

UNIVERSIDADE ESTADUAL DE CAMPINAS  
SISTEMA DE BIBLIOTECAS DA UNICAMP  
REPOSITÓRIO DA PRODUÇÃO CIENTÍFICA E INTELECTUAL DA UNICAMP

**Versão do arquivo anexado / Version of attached file:**

Versão do Editor / Published Version

**Mais informações no site da editora / Further information on publisher's website:**

<https://www.mdpi.com/2673-9984/5/1/19>

**DOI:** <https://doi.org/10.3390/psf2022005019>

**Direitos autorais / Publisher's copyright statement:**

©2022 by MDPI. All rights reserved.

DIRETORIA DE TRATAMENTO DA INFORMAÇÃO

Cidade Universitária Zeferino Vaz Barão Geraldo

CEP 13083-970 – Campinas SP

Fone: (19) 3521-6493

<http://www.repositorio.unicamp.br>

# Information Properties of a Random Variable Decomposition through Lattices <sup>†</sup>

Fábio C. C. Meneghetti , Henrique K. Miyamoto  and Sueli I. R. Costa 

Institute of Mathematics, Statistics and Scientific Computing (IMECC), University of Campinas, Campinas 13083-859, Brazil

\* Correspondence: fabiom@ime.unicamp.br

† Presented at the 41st International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering, Paris, France, 18–22 July 2022.

**Abstract:** A full-rank lattice in the Euclidean space is a discrete set formed by all integer linear combinations of a basis. Given a probability distribution on  $\mathbb{R}^n$ , two operations can be induced by considering the quotient of the space by such a lattice: wrapping and quantization. For a lattice  $\Lambda$ , and a fundamental domain  $\mathcal{D}$ , which tiles  $\mathbb{R}^n$  through  $\Lambda$ , the wrapped distribution over the quotient is obtained by summing the density over each coset, while the quantized distribution over the lattice is defined by integrating over each fundamental domain translation. These operations define wrapped and quantized random variables over  $\mathcal{D}$  and  $\Lambda$ , respectively, which sum up to the original random variable. We investigate information-theoretic properties of this decomposition, such as entropy, mutual information and the Fisher information matrix, and show that it naturally generalizes to the more abstract context of locally compact topological groups.

**Keywords:** Fisher information; information geometry; lattices; mutual information; quantization; topological groups; wrapped distributions



**Citation:** Meneghetti, F.C.C.; Miyamoto, H.K.; Costa, S.I.R. Information Properties of a Random Variable Decomposition through Lattices. *Phys. Sci. Forum* **2022**, *5*, 19. <https://doi.org/10.3390/psf2022005019>

Academic Editors: Frédéric Barbaresco, Ali Mohammad-Djafari, Frank Nielsen and Martino Trassinelli

Published: 5 December 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Lattices are discrete sets in  $\mathbb{R}^n$  formed by all integer linear combinations of a set of independent vectors, and have found different applications, such as in information theory and communications [1–3]. Given a probability distribution in  $\mathbb{R}^n$ , two operations can be induced by considering the quotient of the space by a lattice: wrapping and quantization.

The wrapped distribution over the quotient is obtained by summing the probability density over each coset. It is used to define parameters for lattice coset coding, particularly for AWGN and wiretap channels, such as the flatness factor, which is, up to a constant, the  $L^\infty$  distance from a wrapped probability distribution to a uniform one [4,5]. This factor is equivalent to the smoothing parameter, used in post-quantum lattice-based cryptography [6]. In the context of directional statistics, wrapping has been used as a standard way to construct distributions on a circle and on a torus [7].

The quantized distribution over the lattice can be defined by integrating over each fundamental domain translation, thus corresponding to the distribution of the fundamental domains after lattice-based quantization is applied. Lattice quantization has different uses in signal processing and coding: for instance, it can achieve the optimal rate-distortion trade-off and can be used for shaping in channel coding [2]. A special case of interest is when the distribution on the fundamental region is uniform, which amounts to high-resolution quantization or dithered quantization [8,9].

In this work, we relate these two operations by remarking that the random variables induced by wrapping and quantization sum up to the original one. We study information properties of this decomposition, both from classical information theory [10] and from information geometry [11], and provide some examples for the exponential and Gaussian

distributions. We also propose a generalization of these ideas to locally compact groups. Probability distributions on these groups have been studied in [12], and some information-theoretic properties have been investigated in [13–15]. In addition to probability measures, one can also define the notions of lattice and fundamental domains on them, thereby generalizing the Euclidean case. We show that wrapping and quantization are also well defined, and provide some illustrative examples.

## 2. Lattices, Wrapping and Quantization

### 2.1. Lattices and Fundamental Domains

A lattice  $\Lambda$  in  $\mathbb{R}^n$  is a discrete additive subgroup of  $\mathbb{R}^n$ , or, equivalently, the set  $\Lambda = \{\alpha_1 b_1 + \dots + \alpha_k b_k \mid \alpha_1, \dots, \alpha_k \in \mathbb{Z}\}$  formed by all integer linear combinations of a set of linearly independent vector  $\{b_1, \dots, b_k\} \subset \mathbb{R}^n$ , called a *basis* of  $\Lambda$ . A matrix  $B$  whose column vectors form a basis is called a *generator matrix* of  $\Lambda$ , and we have  $\Lambda = B\mathbb{Z}^k$ . The lattice dimension is  $k$ , and, if  $k = n$ , the lattice is said to be *full-rank*; we henceforth consider full-rank lattices. A lattice  $\Lambda$  defines an equivalence relation in  $\mathbb{R}^n$ :  $x \sim y \iff x - y \in \Lambda$ . The associated equivalence classes are denoted by  $\bar{x}$  or  $x + \Lambda$ . The set of all equivalence classes is the lattice quotient  $\mathbb{R}^n/\Lambda$ , and we denote the standard projection  $\pi: \mathbb{R}^n \rightarrow \mathbb{R}^n/\Lambda$ ,  $\pi(x) = \bar{x}$ .

Let  $\mathcal{D}$  be a Lebesgue-measurable set of  $\mathbb{R}^n$  and  $\Lambda$  a lattice. We say that  $\mathcal{D}$  is a *fundamental domain* or a *fundamental region* of  $\Lambda$ , or that  $\mathcal{D}$  *tiles*  $\mathbb{R}^n$  by  $\Lambda$ , if (1)  $\bigcup_{\lambda \in \Lambda} (\lambda + \mathcal{D}) = \mathbb{R}^n$ , and (2)  $(\lambda + \mathcal{D}) \cap (\tilde{\lambda} + \mathcal{D}) = \emptyset$ , for all  $\lambda \neq \tilde{\lambda}$  in  $\Lambda$  (it is often only asked that this intersection has Lebesgue measure zero, but we require it to be empty). Given a fundamental domain  $\mathcal{D}$ , each coset  $\bar{x} \in \mathbb{R}^n/\Lambda$  has a unique representative in  $\mathcal{D}$ , i.e., the measurable map  $\pi|_{\mathcal{D}}: \mathcal{D} \rightarrow \mathbb{R}^n/\Lambda$  is a bijection. This fact suggests using a fundamental domain to represent the quotient. Each fundamental domain contains exactly one lattice point, which may be chosen as the origin. An example of a fundamental domain is the *fundamental parallelootope* with respect to a basis  $\{b_1, \dots, b_n\}$ , namely  $\mathcal{P}(\Lambda) := \{x = \alpha_1 b_1 + \dots + \alpha_n b_n \mid \alpha_1, \dots, \alpha_n \in [0, 1[ \}$ . Another one is the Voronoi region  $\mathcal{V}(\Lambda)$  of the origin, given by the points that are closer to the origin than to any other lattice point, with an appropriate choice for ties. It is a well-known fact that every fundamental domain has the same volume, denoted by  $\text{covol } \Lambda := \text{vol } \mathcal{D} = |\det B|$ , for any generator matrix  $B$  of  $\Lambda$ .

### 2.2. Wrapping and Quantization

Consider  $\mathbb{R}^n$  with the Lebesgue measure  $\mu$ , and  $P$  a probability measure such that  $P \ll \mu$ . Then the probability density function (pdf) of  $P$  is  $p = \frac{dP}{d\mu}$ , the Radon–Nikodym derivative. For fixed full-rank lattice  $\Lambda$  and fundamental domain  $\mathcal{D}$ , the *wrapping* of  $P$  by  $\Lambda$  is the distribution  $P_\pi := \pi_* P$  on  $\mathbb{R}^n/\Lambda$ , given by  $P_\pi(A) = P(\pi^{-1}A)$ . For simplicity, we identify  $\mathbb{R}^n/\Lambda$  with  $\mathcal{D}$  to regard  $P_\pi$  as a distribution over  $\mathcal{D}$ , and then we have  $\pi: \mathbb{R}^n \rightarrow \mathcal{D}$  given by  $(y + \lambda) \mapsto y$ , for all  $y \in \mathcal{D}, \lambda \in \Lambda$ . Using this identification, the wrapping has density  $p_\pi = \frac{dP_\pi}{d\mu}$  given by

$$p_\pi(y) = \sum_{\lambda \in \Lambda} p(y + \lambda). \quad (1)$$

A construction that is, in some sense, dual to wrapping is quantization. Note that each fundamental domain  $\mathcal{D}$  partitions the space as  $\mathbb{R}^n = \bigsqcup_{\lambda \in \Lambda} (\lambda + \mathcal{D})$ . The quantization function is the measurable map  $\mathcal{Q}: \mathbb{R}^n \rightarrow \Lambda$ , given by  $(y + \lambda) \mapsto \lambda$ , for  $y \in \mathcal{D}$  and  $\lambda \in \Lambda$ . The *quantized probability distribution* of  $P$  on the discrete set  $\Lambda$  is  $P_\mathcal{Q} := \mathcal{Q}_* P$ , given by  $P_\mathcal{Q}(A) := P(\mathcal{Q}^{-1}A)$ . The probability mass function of the quantized distribution is then

$$p_\mathcal{Q}(\lambda) = \int_{\mathcal{D}} p(y + \lambda) dy. \quad (2)$$

Letting  $X$  be a vector random variable in  $\mathbb{R}^n$  with distribution  $p$ , we define  $X_\pi := \pi(X)$  and  $X_\mathcal{Q} := \mathcal{Q}(X)$  the wrapped and quantized random variables, respectively. By definition, they are distributed according to  $p_\pi$  and  $p_\mathcal{Q}$ . Interestingly, they sum up to the original one:

$$X = X_\pi + X_\mathcal{Q}, \quad (3)$$

since  $\pi + \mathcal{Q} = \text{id}_{\mathbb{R}^n}$ . Note also that  $X_\pi + X_{\mathcal{Q}}$  has the same distribution as  $(X_\pi, X_{\mathcal{Q}})$ , by the bimeasurable bijection  $y + \lambda \mapsto (y, \lambda)$ . These factors, however, are not independent, since, in general,  $p(y + \lambda) \neq p_\pi(y)p_{\mathcal{Q}}(\lambda)$ . The difference between  $p(x)$  and  $(p_\pi \otimes p_{\mathcal{Q}})(x) := p_\pi(\pi(x))p_{\mathcal{Q}}(\mathcal{Q}(x))$  shall be illustrated in the following examples. Note that the expression for the quantized distribution depends on the choice of fundamental domain, while the wrapped distribution does not, up to a lattice translation.

We say a random variable  $X$  over  $[0, \infty)$  is *memoryless* if  $\bar{C}(t) = \bar{C}(t+s)/\bar{C}(s)$  for all  $t, s$ , where  $\bar{C}(t) := P[X > t]$  is the tail distribution function. In particular, a memoryless distribution satisfies  $\bar{C}(y + \lambda) = \bar{C}(y)\bar{C}(\lambda)$  for all  $y \in \mathcal{D}$ ,  $\lambda \in \Lambda$ , which implies  $p = p_\pi \otimes p_{\mathcal{Q}}$ . The converse, however, is not true; for example, independence holds whenever  $p$  is constant on each region  $\lambda + \mathcal{D}$ , for  $\lambda \in \Lambda$ .

**Example 1.** The exponential distribution, parametrized by  $\nu > 0$ , is defined as

$$p(x) = \nu e^{-\nu x} \mathbb{1}_{[0, +\infty[}(x),$$

where  $\mathbb{1}_A(x)$  takes value 1 if  $x \in A$ , and 0 otherwise. Choosing the lattice  $\Lambda = \alpha\mathbb{Z}$ ,  $\alpha \in \mathbb{R}_+$ , and the fundamental domain  $\mathcal{D} = [0, \alpha[$ , one can write closed-form expressions for the wrapped and quantized distributions:

$$p_\pi(y) = \frac{\nu e^{-\nu y}}{1 - e^{-\nu \alpha}}, \quad y \in \mathcal{D} \quad \text{and} \quad p_{\mathcal{Q}}(\lambda) = e^{-\nu \lambda} (1 - e^{-\nu \alpha}), \quad \lambda \in \Lambda \cap \mathbb{R}_+. \quad (4)$$

Note that, in this special case,  $p = p_\pi \otimes p_{\mathcal{Q}}$ , as a consequence of memorylessness. The wrapped distribution with  $\alpha = 2\pi$ , which amounts to a distribution on the unitary circle, is well studied in [16].

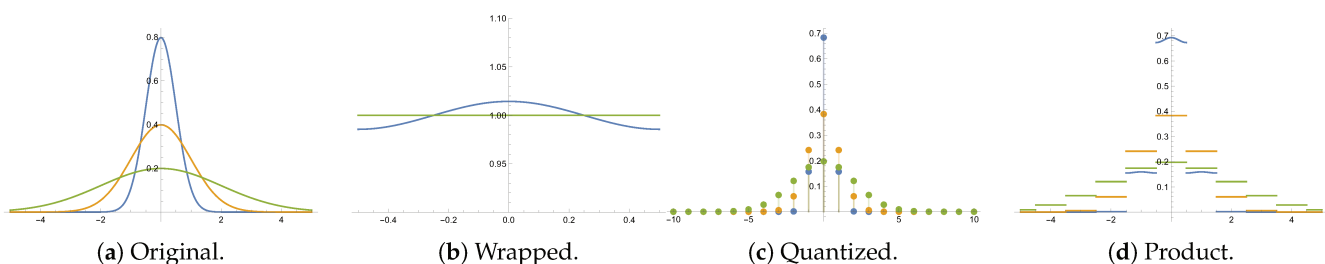
**Example 2.** Consider the univariate Gaussian distribution

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

and the lattice  $\Lambda = \alpha\mathbb{Z}$ , with fundamental domain  $\mathcal{D} = [-\frac{\alpha}{2}, \frac{\alpha}{2}[$ ,  $\alpha \in \mathbb{R}_+$ . The wrapped and quantized distributions are given respectively by

$$p_\pi(y) = \frac{1}{\sqrt{2\pi\sigma^2}} \sum_{i \in \mathbb{Z}} e^{-\frac{(y-\mu+\alpha i)^2}{2\sigma^2}}, \quad y \in \mathcal{D} \quad \text{and} \quad p_{\mathcal{Q}}(\lambda) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{\lambda-\frac{\alpha}{2}}^{\lambda+\frac{\alpha}{2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx, \quad \lambda \in \Lambda.$$

The value  $\alpha = 2\pi$  for the wrapped distribution on a unitary circle is usually considered in directional statistics [7]. Figure 1 illustrates the original, wrapped, quantized and product distributions for different zero-mean Gaussian distributions. As can be seen in the figure, in this case,  $p(x) \neq p_\pi(y)p_{\mathcal{Q}}(\lambda)$ .



**Figure 1.** Example of zero-mean Gaussian distributions and their corresponding wrapped, quantized and product distributions, with  $\Lambda = \mathbb{Z}$  and  $\mathcal{D} = [-\frac{1}{2}, \frac{1}{2}[$  for different variances:  $\sigma^2 = 0.25$  (blue),  $\sigma^2 = 1$  (orange),  $\sigma^2 = 4$  (green).

A straightforward consequence of the decomposition (3) is

1.  $E[X] = E[X_\pi] + E[X_Q]$ ;
2.  $\text{Var}[X] = \text{Var}[X_\pi] + \text{Var}[X_Q] + \text{Cov}[X_\pi, X_Q] + \text{Cov}[X_Q, X_\pi]$ ,

where  $E[\cdot]$ ,  $\text{Var}[\cdot]$  and  $\text{Cov}[\cdot, \cdot]$  denote, respectively, the expectation, the variance and the cross-covariance operators.

We note that different types of discretization have also been studied, other than integrating over a fundamental domain [17]. For instance, in [4,18,19], the discretized distribution is defined by restricting the original pdf  $p(x)$  to the lattice  $\Lambda$ , and then normalizing:

$$D_{\Lambda,c}(\lambda) := \frac{p(c+\lambda)}{\sum_{\tilde{\lambda} \in \Lambda} p(c+\tilde{\lambda})}, \quad (5)$$

for a fixed  $c \in \mathcal{D}$ . This discretization is nothing other than the conditional distribution of  $X_Q$  given that  $X_\pi = c$ , expressed as  $p_{Q|\pi}(\lambda|c) = p(c+\lambda)/p_\pi(c)$ . Moreover, when  $p = p_\pi \otimes p_Q$ , such as in the exponential distribution, cf. example 1, then  $D_{\Lambda,c}(\lambda) = p_Q(\lambda)$ .

### 3. Information Properties

#### 3.1. Information-Theoretic Measures

Let us consider a random variable  $X$  with distribution  $p$  and the induced wrapped and quantized ones, respectively,  $X_\pi \sim p_\pi$  and  $X_Q \sim p_Q$ . The mutual information between  $X_\pi$  and  $X_Q$  is defined as the Kullback–Leibler divergence  $I(X_\pi; X_Q) := D_{\text{KL}}(p \| p_\pi \otimes p_Q)$ , and is a measure of how non-independent the marginal distributions  $p_\pi$  and  $p_Q$  are [10]. Using the theorem of change of variables, we have

$$\begin{aligned} I(X_\pi; X_Q) &= E_X \left[ \log \frac{p(X)}{p_\pi \otimes p_Q(X)} \right] \\ &= E_X[\log p(X)] - E_X[\log p_\pi(X_\pi)] - E_X[\log p_Q(X_Q)] \\ &= E_X[\log p(X)] - E_{X_\pi}[\log p_\pi(X_\pi)] - E_{X_Q}[\log p_Q(X_Q)] \\ &= h(X_\pi) + H(X_Q) - h(X). \end{aligned} \quad (6)$$

Note that, from this decomposition, we have  $h(X) \leq h(X_\pi) + H(X_Q)$ .

**Proposition 1.** Let  $X$  be a real random variable, and  $X_\pi$  and  $X_Q$  the respective wrapped and quantized random variables, using the lattice  $\alpha\mathbb{Z}$ . Denote  $\mu_Q := E[X_Q]$  and  $\sigma_Q^2 := \text{Var}[X_Q]$ . If  $X$  has support  $[0, \infty)$ , then the mutual information  $I(X_\pi; X_Q)$  between  $X_\pi$  and  $X_Q$  is upper-bounded by

$$I(X_\pi; X_Q) < \log(e(\mu_Q + \alpha/2)) - h(X). \quad (7)$$

If  $X$  has support  $\mathbb{R}$ , then  $I(X_\pi; X_Q)$  is upper-bounded by

$$I(X_\pi; X_Q) < \frac{1}{2} \log(2\pi e \sigma_Q^2) + \frac{2 \log e}{\exp(2\pi^2 \alpha^{-2} \sigma_Q^2) - 1} - h(X). \quad (8)$$

**Proof.** First,  $h(X_\pi) \leq \log \alpha$ , since the uniform distribution maximizes entropy on a bounded support. Then, note that the mean and variance of the integer-valued random variable  $\alpha^{-1}X_Q$  are  $\alpha^{-1}\mu_Q$  and  $\alpha^{-2}\sigma_Q^2$ , respectively. For (7), use that, for positive integer random variables,  $H(X_Q) < \log(e(\mu_Q/\alpha + 1/2))$ , as in [20] (Theorem 8); for (8), the upper-bound for integer-valued random variables from [20] (Theorem 10) gives us  $H(X_Q) < \frac{1}{2} \log(2\pi e \alpha^{-2} \sigma_Q^2) + \frac{2 \log e}{\exp(2\pi^2 \alpha^{-2} \sigma_Q^2) - 1}$ . Replacing the corresponding inequalities in (6) yields the desired results.  $\square$

The following lemma can be found in [2] (Appendix 3).

**Lemma 1.**  $h(X_\pi) \leq h(X)$ .

**Proof.**  $h(X) = h(X_\pi) + H(X_Q|X_\pi)$ , and  $H(X_Q|X_\pi) \geq 0$ , since it is a discrete entropy.  $\square$

**Proposition 2.** Let  $\Lambda_\alpha := \alpha\Lambda$ ,  $\alpha > 0$ , be a family of lattices, with fundamental domains  $\mathcal{D}_\alpha := \alpha\mathcal{D}$ .

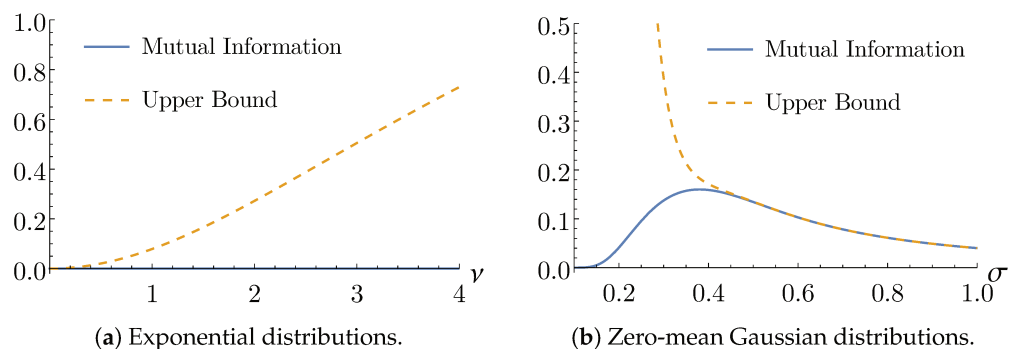
1. If  $\mathcal{D}$  is connected, and  $p$  is continuous and Riemann-integrable, then  $\lim_{\alpha \rightarrow 0} I(X_\pi; X_Q) = 0$ .
2. If 0 is an interior point of  $\mathcal{D}$ , then  $\lim_{\alpha \rightarrow +\infty} I(X_\pi; X_Q) = 0$ .

**Proof.** For  $\alpha \rightarrow 0$ , the proof is an adaptation of [10] (Theorem 8.3.1). Since  $\mathcal{D}$  is connected and  $p$  is continuous, we can use the mean value theorem: for every  $\lambda \in \Lambda$  there exists an  $x_{\lambda,\alpha} \in (\lambda + \mathcal{D}_\alpha)$  such that  $p(x_{\lambda,\alpha}) \text{vol } \mathcal{D}_\alpha = p_Q(\lambda)$ . Therefore, we can write  $H(X_Q) = -\sum_{\lambda \in \Lambda_\alpha} p(x_{\lambda,\alpha}) \log(p(x_{\lambda,\alpha})) \text{vol } \mathcal{D}_\alpha - \log(\text{vol } \mathcal{D}_\alpha)$ , using that  $\sum_{\lambda \in \Lambda_\alpha} p(x_{\lambda,\alpha}) \text{vol } \mathcal{D}_\alpha = 1$ . The first term is an  $n$ -dimensional Riemann sum, and converges to  $h(X)$  when  $\alpha \rightarrow 0$ , while the second term becomes arbitrarily small. Therefore,  $0 \leq I(X_\pi; X_Q) \leq H(X_Q) + \log(\text{vol } \mathcal{D}_\alpha) - h(X) \rightarrow 0$ , so  $I(X_\pi; X_Q) \rightarrow 0$ .

For  $\alpha \rightarrow +\infty$ , note that, from Lemma 1,  $I(X_\pi; X_Q) \leq H(X_Q)$ . However, by choosing  $\alpha$  sufficiently large, we can make  $p_Q(0) = \int_{\mathcal{D}_\alpha} p(x) dx$  arbitrarily close to 1, since 0 is in the interior of  $\mathcal{D}_\alpha$ . Therefore,  $H(X_Q)$  can be made arbitrarily small.  $\square$

**Example 3.** In the case of the exponential distributions, as in Example 1, the distributions of  $X_\pi$  and  $X_Q$  are independent, i.e.,  $p = p_\pi \otimes p_Q$ , therefore  $I(X_\pi; X_Q) = 0$ . The mutual information and the corresponding upper bound (7) are plotted in Figure 2a, as function of the parameter  $v$ .

**Example 4.** In the case of the univariate zero-mean Gaussian distributions, as in Example 2, one can use (6) to numerically compute the mutual information  $I(X_\pi; X_Q)$ , as a function of the standard deviation  $\sigma$ , and compare it with the upper bound (8) (Figure 2b). Interestingly,  $I(X_\pi; X_Q)$  vanishes as  $\sigma \rightarrow 0$  or  $\sigma \rightarrow +\infty$ , which is equivalent to choosing a lattice  $\Lambda = \alpha\mathbb{Z}$  with  $\alpha \rightarrow 0$  or  $\alpha \rightarrow +\infty$ , cf. Proposition 2. The mutual information attains a maximum in  $\sigma \approx 0.38$ , showing this is the value for which  $X_\pi$  and  $X_Q$  are the least independent.



**Figure 2.** Mutual information  $I(X_\pi; X_Q)$  and its upper bound.

### 3.2. Fisher Information

Let  $M = \{p_\theta : \theta \in \Theta\}$  be a family of probability densities  $p_\theta: \mathbb{R}^n \rightarrow \mathbb{R}_+$  smoothly parametrized by  $\theta$  in an open set  $\Theta \subset \mathbb{R}^d$ . The *Fisher information matrix* is defined as the positive semi-definite matrix  $G(\theta)$  with coefficients  $g_{ij}(\theta) = E_{p_\theta}[\partial_i \ell_\theta \partial_j \ell_\theta]$ , where  $\ell_\theta(x) := \log p_\theta(x)$ . When  $M$  is a manifold satisfying certain regularity conditions [11], and  $G$  is positive definite, it becomes a Riemannian manifold with the metric given by  $g_{ij}(\theta)$ , called a *statistical manifold*. Let  $\leq$  denote the Loewner partial order for matrices, given by  $A \leq B$  if, and only if,  $B - A$  is positive semi-definite. The following results justify the name *information matrix* given to this quantity.

**Proposition 3** ([11,21]). Let  $X$  be a random variable distributed according to a distribution parametrized by  $\theta$ , and  $G(\theta)$  its information matrix. The following hold.



1. **Monotonicity:** if  $F: \mathcal{X} \rightarrow \mathcal{Y}$  is a measurable function (i.e., a statistic) and  $G_F(\theta)$  is the information matrix of  $F(X)$ , then  $G_F(\theta) \leq G(\theta)$ , with equality if, and only if,  $F$  is a sufficient statistic for  $\theta$ .
2. **Additivity:** if  $X, Y$  are independent random variables, then the joint information matrix satisfies  $G_{(X,Y)}(\theta) = G_X(\theta) + G_Y(\theta)$ .

Let  $X$  be a random variable on  $\mathbb{R}^n$ , and  $X_\pi$  and  $X_Q$  its wrapped and quantized factors, respectively. We denote their respective Fisher information matrices by  $G(\theta)$ ,  $G_\pi(\theta)$  and  $G_Q(\theta)$ . By additivity, the Fisher information of  $p_\pi \otimes p_Q$  is  $\tilde{G}(\theta) := G_\pi(\theta) + G_Q(\theta)$ , and, by monotonicity, we have both  $G_\pi(\theta) \leq G(\theta)$  and  $G_Q(\theta) \leq G(\theta)$ . It follows immediately that

$$\frac{\tilde{G}(\theta)}{2} = \frac{G_\pi(\theta) + G_Q(\theta)}{2} \leq G(\theta). \quad (9)$$

**Example 5.** In the family of exponential distributions, as in Example 1, the independence of  $X_\pi$  and  $X_Q$  implies that the Fisher information matrix is additive. Indeed, for  $\Lambda = \alpha\mathbb{Z}$ :

$$G(v) = \frac{1}{v^2}, \quad G_\pi(v) = \frac{1}{v^2} + \frac{\alpha^2}{2(1 - \cosh(\alpha v))}, \quad \text{and} \quad G_Q(v) = \frac{\alpha^2}{2(\cosh(\alpha v) - 1)}.$$

#### 4. A Generalization to Topological Groups

A topological group is a topological space  $(G, \tau_G)$  that is also a group with respect to some operation  $\cdot$  called *product*, and such that the inverse  $g^{-1}$  and product  $g \cdot h$  are continuous. As additional requisites, we ask  $G$  to be locally compact, Hausdorff and second-countable (i.e., has a countable basis) [22]. Let  $\mathcal{B}_G$  be the Borel  $\sigma$ -algebra of  $G$ . Haar's theorem says there is a unique (up to a constant) Radon measure on  $G$  that is invariant by left translations—we will suppose a fixed normalization, and denote both the measure and integration with respect to it by  $dg$ . The group  $G$  is said to be *unimodular* if  $dg$  is also invariant by right translations. Since  $G$  is  $\sigma$ -compact, the Haar measure is  $\sigma$ -finite [12].

Let  $\Gamma$  be a discrete subgroup of  $G$ , which is necessarily closed, since  $G$  is Hausdorff, and countable, since  $G$  is second-countable. Let us also consider the quotient space of left cosets  $G/\Gamma = \{\bar{g} = g\Gamma \mid g \in G\}$ , which has a natural projection  $\pi: G \rightarrow G/\Gamma$ , given by  $\pi(g) = \bar{g}$ . We call  $\Gamma$  a *lattice* if the induced Haar measure on  $G/\Gamma$  is finite and bi-invariant. A particular case is when the quotient  $G/\Gamma$  is compact; then  $\Gamma$  is said to be a *uniform lattice*. A *cross-section* is defined as a set  $\mathcal{D} \subset G$  of representatives of  $G/\Gamma$  such that all cosets are uniquely represented. A *fundamental domain* is a measurable cross-section. It can be shown that  $\Gamma$  is a lattice if, and only if, it admits a fundamental domain. Furthermore, every fundamental domain has the same measure [23,24].

Let  $P$  be a probability measure on the space  $(G, \mathcal{B}_G)$  that is absolutely continuous with respect to the Haar measure  $dg$ . By the Radon–Nikodym theorem, we can define a density function  $p = \frac{dP}{dg} \in L^1(G)$ , such that  $p \geq 0$  and  $P(A) = \int_A p(g) dg$ , for all  $A \in \mathcal{B}_G$ . The original measure can be represented as  $P = p dg$ , and we consider the family of all such densities

$$\mathcal{P}(G) = \left\{ p \in L^1(G) \mid p \geq 0 \text{ dg-a.e., } \int p dg = 1 \right\}.$$

Probability distributions on locally compact groups have been studied in [12], and some information-theoretic properties have been investigated in [13–15]. The result that allows us to consider wrapped distributions in this context is the *Weil formula*, taken as a particular case of [24] (Theorem 3.4.6):

**Theorem 1.** For any  $f \in L^1(G)$ , the wrapping  $f_\pi \in L^1(G/\Gamma)$ ,  $f_\pi(\bar{g}) := \sum_{\lambda \in \Gamma} f(g\lambda)$  is well defined  $d\bar{g}$ -almost everywhere, belongs to  $L^1(G/\Gamma)$ , and

$$\int_G f(g) dg = \int_{G/\Gamma} \sum_{\lambda \in \Gamma} f(g\lambda) d\bar{g}. \quad (10)$$

As a consequence, for every probability density  $p \in \mathcal{P}(G)$ , we can consider its wrapping  $p_\pi(\bar{g}) = \sum_{\lambda \in \Gamma} p(g\lambda)$ , which is  $L^1(G/\Gamma)$ , non-negative and is also a probability density:  $\int_{G/\Gamma} p_\pi d\bar{g} = 1$ . The associated probability measure over  $(G/\Gamma, \mathcal{B}_{G/\Gamma})$  is  $P_\pi = p_\pi d\bar{g}$ . This notation, suggesting  $P_\pi$  as the push-forward measure by  $\pi$ , is not a coincidence, since, from Theorem 1,

$$\begin{aligned}\pi_* P(A) &= \int_G \mathbb{1}_A(\pi(g)) p(g) dg \\ &= \int_{G/\Gamma} \sum_{\lambda \in \Gamma} \mathbb{1}_A(\pi(g)) p(g\lambda) d\bar{g} \\ &= \int_{G/\Gamma} \mathbb{1}_A(\bar{g}) p_\pi(\bar{g}) d\bar{g} = P_\pi(A).\end{aligned}$$

Analogously, given a fundamental domain  $\mathcal{D}$ , it is possible to define a quantization map  $\mathcal{Q}: G \rightarrow \Lambda$  by  $\mathcal{Q}(g\lambda) = \lambda$ , for every  $g \in \mathcal{D}, \lambda \in \Gamma$ , which is unique since  $G = \bigsqcup_{g \in \mathcal{D}} g\Gamma$ . The quantized probability distribution is the discrete probability measure  $P_\mathcal{Q}$  over  $\Lambda$ , defined by the mass function  $p_\mathcal{Q}(\lambda) = \int_{\mathcal{D}} p(g\lambda) dg$ , or as the push-forward measure  $\mathcal{Q}_* P$ .

If  $X$  is distributed according to  $p$ , and  $X_\pi = \pi(X) \sim p_\pi$ ,  $X_\mathcal{Q} = \mathcal{Q}(X) \sim p_\mathcal{Q}$ , then  $X = X_\pi \cdot X_\mathcal{Q}$ , again, as a consequence of  $g \mapsto (\pi(g), \mathcal{Q}(g))$  being a measurable bijection whose inverse is the product  $\pi(g) \cdot \mathcal{Q}(g)$ . Despite being an abstract definition, this framework expands the scope of the previous approach, cf. examples below. In the following, let  $\Lambda \subset \mathbb{R}^n$  be a full-rank lattice, and  $\Lambda_s \subset \Lambda$  be a full-rank sublattice, as defined in Section 2.

**Example 6.** Let  $G = \mathbb{R}^n$  and  $\Gamma = \Lambda$ . This recovers the approach from Section 2 as a particular case.

**Example 7.** Let  $G = \Lambda$  and  $\Gamma = \Lambda_s$ . A fundamental domain is a choice  $\mathcal{D} = \{d_1, \dots, d_k\}$  of  $k = |\Lambda/\Lambda_s|$  points, where each point corresponds to a coset  $\bar{\lambda} = (\lambda + \Lambda_s) \in \Lambda/\Lambda_s$ . Of particular interest are Voronoi constellations [25,26] where the coset leaders are selected, with some choice made for ties. Since  $\Lambda$  is discrete, the Haar measure is the counting measure  $\mu(A) = |A|$ , and  $p: \Lambda \rightarrow [0, 1]$ . The wrapped and quantized distributions are  $p_\pi(\bar{\lambda}) = \sum_{\lambda_s \in \Lambda_s} p(\lambda + \lambda_s)$ , and  $p_\mathcal{Q}(\lambda_s) = \sum_{i=1}^k p(d_i + \lambda_s)$ .

**Example 8.** Let  $G = \mathbb{R}^n/\Lambda_s$  (a torus) and  $\Gamma = \pi_s(\Lambda)$  (the projection of  $\Lambda$  to  $G$ ). Then  $\pi_s(\Lambda)$  consists of a finite family of cosets  $\bar{\lambda}_1, \dots, \bar{\lambda}_k$ , for  $k = |\Lambda/\Lambda_s|$ , and a choice of fundamental domain  $\bar{\mathcal{D}}$  is the projection of a fundamental domain  $\mathcal{D}$  of  $\Lambda$ . There are some standard choices for the distribution on  $G$ , such as a wrapping from the Euclidean space and the bivariate von Mises distribution [7] (Section 11.4). Then,  $p_\pi(\bar{x}) = \sum_{i=1}^k p(\bar{x} + \bar{\lambda}_i)$  and  $p_\mathcal{Q}(\bar{\lambda}_i) = \int_{\bar{\mathcal{D}}} p(\bar{x} + \bar{\lambda}_k) d\bar{x}$ , and, in the particular case where  $p(\bar{x}) = \sum_{\lambda_s \in \Lambda_s} p(x + \lambda_s)$  is a  $\Lambda_s$ -wrapped distribution, they become  $p_\pi(\bar{x}) = \sum_{i=1}^k \sum_{\lambda_s \in \Lambda_s} p(x + \lambda_s + \lambda_i)$  and  $p_\mathcal{Q}(\bar{\lambda}_i) = \sum_{\lambda_s \in \Lambda_s} \int_{\mathcal{D}} p(x + \lambda_s + \lambda_i) dx$ .

**Example 9.** Let  $G = \mathbb{F}_q^n$  (a finite field) or  $G = \mathbb{Z}_q^n$ , and  $\Gamma = \mathcal{C}$  (any linear block code). A fundamental domain can be a finite set of points that tiles the space by  $\mathcal{C}$ . The distributions then become finite sums, such as in Example 7.

**Example 10.** Let  $G = \text{SL}(n, \mathbb{R})$ , the Lie group of square matrices with determinant 1, and  $\Gamma = \text{SL}(n, \mathbb{Z})$  (the subgroup of integer matrices). This is, in fact, a lattice, since for  $n = 2$ ,  $\text{vol}(G/\Gamma) = \sqrt{2}\zeta(2)$  where  $\zeta$  is the Riemann zeta function, and for  $n > 2$  the finite covolume is calculated in [27], where descriptions of fundamental domains are also given.

## 5. Conclusions

In this work, we have studied the decomposition of a random variable through lattices into its wrapping and quantization terms. Generalization of examples and of Proposition 1 to higher dimensions constitutes a work in progress. We have also proposed a generalization of this decomposition to topological groups; in particular, this allows



one to study information theory on such abstract spaces, which is another perspective for future work.

**Author Contributions:** Conceptualization, all authors; investigation, F.C.C.M. and H.K.M.; writing—original draft preparation, F.C.C.M. and H.K.M.; writing—review and editing, all authors; supervision, S.I.R.C. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was partly supported by Brazilian National Council for Scientific and Technological Development (CNPq) grants 141407/2020-4 and 314441/2021-2, and by São Paulo Research Foundation (FAPESP) grant 2021/04516-8.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

**Acknowledgments:** The authors are grateful to Max Costa for fruitful discussions, and to the reviewer for the relevant suggestions.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Conway, J.H.; Sloane, N.J.A. *Sphere Packings, Lattices and Groups*; Springer: New York, NY, USA, 1999.
2. Zamir, R. *Lattice Coding for Signals and Networks: A Structured Coding Approach to Quantization, Modulation and Multiuser Information Theory*; Cambridge University Press: Cambridge, UK, 2014.
3. Costa, S.I.R.; Oggier, F.; Campello, A.; Belfiore, J.C.; Viterbo, E. *Lattices Applied to Coding for Reliable and Secure Communications*; Springer: Cham, Switzerland, 2017.
4. Ling, C.; Belfiore, J.C. Achieving AWGN channel capacity with lattice Gaussian coding. *IEEE Trans. Inf. Theory* **2014**, *60*, 5918–5929. [[CrossRef](#)]
5. Damir, M.T.; Karrila, A.; Amorós, L.; Gnille, O.W.; Karpuk, D.; Hollanti, C. Well-rounded lattices: Towards optimal coset codes for Gaussian and fading wiretap channels. *IEEE Trans. Inf. Theory* **2021**, *67*, 3645–3663. [[CrossRef](#)]
6. Chung, K.M.; Dadush, D.; Liu, F.H.; Peikert, C. On the lattice smoothing parameter problem. In Proceedings of the 2013 IEEE Conference on Computational Complexity, Stanford, CA, USA, 5–7 June 2013; pp. 230–241.
7. Mardia, K.V.; Jupp, P.E. *Directional Statistics*; Wiley: New York, NY, USA, 2000.
8. Zamir, R.; Feder, M. On lattice quantization noise. *IEEE Trans. Inf. Theory* **1996**, *42*, 1152–1159. [[CrossRef](#)]
9. Ling, C.; Gan, L. Lattice quantization noise revisited. In Proceedings of the 2013 IEEE Information Theory Workshop (ITW), Sevilla, Spain, 9–13 September 2013; pp. 1–5.
10. Cover, T.M.; Thomas, J.A. *Elements of Information Theory*, 2nd ed.; Wiley: Hoboken, NJ, USA, 2006.
11. Amari, S.; Nagaoka, H. *Methods of Information Geometry*; American Mathematical Society: Providence, RI, USA, 2000.
12. Heyer, H. *Probability Measures on Locally Compact Groups*; Springer: Berlin/Heidelberg, Germany, 1977.
13. Chirikjian, G.S. *Stochastic Models, Information Theory, and Lie Groups*; Birkhäuser: Boston, MA, USA, 2009.
14. Johnson, O.; Suhov, Y. Entropy and convergence on compact groups. *J. Theor. Probab.* **2000**, *13*, 843–857. [[CrossRef](#)]
15. Chirikjian, G.S. Information-theoretic inequalities on unimodular Lie groups. *J. Geom. Mech.* **2010**, *2*, 119–158. [[CrossRef](#)] [[PubMed](#)]
16. Jammalamadaka, S.R.; Kozubowski, T.J. New families of wrapped distributions for modeling skew circular data. *Commun. Stat.—Theory Methods* **2004**, *33*, 2059–2074. [[CrossRef](#)]
17. Chakraborty, S. Generating discrete analogues of continuous probability distributions—A survey of methods and constructions. *J. Stat. Distrib. Appl.* **2015**, *2*, 6. [[CrossRef](#)]
18. Nielsen, F. The Kullback–Leibler divergence between lattice Gaussian distributions. *J. Indian Inst. Sci.* **2022**. . 1007/s41745-021-00279-5. [[CrossRef](#)]
19. Luzzi, L.; Vehkalahti, R.; Ling, C. Almost universal codes for MIMO wiretap channels. *IEEE Trans. Inf. Theory* **2018**, *64*, 7218–7241. [[CrossRef](#)]
20. Rioul, O. Variations on a Theme by Massey. *IEEE Trans. Inf. Theory* **2022**, *68*, 2813–2828. [[CrossRef](#)]
21. Kagan, A.; Smith, P.J. Multivariate normal distributions, Fisher information and matrix inequalities. *Int. J. Math. Educ. Sci. Technol.* **2001**, *32*, 91–96. [[CrossRef](#)]
22. Pontryagin, L.S. *Topological Groups*, 3rd ed.; Gordon and Breach Science Publishers: Montreux, Switzerland, 1986.
23. Raghunathan, M.S. *Discrete Subgroups of Lie Groups*; Springer: New York, NY, USA, 1972.
24. Reiter, H.; Stegeman, J.D. *Classical Harmonic Analysis and Locally Compact Groups*, 2nd ed.; Clarendon Press: Oxford, UK, 2000.

- 
25. Forney, G.D. Multidimensional constellations. II. Voronoi constellations. *IEEE J. Sel. Areas Commun.* **1989**, *7*, 941–958. [[CrossRef](#)]
  26. Boutros, J.J.; Jardi, F.; Méasson, C. Probabilistic shaping and non-binary codes. In Proceedings of the 2017 IEEE International Symposium on Information Theory (ISIT), Aachen, Germany, 25–30 June 2017; pp. 2308–2312.
  27. Paula, G.T. Comparison of volumes of Siegel sets and fundamental domains for  $SL_n(\mathbb{Z})$ . *Geom. Dedicata* **2019**, *199*, 291–306. [[CrossRef](#)]