - \boldsymbol{\mu}) = f_{\tilde{\sigma}}(\mathbf{x}_Q - \mathbf{u} - \boldsymbol{\mu})f_{\tilde{\sigma}}(\mathbf{x} - \bar{\mathbf{c}}(\mathbf{x}_Q, \mathbf{u}, \boldsymbol{\mu})), \quad (13)$$

where  $\frac{1}{\tilde{\sigma}^2} = \frac{1}{\sigma^2} + \frac{1}{\sigma_Q^2}$  and  $\bar{\mathbf{c}}(\mathbf{x}_Q, \mathbf{u}, \boldsymbol{\mu}) = \frac{\tilde{\sigma}^2}{\sigma_Q^2}(\mathbf{x}_Q - \mathbf{u}) + \frac{\tilde{\sigma}^2}{\sigma^2}\boldsymbol{\mu}$ . Then we can write

$$\begin{aligned} & \sum_{\mathbf{x}_Q \in \Lambda} \int_{\mathcal{R}(\Lambda)} \left| \int_{\mathbb{R}^n} \frac{f_{\sigma}(\mathbf{x} - \boldsymbol{\mu})f_{\sigma_Q}(\mathbf{x}_Q - \mathbf{x} - \mathbf{u})}{V(\Lambda)f_{\sigma_Q}(\Lambda - \mathbf{x} - \mathbf{u})} d\mathbf{x} - f_{\tilde{\sigma}}(\mathbf{x}_Q - \mathbf{u} - \boldsymbol{\mu}) \right| d\mathbf{u} \\ & \stackrel{(a)}{=} \sum_{\mathbf{x}_Q \in \Lambda} \int_{\mathcal{R}(\Lambda)} \left| \int_{\mathbb{R}^n} \frac{f_{\sigma}(\mathbf{x} - \boldsymbol{\mu})f_{\sigma_Q}(\mathbf{x}_Q - \mathbf{x} - \mathbf{u})}{V(\Lambda)f_{\sigma_Q}(\Lambda - \mathbf{x} - \mathbf{u})} d\mathbf{x} - f_{\tilde{\sigma}}(\mathbf{x}_Q - \mathbf{u} - \boldsymbol{\mu}) \int_{\mathbb{R}^n} f_{\tilde{\sigma}}(\mathbf{x} - \bar{\mathbf{c}}(\mathbf{x}_Q, \mathbf{u}, \boldsymbol{\mu})) d\mathbf{x} \right| d\mathbf{u} \\ & \stackrel{(b)}{=} \sum_{\mathbf{x}_Q \in \Lambda} \int_{\mathcal{R}(\Lambda)} \left| \int_{\mathbb{R}^n} \frac{f_{\sigma}(\mathbf{x} - \boldsymbol{\mu})f_{\sigma_Q}(\mathbf{x}_Q - \mathbf{x} - \mathbf{u})}{V(\Lambda)f_{\sigma_Q}(\Lambda - \mathbf{x} - \mathbf{u})} d\mathbf{x} - \int_{\mathbb{R}^n} f_{\sigma_Q}(\mathbf{x}_Q - \mathbf{x} - \mathbf{u})f_{\sigma}(\mathbf{x} - \boldsymbol{\mu}) d\mathbf{x} \right| d\mathbf{u} \\ & \leq \int_{\mathcal{R}(\Lambda)} \int_{\mathbb{R}^n} \sum_{\mathbf{x}_Q \in \Lambda} \frac{f_{\sigma}(\mathbf{x} - \boldsymbol{\mu})f_{\sigma_Q}(\mathbf{x}_Q - \mathbf{x} - \mathbf{u})}{f_{\sigma_Q}(\Lambda - \mathbf{x} - \mathbf{u})} \left| \frac{1}{V(\Lambda)} - f_{\sigma_Q}(\Lambda - \mathbf{x} - \mathbf{u}) \right| d\mathbf{x} d\mathbf{u} \\ & = \int_{\mathbb{R}^n} f_{\sigma}(\mathbf{x} - \boldsymbol{\mu}) \int_{\mathcal{R}(\Lambda)} \left| \frac{1}{V(\Lambda)} - f_{\Lambda, \sigma_Q}(\mathbf{x} + \mathbf{u}) \right| d\mathbf{u} d\mathbf{x} \\ & = \int_{\mathbb{R}^n} f_{\sigma}(\mathbf{x} - \boldsymbol{\mu}) \int_{\mathcal{R}(\Lambda)} \left| \frac{1}{V(\Lambda)} - f_{\Lambda, \sigma_Q}(\mathbf{u}) \right| d\mathbf{u} d\mathbf{x} = \epsilon_{\Lambda}^1(\sigma_Q), \end{aligned}$$

where (a) follows from the fact that  $\int_{\mathbb{R}^n} f_{\tilde{\sigma}}(\mathbf{x} - \bar{\mathbf{c}}(\mathbf{x}_Q, \mathbf{u}, \boldsymbol{\mu})) d\mathbf{x} = 1$ , and (b) follows from (13).  $\square$

*Proof of Lemma 3:* We have

$$\begin{aligned} & \mathbb{E}_{\mathbf{U}} [\mathbb{V}(p_{\mathbf{X}_Q|\mathbf{U}}, D_{\Lambda, \tilde{\sigma}, \mathbf{U} + \boldsymbol{\mu}})] = \\ & = \int_{\mathcal{R}(\Lambda)} \frac{1}{V(\Lambda)} \sum_{\mathbf{x}_Q \in \Lambda} |p_{\mathbf{X}_Q|\mathbf{U}}(\mathbf{x}_Q|\mathbf{u}) - D_{\Lambda, \tilde{\sigma}, \mathbf{u} + \boldsymbol{\mu}}(\mathbf{x}_Q)| d\mathbf{u} \\ & = \int_{\mathcal{R}(\Lambda)} \frac{1}{V(\Lambda)} \sum_{\mathbf{x}_Q \in \Lambda} \left| \int_{\mathbb{R}^n} p_{\mathbf{X}_Q|\mathbf{X}^n, \mathbf{U}}(\mathbf{x}_Q|\mathbf{x}, \mathbf{u}) p_{\mathbf{X}^n}(\mathbf{x}) d\mathbf{x} - \frac{f_{\tilde{\sigma}}(\mathbf{x}_Q - \mathbf{u} - \boldsymbol{\mu})}{f_{\tilde{\sigma}}(\Lambda - \mathbf{u} - \boldsymbol{\mu})} \right| d\mathbf{u} \\ & = \sum_{\mathbf{x}_Q \in \Lambda} \int_{\mathcal{R}(\Lambda)} \frac{1}{V(\Lambda)} \left| \int_{\mathbb{R}^n} \frac{f_{\sigma}(\mathbf{x})f_{\sigma_Q}(\mathbf{x}_Q - \mathbf{x} - \mathbf{u} - \boldsymbol{\mu})}{f_{\sigma_Q}(\Lambda - \mathbf{x} - \mathbf{u} - \boldsymbol{\mu})} d\mathbf{x} - \frac{f_{\tilde{\sigma}}(\mathbf{x}_Q - \mathbf{u} - \boldsymbol{\mu})}{f_{\tilde{\sigma}}(\Lambda - \mathbf{u} - \boldsymbol{\mu})} \right| d\mathbf{u} \\ & \stackrel{(a)}{\leq} \sum_{\mathbf{x}_Q \in \Lambda} \int_{\mathcal{R}(\Lambda)} \frac{1}{V(\Lambda)} \left| \int_{\mathbb{R}^n} \frac{f_{\sigma}(\mathbf{x} - \boldsymbol{\mu})f_{\sigma_Q}(\mathbf{x}_Q - \mathbf{x} - \mathbf{u})}{f_{\sigma_Q}(\Lambda - \mathbf{x} - \mathbf{u})} d\mathbf{x} - f_{\tilde{\sigma}}(\mathbf{x}_Q - \mathbf{u} - \boldsymbol{\mu})V(\Lambda) \right| d\mathbf{u} \quad (14) \end{aligned}$$

$$+ \sum_{\mathbf{x}_Q \in \Lambda} \int_{\mathcal{R}(\Lambda)} \frac{1}{V(\Lambda)} \left| f_{\tilde{\sigma}}(\mathbf{x}_Q - \mathbf{u} - \boldsymbol{\mu})V(\Lambda) - \frac{f_{\tilde{\sigma}}(\mathbf{x}_Q - \mathbf{u} - \boldsymbol{\mu})}{f_{\tilde{\sigma}}(\Lambda - \mathbf{u} - \boldsymbol{\mu})} \right| d\mathbf{u}, \quad (15)$$

where (a) follows from the triangle inequality and the change of variables  $\mathbf{x} \mapsto \mathbf{x} + \boldsymbol{\mu}$ . The term (14) is bounded by  $\epsilon_{\Lambda}^1(\sigma_Q)$  because of Lemma 4. The term (15) is equal to

$$\begin{aligned} & \sum_{\mathbf{x}_Q \in \Lambda} \int_{\mathcal{R}(\Lambda)} \frac{f_{\tilde{\sigma}}(\mathbf{x}_Q - \mathbf{u} - \boldsymbol{\mu})}{f_{\tilde{\sigma}}(\Lambda - \mathbf{u} - \boldsymbol{\mu})} \left| f_{\tilde{\sigma}}(\Lambda - \mathbf{u} - \boldsymbol{\mu}) - \frac{1}{V(\Lambda)} \right| d\mathbf{u} \\ & = \int_{\mathcal{R}(\Lambda)} \left| f_{\tilde{\sigma}}(\Lambda - \mathbf{u} - \boldsymbol{\mu}) - \frac{1}{V(\Lambda)} \right| d\mathbf{u} = \epsilon_{\Lambda}^1(\tilde{\sigma}) \stackrel{(b)}{\leq} \epsilon_{\Lambda}^1(\sigma_Q), \end{aligned}$$

where (b) follows from Lemma 8.  $\square$

### III. SECRET KEY GENERATION

In this section, we present our system model for secret key generation from correlated Gaussian sources with one-way rate limited communication, in the presence of an eavesdropper, and our proposed key generation protocol based on nested lattices.

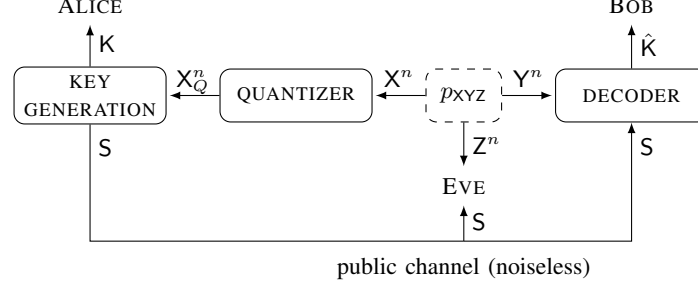


Fig. 1. Secret key generation in the presence of an eavesdropper with communication over a public channel.

#### A. System model

We consider the same model as in [1], illustrated in Fig. 1, in which Alice, Bob and Eve observe the random variables  $X^n$ ,  $Y^n$ ,  $Z^n$  respectively, generated by an i.i.d. memoryless Gaussian source  $p_{XYZ}$  whose components are jointly Gaussian with zero mean. The distribution is fully described by the variances  $\sigma_x^2$ ,  $\sigma_y^2$ ,  $\sigma_z^2$  and the correlation coefficients  $\rho_{xy}$ ,  $\rho_{xz}$ ,  $\rho_{yz}$ . We can write [23, Eq. (6)]:

$$\begin{cases} X^n = \rho_{xy} \frac{\sigma_x}{\sigma_y} Y^n + W_1^n, \\ X^n = \rho_{xz} \frac{\sigma_x}{\sigma_z} Z^n + W_2^n, \end{cases} \quad (16)$$

where  $W_1^n$  and  $W_2^n$  are i.i.d. zero-mean Gaussian noise vectors of variances

$$\sigma_1^2 = \sigma_x^2(1 - \rho_{xy}^2), \quad \sigma_2^2 = \sigma_x^2(1 - \rho_{xz}^2), \quad (17)$$

respectively, such that  $\sigma_2 > \sigma_1$ . Further,  $W_1^n$  is independent of  $Y^n$ , and  $W_2^n$  is independent of  $Z^n$ .

We assume that only one round of one-way public communication takes place from Alice to Bob. More precisely, Alice computes a public message  $S$  and a secret key  $K$  from her observation  $X^n$ ; she then transmits  $S$  over the public channel (see Fig. 1). From this message and his own observation  $Y^n$ , Bob reconstructs a key  $\hat{K}$ .

Let  $\mathcal{K}_n$  and  $\mathcal{S}_n$  be the sets of secret keys and public messages respectively. A *secret key rate - public rate pair*  $(R_K, R_P)$  is achievable if there exists a sequence of protocols with

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \log |\mathcal{K}_n| \geq R_K, \quad \limsup_{n \rightarrow \infty} \frac{1}{n} \log |\mathcal{S}_n| \leq R_P,$$

such that the following properties hold:

$$\begin{aligned} \lim_{n \rightarrow \infty} \log |\mathcal{K}_n| - \mathbb{H}(K) &= 0 && \text{(uniformity)} \\ \lim_{n \rightarrow \infty} \mathbb{P} \{K \neq \hat{K}\} &= 0 && \text{(reliability)} \\ \lim_{n \rightarrow \infty} \mathbb{I}(K; S, Z^n) &= 0 && \text{(strong secrecy)}. \end{aligned}$$

Following [23], we denote

$$\mathcal{R}(X, Y, Z) = \{(R_P, R_K) : (R_P, R_K) \text{ is achievable}\}.$$

The optimal trade-off between secret key rate and public rate was derived in [23]. For the source model (16), given public rate  $R_P$ , the secret key rate is upper bounded by

$$R_K \leq \bar{R}_K(R_P) = \frac{1}{2} \log \left( e^{-2R_P} + \frac{\sigma_2^2}{\sigma_1^2} (1 - e^{-2R_P}) \right). \quad (18)$$

See Appendix E for details.

We recall that the secret key capacity of the Gaussian source model (16) is defined as the maximum achievable



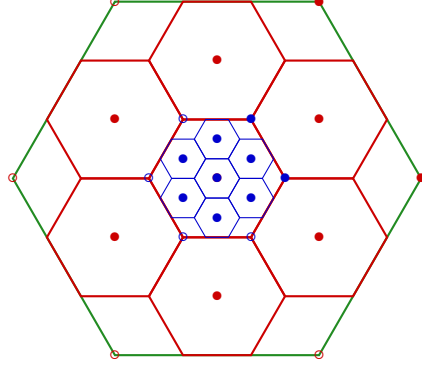


Fig. 2. A schematic representation of the chain of nested lattices  $\Lambda_1 \supset \Lambda_2 \supset \Lambda_3$ . The fundamental regions of  $\Lambda_1$ ,  $\Lambda_2$  and  $\Lambda_3$  are pictured in blue, red and green respectively. The quotient groups  $\Lambda_1/\Lambda_2$  and  $\Lambda_2/\Lambda_3$  are represented by the blue and red points respectively.

secret key rate with unlimited public communication and is given by [53, 23]

$$C_s = \sup \{R_K \text{ such that } \exists R_P \geq 0 : (R_P, R_K) \in \mathcal{R}(X, Y, Z)\} = \frac{1}{2} \log \frac{\sigma_2^2}{\sigma_1^2}. \quad (19)$$

*Additional notation.* To simplify notation, we define  $\hat{Y}^n = \rho_{xy} \frac{\sigma_x}{\sigma_y} Y^n$  and  $\hat{Z}^n = \rho_{xz} \frac{\sigma_x}{\sigma_z} Z^n$ , so that

$$\begin{cases} X^n = \hat{Y}^n + W_1^n, \\ X^n = \hat{Z}^n + W_2^n, \end{cases} \quad (20)$$

where  $\hat{Y}^n$  and  $W_1^n$  are independent, and  $\hat{Z}^n$  and  $W_2^n$  are independent. We denote the variances of  $\hat{Y}^n$  and  $\hat{Z}^n$  by  $\hat{\sigma}_y = \rho_{xy} \sigma_x = \sqrt{\sigma_x^2 - \sigma_1^2}$  and  $\hat{\sigma}_z = \rho_{xz} \sigma_x = \sqrt{\sigma_x^2 - \sigma_2^2}$  respectively.

### B. Secret key generation protocol

To define our key generation scheme, we use the lattice partition chain  $\Lambda_1/\Lambda_2/\Lambda_3$ , where

- $\Lambda_1$  is  $L^1$  secrecy-good with respect to  $\sigma_Q$ , and serves as the “source-code” component of Wyner-Ziv coding;
- $\Lambda_2$  is AWGN-good with respect to  $\tilde{\sigma}_1 = \sqrt{\sigma_1^2 + \sigma_Q^2}$ , and serves as the “channel-code” component in Wyner-Ziv coding;
- $\Lambda_3$  is  $L^1$  secrecy-good with respect to  $\tilde{\sigma}_2 = \sqrt{\sigma_2^2 + \sigma_Q^2}$ , and serves as the extractor of randomness.

The existence of such a chain of lattices will be established in Appendix D.

In addition, we assume that  $U$  is a uniform dither over a fundamental region  $\mathcal{R}(\Lambda_1)$ , which is known by Alice, Bob and Eve<sup>6</sup>.

Our protocol is similar to the secret key generation scheme in our previous work [1] with some notable differences due to switching from an  $L^\infty$  flatness factor criterion to an  $L^1$  flatness factor criterion:

- As in [1], the modulo  $\mathcal{R}(\Lambda_3)$  operation is used for privacy amplification. Since the flatness factor  $\epsilon_{\Lambda_3}^1(\sigma)$  only depends on  $f_{\sigma, \Lambda_3}$  which is periodic mod  $\Lambda_3$ , nearest-neighbor quantization is not needed and we can choose any fundamental region  $\mathcal{R}(\Lambda_3)$ . Note that the mod  $\mathcal{R}(\Lambda)$  operation can be performed in polynomial time for many fundamental regions. In particular, we can choose the fundamental parallelepiped.
- Nearest-neighbor quantization with respect to the intermediate lattice  $\Lambda_2$  is performed for information reconciliation.
- As in [1], quantization with respect to the fine lattice  $\Lambda_1$  is performed to obtain a discrete key. However, deterministic quantization is replaced with randomized rounding (using local randomness at Alice’s side), which allows to achieve the optimal trade-off between secret key rate and public rate. Since the  $L^1$  flatness

<sup>6</sup>If Alice and Bob already share a secret source of randomness, there is no need for secret key generation. Hence, Eve should know  $U$  to avoid trivializing the problem.

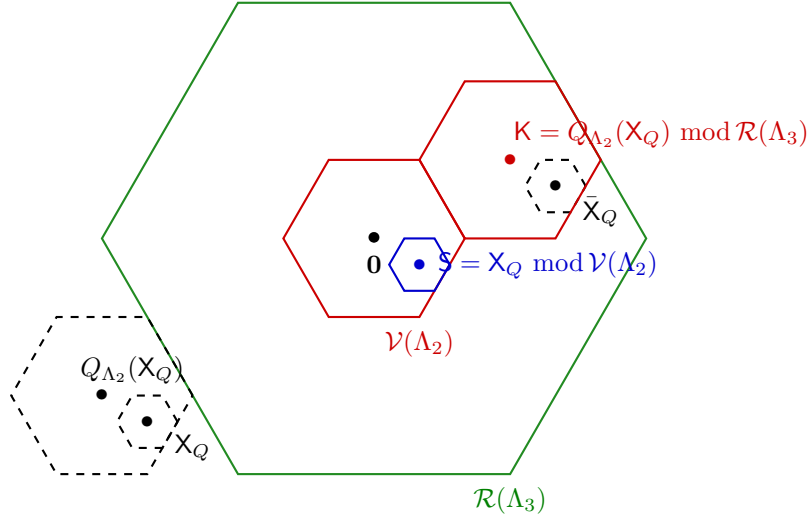


Fig. 3. A schematic representation of the quantized signal  $X_Q$ , the secret key  $K$  and the public message  $S$ .

factor is only an average condition, dithering is required in order to obtain almost uniform keys. Again, since an  $L^1$  flatness factor criterion is used, the dither can be generated uniformly over any fundamental region  $\mathcal{R}(\Lambda_1)$ .

More precisely, the secret key generation proceeds as follows (see Figure 3):

- Alice quantizes  $X^n$  to

$$X_Q = \lfloor X^n + U \rfloor_{\Lambda_1, \sigma_Q}.$$

That is,  $X_Q \sim D_{\Lambda_1, \sigma_Q, \mathbf{x} + \mathbf{u}}$  if  $X^n = \mathbf{x}$ ,  $U = \mathbf{u}$ , or equivalently

$$p_{X_Q|X^n, U}(\mathbf{x}_Q | \mathbf{x}, \mathbf{u}) = \frac{f_{\sigma_Q}(\mathbf{x}_Q - \mathbf{x} - \mathbf{u})}{f_{\sigma_Q}(\Lambda_1 - \mathbf{x} - \mathbf{u})}. \quad (21)$$

Alice then computes the public message  $S$  and the key  $K$  as follows:

$$\begin{aligned} S &= X_Q \bmod \mathcal{V}(\Lambda_2), \\ K &= Q_{\Lambda_2}(X_Q) \bmod \mathcal{R}(\Lambda_3), \end{aligned}$$

and transmits  $S$  to Bob over the public channel.

- Upon receiving  $S$ , Bob reconstructs

$$\hat{X}_Q = S + Q_{\Lambda_2} \left( \rho_{xy} \frac{\sigma_x}{\sigma_y} Y^n + U - S \right).$$

He then computes his version of the key:

$$\hat{K} = Q_{\Lambda_2}(\hat{X}_Q) \bmod \mathcal{R}(\Lambda_3).$$

Let  $\bar{X}_Q = X_Q \bmod \mathcal{R}(\Lambda_3) \in \Lambda_1/\Lambda_3$ , where the quotient  $\Lambda_1/\Lambda_3$  is identified with the set of coset representatives  $\Lambda_1 \cap \mathcal{R}(\Lambda_3)$ . By definition,  $\bar{X}_Q = S + K$ . Note that  $K$  and  $S$  are both functions of  $\bar{X}_Q$ :

$$K = Q_{\Lambda_2}(X_Q) \bmod \mathcal{R}(\Lambda_3) \stackrel{(a)}{=} Q_{\Lambda_2}(X_Q \bmod \mathcal{R}(\Lambda_3)) \bmod \mathcal{R}(\Lambda_3) = Q_{\Lambda_2}(\bar{X}_Q) \bmod \mathcal{R}(\Lambda_3) = f(\bar{X}_Q). \quad (22)$$

where (a) follows from equation (3). Similarly,

$$\begin{aligned} \bar{X}_Q \bmod \Lambda_2 &= \bar{X}_Q - Q_{\Lambda_2}(\bar{X}_Q) = X_Q - Q_{\mathcal{R}(\Lambda_3)}(X_Q) - Q_{\Lambda_2}(X_Q - Q_{\mathcal{R}(\Lambda_3)}(X_Q)) = X_Q - Q_{\Lambda_2}(X_Q) \\ &= X_Q \bmod \Lambda_2 = S = g(\bar{X}_Q). \end{aligned} \quad (23)$$

*Remark 5:* Because of the previous relations, we can conclude that there exists a bijection  $(f, g) : \Lambda_1/\Lambda_3 \rightarrow$

$\Lambda_1/\Lambda_2 \times \Lambda_2/\Lambda_3$  that sends  $\bar{X}_Q$  into the corresponding pair  $(S, K)$ .

We now state the main result of the paper, which will be proven in the following sections:

*Theorem 2:* For the Gaussian source model (16), there exists a sequence of nested lattices  $\Lambda_3^{(n)} \subset \Lambda_2^{(n)} \subset \Lambda_1^{(n)}$  such that for any public rate  $R_P > 0$ , the previous secret key generation protocol asymptotically achieves the optimal secret key rate  $\bar{R}_K(R_P)$  in (18). In particular, any secret key rate  $R_K < C_s = \frac{1}{2} \log \frac{\sigma_2^2}{\sigma_1^2}$  is achievable.

### C. Reliability

We want to show that the error probability  $P_e = \mathbb{P}\{K \neq \hat{K}\} \rightarrow 0$  as  $n \rightarrow \infty$ .

Note that  $K = \hat{K}$  if  $\hat{X}_Q = X_Q$ . Since  $X_Q = S + Q_{\Lambda_2}(X_Q)$ , we have

$$\hat{X}_Q = X_Q \Leftrightarrow Q_{\Lambda_2}(\hat{Y}^n + U - S) = Q_{\Lambda_2}(X_Q).$$

Observe that

$$Q_{\Lambda_2}(\hat{Y}^n + U - S) = Q_{\Lambda_2}(\hat{Y}^n + U - X_Q + Q_{\Lambda_2}(X_Q)) = Q_{\Lambda_2}(\hat{Y}^n + U - X_Q) + Q_{\Lambda_2}(X_Q).$$

Therefore

$$\hat{X}_Q = X_Q \Leftrightarrow Q_{\Lambda_2}(\hat{Y}^n + U - X_Q) = 0 \Leftrightarrow \hat{Y}^n \in X_Q - U + \mathcal{V}(\Lambda_2). \quad (24)$$

The error probability is bounded by

$$\begin{aligned} P_e &\leq \mathbb{P}\{\hat{X}_Q \neq X_Q\} = \int_{\mathbb{R}^n} \int_{\mathcal{R}(\Lambda_1)} \mathbb{P}\{\hat{X}_Q \neq X_Q \mid \hat{Y}^n = \mathbf{y}, U = \mathbf{u}\} \frac{p_{\hat{Y}^n}(\mathbf{y})}{V(\Lambda_1)} d\mathbf{u} d\mathbf{y} \\ &= \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} \int_{\mathcal{R}(\Lambda_1)} \mathbb{P}\{\hat{X}_Q \neq X_Q \mid \hat{Y}^n = \mathbf{y}, X^n = \mathbf{x}, U = \mathbf{u}\} \frac{p_{X^n|\hat{Y}^n}(\mathbf{x}|\mathbf{y}) p_{\hat{Y}^n}(\mathbf{y})}{V(\Lambda_1)} d\mathbf{u} d\mathbf{y} d\mathbf{x} \\ &= \sum_{\mathbf{x}_Q \in \Lambda_1} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} \int_{\mathcal{R}(\Lambda_1)} p_{X_Q|X^n U}(\mathbf{x}_Q|\mathbf{x}, \mathbf{u}) \mathbb{P}\{\hat{X}_Q \neq \mathbf{x}_Q \mid \hat{Y}^n = \mathbf{y}, U = \mathbf{u}, X_Q = \mathbf{x}_Q\} \frac{p_{X^n|\hat{Y}^n}(\mathbf{x}|\mathbf{y}) p_{\hat{Y}^n}(\mathbf{y})}{V(\Lambda_1)} d\mathbf{u} d\mathbf{y} d\mathbf{x}. \end{aligned}$$

In the last step we have used the Markov chain  $X^n - (\hat{Y}^n, X_Q, U) - \hat{X}_Q$ . Replacing the expression for the conditional distribution in equation (21), we obtain

$$\begin{aligned} P_e &= \sum_{\mathbf{x}_Q \in \Lambda_1} \int_{\mathbb{R}^n} \int_{\mathcal{R}(\Lambda_1)} \left( \int_{\mathbb{R}^n} \frac{f_{\sigma_Q}(\mathbf{x}_Q - \mathbf{x} - \mathbf{u}) f_{\sigma_1}(\mathbf{x} - \mathbf{y})}{f_{\sigma_Q}(\Lambda_1 - \mathbf{x} - \mathbf{u}) V(\Lambda_1)} d\mathbf{x} \right) \mathbb{1}_{\{\mathbf{y} \notin \mathbf{x}_Q - \mathbf{u} + \mathcal{V}(\Lambda_2)\}} f_{\hat{\sigma}_y}(\mathbf{y}) d\mathbf{u} d\mathbf{y} \\ &\stackrel{(a)}{\leq} \sum_{\mathbf{x}_Q \in \Lambda_1} \int_{\mathbb{R}^n} \int_{\mathcal{R}(\Lambda_1)} \left| \int_{\mathbb{R}^n} \frac{f_{\sigma_Q}(\mathbf{x}_Q - \mathbf{x} - \mathbf{u}) f_{\sigma_1}(\mathbf{x} - \mathbf{y})}{f_{\sigma_Q}(\Lambda_1 - \mathbf{x} - \mathbf{u}) V(\Lambda_1)} d\mathbf{x} - f_{\hat{\sigma}_1}(\mathbf{x}_Q - \mathbf{u} - \mathbf{y}) \right| \mathbb{1}_{\{\mathbf{y} \notin \mathbf{x}_Q - \mathbf{u} + \mathcal{V}(\Lambda_2)\}} f_{\hat{\sigma}_y}(\mathbf{y}) d\mathbf{u} d\mathbf{y} \quad (25) \end{aligned}$$

$$+ \sum_{\mathbf{x}_Q \in \Lambda_1} \int_{\mathbb{R}^n} \int_{\mathcal{R}(\Lambda_1)} f_{\hat{\sigma}_1}(\mathbf{x}_Q - \mathbf{u} - \mathbf{y}) \mathbb{1}_{\{\mathbf{y} \notin \mathbf{x}_Q - \mathbf{u} + \mathcal{V}(\Lambda_2)\}} f_{\hat{\sigma}_y}(\mathbf{y}) d\mathbf{u} d\mathbf{y} \quad (26)$$

where (a) follows from the triangle inequality.

The term (25) is upper bounded by

$$\begin{aligned} &\int_{\mathbb{R}^n} \sum_{\mathbf{x}_Q \in \Lambda_1} \int_{\mathcal{R}(\Lambda_1)} \left| \int_{\mathbb{R}^n} \frac{f_{\sigma_Q}(\mathbf{x}_Q - \mathbf{x} - \mathbf{u}) f_{\sigma_1}(\mathbf{x} - \mathbf{y})}{f_{\sigma_Q}(\Lambda_1 - \mathbf{x} - \mathbf{u}) V(\Lambda_1)} d\mathbf{x} - f_{\hat{\sigma}_1}(\mathbf{x}_Q - \mathbf{u} - \mathbf{y}) \right| d\mathbf{u} f_{\hat{\sigma}_y}(\mathbf{y}) d\mathbf{y} \\ &\leq \int_{\mathbb{R}^n} \epsilon_{\Lambda_1}^1(\sigma_Q) f_{\hat{\sigma}_y}(\mathbf{y}) d\mathbf{y} = \epsilon_{\Lambda_1}^1(\sigma_Q) \end{aligned}$$

using Lemma 4. This tends to 0 provided that  $\Lambda_1$  is  $L^1$  secrecy-good and

$$\frac{V(\Lambda_1)^{2/n}}{\sigma_Q^2} < 2\pi e. \quad (27)$$

With the change of variables  $\mathbf{y}' = \mathbf{y} - \mathbf{x}_Q + \mathbf{u}$ , the term (26) can be rewritten as

$$\begin{aligned}
& \sum_{\mathbf{x}_Q \in \Lambda_1} \int_{\mathcal{R}(\Lambda_1)} \int_{\mathbb{R}^n} f_{\tilde{\sigma}_1}(\mathbf{y}') \mathbb{1}_{\{\mathbf{y}' \notin \mathcal{V}(\Lambda_2)\}} f_{\tilde{\sigma}_y}(\mathbf{y}' + \mathbf{x}_Q - \mathbf{u}) d\mathbf{y}' d\mathbf{u} \\
&= \sum_{\mathbf{x}_Q \in \Lambda_1} \int_{\mathcal{R}(\Lambda_1)} \int_{\mathbb{R}^n \setminus \mathcal{V}(\Lambda_2)} f_{\tilde{\sigma}_1}(\mathbf{y}') f_{\tilde{\sigma}_y}(\mathbf{y}' + \mathbf{x}_Q - \mathbf{u}) d\mathbf{y}' d\mathbf{u} \\
&= \int_{\mathbb{R}^n \setminus \mathcal{V}(\Lambda_2)} f_{\tilde{\sigma}_1}(\mathbf{y}') \int_{\mathcal{R}(\Lambda_1)} f_{\tilde{\sigma}_y, \Lambda_1}(\mathbf{y}' - \mathbf{u}) d\mathbf{u} d\mathbf{y}' \\
&\stackrel{(b)}{=} \int_{\mathbb{R}^n \setminus \mathcal{V}(\Lambda_2)} f_{\tilde{\sigma}_1}(\mathbf{y}') d\mathbf{y}'
\end{aligned}$$

where (b) follows from the fact that  $\int_{\mathcal{R}(\Lambda_1)} f_{\tilde{\sigma}_y, \Lambda_1}(\mathbf{y}' - \mathbf{u}) d\mathbf{u} = 1$ . This tends to 0 provided that  $\Lambda_2$  is AWGN-good and

$$\frac{V(\Lambda_2)^{2/n}}{\tilde{\sigma}_1^2} > 2\pi e. \quad (28)$$

#### D. Uniformity

We want to show that the key is asymptotically uniform when  $n \rightarrow \infty$ . First, we want to bound the  $L^1$  distance between  $p_{\bar{\mathbf{x}}_Q}$  and the uniform distribution over  $\Lambda_1/\Lambda_3$ . Given  $\mathbf{x} \in \mathbb{R}^n$ ,  $\bar{\mathbf{x}}_Q \in \Lambda_1/\Lambda_3$ , we have

$$p_{\bar{\mathbf{x}}_Q | \mathbf{x}^n \cup}(\bar{\mathbf{x}}_Q | \mathbf{x}, \mathbf{u}) = \sum_{\lambda_3 \in \Lambda_3} p_{\mathbf{x}_Q | \mathbf{x}^n, \cup}(\bar{\mathbf{x}}_Q + \lambda_3 | \mathbf{x}, \mathbf{u}) = \sum_{\lambda_3 \in \Lambda_3} \frac{f_{\sigma_Q}(\bar{\mathbf{x}}_Q + \lambda_3 - \mathbf{x} - \mathbf{u})}{f_{\sigma_Q}(\Lambda_1 - \mathbf{x} - \mathbf{u})}. \quad (29)$$

Then

$$p_{\bar{\mathbf{x}}_Q}(\bar{\mathbf{x}}_Q) = \int_{\mathcal{R}(\Lambda_1)} \int_{\mathbb{R}^n} p_{\bar{\mathbf{x}}_Q | \mathbf{x}^n, \cup}(\bar{\mathbf{x}}_Q | \mathbf{x}, \mathbf{u}) \frac{p_{\mathbf{x}^n}(\mathbf{x})}{V(\Lambda_1)} d\mathbf{x} d\mathbf{u} = \sum_{\lambda_3 \in \Lambda_3} \int_{\mathcal{R}(\Lambda_1)} \int_{\mathbb{R}^n} \frac{f_{\sigma_Q}(\bar{\mathbf{x}}_Q + \lambda_3 - \mathbf{x} - \mathbf{u}) f_{\sigma_x}(\mathbf{x})}{f_{\sigma_Q}(\Lambda_1 - \mathbf{x} - \mathbf{u}) V(\Lambda_1)} d\mathbf{x} d\mathbf{u}.$$

Using the previous expression, we find

$$\begin{aligned}
& \left\| p_{\bar{\mathbf{x}}_Q} - \mathcal{U}_{\Lambda_1/\Lambda_3} \right\|_{L^1} = \sum_{\bar{\mathbf{x}}_Q \in \Lambda_1/\Lambda_3} \left| p_{\bar{\mathbf{x}}_Q}(\bar{\mathbf{x}}_Q) - \frac{V(\Lambda_1)}{V(\Lambda_3)} \right| \\
&= \sum_{\bar{\mathbf{x}}_Q \in \Lambda_1/\Lambda_3} \left| \int_{\mathcal{R}(\Lambda_1)} \sum_{\lambda_3 \in \Lambda_3} \int_{\mathbb{R}^n} \frac{f_{\sigma_Q}(\bar{\mathbf{x}}_Q + \lambda_3 - \mathbf{x} - \mathbf{u}) f_{\sigma_x}(\mathbf{x})}{f_{\sigma_Q}(\Lambda_1 - \mathbf{x} - \mathbf{u}) V(\Lambda_1)} d\mathbf{x} d\mathbf{u} - \frac{V(\Lambda_1)}{V(\Lambda_3)} \right| \\
&\stackrel{(a)}{\leq} \sum_{\bar{\mathbf{x}}_Q \in \Lambda_1/\Lambda_3} \left| \int_{\mathcal{R}(\Lambda_1)} \sum_{\lambda_3 \in \Lambda_3} \int_{\mathbb{R}^n} \frac{f_{\sigma_Q}(\bar{\mathbf{x}}_Q + \lambda_3 - \mathbf{x} - \mathbf{u}) f_{\sigma_x}(\mathbf{x})}{f_{\sigma_Q}(\Lambda_1 - \mathbf{x} - \mathbf{u}) V(\Lambda_1)} d\mathbf{x} d\mathbf{u} - \sum_{\lambda_3 \in \Lambda_3} \int_{\mathcal{R}(\Lambda_1)} f_{\tilde{\sigma}_x}(\bar{\mathbf{x}}_Q + \lambda_3 - \mathbf{u}) d\mathbf{u} \right| \quad (30)
\end{aligned}$$

$$+ \sum_{\bar{\mathbf{x}}_Q \in \Lambda_1/\Lambda_3} \left| \int_{\mathcal{R}(\Lambda_1)} \sum_{\lambda_3 \in \Lambda_3} f_{\tilde{\sigma}_x}(\bar{\mathbf{x}}_Q + \lambda_3 - \mathbf{u}) d\mathbf{u} - \frac{V(\Lambda_1)}{V(\Lambda_3)} \right| \quad (31)$$

where (a) follows from the triangle inequality, and  $\tilde{\sigma}_x^2 = \sigma_x^2 + \sigma_Q^2$ . The term (30) is upper bounded by

$$\begin{aligned}
& \sum_{\bar{\mathbf{x}}_Q \in \Lambda_1/\Lambda_3} \sum_{\lambda_3 \in \Lambda_3} \int_{\mathcal{R}(\Lambda_1)} \left| \int_{\mathbb{R}^n} \frac{f_{\sigma_Q}(\bar{\mathbf{x}}_Q + \lambda_3 - \mathbf{x} - \mathbf{u}) f_{\sigma_x}(\mathbf{x})}{f_{\sigma_Q}(\Lambda_1 - \mathbf{x} - \mathbf{u}) V(\Lambda_1)} d\mathbf{x} - f_{\tilde{\sigma}_x}(\bar{\mathbf{x}}_Q + \lambda_3 - \mathbf{u}) \right| d\mathbf{u} \\
&\leq \sum_{\mathbf{x}_Q \in \Lambda_1} \int_{\mathcal{R}(\Lambda_1)} \left| \int_{\mathbb{R}^n} \frac{f_{\sigma_Q}(\mathbf{x}_Q - \mathbf{x} - \mathbf{u}) f_{\sigma_x}(\mathbf{x})}{f_{\sigma_Q}(\Lambda_1 - \mathbf{x} - \mathbf{u}) V(\Lambda_1)} d\mathbf{x} - f_{\tilde{\sigma}_x}(\mathbf{x}_Q - \mathbf{u}) \right| d\mathbf{u} \leq \epsilon_{\Lambda_1}^1(\sigma_Q)
\end{aligned}$$

by Lemma 4. This vanishes as  $o\left(\frac{1}{n}\right)$  if  $\Lambda_1$  is  $L^1$  secrecy-good and the condition (27) is satisfied. The term (31) is equal to

$$\begin{aligned} & \sum_{\bar{\mathbf{x}}_Q \in \Lambda_1/\Lambda_3} \left| \int_{\mathcal{R}(\Lambda_1)} f_{\tilde{\sigma}_x, \Lambda_3}(\bar{\mathbf{x}}_Q - \mathbf{u}) d\mathbf{u} - \int_{\mathcal{R}(\Lambda_1)} \frac{1}{V(\Lambda_1)} d\mathbf{u} \frac{V(\Lambda_1)}{V(\Lambda_3)} \right| \\ & \leq \sum_{\bar{\mathbf{x}}_Q \in \Lambda_1/\Lambda_3} \int_{\mathcal{R}(\Lambda_1)} \left| f_{\tilde{\sigma}_x, \Lambda_3}(\bar{\mathbf{x}}_Q - \mathbf{u}) - \frac{1}{V(\Lambda_3)} \right| d\mathbf{u}. \end{aligned}$$

Setting  $\mathbf{v} = \bar{\mathbf{x}}_Q - \mathbf{u} \bmod \mathcal{R}(\Lambda_3)$ , and recalling that  $\mathcal{R}(\Lambda_3) = \bigcup_{\bar{\mathbf{x}}_Q \in \Lambda_1/\Lambda_3} ([\bar{\mathbf{x}}_Q + \mathcal{R}(\Lambda_1)] \bmod \mathcal{R}(\Lambda_3))$  by (4), where the union is disjoint, the last expression is equal to

$$\int_{\mathcal{R}(\Lambda_3)} \left| f_{\tilde{\sigma}_x, \Lambda_3}(\mathbf{v}) - \frac{1}{V(\Lambda_3)} \right| d\mathbf{v} = \epsilon_{\Lambda_3}^1(\tilde{\sigma}_x) \leq \epsilon_{\Lambda_3}^1(\tilde{\sigma}_2)$$

where  $\tilde{\sigma}_2^2 = \sigma_2^2 + \sigma_Q^2 \leq \sigma_x^2 + \sigma_Q^2 = \tilde{\sigma}_x^2$ . Thus, the term (31) vanishes as  $o\left(\frac{1}{n}\right)$  if both  $\Lambda_1$  and  $\Lambda_3$  are  $L^1$ -secrecy good and satisfy the volume conditions (27) and

$$\frac{V(\Lambda_3)^{2/n}}{\tilde{\sigma}_2^2} < 2\pi e. \quad (32)$$

We now show that the distribution of the key is close to the uniform distribution  $\mathcal{U}_{\mathcal{K}}$  over  $\mathcal{K} = \Lambda_2/\Lambda_3$ :

$$\begin{aligned} \mathbb{V}(p_{\mathcal{K}}, \mathcal{U}_{\mathcal{K}}) &= \sum_{k \in \mathcal{K}} \left| p_{\mathcal{K}}(k) - \frac{V(\Lambda_2)}{V(\Lambda_3)} \right| = \sum_{k \in \mathcal{K}} \left| \sum_{s \in \Lambda_1/\Lambda_2} p_{\bar{\mathbf{x}}_Q}(s \oplus k) - \sum_{s \in \Lambda_1/\Lambda_2} \frac{V(\Lambda_1)}{V(\Lambda_3)} \right| \\ &\leq \sum_{k \in \mathcal{K}} \sum_{s \in \Lambda_1/\Lambda_2} \left| p_{\bar{\mathbf{x}}_Q}(s \oplus k) - \frac{V(\Lambda_1)}{V(\Lambda_3)} \right| = \sum_{\bar{\mathbf{x}}_Q \in \Lambda_1/\Lambda_3} \left| p_{\bar{\mathbf{x}}_Q}(\bar{\mathbf{x}}_Q) - \frac{V(\Lambda_1)}{V(\Lambda_3)} \right| = \mathbb{V}(p_{\bar{\mathbf{x}}_Q}, \mathcal{U}_{\Lambda_1/\Lambda_3}) \end{aligned}$$

which vanishes as  $o\left(\frac{1}{n}\right)$  as shown previously. Using [54, Lemma 2.7], we have that if  $\mathbb{V}(p_{\mathcal{K}}, \mathcal{U}_{\mathcal{K}}) \leq \frac{1}{2}$ ,

$$\begin{aligned} |\mathbb{H}(p_{\mathcal{K}}) - \mathbb{H}(\mathcal{U}_{\mathcal{K}})| &\leq -\mathbb{V}(p_{\mathcal{K}}, \mathcal{U}_{\mathcal{K}}) \log \frac{\mathbb{V}(p_{\mathcal{K}}, \mathcal{U}_{\mathcal{K}})}{|\mathcal{K}|} = \mathbb{V}(p_{\mathcal{K}}, \mathcal{U}_{\mathcal{K}}) \log \frac{2^{nR_K}}{\mathbb{V}(p_{\mathcal{K}}, \mathcal{U}_{\mathcal{K}})} \\ &= nR_K \mathbb{V}(p_{\mathcal{K}}, \mathcal{U}_{\mathcal{K}}) - \mathbb{V}(p_{\mathcal{K}}, \mathcal{U}_{\mathcal{K}}) \log \mathbb{V}(p_{\mathcal{K}}, \mathcal{U}_{\mathcal{K}}). \end{aligned}$$

This vanishes as long as  $\mathbb{V}(p_{\mathcal{K}}, \mathcal{U}_{\mathcal{K}}) \sim o\left(\frac{1}{n}\right)$ , which is indeed the case.

### E. Strong secrecy

Using [29, Lemma 1], we can bound the leakage as follows:

$$\mathbb{I}(\mathbf{K}; \mathbf{S}, \mathbf{Z}^n, \mathbf{U}) = \mathbb{I}(\mathbf{K}; \mathbf{S}, \hat{\mathbf{Z}}^n, \mathbf{U}) \leq d_{\text{av}} \log \frac{|\mathcal{K}|}{d_{\text{av}}}, \quad (33)$$

where

$$\begin{aligned} d_{\text{av}} &= \sum_{k \in \mathcal{K}} p_{\mathcal{K}}(k) \mathbb{V}(p_{\mathbf{S}\hat{\mathbf{Z}}^n \mathbf{U} | \mathbf{K}=k}, p_{\mathbf{S}\hat{\mathbf{Z}}^n \mathbf{U}}) \\ &\leq \sum_{k \in \mathcal{K}} p_{\mathcal{K}}(k) \sum_{s \in \Lambda_1/\Lambda_2} \int_{\mathcal{R}(\Lambda_1)} \int_{\mathbb{R}^n} \left| p_{\mathbf{S}\hat{\mathbf{Z}}^n \mathbf{U} | \mathbf{K}=k}(s, \mathbf{z}, \mathbf{u} | k) - \frac{p_{\hat{\mathbf{Z}}^n}(\mathbf{z})}{V(\Lambda_2)} \right| d\mathbf{z} d\mathbf{u} \end{aligned} \quad (34)$$

$$+ \sum_{k \in \mathcal{K}} p_{\mathcal{K}}(k) \sum_{s \in \Lambda_1/\Lambda_2} \int_{\mathcal{R}(\Lambda_1)} \int_{\mathbb{R}^n} \left| \frac{p_{\hat{\mathbf{Z}}^n}(\mathbf{z})}{V(\Lambda_2)} - p_{\mathbf{S}\hat{\mathbf{Z}}^n \mathbf{U}}(s, \mathbf{z}, \mathbf{u}) \right| d\mathbf{z} d\mathbf{u} \quad (35)$$

by the triangle inequality.

Due to Remark 5, we can write

$$\begin{aligned} p_{S\hat{Z}^n \cup K}(s, \mathbf{z}, \mathbf{u}|k) &= \frac{p_{S\hat{Z}^n \cup K}(s, \mathbf{z}, \mathbf{u}, k)}{p_K(k)} = \frac{p_{\hat{Z}^n}(\mathbf{z})}{V(\Lambda_1)} \frac{p_{\bar{X}_Q|\hat{Z}^n, U}(k+s|\mathbf{z}, \mathbf{u})}{p_K(k)} \\ &= \frac{p_{\hat{Z}^n}(\mathbf{z})}{V(\Lambda_1)p_K(k)} \sum_{\lambda_3 \in \Lambda_3} p_{X_Q|\hat{Z}^n, U}(k+s+\lambda_3|\mathbf{z}, \mathbf{u}), \end{aligned}$$

and so the term (34) is equal to

$$\begin{aligned} &\sum_{k \in \mathcal{K}} \sum_{s \in \Lambda_1/\Lambda_2} \int_{\mathcal{R}(\Lambda_1)} \int_{\mathbb{R}^n} p_{\hat{Z}^n}(\mathbf{z}) \left| \sum_{\lambda_3 \in \Lambda_3} \frac{p_{X_Q|\hat{Z}^n, U}(k+s+\lambda_3|\mathbf{z}, \mathbf{u})}{V(\Lambda_1)} - \frac{p_K(k)}{V(\Lambda_2)} \right| d\mathbf{z} d\mathbf{u} \\ &\leq \sum_{k \in \mathcal{K}} \sum_{s \in \Lambda_1/\Lambda_2} \int_{\mathcal{R}(\Lambda_1)} \int_{\mathbb{R}^n} p_{\hat{Z}^n}(\mathbf{z}) \left| \sum_{\lambda_3 \in \Lambda_3} \frac{p_{X_Q|\hat{Z}^n, U}(k+s+\lambda_3|\mathbf{z}, \mathbf{u})}{V(\Lambda_1)} - \sum_{\lambda_3 \in \Lambda_3} f_{\tilde{\sigma}_2}(k+s+\lambda_3-\mathbf{u}-\mathbf{z}) \right| d\mathbf{z} d\mathbf{u} \quad (36) \end{aligned}$$

$$+ \sum_{k \in \mathcal{K}} \sum_{s \in \Lambda_1/\Lambda_2} \int_{\mathcal{R}(\Lambda_1)} \int_{\mathbb{R}^n} p_{\hat{Z}^n}(\mathbf{z}) \left| \sum_{\lambda_3 \in \Lambda_3} f_{\tilde{\sigma}_2}(k+s+\lambda_3-\mathbf{u}-\mathbf{z}) - \frac{p_K(k)}{V(\Lambda_2)} \right| d\mathbf{z} d\mathbf{u}, \quad (37)$$

where  $\tilde{\sigma}_2^2 = \sigma_2^2 + \sigma_Q^2$ . Recalling that

$$p_{X_Q|\hat{Z}^n, U}(\mathbf{x}_Q|\mathbf{z}, \mathbf{u}) = \int_{\mathbb{R}^n} p_{X_Q|X^n, U}(\mathbf{x}_Q|\mathbf{x}, \mathbf{u}) p_{X^n|\hat{Z}^n}(\mathbf{x}|\mathbf{z}) d\mathbf{x} = \int_{\mathbb{R}^n} \frac{f_{\sigma_Q}(\mathbf{x}_Q - \mathbf{x} - \mathbf{u})}{f_{\sigma_Q}(\Lambda_1 - \mathbf{x} - \mathbf{u})} f_{\sigma_2}(\mathbf{x} - \mathbf{z}) d\mathbf{x},$$

the term (36) can be upper bounded by

$$\begin{aligned} &\int_{\mathbb{R}^n} p_{\hat{Z}^n}(\mathbf{z}) \int_{\mathcal{R}(\Lambda_1)} \sum_{k \in \mathcal{K}} \sum_{s \in \Lambda_1/\Lambda_2} \sum_{\lambda_3 \in \Lambda_3} \left| \int_{\mathbb{R}^n} \frac{f_{\sigma_Q}(k+s+\lambda_3-\mathbf{x}-\mathbf{u})}{V(\Lambda_1)f_{\sigma_Q}(\Lambda_1-\mathbf{x}-\mathbf{u})} f_{\sigma_2}(\mathbf{x}-\mathbf{z}) d\mathbf{x} - f_{\tilde{\sigma}_2}(k+s+\lambda_3-\mathbf{u}-\mathbf{z}) \right| d\mathbf{u} d\mathbf{z} \\ &= \int_{\mathbb{R}^n} p_{\hat{Z}^n}(\mathbf{z}) \int_{\mathcal{R}(\Lambda_1)} \sum_{\mathbf{x}_Q \in \Lambda_1} \left| \int_{\mathbb{R}^n} \frac{f_{\sigma_Q}(\mathbf{x}_Q - \mathbf{x} - \mathbf{u})}{V(\Lambda_1)f_{\sigma_Q}(\Lambda_1-\mathbf{x}-\mathbf{u})} f_{\sigma_2}(\mathbf{x}-\mathbf{z}) d\mathbf{x} - f_{\tilde{\sigma}_2}(\mathbf{x}_Q - \mathbf{u} - \mathbf{z}) \right| d\mathbf{u} d\mathbf{z} \leq \epsilon_{\Lambda_1}^1(\sigma_Q) \end{aligned}$$

by Lemma 4. This vanishes as  $o(\frac{1}{n})$  assuming the condition (27).

On the other hand, by the triangle inequality the term (37) can be bounded by

$$\sum_{k \in \mathcal{K}} \sum_{s \in \Lambda_1/\Lambda_2} \int_{\mathcal{R}(\Lambda_1)} \int_{\mathbb{R}^n} p_{\hat{Z}^n}(\mathbf{z}) \left| \sum_{\lambda_3 \in \Lambda_3} f_{\tilde{\sigma}_2}(k+s+\lambda_3-\mathbf{u}-\mathbf{z}) - \frac{1}{V(\Lambda_3)} \right| d\mathbf{z} d\mathbf{u} \quad (38)$$

$$+ \sum_{k \in \mathcal{K}} \sum_{s \in \Lambda_1/\Lambda_2} \int_{\mathcal{R}(\Lambda_1)} \int_{\mathbb{R}^n} p_{\hat{Z}^n}(\mathbf{z}) \left| \frac{1}{V(\Lambda_3)} - \frac{p_K(k)}{V(\Lambda_2)} \right| d\mathbf{z} d\mathbf{u}. \quad (39)$$

Setting  $\mathbf{v} = k+s-\mathbf{u} \bmod \mathcal{R}(\Lambda_3)$  and using the property (4), the term (38) can be written as

$$\int_{\mathbb{R}^n} p_{\hat{Z}^n}(\mathbf{z}) \int_{\mathcal{R}(\Lambda_3)} \left| f_{\tilde{\sigma}_2, \Lambda_3}(\mathbf{v}-\mathbf{z}) - \frac{1}{V(\Lambda_3)} \right| d\mathbf{v} d\mathbf{z} = \epsilon_{\Lambda_3}^1(\tilde{\sigma}_2),$$

which vanishes as  $o(\frac{1}{n})$  assuming the condition (32). Finally, (39) is equal to

$$V(\Lambda_2) \sum_{k \in \mathcal{K}} \left| \frac{1}{V(\Lambda_3)} - \frac{p_K(k)}{V(\Lambda_2)} \right| = \sum_{k \in \mathcal{K}} |\mathcal{U}_K - p_K(k)| = \mathbb{V}(\mathcal{U}_K, p_K) = o\left(\frac{1}{n}\right) \rightarrow 0$$

as already shown in Section III-D.

We now come back to the expression (35), which is equal to

$$\begin{aligned} & \sum_{s \in \Lambda_1/\Lambda_2} \int_{\mathcal{R}(\Lambda_1)} \int_{\mathbb{R}^n} \left| \frac{p_{\hat{\mathbf{Z}}^n}(\mathbf{z})}{V(\Lambda_2)} - p_{S\hat{\mathbf{Z}}^n\mathbf{U}}(s, \mathbf{z}, \mathbf{u}) \right| d\mathbf{z} d\mathbf{u} \\ & \leq \sum_{s \in \Lambda_1/\Lambda_2} \int_{\mathcal{R}(\Lambda_1)} \int_{\mathbb{R}^n} \left| \frac{p_{\hat{\mathbf{Z}}^n}(\mathbf{z})}{V(\Lambda_2)} - \sum_{k' \in \mathcal{K}} \sum_{\lambda_3 \in \Lambda_3} p_{\hat{\mathbf{Z}}^n}(\mathbf{z}) f_{\tilde{\sigma}_2}(k' + s + \lambda_3 - \mathbf{u} - \mathbf{z}) \right| d\mathbf{z} d\mathbf{u} \end{aligned} \quad (40)$$

$$+ \sum_{s \in \Lambda_1/\Lambda_2} \int_{\mathcal{R}(\Lambda_1)} \int_{\mathbb{R}^n} \left| \sum_{k' \in \mathcal{K}} \sum_{\lambda_3 \in \Lambda_3} p_{\hat{\mathbf{Z}}^n}(\mathbf{z}) f_{\tilde{\sigma}_2}(k' + s + \lambda_3 - \mathbf{u} - \mathbf{z}) - p_{S\hat{\mathbf{Z}}^n\mathbf{U}}(s, \mathbf{z}, \mathbf{u}) \right| d\mathbf{z} d\mathbf{u} \quad (41)$$

by the triangle inequality.

The term (40) is upper bounded by

$$\begin{aligned} & \int_{\mathbb{R}^n} p_{\hat{\mathbf{Z}}^n}(\mathbf{z}) \sum_{s \in \Lambda_1/\Lambda_2} \sum_{k' \in \mathcal{K}} \int_{\mathcal{R}(\Lambda_1)} \left| \frac{1}{V(\Lambda_3)} - f_{\tilde{\sigma}_2, \Lambda_3}(k' + s - \mathbf{u} - \mathbf{z}) \right| d\mathbf{u} d\mathbf{z} \\ & = \int p_{\hat{\mathbf{Z}}^n}(\mathbf{z}) \int_{\mathcal{R}(\Lambda_3)} \left| \frac{1}{V(\Lambda_3)} - f_{\tilde{\sigma}_2, \Lambda_3}(\mathbf{v} - \mathbf{z}) \right| d\mathbf{v} d\mathbf{z} = \epsilon_{\Lambda_3}^1(\tilde{\sigma}_2) \end{aligned}$$

and vanishes as  $o(\frac{1}{n})$  if condition (32) is satisfied.

Observe that

$$\begin{aligned} p_{S\hat{\mathbf{Z}}^n\mathbf{U}}(s, \mathbf{z}, \mathbf{u}) &= \sum_{k' \in \mathcal{K}} p_{SK\hat{\mathbf{Z}}^n\mathbf{U}}(s, k', \mathbf{z}, \mathbf{u}) = \sum_{k' \in \mathcal{K}} \frac{p_{\hat{\mathbf{Z}}^n}(\mathbf{z})}{V(\Lambda_1)} p_{\mathbf{X}_Q|\hat{\mathbf{Z}}^n\mathbf{U}}(s + k'|\mathbf{z}, \mathbf{u}) \\ &= \sum_{k' \in \mathcal{K}} \frac{p_{\hat{\mathbf{Z}}^n}(\mathbf{z})}{V(\Lambda_1)} \sum_{\lambda_3 \in \Lambda_3} p_{\mathbf{X}_Q|\hat{\mathbf{Z}}^n\mathbf{U}}(s + k' + \lambda_3|\mathbf{z}, \mathbf{u}) = \sum_{k' \in \mathcal{K}} \frac{p_{\hat{\mathbf{Z}}^n}(\mathbf{z})}{V(\Lambda_1)} \sum_{\lambda_3 \in \Lambda_3} \int_{\mathbb{R}^n} \frac{f_{\sigma_Q}(s + k' + \lambda_3 - \mathbf{x} - \mathbf{u}) f_{\sigma_2}(\mathbf{x} - \mathbf{z})}{f_{\sigma_Q}(\Lambda_1 - \mathbf{x} - \mathbf{u})} d\mathbf{x}. \end{aligned}$$

Thus the term (41) can be bounded by

$$\begin{aligned} & \int_{\mathbb{R}^n} p_{\hat{\mathbf{Z}}^n}(\mathbf{z}) \int_{\mathcal{R}(\Lambda_1)} \sum_{k' \in \mathcal{K}} \sum_{s \in \Lambda_1/\Lambda_2} \sum_{\lambda_3 \in \Lambda_3} \left| f_{\tilde{\sigma}_2}(k' + s + \lambda_3 - \mathbf{u} - \mathbf{z}) - \int_{\mathbb{R}^n} \frac{f_{\sigma_Q}(s + k' + \lambda_3 - \mathbf{x} - \mathbf{u}) f_{\sigma_2}(\mathbf{x} - \mathbf{z})}{V(\Lambda_1) f_{\sigma_Q}(\Lambda_1 - \mathbf{x} - \mathbf{u})} d\mathbf{x} \right| d\mathbf{u} d\mathbf{z} \\ &= \int_{\mathbb{R}^n} p_{\hat{\mathbf{Z}}^n}(\mathbf{z}) \int_{\mathcal{R}(\Lambda_1)} \sum_{\mathbf{x}_Q \in \Lambda_1} \left| f_{\tilde{\sigma}_2}(\mathbf{x}_Q - \mathbf{u} - \mathbf{z}) - \int_{\mathbb{R}^n} \frac{f_{\sigma_Q}(\mathbf{x}_Q - \mathbf{x} - \mathbf{u}) f_{\sigma_2}(\mathbf{x} - \mathbf{z})}{V(\Lambda_1) f_{\sigma_Q}(\Lambda_1 - \mathbf{x} - \mathbf{u})} d\mathbf{x} \right| d\mathbf{u} d\mathbf{z} \leq \epsilon_{\Lambda_1}^1(\sigma_Q) \end{aligned}$$

by Lemma 4, which again vanishes as  $o(\frac{1}{n})$  under condition (27).

In conclusion,  $d_{\text{av}} \sim o(\frac{1}{n})$  and thus from (33), we find that the leakage vanishes asymptotically as  $n \rightarrow \infty$ .

*Remark 6:* Although in Section III-D we only showed that the key is close to uniform on average over the dither  $\mathbf{U}$ , using the results in this section we see that

$$\mathbb{H}(\mathcal{U}_{\mathcal{K}}) - \mathbb{H}(\mathbf{K}|\mathbf{U}) = \mathbb{H}(\mathcal{U}_{\mathcal{K}}) - \mathbb{H}(\mathbf{K}) + \mathbb{I}(\mathbf{K}; \mathbf{U}) \leq \mathbb{H}(\mathcal{U}_{\mathcal{K}}) - \mathbb{H}(\mathbf{K}) + \mathbb{I}(\mathbf{K}; \mathbf{S}, \mathbf{Z}^n, \mathbf{U}) \rightarrow 0.$$

#### F. Achievable strong secrecy rate and optimal trade-off

Recall that in the previous sections we have imposed the conditions (27), (28) and (32) on the volumes of  $\Lambda_1$ ,  $\Lambda_2$  and  $\Lambda_3$  respectively, i.e.

$$\frac{V(\Lambda_1)^{2/n}}{\sigma_Q^2} < 2\pi e, \quad \frac{V(\Lambda_2)^{2/n}}{\tilde{\sigma}_1^2} > 2\pi e, \quad \frac{V(\Lambda_3)^{2/n}}{\tilde{\sigma}_2^2} < 2\pi e.$$

Therefore, the achievable secret key rate is upper bounded by

$$R_K = \frac{1}{n} \log \frac{V(\Lambda_3)}{V(\Lambda_2)} < \frac{1}{2} \log \frac{\tilde{\sigma}_2^2}{\tilde{\sigma}_1^2} = \frac{1}{2} \log \frac{\sigma_2^2 + \sigma_Q^2}{\sigma_1^2 + \sigma_Q^2} \quad (42)$$

As  $\sigma_Q \rightarrow 0$ ,

$$R_K \rightarrow \frac{1}{2} \log \frac{\sigma_2^2}{\sigma_1^2},$$

which is the optimal secret key rate. This improves upon our previous work [1] in which the achievable secrecy rate had a  $1/2$  nat gap compared to the optimal.

*Remark 7:* The optimal scaling of the lattice  $\Lambda_3$  requires the noise variance  $\sigma_2$  to be known by Alice; if only a lower bound for  $\sigma_2$  is available, positive secret key rates can still be attained.

The public communication rate is lower bounded by

$$R_P = \frac{1}{n} \log \frac{V(\Lambda_2)}{V(\Lambda_1)} > \frac{1}{2} \log \frac{\sigma_1^2 + \sigma_Q^2}{\sigma_Q^2}.$$

Equivalently, we have  $\sigma_Q^2 > \frac{\sigma_1^2}{e^{2R_P} - 1}$ . Replacing this expression in the bound (42) for  $R_K$ , and observing that (42) is a decreasing function of  $\sigma_Q^2$ , we find

$$R_K < \frac{1}{2} \log \left( e^{-2R_P} + \frac{\sigma_2^2}{\sigma_1^2} (1 - e^{-2R_P}) \right).$$

which corresponds to the optimal public rate / secret key rate trade-off (18).

#### IV. CONCLUSIONS AND PERSPECTIVES

To conclude, we have proposed a new lattice-based technique to extract a secret key from correlated Gaussian sources against an eavesdropper. Using  $L^1$  distance and KL divergence, we have proved the existence of lattices with a vanishing flatness factor for all VNRs up to  $2\pi e$ . This improves upon the previous result for VNRs up to  $2\pi$ , based on  $L^\infty$  distance. Together with dithering and randomized rounding, it has enabled us to achieve the optimal trade-off with one-way public communication. In the same way, it is possible to remove the  $\frac{1}{2}$ -nat gap to the secrecy capacity of wiretap channels [35] associated to the use of the  $L^\infty$  flatness factor [39, p. 1656].

An immediate step for future work is to turn the existence result of this paper into a practical scheme. There are avenues for replacing random nested lattices for Wyner-Ziv coding with lower-complexity techniques, such as superposition coding or residual quantization [55, 56]. However such techniques do not address privacy amplification. In order to implement the approach proposed in this paper based on the notion of flatness factor of a lattice, a promising option is to instantiate the lattices using polar codes (aka polar lattices), which have been shown to be good for quantization, channel coding [57] and secrecy. A polar lattice has been constructed in [39] to achieve the secrecy capacity of Gaussian wiretap channels. It can be shown that the secrecy-good lattice in [39] enjoys a vanishing  $L^1$  flatness factor. Since the encoding and decoding complexity of a polar lattice is quasi-linear in blocklength  $n$ , it is an excellent candidate to build a practical scheme for secret key generation. It is also possible to implement the randomized rounding algorithm over a polar lattice.

Another problem is to see if it is possible to modify the design of this paper to yield a fuzzy extractor, which would require redesigning a lattice with respect to other entropy measures. Other open problems include identifying whether it is possible to remove dithering and/or randomized quantization, characterizing the second-order asymptotics and the extension of the proposed key-agreement protocol to multi-terminal systems. Furthermore, the reconciliation technique based on Wyner-Ziv coding may be extended to key-encapsulation mechanisms (KEM) in lattice-based cryptography, due to the similarity between KEM and secret key agreement. Finally, it is interesting to explore the applications of  $L^1$  and KL smoothing parameters in other cryptographic and mathematical problems [37, 38].

#### ACKNOWLEDGMENTS

The second author is grateful to Antonio Campello, Daniel Dadush and Ling Liu for helpful discussions. The authors would like to thank the two anonymous reviewers for their detailed comments and suggestions which helped improve the paper.



## APPENDIX A RESOLVABILITY CODES

In this section we review some results from [36] about resolvability codes for regular channels, which are needed for the proof of Theorem 1.

First, we need some preliminary definitions. In the following, we assume  $\mathcal{X}$  is a finite abelian group and  $\mathcal{Y}$  is a measurable space. Given a channel  $W : \mathcal{X} \rightarrow \mathcal{Y}$ , we use the notation  $W_x(y) = W(y|x)$  for  $x \in \mathcal{X}, y \in \mathcal{Y}$ .

*Definition 7 (Rényi Entropy):* Given a discrete distribution  $p_A$  on  $\mathcal{A}$  and  $\rho \geq 0$ , we define

$$H_{1+\rho}(A) = -\frac{1}{\rho} \log \sum_{a \in \mathcal{A}} p_A(a)^{1+\rho}.$$

*Definition 8:* Given a channel  $W : \mathcal{X} \rightarrow \mathcal{Y}$  and a probability distribution  $p_X$  on  $\mathcal{X}$ , we define  $\forall \rho \geq 0$

$$\psi(\rho|W, p_X) = \log \sum_{x \in \mathcal{X}} p_X(x) \int_{\mathcal{Y}} W_x(y)^{1+\rho} (W \circ p_X)(y)^{-\rho} dy.$$

This function has the following properties:

$$\psi(0|W, p_X) = 0, \tag{43}$$

$$\psi(\rho|W^n, p_X^{\otimes n}) = n\psi(\rho|W, p_X), \tag{44}$$

$$\lim_{\rho \rightarrow 0} \frac{\psi(\rho|W, p_X)}{\rho} = \mathbb{I}(X; Y). \tag{45}$$

We also compute the second derivative in 0 which will be needed in the next section.

*Lemma 5:*

$$\psi''(0) = \sum_{x \in \mathcal{X}} p_X(x) \int_{\mathcal{Y}} W_x(y) \left( \log \frac{W_x(y)}{(W \circ p_X)(y)} \right)^2 dy - \left( \sum_{x \in \mathcal{X}} p_X(x) \int_{\mathcal{Y}} W_x(y) \log \frac{W_x(y)}{(W \circ p_X)(y)} dy \right)^2.$$

The proof of Lemma 5 can be found in Appendix F.

*Definition 9 (Regular channel):* The channel  $W : \mathcal{X} \rightarrow \mathcal{Y}$  is called *regular* if  $\mathcal{X}$  acts on  $\mathcal{Y}$  by permutations  $\{\pi_x\}_{x \in \mathcal{X}}$  such that  $\pi_x(\pi_{x'}'(y)) = \pi_{x+x'}(y) \forall x, x' \in \mathcal{X}$ , and there exists a probability density  $p_Y$  on  $\mathcal{Y}$  such that  $W_x(y) = p_Y(\pi_x(y)) \forall x \in \mathcal{X}, \forall y \in \mathcal{Y}$ .

In particular, a regular channel is symmetric [58, 59] in the sense of Gallager [60], and its capacity is achieved by the uniform distribution.

The following theorem was stated for discrete memoryless channels [36, Corollary 18] but can be extended to continuous outputs [36, Appendix D] as follows:

*Theorem 3:* Let  $\mathcal{M}$  and  $\mathcal{X}$  be a finite-dimensional vector spaces over  $\mathbb{F}_p$  and  $\mathcal{Y}$  a measurable space. Consider a uniform random variable  $F$  taking values over the set of linear mappings  $f : \mathcal{M} \rightarrow \mathcal{X}$  and a distribution  $p_M$  on  $\mathcal{M}$ . If  $W : \mathcal{X} \rightarrow \mathcal{Y}$  is regular, then  $\forall \rho \in (0, 1]$ ,

$$\mathbb{E}_F \left[ e^{\rho \mathbb{D}(W \circ F \circ p_M || W \circ \mathcal{U}_X)} \right] \leq 1 + e^{-\rho H_{1+\rho}(M)} e^{\psi(\rho|W, \mathcal{U}_X)}.$$

Theorem 3 is a one-shot result, but we can apply it to  $n$  uses of an i.i.d. channel to get the following.

*Corollary 1:* Let  $\mathcal{X}$  be a finite-dimensional vector space over  $\mathbb{F}_p$  and  $\mathcal{Y}$  a measurable space, and  $W : \mathcal{X} \rightarrow \mathcal{Y}$  a regular channel. Let  $R > \mathbb{I}(X; Y)$ , where  $X \sim \mathcal{U}_X$  and  $Y \sim W \circ \mathcal{U}_X$ . Consider  $C_n \subset \mathcal{X}^n$  chosen uniformly at random in the set of  $(n, k)$  linear codes in  $\mathcal{X}^n$ , where  $k = \frac{\lceil nR \rceil}{\log p}$ . Denote by  $\mathcal{U}_{C_n}$  the uniform distribution over the codewords in  $C_n$ . Then

$$\mathbb{E}_{C_n} [\mathbb{D}(W^n \circ \mathcal{U}_{C_n} || W^n \circ \mathcal{U}_X^{\otimes n})] \rightarrow 0$$

exponentially fast as  $n \rightarrow \infty$ .

*Proof:* Note that  $W^n : \mathcal{X}^n \rightarrow \mathcal{Y}^n$  is still a regular channel with respect to the set of permutations  $\{\pi_{\mathbf{x}}\}_{\mathbf{x} \in \mathcal{X}^n}$ , where we define  $\pi_{\mathbf{x}}(y_1, \dots, y_n) = (\pi_{x_1}(y_1), \dots, \pi_{x_n}(y_n))$  for  $\mathbf{x} = (x_1, \dots, x_n)$ .

Applying Theorem 3 to this channel, and taking  $\mathcal{M} = \mathbb{F}_p^k$  with  $k = \frac{\lceil nR \rceil}{\log p}$  and  $p_M = \mathcal{U}_{\mathcal{M}}$ , for  $F_n$  representing a random linear encoder  $f_n : \mathcal{M} \rightarrow \mathcal{X}^n$  we have

$$\mathbb{E}_{F_n} \left[ e^{\rho \mathbb{D}(W^n \circ F_n \circ \mathcal{U}_{\mathcal{M}} \| W^n \circ \mathcal{U}_{\mathcal{X}^n})} \right] \leq 1 + e^{-\rho H_{1+\rho}(M)} e^{\psi(\rho | W^n, \mathcal{U}_{\mathcal{X}^n})}.$$

By Jensen's inequality,

$$\mathbb{E}_{F_n} [\mathbb{D}(W^n \circ F_n \circ \mathcal{U}_{\mathcal{M}} \| W^n \circ \mathcal{U}_{\mathcal{X}^n})] \leq \frac{1}{\rho} \log \left( 1 + e^{-\rho H_{1+\rho}(M)} e^{\psi(\rho | W^n, \mathcal{U}_{\mathcal{X}^n})} \right) \leq \frac{1}{\rho} e^{-\rho H_{1+\rho}(M) + \psi(\rho | W^n, \mathcal{U}_{\mathcal{X}^n})}.$$

Note that  $H_{1+\rho}(M) = nR$  since  $M$  is uniform. Using (44), we find that  $\forall \rho \in (0, 1]$ ,

$$\mathbb{E}_{F_n} [\mathbb{D}(W^n \circ F_n \circ \mathcal{U}_{\mathcal{M}} \| W^n \circ \mathcal{U}_{\mathcal{X}^n})] \leq \frac{1}{\rho} e^{-n(\rho R - \psi(\rho | W, \mathcal{U}_{\mathcal{X}}))}. \quad (46)$$

From (43) and (45), we have  $\psi(\rho | W, p_X) = \rho \mathbb{I}(X; Y) + \eta(\rho)$ , where  $\lim_{\rho \rightarrow 0} \frac{\eta(\rho)}{\rho} = 0$ . Given  $R > \mathbb{I}(X; Y)$ ,  $\exists \bar{\rho}$  sufficiently small such that  $\delta = R - \mathbb{I}(X; Y) - \frac{\eta(\bar{\rho})}{\bar{\rho}} > 0$ . Therefore

$$\mathbb{E}_{F_n} [\mathbb{D}(W^n \circ F_n \circ \mathcal{U}_{\mathcal{M}} \| W^n \circ \mathcal{U}_{\mathcal{X}^n})] \leq \frac{1}{\bar{\rho}} e^{-n\bar{\rho}\delta} \rightarrow 0 \quad (47)$$

as  $n \rightarrow \infty$ . □

## APPENDIX B

### MODULO LATTICE CHANNELS AND THE KL FLATNESS FACTOR

In this section, we review some properties of modulo lattice channels and introduce another notion of flatness factor based on the KL divergence, which will be used in the proof of Theorem 1.

#### A. The mod- $\Lambda$ channel and the mod- $\Lambda/\Lambda'$ channel

Following Forney et al. [58], given a fundamental region  $\mathcal{R}(\Lambda)$  of a lattice  $\Lambda$  we can define the mod- $\Lambda$  channel with input  $X^n \in \mathcal{R}(\Lambda)$  and output

$$Y^n = [X^n + W^n] \bmod \mathcal{R}(\Lambda),$$

where  $W^n$  is a noise vector. When  $W^n$  is i.i.d. Gaussian with variance  $\sigma^2$ , this channel has capacity

$$C(\Lambda, \sigma^2) = \log V(\Lambda) - h(f_{\sigma, \Lambda}).$$

In the above expression, with slight abuse of notation we denote by  $h(f_{\sigma, \Lambda})$  the differential entropy of  $f_{\sigma, \mathcal{R}(\Lambda)}$ , which does not depend on the choice of the region  $\mathcal{R}(\Lambda)$ .

The following result [57, Lemma 1] relates the  $L^\infty$  flatness factor to the capacity of the mod- $\Lambda$  channel.

*Lemma 6:* The capacity  $C(\Lambda, \sigma^2)$  of the mod- $\Lambda$  channel is bounded by  $C(\Lambda, \sigma^2) \leq \log(1 + \epsilon_\Lambda(\sigma)) \leq \epsilon_\Lambda(\sigma)$ .

Given two nested lattices  $\Lambda' \subset \Lambda$  and a fundamental region  $\mathcal{R}(\Lambda')$ , we can define the mod- $\Lambda/\Lambda'$  channel with discrete input  $X^n \in \Lambda \cap \mathcal{R}(\Lambda')$  and output

$$Y^n = [X^n + W^n] \bmod \mathcal{R}(\Lambda').$$

It was shown in [58] that this channel has capacity

$$C(\Lambda/\Lambda', \sigma^2) = \log |\Lambda/\Lambda'| + h(f_{\sigma, \Lambda}) - h(f_{\sigma, \Lambda'}).$$

In particular, the following relation holds:

$$C(\Lambda/\Lambda', \sigma^2) = C(\Lambda', \sigma^2) - C(\Lambda, \sigma^2). \quad (48)$$

*Lemma 7:* For any  $\sigma > 0$ ,

$$C(\Lambda/\Lambda', \sigma^2) = \mathbb{D} \left( f_{\sigma, \mathcal{R}(\Lambda')} \left\| \frac{1}{|\Lambda/\Lambda'|} f_{\sigma, \Lambda | \mathcal{R}(\Lambda')} \right. \right).$$

*Proof:* By definition,

$$\begin{aligned}
\mathbb{D}\left(f_{\sigma, \mathcal{R}(\Lambda')} \left\| \frac{1}{|\Lambda/\Lambda'|} f_{\sigma, \Lambda|\mathcal{R}(\Lambda')} \right.\right) &= \int_{\mathcal{R}(\Lambda')} f_{\sigma, \Lambda'}(\mathbf{y}) \log \frac{f_{\sigma, \Lambda'}(\mathbf{y}) |\Lambda/\Lambda'|}{f_{\sigma, \Lambda}(\mathbf{y})} d\mathbf{y} \\
&= -h(f_{\sigma, \Lambda'}) + \int_{\mathcal{R}(\Lambda')} f_{\sigma, \Lambda'}(\mathbf{y}) \log \frac{|\Lambda/\Lambda'|}{f_{\sigma, \Lambda}(\mathbf{y})} d\mathbf{y} \\
&= -h(f_{\sigma, \Lambda'}) + \log |\Lambda/\Lambda'| - \int_{\mathcal{R}(\Lambda')} f_{\sigma, \Lambda'}(\mathbf{y}) \log f_{\sigma, \Lambda}(\mathbf{y}) d\mathbf{y}.
\end{aligned}$$

The conclusion follows by observing that

$$\begin{aligned}
- \int_{\mathcal{R}(\Lambda')} f_{\sigma, \Lambda'}(\mathbf{y}) \log f_{\sigma, \Lambda}(\mathbf{y}) d\mathbf{y} &= - \sum_{\lambda \in \Lambda/\Lambda'} \int_{\mathcal{R}(\Lambda) + \lambda} f_{\sigma, \Lambda'}(\mathbf{y}) \log f_{\sigma, \Lambda}(\mathbf{y}) d\mathbf{y} \\
&= - \sum_{\lambda \in \Lambda/\Lambda'} \int_{\mathcal{R}(\Lambda)} f_{\sigma, \Lambda'}(\mathbf{y} - \lambda) \log f_{\sigma, \Lambda}(\mathbf{y}) d\mathbf{y} = - \int_{\mathcal{R}(\Lambda)} f_{\sigma, \Lambda}(\mathbf{y}) \log f_{\sigma, \Lambda}(\mathbf{y}) d\mathbf{y} = h(f_{\sigma, \Lambda}). \quad \square
\end{aligned}$$

### B. The KL flatness factor

We can now introduce a notion of flatness factor based on KL divergence.

*Definition 10:* Given a lattice  $\Lambda$ , a fundamental region  $\mathcal{R}(\Lambda)$  and  $\sigma > 0$ , we define the *KL flatness factor* as follows:

$$\epsilon_{\Lambda}^{KL}(\sigma) = \mathbb{D}(f_{\sigma, \mathcal{R}(\Lambda)} \| \mathcal{U}_{\mathcal{R}(\Lambda)}). \quad (49)$$

Note that as before, the definition does not depend on the choice of the fundamental region.

*Remark 8:* By Pinsker's inequality,  $\forall \sigma > 0$ ,

$$\epsilon_{\Lambda}^1(\sigma) \leq \sqrt{2\epsilon_{\Lambda}^{KL}(\sigma)}.$$

*Remark 9 (Relation to the capacity of the mod- $\Lambda$  channel):* Note that [39, p.1656]

$$\mathbb{D}(f_{\sigma, \mathcal{R}(\Lambda)} \| \mathcal{U}_{\mathcal{R}(\Lambda)}) = \log V(\Lambda) - h(f_{\sigma, \Lambda}) = C(\Lambda, \sigma^2).$$

By shift-invariance of the differential entropy, the KL flatness factor is also shift-invariant, i.e.

$$\epsilon_{\Lambda}^{KL}(\sigma) = \mathbb{D}(f_{\sigma, \Lambda, \mathbf{c}|\mathcal{R}(\Lambda)} \| \mathcal{U}_{\mathcal{R}(\Lambda)})$$

for all  $\mathbf{c} \in \mathbb{R}^n$ .

Thanks to Remark 9, we are able to prove that the KL flatness factor is monotonic:

*Lemma 8:* For any lattice  $\Lambda$ ,  $\forall \sigma' > \sigma$ ,  $\epsilon_{\Lambda}^{KL}(\sigma') \leq \epsilon_{\Lambda}^{KL}(\sigma)$ .

*Proof:* With the same notation as in the proof of Lemma 2, from the data processing inequality for the KL divergence [54, Lemma 3.11] we have

$$\epsilon_{\Lambda}^{KL}\left(\sqrt{\sigma^2 + \sigma_0^2}\right) = \mathbb{D}\left(f_{\sqrt{\sigma^2 + \sigma_0^2}, \mathcal{R}(\Lambda)} \| \mathcal{U}_{\mathcal{R}(\Lambda)}\right) = \mathbb{D}(Y^n \| U^n) \leq \mathbb{D}(X^n \| U^n) = \mathbb{D}(f_{\sigma, \mathcal{R}(\Lambda)} \| \mathcal{U}_{\mathcal{R}(\Lambda)}) = \epsilon_{\Lambda}^{KL}(\sigma). \quad \square$$

Similarly to Definition 4, we can introduce a notion of secrecy goodness based on the KL flatness factor.

*Definition 11:* A sequence of lattices  $\{\Lambda^{(n)}\}$  is *KL secrecy-good* if  $\epsilon_{\Lambda^{(n)}}^{KL}(\sigma) = o\left(\frac{1}{n^c}\right)$ .

By Remark 8, a sequence of KL secrecy-good lattices is also  $L^1$  secrecy-good.

### C. Extracting the intrinsic randomness of additive Gaussian channels with a modulo lattice operation

In this section, we show that under the assumption of a small KL flatness factor, the modulo lattice operation allows to extract the intrinsic randomness of the additive Gaussian channel (in the sense of [30]). For simplicity, we only focus on privacy amplification and we assume that Alice and Bob observe the same i.i.d. Gaussian random



Fig. 4. Intrinsic randomness extractor.

variable  $X^n = Y^n$  of variance  $\sigma_x^2$  per dimension. Eve observes a correlated i.i.d. random variable  $Z^n$ . We assume that  $X^n$  and  $Z^n$  are jointly Gaussian, according to the model

$$X^n = Z^n + W^n, \quad (50)$$

where  $W^n$  is an i.i.d. zero-mean Gaussian random vector of variance  $\sigma^2$  per dimension, and is independent of  $Z^n$ .

Our aim is to extract from  $X^n$  a random variable that is almost uniform on  $\mathcal{R}(\Lambda)$  and almost independent of  $Z^n$  (cf. Fig. 4). More precisely, we give the following definition.

*Definition 12 (Continuous lattice extractor for channel intrinsic randomness under a KL divergence criterion):* Given the model (50) and  $\varepsilon > 0$ , an  $\varepsilon$ -lattice extractor with lattice  $\Lambda$  for channel intrinsic randomness (under a KL divergence criterion) is a device that takes  $X^n$  as input and outputs  $\bar{X}^n \in \mathcal{R}(\Lambda)$  such that  $\mathbb{D}(p_{\bar{X}^n} || \mathcal{U}_{\mathcal{R}(\Lambda)}) \leq \varepsilon$  and  $\mathbb{I}(X^n; Z^n) = \mathbb{D}(p_{Z^n \bar{X}^n} || p_{Z^n} p_{\bar{X}^n}) \leq \varepsilon$ .

We will show that if  $\Lambda$  is a lattice in  $\mathbb{R}^n$  such that  $\epsilon_{\Lambda}^{KL}(\sigma) = \epsilon$ , and  $\mathcal{R}(\Lambda)$  is a fundamental region of  $\Lambda$ , then  $\bar{X}^n = X^n \bmod \mathcal{R}(\Lambda)$  satisfies Definition 12. Recall that

$$p_{\bar{X}^n}(\bar{\mathbf{x}}) = f_{\sigma_x, \Lambda}(\bar{\mathbf{x}}) \mathbb{1}_{\mathcal{R}(\Lambda)}(\bar{\mathbf{x}}).$$

The conditional density of  $\bar{X}^n$  given  $Z^n$  is

$$\begin{aligned} p_{\bar{X}^n|Z^n}(\bar{\mathbf{x}}|\mathbf{z}) &= \sum_{\mathbf{x}: \bar{\mathbf{x}} = \mathbf{x} \bmod \mathcal{R}(\Lambda)} p_{X^n|Z^n}(\mathbf{x}|\mathbf{z}) = \sum_{\mathbf{x} \in \bar{\mathbf{x}} + \Lambda} p_{X^n|Z^n}(\mathbf{x}|\mathbf{z}) = \sum_{\lambda \in \Lambda} \frac{1}{(\sqrt{2\pi}\sigma)^n} e^{-\frac{\|\bar{\mathbf{x}} + \lambda - \mathbf{z}\|^2}{2\sigma^2}} \\ &= f_{\sigma, \Lambda}(\bar{\mathbf{x}} - \mathbf{z}) \mathbb{1}_{\mathcal{R}(\Lambda)}(\bar{\mathbf{x}}) = f_{\sigma, \Lambda, \mathbf{z}}(\bar{\mathbf{x}}) \mathbb{1}_{\mathcal{R}(\Lambda)}(\bar{\mathbf{x}}). \end{aligned}$$

By monotonicity of the KL flatness factor (Lemma 8),  $\mathbb{D}(p_{\bar{X}^n} || \mathcal{U}_{\mathcal{R}(\Lambda)}) = \epsilon_{\Lambda}^{KL}(\sigma_x) \leq \epsilon_{\Lambda}^{KL}(\sigma)$ , and moreover

$$\begin{aligned} h(\bar{X}^n) &= h(f_{\sigma_x, \Lambda}), \\ h(\bar{X}^n|Z^n = \mathbf{z}) &= - \int_{\mathcal{R}(\Lambda)} p_{\bar{X}^n|Z^n}(\bar{\mathbf{x}}|\mathbf{z}) \log p_{\bar{X}^n|Z^n}(\bar{\mathbf{x}}|\mathbf{z}) d\bar{\mathbf{x}} = - \int_{\mathcal{R}(\Lambda)} f_{\sigma, \Lambda}(\bar{\mathbf{x}} - \mathbf{z}) \log f_{\sigma, \Lambda}(\bar{\mathbf{x}} - \mathbf{z}) d\bar{\mathbf{x}} \\ &= - \int_{\mathcal{R}(\Lambda) - \mathbf{z}} f_{\sigma, \Lambda}(\bar{\mathbf{x}}) \log f_{\sigma, \Lambda}(\bar{\mathbf{x}}) d\bar{\mathbf{x}} = h(f_{\sigma, \Lambda}) \end{aligned}$$

since  $f_{\sigma, \Lambda}$  is  $\Lambda$ -periodic and  $\mathcal{R}(\Lambda) - \mathbf{z}$  is a fundamental region of  $\Lambda$ . We can now bound the mutual information as follows:

$$\begin{aligned} \mathbb{I}(\bar{X}^n; Z^n) &= h(\bar{X}^n) - h(\bar{X}^n|Z^n) \\ &= h(f_{\sigma_x, \Lambda}) - \int_{\mathbb{R}^n} p_{Z^n}(\mathbf{z}) h(\bar{X}^n|Z^n = \mathbf{z}) d\mathbf{z} = h(f_{\sigma_x, \Lambda}) - h(f_{\sigma, \Lambda}) \\ &= \log V(\Lambda) - h(f_{\sigma, \Lambda}) = (\log V(\Lambda) - h(f_{\sigma_x, \Lambda})) = \epsilon_{\Lambda}^{KL}(\sigma) - \epsilon_{\Lambda}^{KL}(\sigma_x) \leq \epsilon_{\Lambda}^{KL}(\sigma). \end{aligned}$$

The differential entropy of  $\bar{X}^n$  is lower bounded by

$$h(\bar{X}^n) = h(f_{\sigma_x, \Lambda}) = \log V(\Lambda) - \epsilon_{\Lambda}^{KL}(\sigma_x) = h(\mathcal{U}_{\mathcal{R}(\Lambda)}) - \epsilon_{\Lambda}^{KL}(\sigma_x) \geq h(\mathcal{U}_{\mathcal{R}(\Lambda)}) - \epsilon_{\Lambda}^{KL}(\sigma)$$

since, by monotonicity of the KL flatness factor,  $\epsilon_{\Lambda}^{KL}(\sigma_x) \leq \epsilon_{\Lambda}^{KL}(\sigma)$ .

*Remark 10:* Taking a sequence of KL secrecy-good lattices  $\{\Lambda^{(n)}\}$  such that  $\gamma_{\Lambda^{(n)}}(\sigma) \rightarrow 2\pi e$  and  $\epsilon_{\Lambda^{(n)}}^{KL}(\sigma) \rightarrow 0$  as  $n \rightarrow \infty$ , the asymptotic differential entropy rate of  $\bar{X}^n$  is

$$r = \liminf_{n \rightarrow \infty} \frac{1}{n} h(\bar{X}^n) = \liminf_{n \rightarrow \infty} \frac{1}{n} \log V(\Lambda) = \frac{1}{2} \log 2\pi e \sigma^2,$$

which is equal to the asymptotic differential entropy rate of the Gaussian noise  $W^n$ . Thus, the modulo lattice

operation can asymptotically extract the intrinsic randomness of the channel.

*Remark 11:* It is worth mentioning that our continuous lattice extractor scheme does not require dithering. Moreover, since the flatness factor  $\epsilon_{\Lambda}^{KL}$  only depends on  $f_{\sigma, \Lambda}$  which is periodic mod  $\Lambda$ , we can choose any fundamental region  $\mathcal{R}(\Lambda)$ .

*Remark 12:* We note that if we replace the KL flatness factor condition  $\epsilon_{\Lambda}^{KL}(\sigma) = \varepsilon$  with the  $L^1$  flatness factor condition  $\epsilon_{\Lambda}^1(\sigma) = \varepsilon$ , one can easily show that  $\mathbb{V}(p_{\bar{X}^n}, \mathcal{U}_{\mathcal{R}(\Lambda)}) \leq \varepsilon$  and  $\mathbb{V}(p_{Z^n \bar{X}^n}, p_{Z^n} p_{\bar{X}^n}) \leq 2\varepsilon$ , but on the other hand we cannot guarantee that the differential entropy  $h(\bar{X}^n)$  is close to  $h(\mathcal{U}_{\mathcal{R}(\Lambda)})$ . In fact, for general continuous distributions, convergence in variational distance does not necessarily imply convergence in differential entropy [61]<sup>7</sup>.

Finally, we note that since the  $L^1$  flatness factor and KL flatness factor are *average* conditions, we can't obtain a discrete key from  $\bar{X}^n$  simply by quantizing it with respect to a fine lattice  $\Lambda_1 \supset \Lambda$ . In order to obtain good discrete keys, one can combine the modulo lattice operation with uniform dithering over a fundamental region  $\mathcal{R}(\Lambda_1)$ , as shown in Section III.

## APPENDIX C PROOF OF THEOREM 1

In order to prove Theorem 1, we will actually show a stronger result:

*Proposition 2:* If  $\gamma_{\Lambda}(\sigma) < 2\pi e$  is fixed, then there exists a sequence  $\{\Lambda^{(n)}\}$  of lattices which are KL secrecy-good. Theorem 1 then follows from Proposition 2 by Remark 8.

Before proceeding with the proof, we summarize the main idea here. We use the standard Construction A to find the sought-after lattice  $\Lambda$ , by choosing a coarse lattice  $\Lambda_c = \alpha p\mathbb{Z}$ , a fine lattice  $\Lambda_f = \alpha\mathbb{Z}$ , an  $(n, k)$  linear code  $\mathcal{C}$  over  $\mathbb{F}_p$ , and  $\Lambda_c^n \subseteq \Lambda = \alpha(p\mathbb{Z}^n + \mathcal{C}) \subseteq \Lambda_f^n$ . Using the chain rule (48), we have

$$\mathbb{D}(f_{\mathcal{R}(\Lambda), \sigma} || \mathcal{U}_{\mathcal{R}(\Lambda)}) = C(\Lambda, \sigma^2) = C(\Lambda_f^n, \sigma^2) + C(\Lambda_f^n / \Lambda, \sigma^2).$$

Now, using a sufficiently fine lattice  $\Lambda_f$ , we can easily make  $C(\Lambda_f^n, \sigma^2) \rightarrow 0$  thanks to the flatness phenomenon (cf. Lemma 6). The non-trivial part of the proof is to exhibit a lattice  $\Lambda$  such that  $C(\Lambda_f^n / \Lambda, \sigma^2) \rightarrow 0$  as well. It turns out that if the linear code  $\mathcal{C}$  is a *resolvability code* for the mod- $\Lambda_f / \Lambda_c$  channel  $W$ , i.e. if the output of the code is close to the output of uniform input, then  $\mathcal{C}$  provides the desired solution. In fact, we show that

$$\mathbb{D}(W^n \circ \mathcal{U}_{\mathcal{C}} || W^n \circ \mathcal{U}_{(\Lambda_f / \Lambda_c)^n}) = C(\Lambda_f^n / \Lambda, \sigma^2),$$

which tends to 0 if  $\mathcal{C}$  is a resolvability code. The existence of such linear resolvability codes follows from the results of [36] (see Appendix A). However, making the above argument rigorous involve certain technicalities, as seen in the following.

*Proof of Proposition 2:*

For a given dimension  $n$ , we will construct  $\Lambda$  as a scaled mod- $p$  lattice [63] of the form  $\Lambda = \alpha(p\mathbb{Z}^n + \mathcal{C}_n)$ , where  $\mathcal{C}_n$  is an  $(n, k)$ -linear code over  $\mathbb{F}_p$ .

We will consider the asymptotic behavior as  $n \rightarrow \infty$ ,  $\alpha \rightarrow 0$ ,  $p \rightarrow \infty$  while satisfying the volume condition  $\alpha^n p^{n-k} = V(\Lambda) = (\gamma\sigma^2)^{n/2}$ . Here,  $\gamma$  is the volume-to-noise ratio, which is assumed to be fixed.

By construction,  $\Lambda_c^n \subset \Lambda \subset \Lambda_f^n$ , where  $\Lambda_c = \alpha p\mathbb{Z}$  and  $\Lambda_f = \alpha\mathbb{Z}$  are one-dimensional lattices.

From Remark 9 and the relation (48), we have

$$\mathbb{D}(f_{\sigma, \mathcal{R}(\Lambda)} || \mathcal{U}_{\mathcal{R}(\Lambda)}) = C(\Lambda, \sigma^2) = C(\Lambda_f^n, \sigma^2) + C(\Lambda_f^n / \Lambda, \sigma^2).$$

We want to show that both terms in the sum tend to zero when  $n \rightarrow \infty$ .

- 1) First, we will show that  $C(\Lambda_f^n, \sigma^2) = C((\alpha\mathbb{Z})^n, \sigma^2) \rightarrow 0$  if  $\alpha = o(\frac{1}{n^c})$  for some  $c > 0$ . We follow the same approach as in [57, Appendix A]. From Lemma 6 we have that  $C(\Lambda_f^n, \sigma^2) \leq \epsilon_{\Lambda_f^n}(\sigma)$ . Furthermore, it was shown in [64, Lemma 3] that

$$\epsilon_{\Lambda_f^n}(\sigma) = (1 + \epsilon_{\Lambda_f}(\sigma))^n - 1. \quad (51)$$

<sup>7</sup>More precisely, [61, Definition 3] identifies a class of probability distributions for which convergence in total variation leads to convergence in entropy. One of the requirements is that the densities should be bounded, which is not guaranteed under an  $L^1$  flatness factor criterion. See also the discussion in [62].

Finally, one can show that [57, Appendix A]

$$\epsilon_{\Lambda_f}(\sigma) = \epsilon_{\alpha\mathbb{Z}}(\sigma) \leq 4e^{-\frac{2\pi^2\sigma^2}{\alpha^2}}. \quad (52)$$

Then

$$\epsilon_{\Lambda_f^n}(\sigma) \leq \left(1 + 4e^{-\frac{2\pi^2\sigma^2}{\alpha^2}}\right)^n - 1 \leq 4ne^{-\frac{2\pi^2\sigma^2}{\alpha^2}} + o(e^{-\frac{2\pi^2\sigma^2}{\alpha^2}}) \rightarrow 0.$$

since  $(1+x)^n = 1 + nx + o(x)$  when  $x \rightarrow 0$ .

- 2) Next, we want to show that there exists a sequence of lattices  $\Lambda$  of the form  $\alpha(p\mathbb{Z}^n + \mathcal{C}_n)$  such that  $C(\Lambda_f^n/\Lambda, \sigma^2) \rightarrow 0$  as  $n \rightarrow \infty$ .

Consider the mod- $(\Lambda_f/\Lambda_c)$  channel  $W : \Lambda_f \cap \mathcal{R}(\Lambda_c) \rightarrow \mathcal{R}(\Lambda_c)$ . This channel is regular (see Definition 9 in Appendix A) with respect to the set of permutations  $\pi_x(y) = [y - x] \bmod \Lambda_c$  for  $x \in \mathcal{X} = \Lambda_f \cap \mathcal{R}(\Lambda_c)$ ,  $y \in \mathcal{R}(\Lambda_c)$ . In fact,

$$W_x(y) = W(y|x) = f_{\sigma, \Lambda_c}(y - x) = f_{\sigma, \Lambda_c}([y - x] \bmod \Lambda_c) = f_{\sigma, \Lambda_c}(\pi_x(y)).$$

Being regular, the mod  $\Lambda_f/\Lambda_c$  channel is symmetric and the uniform distribution over  $\mathcal{X}$  achieves capacity (see Appendix A). Moreover,  $\Lambda_f/\Lambda_c \cong \mathbb{F}_p$  as abelian groups. We consider the required rate condition in Corollary 1:

$$R = \frac{1}{n} \log |\mathcal{C}_n| = \frac{1}{n} \log |\Lambda/\Lambda_c^n| = \frac{1}{n} \log \frac{\alpha^n p^n}{V(\Lambda)} > \mathbb{I}(X; Y) = C(\Lambda_f/\Lambda_c, \sigma^2). \quad (53)$$

We have

$$\begin{aligned} C(\Lambda_f/\Lambda_c, \sigma^2) &= \log |\Lambda_f/\Lambda_c| + h(f_{\sigma, \Lambda_f}) - h(f_{\sigma, \Lambda_c}) = \log p + h(f_{\sigma, \Lambda_f}) - h(f_{\sigma, \Lambda_c}) \\ &= \log p + \log \alpha - C(\Lambda_f, \sigma^2) - h(f_{\sigma, \Lambda_c}). \end{aligned}$$

Therefore, the condition (53) is equivalent to

$$\frac{1}{n} \log V(\Lambda) < h(f_{\sigma, \Lambda_c}) + C(\Lambda_f, \sigma^2).$$

In the asymptotic limit for  $\alpha \rightarrow 0$ ,  $p \rightarrow \infty$  while keeping  $\alpha^n p^{n-k} = V(\Lambda) = (\gamma\sigma^2)^{n/2}$ , we have  $C(\Lambda_f, \sigma^2) \rightarrow 0$ . Moreover,  $\alpha p \rightarrow \infty$ , and so  $h(\Lambda_c, \sigma^2) \rightarrow \frac{1}{2} \log 2\pi e \sigma^2$ . So asymptotically, the rate condition is satisfied when

$$\frac{V(\Lambda)^{2/n}}{2\pi e \sigma^2} < 1. \quad (54)$$

In this case we have

$$R - \mathbb{I}(X; Y) = -\frac{1}{n} \log V(\Lambda) + C(\Lambda_f, \sigma^2) - h(f_{\sigma, \Lambda_c}) \rightarrow \delta_0 = \frac{1}{2} \log \frac{2\pi e \sigma^2}{V(\Lambda)^{2/n}} = \frac{1}{2} \log \frac{2\pi e}{\gamma\Lambda(\sigma)} > 0 \quad (55)$$

as  $n \rightarrow \infty$ .

*Remark 13:* Note that we cannot directly apply Corollary 1 in Appendix A to this setting, since the definition of the channel  $W$  depends on  $\alpha$  and  $p$  which are not fixed but are a function of  $n$ . However, we will show that the proof of the Corollary can be extended to this channel since the convergence in (47) is uniform.

*Proof of Remark 13:* Let  $X$  be a uniformly distributed variable on  $\Lambda_f \cap \mathcal{R}(\Lambda_c)$  and  $Y$  the corresponding output distribution. Consider the function  $\psi(\rho) = \psi(\rho|W, \mathcal{U}_X)$  in Definition 8. From (43) and (45), it follows that its Taylor expansion in 0 is given by

$$\psi(\rho) = \rho \mathbb{I}(X; Y) + \rho^2 \psi''(0) + o(\rho^2), \quad (56)$$

where  $\psi''(0)$  is given in Lemma 5. Noting that

$$(W \circ \mathcal{U}_X)(y) = \sum_{x \in \mathcal{X}} \frac{1}{|\mathcal{X}|} W_x(y) = \sum_{x \in \Lambda_f \cap \mathcal{R}(\Lambda_c)} \frac{1}{|\Lambda_f/\Lambda_c|} f_{\sigma, \Lambda_c}(y - x) = \frac{1}{|\Lambda_f/\Lambda_c|} f_{\sigma, \Lambda_f}(y),$$

we find

$$\begin{aligned}
\psi''(0) &= \sum_{x \in \Lambda_f \cap \mathcal{R}(\Lambda_c)} \frac{1}{|\Lambda_f/\Lambda_c|} \int_{\mathcal{R}(\Lambda_c)} f_{\sigma, \Lambda_c}(y-x) \left( \log \frac{f_{\sigma, \Lambda_c}(y-x)}{\frac{1}{|\Lambda_f/\Lambda_c|} f_{\sigma, \Lambda_f}(y)} \right)^2 dy \\
&\quad - \left( \sum_{x \in \Lambda_f \cap \mathcal{R}(\Lambda_c)} \frac{1}{|\Lambda_f/\Lambda_c|} \int_{\mathcal{R}(\Lambda_c)} f_{\sigma, \Lambda_c}(y-x) \log \frac{f_{\sigma, \Lambda_c}(y-x)}{\frac{1}{|\Lambda_f/\Lambda_c|} f_{\sigma, \Lambda_f}(y)} dy \right)^2 \\
&\leq \sum_{x \in \Lambda_f \cap \mathcal{R}(\Lambda_c)} \frac{1}{|\Lambda_f/\Lambda_c|} \int_{\mathcal{R}(\Lambda_c)} f_{\sigma, \Lambda_c}(y-x) \left( \log \frac{f_{\sigma, \Lambda_c}(y-x)}{\frac{1}{|\Lambda_f/\Lambda_c|} f_{\sigma, \Lambda_f}(y)} \right)^2 dy \\
&= \int_{\mathcal{R}(\Lambda_c)} f_{\sigma, \Lambda_c}(y') \left( \log \frac{f_{\sigma, \Lambda_c}(y')}{\frac{1}{|\Lambda_f/\Lambda_c|} f_{\sigma, \Lambda_f}(y')} \right)^2 dy'
\end{aligned}$$

with the change of variables  $y' = y - x \bmod \mathcal{R}(\Lambda_c)$ . From the definition of flatness factor and the bound (52), we find that  $\forall y' \in \mathcal{R}(\Lambda_c)$ ,

$$f_{\sigma, \Lambda_f}(y') \geq \frac{1 - \epsilon_{\Lambda_f}(\sigma)}{V(\Lambda_f)} \geq \frac{1 - 4e^{-\frac{2\pi^2\sigma^2}{\alpha^2}}}{\alpha}.$$

Recalling the definition of the theta series of a lattice in (6) and the relation (8), we have  $\epsilon_{\Lambda}(\sigma) = \Theta_{\Lambda^*}(2\pi\sigma^2) - 1$ , where  $\Lambda^*$  is the dual lattice of  $\Lambda$ . Then by [35, Remark 1],  $\forall y' \in \mathcal{V}(\Lambda_c)$

$$f_{\sigma, \Lambda_c}(y') \leq f_{\sigma, \Lambda_c}(0) = \frac{1}{\sqrt{2\pi}\sigma} \Theta_{\Lambda_c} \left( \frac{1}{2\pi\sigma^2} \right) = \frac{1}{\sqrt{2\pi}\sigma} \left( 1 + \epsilon_{\Lambda_c^*} \left( \frac{1}{2\pi\sigma} \right) \right).$$

Again using the bound (52), we have

$$\epsilon_{\Lambda_c^*} \left( \frac{1}{2\pi\sigma} \right) = \epsilon_{\frac{1}{\alpha p} \mathbb{Z}} \left( \frac{1}{2\pi\sigma} \right) \leq 4e^{-\frac{\alpha^2 p^2}{2\sigma^2}}.$$

Then, since  $\alpha \rightarrow 0$  and  $\alpha p \rightarrow \infty$  when  $n \rightarrow \infty$ , for sufficiently large  $n$  we have

$$\frac{f_{\sigma, \Lambda_c}(y')}{\frac{1}{|\Lambda_f/\Lambda_c|} f_{\sigma, \Lambda_f}(y')} \leq \frac{1}{\sqrt{2\pi}\sigma} \frac{\alpha p (1 + 4e^{-\frac{\alpha^2 p^2}{2\sigma^2}})}{1 - 4e^{-\frac{2\pi^2\sigma^2}{\alpha^2}}} \leq C\alpha p$$

for some constant  $C > 0$ . Consequently, for large enough  $n$ ,  $\exists C' > 0$  such that

$$\psi''(0) \leq C'(\log \alpha p)^2.$$

Then, from the Taylor expansion (56) we obtain the bound

$$\psi(\rho) \leq \rho \mathbb{I}(\mathbf{X}; \mathbf{Y}) + \rho^2 C'' (\log \alpha p)^2$$

for another suitable constant  $C'' > 0$ . In particular, we can bound the exponent in equation (46) as follows:

$$\rho R - \psi(\rho | W, \mathcal{U}_{\mathcal{X}}) \geq \rho(R - \mathbb{I}(\mathbf{X}; \mathbf{Y}) - \rho C'' (\log \alpha p)^2) > \rho \frac{\delta_0}{2}$$

for sufficiently large  $n$ , where  $\delta_0$  is defined in (55), as long as  $\rho = o\left(\frac{1}{(\log \alpha p)^2}\right)$  and the VNR condition (54) is satisfied. In particular if we choose the scaling<sup>8</sup>

$$p = \xi n^{3/2}, \quad \alpha p = 2\sqrt{n}, \quad (57)$$

where  $\xi$  is the smallest number in the interval  $[1, 2)$  such that  $p$  is prime [65, Section IV], we have convergence

<sup>8</sup>This choice of scaling is compatible with the existence of a suitable sequence of nested lattices, see Appendix D.

in (47) with  $\bar{\rho} = \frac{1}{(\log 2\sqrt{n})^{2+\eta}}$  for some  $\eta > 0$  since

$$\frac{1}{\bar{\rho}} e^{-n\bar{\rho}\frac{\delta_0}{2}} = (\log 2\sqrt{n})^{2+\eta} e^{-\frac{n\delta_0}{2(\log 2\sqrt{n})^{2+\eta}}} \rightarrow 0.$$

This concludes the proof of Remark 13.  $\square$

Then according to Corollary 1, for  $\mathcal{C}_n$  chosen uniformly in the set of  $(n, k)$  linear codes over  $\mathbb{F}_p$  of rate  $R = \frac{k}{n} \log p$ ,

$$\mathbb{E}_{\mathcal{C}_n} [\mathbb{D}(W^n \circ \mathcal{U}_{\mathcal{C}_n} \| W^n \circ \mathcal{U}_{\mathcal{X}}^{\otimes n})] \leq \frac{1}{\bar{\rho}} e^{-n\bar{\rho}\frac{\delta_0}{2}} \rightarrow 0$$

as  $n \rightarrow \infty$ . In particular, there exists at least one code  $\mathcal{C}_n$  such that  $\mathbb{D}(W^n \circ \mathcal{U}_{\mathcal{C}_n} \| W^n \circ \mathcal{U}_{\mathcal{X}}^{\otimes n}) \rightarrow 0$ . Note that

$$\begin{aligned} (W^n \circ \mathcal{U}_{\mathcal{C}_n})(\mathbf{y}) &= \sum_{\mathbf{c} \in \mathcal{C}_n} \frac{1}{|\mathcal{C}_n|} f_{\sigma, \Lambda_c^n}(\mathbf{y} - \alpha \mathbf{c}) = \sum_{\mathbf{c} \in \mathcal{C}_n} \sum_{\lambda_c \in \Lambda_c^n} \frac{1}{p^k} f_{\sigma}(\mathbf{y} - \alpha \mathbf{c} - \lambda_c) = \frac{1}{p^k} \sum_{\lambda \in \Lambda} f_{\sigma}(\mathbf{y} - \lambda) \\ &= \frac{1}{p^k} f_{\sigma, \Lambda}(\mathbf{y}), \end{aligned} \quad (58)$$

$$(W^n \circ \mathcal{U}_{\mathcal{X}}^{\otimes n})(\mathbf{y}) = \sum_{\mathbf{x} \in \Lambda_f^n \cap \mathcal{R}(\Lambda_c^n)} \frac{1}{p^n} f_{\sigma, \Lambda_c^n}(\mathbf{y} - \mathbf{x}) = \frac{1}{p^n} f_{\sigma, \Lambda_f^n}(\mathbf{y}). \quad (59)$$

Since both  $(W^n \circ \mathcal{U}_{\mathcal{C}_n})$  and  $(W^n \circ \mathcal{U}_{\mathcal{X}}^{\otimes n})$  are  $\Lambda$ -periodic, we can write

$$\begin{aligned} \mathbb{D}(W^n \circ \mathcal{U}_{\mathcal{C}_n} \| W^n \circ \mathcal{U}_{\mathcal{X}}^{\otimes n}) &= \int_{\mathcal{R}(\Lambda_c^n)} p^{-k} f_{\sigma, \Lambda}(\mathbf{y}) \log \frac{p^{-k} f_{\sigma, \Lambda}(\mathbf{y})}{p^{-n} f_{\sigma, \Lambda_f^n}(\mathbf{y})} d\mathbf{y} \\ &= \int_{\mathcal{R}(\Lambda)} f_{\sigma, \Lambda}(\mathbf{y}) \log \frac{f_{\sigma, \Lambda}(\mathbf{y})}{p^{-(n-k)} f_{\sigma, \Lambda_f^n}(\mathbf{y})} d\mathbf{y} = \mathbb{D}(f_{\sigma, \mathcal{R}(\Lambda)} \| p^{-(n-k)} f_{\sigma, \Lambda_f^n | \mathcal{R}(\Lambda)}) = C(\Lambda_f^n / \Lambda, \sigma^2) \rightarrow 0 \end{aligned}$$

using Lemma 7. This concludes the proof.  $\square$

*Remark 14:* With a standard argument based on Markov's inequality, we can also show that the set of KL-secrecy good lattices has large measure, since  $\forall \xi > 0$ ,

$$\mathbb{P} \{ \mathbb{D}(W^n \circ \mathcal{U}_{\mathcal{C}_n} \| W^n \circ \mathcal{U}_{\mathcal{X}}^{\otimes n}) > \xi \} \leq \frac{1}{\xi} \mathbb{E}_{\mathcal{C}_n} [\mathbb{D}(W^n \circ \mathcal{U}_{\mathcal{C}_n} \| W^n \circ \mathcal{U}_{\mathcal{X}}^{\otimes n})].$$

Given  $0 < c < 1/2$ , we can take  $\xi = \frac{1}{c} \frac{e^{-n\bar{\rho}\delta_0}}{\bar{\rho}}$  and we obtain

$$\mathbb{P} \{ \mathbb{D}(W^n \circ \mathcal{U}_{\mathcal{C}_n} \| W^n \circ \mathcal{U}_{\mathcal{X}}^{\otimes n}) > \xi \} \leq c.$$

## APPENDIX D

### EXISTENCE OF A SEQUENCE OF NESTED LATTICES FOR SECRET KEY GENERATION

In this section, we show the existence of a sequence of nested lattices  $\Lambda_3^{(n)} \subset \Lambda_2^{(n)} \subset \Lambda_1^{(n)}$  such that  $\Lambda_3$  is KL secrecy-good,  $\Lambda_2$  is AWGN-good and  $\Lambda_1$  is KL secrecy-good. Note that we don't need covering-goodness, which requires more stringent conditions on the parameters [66].

We will follow the construction in [65]. We denote by  $V_{\mathcal{B}, n}$  the volume of the  $n$ -dimensional ball of radius 1. Given  $P_3 > P_2 > P_1 > 0$ , let  $a_i = \log \frac{1}{P_i}$  for  $i = 1, 2, 3$ . We consider the dimensions  $k_3 < k_2 < k_1 \leq n$  defined as follows:

$$k_i = \left\lfloor \frac{n}{2 \log p} \left( \log \left( \frac{4}{V_{\mathcal{B}, n}^{2/n}} \right) + a_i \right) \right\rfloor, \quad i = 1, 2, 3,$$

where  $p = \xi n^{3/2}$ , and  $\xi$  is taken to be the smallest number in the interval  $[1, 2)$  such that  $p$  is prime [65, Section IV]<sup>9</sup>. Let  $\mathcal{C}_1$  be uniformly sampled from the set of all linear  $(n, k_1)$  codes over  $\mathbb{F}_p$ , with generator matrix  $G_1$  (in column notation). We denote by  $G_2$  and  $G_3$  the submatrices of  $G_1$  corresponding to the first  $k_2$  and  $k_3$  columns

<sup>9</sup>Note that the conclusions of [65] still hold for any  $p = \Theta(n^{\frac{1}{2}+\delta})$  with  $\delta > 0$ , see Remark 7 in that paper.



respectively, and by  $\mathcal{C}_2, \mathcal{C}_3$  the corresponding linear codes. Finally, we define the lattices  $\tilde{\Lambda}_i = \frac{1}{p}\mathcal{C}_i + \mathbb{Z}^n$  and  $\Lambda_i = \alpha p \tilde{\Lambda}_i$  for  $i = 1, 2, 3$ , where  $\alpha = \frac{2\sqrt{n}}{p}$ . Then by [65, Theorem 1 and Theorem 6], the matrices  $G_1, G_2, G_3$  are full rank and the nested lattices  $\Lambda_3^{(n)} \subset \Lambda_2^{(n)} \subset \Lambda_1^{(n)}$  obtained in this way are good for quantization and coding with probability that tends to 1 when  $n \rightarrow \infty$  and

$$\lim_{n \rightarrow \infty} V^{2/n}(\Lambda_i^{(n)}) = 2\pi e P_i, \quad i = 1, 2, 3.$$

Note that we have taken the same scaling as in (57). In particular, when  $n \rightarrow \infty$  we have  $p \rightarrow \infty, \alpha \rightarrow 0$  and  $\alpha p \rightarrow \infty$ .

Moreover,  $\alpha = \frac{2}{\xi n}$  satisfies the condition  $\alpha = o\left(\frac{1}{n^c}\right)$  in Appendix C. Therefore, due to Remark 14 the lattices  $\Lambda_3$  and  $\Lambda_1$  are also KL secrecy-good with probability close to 1, which concludes the proof.

## APPENDIX E

### OPTIMAL PUBLIC RATE / SECRET KEY RATE TRADE-OFF

In this section, we derive the optimal trade-off between public rate and secret key rate from [23] for the setting in our paper. Note that Theorem 4 in [23] doesn't directly apply to our model because our source doesn't necessarily satisfy  $X \rightarrow Y \rightarrow Z$ . However, the proof of Lemma 6 in [23] shows how to obtain a new source  $(\bar{X}, \bar{Y}, \bar{Z})$  which is degraded ( $\bar{X} \rightarrow \bar{Y} \rightarrow \bar{Z}$ ) and has the same achievable region ( $\mathcal{R}(X, Y, Z) = \mathcal{R}(\bar{X}, \bar{Y}, \bar{Z})$ ). In particular, translating the proof into our notation, we can take  $\bar{X} = X, \bar{Y} = Y$  and

$$\bar{Z} = \frac{\sigma_z \rho_{xz}}{\sigma_y \rho_{xy}} Y + \hat{N},$$

where  $\hat{N}$  is independent of all other random variables and has variance  $\sigma_z^2 \left(1 - \frac{\rho_{xz}^2}{\rho_{xy}^2}\right)$ .

From elementary computations we see that  $\sigma_{\bar{z}} = \sigma_z, \rho_{x\bar{z}} = \rho_{xz}$  and  $\rho_{y\bar{z}} = \frac{\rho_{xz}}{\rho_{xy}}$ .

In our notation, the optimal trade-off given by Theorem 4 of [23] is given by

$$R_K \leq \frac{1}{2} \log \frac{(1 - \rho_{y\bar{z}}^2)(1 - \rho_{x\bar{z}}^2) - (\rho_{x\bar{y}} - \rho_{y\bar{z}}\rho_{x\bar{z}})^2 e^{-2R_P}}{(1 - \rho_{y\bar{z}}^2)(1 - \rho_{x\bar{z}}^2) - (\rho_{x\bar{y}} - \rho_{y\bar{z}}\rho_{x\bar{z}})^2}.$$

In terms of the original variables  $X, Y, Z$ , after simplifying the expression we obtain the optimal trade-off

$$R_K \leq \frac{1}{2} \log \frac{(1 - \rho_{xz}^2) - (\rho_{xy}^2 - \rho_{xz}^2)e^{-2R_P}}{1 - \rho_{xy}^2}.$$

(Recall that  $\rho_{xy} > \rho_{xz}$  in our setting). Using the notation  $\sigma_1^2 = \sigma_x^2(1 - \rho_{xy}^2), \sigma_2^2 = \sigma_x^2(1 - \rho_{xz}^2)$  from our paper, this is equal to

$$R_K \leq \frac{1}{2} \log \left( e^{-2R_P} + \frac{\sigma_2^2}{\sigma_1^2} (1 - e^{-2R_P}) \right). \quad (60)$$

## APPENDIX F

### PROOF OF LEMMA 5

The first derivative of the function  $\psi(\rho) = \psi(\rho|W, p_X)$  is

$$\psi'(\rho) = \frac{\sum_{x \in \mathcal{X}} p_X(x) \int_{\mathcal{Y}} \frac{W_x(y)^{1+\rho}}{((W \circ p_X)(y))^\rho} \log \frac{W_x(y)}{(W \circ p_X)(y)} dy}{\sum_{x \in \mathcal{X}} p_X(x) \int_{\mathcal{Y}} \frac{W_x(y)^{1+\rho}}{((W \circ p_X)(y))^\rho} dy} = \frac{f(\rho)}{g(\rho)}.$$

Then we have

$$\begin{aligned} g(0) &= 1, \\ f(0) &= \sum_{x \in \mathcal{X}} p_X(x) \int_{\mathcal{Y}} W_x(y) \log \frac{W_x(y)}{(W \circ p_X)(y)} dy = g'(0), \\ f'(0) &= \sum_{x \in \mathcal{X}} p_X(x) \int_{\mathcal{Y}} W_x(y) \left( \log \frac{W_x(y)}{(W \circ p_X)(y)} \right)^2 dy. \end{aligned}$$

The conclusion follows since

$$\psi''(0) = \frac{f'(0)g(0) - f(0)g'(0)}{g^2(0)}. \quad \square$$

## REFERENCES

- [1] C. Ling, L. Luzzi, and M. Bloch, “Secret key generation from Gaussian sources using lattice hashing,” in *Proc. IEEE Int. Symp. Inf. Theory*, July 2013.
- [2] C. Ling, A. Campello, and L. Liu, “On the  $L^1$  flatness factor of lattices,” 2018, poster presented at the recent results session of the *International Zurich Seminar on Communications*.
- [3] U. Maurer, “Secret key agreement by public discussion from common information,” *IEEE Trans. Inf. Theory*, vol. 39, no. 3, pp. 733–742, May 1993.
- [4] R. Ahlswede and I. Csiszár, “Common randomness in information theory and cryptography—Part I: Secret sharing,” *IEEE Trans. Inf. Theory*, vol. 39, no. 4, pp. 1121–1132, Jul. 1993.
- [5] T.-H. Chou, V. Y. F. Tan, and S. C. Draper, “The sender-excited secret key agreement model: Capacity, reliability, and secrecy exponents,” *IEEE Trans. Inf. Theory*, vol. 61, no. 1, pp. 609–627, 2015.
- [6] C. Chan, M. Mukherjee, N. Kashyap, and Q. Zhou, “On the optimality of secret key agreement via omniscience,” *IEEE Trans. Inf. Theory*, vol. 64, no. 4, pp. 2371–2389, 2018.
- [7] H. Tyagi and S. Watanabe, “Converses for secret key agreement and secure computing,” *IEEE Trans. Inf. Theory*, vol. 61, no. 9, pp. 4809–4827, 2015.
- [8] M. Iwamoto, K. Ohta, and J. Shikata, “Security formalizations and their relationships for encryption and key agreement in information-theoretic cryptography,” *IEEE Trans. Inf. Theory*, vol. 64, no. 1, pp. 654–685, 2018.
- [9] A. Gohari, O. Günlü, and G. Kramer, “Coding for positive rate in the source model key agreement problem,” *IEEE Trans. Inf. Theory*, vol. 66, no. 10, pp. 6303–6323, 2020.
- [10] C. T. Li and V. Anantharam, “One-shot variable-length secret key agreement approaching mutual information,” *IEEE Trans. Inf. Theory*, vol. 67, no. 8, pp. 5509–5525, 2021.
- [11] J. Liu, P. Cuff, and S. Verdú, “Key capacity for product sources with application to stationary Gaussian processes,” *IEEE Trans. Inf. Theory*, vol. 62, no. 2, pp. 984–1005, 2016.
- [12] A. Khisti, “Secret-key agreement over non-coherent block-fading channels with public discussion,” *IEEE Trans. Inf. Theory*, vol. 62, no. 12, pp. 7164–7178, 2016.
- [13] C. Chan, M. Mukherjee, N. Kashyap, and Q. Zhou, “Upper bounds via lamination on the constrained secrecy capacity of hypergraphical sources,” *IEEE Trans. Inf. Theory*, vol. 65, no. 8, pp. 5080–5093, 2019.
- [14] H. Tyagi and S. Watanabe, “Universal multiparty data exchange and secret key agreement,” *IEEE Trans. Inf. Theory*, vol. 63, no. 7, pp. 4057–4074, 2017.
- [15] C. Chan and L. Zheng, “Multiterminal secret key agreement,” *IEEE Trans. Inf. Theory*, vol. 60, no. 6, pp. 3379–3412, 2014.
- [16] A. A. Gohari and V. Anantharam, “Information-theoretic key agreement of multiple terminals—Part I,” *IEEE Trans. Inf. Theory*, vol. 56, no. 8, pp. 3973–3996, 2010.
- [17] K. P. Seshadreesan, M. Takeoka, and M. M. Wilde, “Bounds on entanglement distillation and secret key agreement for quantum broadcast channels,” *IEEE Trans. Inf. Theory*, vol. 62, no. 5, pp. 2849–2866, 2016.
- [18] G. Bassi, P. Piantanida, and S. Shamai Shitz, “The wiretap channel with generalized feedback: Secure communication and key generation,” *IEEE Trans. Inf. Theory*, vol. 65, no. 4, pp. 2213–2233, 2019.
- [19] M. Hayashi, H. Tyagi, and S. Watanabe, “Secret key agreement: General capacity and second-order asymptotics,” *IEEE Trans. Inf. Theory*, vol. 62, no. 7, pp. 3796–3810, 2016.
- [20] A. Poostindouz and R. Safavi-Naini, “Second-order asymptotics for one-way secret key agreement,” in *2021 IEEE International Symposium on Information Theory (ISIT)*, 2021, pp. 1254–1259.
- [21] J. Muramatsu and S. Miyake, “Construction of codes for the wiretap channel and the secret key agreement from correlated source outputs based on the hash property,” *IEEE Trans. Inf. Theory*, vol. 58, no. 2, pp. 671–692, 2012.
- [22] R. A. Chou, M. R. Bloch, and E. Abbe, “Polar coding for secret-key generation,” *IEEE Trans. Inf. Theory*, vol. 61, no. 11, pp. 6213–6237, 2015.
- [23] S. Watanabe and Y. Oohama, “Secret key agreement from correlated Gaussian sources by rate limited public communication,” *IEICE Trans. Fundamentals*, vol. E93-A, pp. 1976–1983, Nov. 2010.
- [24] A. Khisti, S. N. Diggavi, and G. W. Wornell, “Secret-key generation using correlated sources and channels,” *IEEE Trans. Inf. Theory*, vol. 58, no. 2, pp. 652–670, 2012.
- [25] S. Nitinawarat and P. Narayan, “Secret key generation for correlated Gaussian sources,” *IEEE Trans. Inf. Theory*, vol. 58, no. 6, pp. 3373–3391, June 2012.
- [26] M. Navascués, J. Bae, J. I. Cirac, M. Lewenstein, A. Sanpera, and A. Acín, “Quantum key distillation from Gaussian states by Gaussian operations,” *Physical review letters*, vol. 94, no. 1, p. 010502, 2005.

- [27] L. Lami, L. Mišta Jr, and G. Adesso, “Fundamental limitations to key distillation from Gaussian states with Gaussian operations,” *arXiv preprint arXiv:2010.15729*, 2020.
- [28] U. Maurer and S. Wolf, “Information-theoretic key agreement: From weak to strong secrecy for free,” *Lecture Notes in Computer Science*, pp. 351–368, 2000.
- [29] I. Csiszár, “Almost independence and secrecy capacity,” *Problems of Information Transmission*, vol. 32, pp. 40–47, 1996.
- [30] M. Bloch, “Channel intrinsic randomness,” in *Proc. Int. Symp. Inf. Theory (ISIT 2010)*, June 2010, pp. 2607–2611.
- [31] J. Muramatsu, H. Koga, and T. Mukouchi, “On the problem of generating mutually independent random sequences,” *IEICE Trans. Fundamentals*, vol. E86-A, no. 5, pp. 1275–1284, May 2003.
- [32] M. Hayashi, “Exponential decreasing rate of leaked information in universal random privacy amplification,” *IEEE Trans. Inf. Theory*, vol. 57, no. 6, pp. 3989–4001, Jun. 2011.
- [33] D. Micciancio and O. Regev, “Worst-case to average-case reductions based on Gaussian measures,” in *Proc. Ann. Symp. Found. Computer Science*, Rome, Italy, Oct. 2004, pp. 372–381.
- [34] J.-C. Belfiore, “Lattice codes for the compute-and-forward protocol: The flatness factor,” in *Proc. ITW 2011*, Paraty, Brazil, 2011.
- [35] C. Ling, L. Luzzi, J.-C. Belfiore, and D. Stehlé, “Semantically secure lattice codes for the Gaussian wiretap channel,” *IEEE Trans. Inf. Theory*, vol. 60, no. 10, pp. 6399–6416, Oct. 2014.
- [36] M. Hayashi and R. Matsumoto, “Secure multiplex coding with dependent and non-uniform multiple messages,” *IEEE Trans. Inf. Theory*, vol. 62, no. 5, pp. 2355–2409, 2016.
- [37] K.-M. Chung, D. Dadush, F.-H. Liu, and C. Peikert, “On the lattice smoothing parameter problem,” in *IEEE Conference on Computational Complexity*, 2013.
- [38] D. Dadush and O. Regev, “Towards strong reverse Minkowski-type inequalities for lattices,” in *2016 IEEE 57th Annual Symposium on Foundations of Computer Science (FOCS)*. IEEE, 2016, pp. 447–456.
- [39] L. Liu, Y. Yan, and C. Ling, “Achieving secrecy capacity of the Gaussian wiretap channel with polar lattices,” *IEEE Trans. Inf. Theory*, vol. 64, no. 3, pp. 1647–1665, 2018.
- [40] H. Mirghasemi and J. Belfiore, “The semantic secrecy rate of the lattice Gaussian coding for the Gaussian wiretap channel,” in *2014 IEEE Information Theory Workshop (ITW 2014)*, Nov 2014, pp. 112–116.
- [41] T. Debris-Alazard, L. Ducas, N. Resch, and J.-P. Tillich, “Smoothing codes and lattices: Systematic study and new bounds,” 2022. [Online]. Available: <https://arxiv.org/abs/2205.10552>
- [42] R. Zamir, S. Shamai, and U. Erez, “Nested linear/lattice codes for structured multiterminal binning,” *IEEE Trans. Inf. Theory*, vol. 48, pp. 1250–1276, Jun. 2002.
- [43] Z. Liu, S. Cheng, A. Liveris, and Z. Xiong, “Slepian-Wolf coded nested lattice quantization for Wyner-Ziv coding: High-rate performance analysis and code design,” *IEEE Trans. Inf. Theory*, vol. 52, no. 10, pp. 4358–4379, Oct. 2006.
- [44] C. Ling, S. Gao, and J.-C. Belfiore, “Wyner-Ziv coding based on multidimensional nested lattices,” *IEEE Trans. Commun.*, vol. 60, no. 5, pp. 1328–1335, May 2012.
- [45] C. Peikert, “An efficient and parallel Gaussian sampler for lattices,” in *Proc. CRYPTO*, vol. 6223. Springer-Verlag, 2010, pp. 80–97.
- [46] U. Erez and R. Zamir, “Achieving  $1/2 \log(1+\text{SNR})$  on the AWGN channel with lattice encoding and decoding,” *IEEE Trans. Inf. Theory*, vol. 50, no. 10, pp. 2293–2314, Oct. 2004.
- [47] A. Campello, D. Dadush, and C. Ling, “AWGN-Goodness is enough: Capacity-achieving lattice codes based on dithered probabilistic shaping,” *IEEE Trans. Inf. Theory*, vol. 65, no. 3, pp. 1961–1971, 2019.
- [48] Y. Dodis, R. Ostrovsky, L. Reyzin, and A. Smith, “Fuzzy extractors: How to generate strong keys from biometrics and other noisy data,” *SIAM Journal on Computing*, vol. 38, no. 1, pp. 97–139, 2008.
- [49] E. A. Verbitskiy, P. Tuyls, C. Obi, B. Schoenmakers, and B. Skoric, “Key extraction from general nondiscrete signals,” *IEEE Trans. Inf. Forensics Secur.*, vol. 5, no. 2, pp. 269–279, 2010.
- [50] J.-P. Linnartz and P. Tuyls, “New shielding functions to enhance privacy and prevent misuse of biometric templates,” in *Audio- and Video-Based Biometric Person Authentication*, J. Kittler and M. S. Nixon, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2003, pp. 393–402.
- [51] R. Zamir, *Lattice Coding for Signals and Networks: A Structured Coding Approach to Quantization, Modulation, and Multiuser Information Theory*. Cambridge University Press, 2014.
- [52] B. Nazer and M. Gastpar, “Compute-and-forward: Harnessing interference through structured codes,” *IEEE Trans. Inf. Theory*, vol. 57, pp. 6463–6486, Oct. 2011.
- [53] I. Csiszár and P. Narayan, “Common randomness and secret key generation with a helper,” *IEEE Trans. Inf. Theory*, vol. 46, no. 2, pp. 344–366, 2000.
- [54] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*. Cambridge University Press, 2011.
- [55] A. Bennatan, D. Burshtein, G. Caire, and S. Shamai, “Superposition coding for side-information channels,” *IEEE Transactions on Information Theory*, vol. 52, no. 5, pp. 1872–1889, 2006.

- [56] Y.-P. Wei, S.-C. Lin, S.-J. Lin, H.-J. Su, and H. V. Poor, "Residual-quantization based code design for compressing noisy sources with arbitrary decoder side information," *IEEE Transactions on Communications*, vol. 64, no. 4, pp. 1711–1725, 2016.
- [57] L. Liu, Y. Yan, C. Ling, and X. Wu, "Construction of capacity-achieving lattice codes: Polar lattices," *IEEE Trans. Commun.*, vol. 67, no. 2, pp. 915–928, 2019.
- [58] G. Forney, M. Trott, and S.-Y. Chung, "Sphere-bound-achieving coset codes and multilevel coset codes," *IEEE Trans. Inf. Theory*, vol. 46, no. 3, pp. 820–850, May 2000.
- [59] P. Delsarte and P. Piret, "Algebraic constructions of Shannon codes for regular channels," *IEEE Trans. Inf. Theory*, vol. 28, no. 4, pp. 593–599, 1982.
- [60] R. G. Gallager, *Information Theory and Reliable Communication*. Wiley, 1968.
- [61] H. Ghourchian, A. Gohari, and A. Amini, "Existence and continuity of differential entropy for a class of distributions," *IEEE Communications Letters*, vol. 21, no. 7, pp. 1469–1472, 2017.
- [62] F. J. Píera and P. Parada, "On convergence properties of Shannon entropy," *Problems of Information Transmission*, vol. 45, no. 2, pp. 75–94, 2009.
- [63] H. A. Loeliger, "Averaging bounds for lattices and linear codes," *IEEE Trans. Inf. Theory*, vol. 43, pp. 1767–1773, Nov. 1997.
- [64] C. Ling and J.-C. Belfiore, "Achieving AWGN channel capacity with lattice Gaussian coding," *IEEE Trans. Inf. Theory*, vol. 60, no. 10, pp. 5918–5929, Oct 2014.
- [65] O. Ordentlich and U. Erez, "A simple proof for the existence of "good" pairs of nested lattices," *IEEE Trans. Inf. Theory*, vol. 62, no. 8, pp. 4439–4453, Aug 2016.
- [66] 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.