THE GAUSSIAN ENTROPY MAP IN VALUED FIELDS

YASSINE EL MAAZOUZ
THE GAUSSIAN ENTROPY MAP IN VALUED FIELDS

YASSINE EL MAAZOUZ

The entropy map for multivariate real valued Gaussian distributions is the map that sends a positive definite matrix $\Sigma$ to the sequence of logarithms of its principal minors $\left(\log(\det \Sigma_i)\right)_i$. We exhibit the analog of this map in the nonarchimedean local fields setting, like the field of $p$-adic numbers for example. As in the real case, the image of this map lies in the supermodular cone. Moreover, given a multivariate Gaussian measure on a local field, its image under the entropy map determines its pushforward under valuation. In general, this map can be defined for nonarchimedean valued fields whose valuation group is an additive subgroup of the real line, and it remains supermodular. We also explicitly compute the image of this map in dimension 3.

1. Introduction and notation

Gaussian measures on local fields are introduced in [9]. In this text, we aim to exhibit the entropy map of these measures and discuss the properties this map satisfies. Our aim is to highlight the similarities with the real case. Before we discuss Gaussian measures on local fields (see Section 2), we begin by reviewing the entropy map in the real setting.

1.1. Entropy of real multivariate Gaussian distributions. For a positive integer $d$, multivariate Gaussian distributions on $\mathbb{R}^d$ are determined by their mean $\mu \in \mathbb{R}^d$ and their positive semidefinite covariance matrix $\Sigma \in \mathbb{R}^{d \times d}$. Hence the natural parameter space for centered (i.e., with zero mean) Gaussian distributions on $\mathbb{R}^d$ is the positive semidefinite cone in $\mathbb{R}^{d \times d}$, which we denote by

$$\text{PSD}_d := \{ \Sigma \in \text{Sym}_d(\mathbb{R}) : \langle x, \Sigma x \rangle \geq 0 \text{ for all } x \in \mathbb{R}^d \},$$

where $\text{Sym}_d(\mathbb{R})$ is the space of real symmetric matrices in $\mathbb{R}^{d \times d}$ and $\langle \cdot, \cdot \rangle$ is the usual inner product on $\mathbb{R}^d$. Nondegenerate Gaussian distributions are those whose covariance matrix $\Sigma$ is positive definite, i.e., $\Sigma \in \text{PD}_d$ where

$$\text{PD}_d := \text{PSD}_d^\circ = \{ \Sigma \in \text{Sym}_d(\mathbb{R}) : \langle x, \Sigma x \rangle > 0 \text{ for all nonzero } x \in \mathbb{R}^d \}.$$

There is no shortage of instances where the PSD cone appears in probability and statistics [25], optimization [19, Chapter 12] and combinatorics [13].

The positive definite cone has a pleasant group-theoretic structure in the sense that its elements are in one-to-one correspondence with left cosets of the orthogonal group $O_d(\mathbb{R})$ in the general linear
group $\text{GL}_d(\mathbb{R})$. The map sending the coset $A O_d(\mathbb{R}) \in \text{GL}_d(\mathbb{R})/O_d(\mathbb{R})$ to $AA^T \in \text{PD}_d$ is a bijection. This underscores the fact that multivariate Gaussians are tightly linked to the linearity and orthogonality structures that the Euclidean space $\mathbb{R}^d$ enjoys.

An important concept in statistics, probability, and information theory is the notion of entropy, which is a measure of uncertainty and disorder in a distribution; see [18]. The entropy of a centered multivariate Gaussian with covariance matrix $\Sigma$ is given up to an additive constant by

$$h(\Sigma) = -\log |\det \Sigma| = -\log(\det \Sigma).$$

If $X$ is a random vector in $\mathbb{R}^d$ with nondegenerate centered Gaussian distribution given by a covariance matrix $\Sigma \in \text{PD}_d$, then for any subset $I$ of $[d] := \{1, 2, \ldots, d\}$, the vector $X_I$ of coordinates of $X$ indexed by $I$ is also a random vector with nondegenerate Gaussian measure on $\mathbb{R}^{|I|}$. Moreover, its covariance matrix is $\Sigma_I = (\Sigma_{i,j})_{i,j \in I} \in \mathbb{R}^{|I| \times |I|}$, so we can define the entropy $h_I(\Sigma)$ of $X_I$ as

$$h_I(\Sigma) := h(\Sigma_I) = -\log(\det \Sigma_I).$$

The collection of entropy values $(h_I(\Sigma))_{I \subseteq [d]}$ satisfies the inequalities

$$h_I(\Sigma) + h_J(\Sigma) \leq h_{I \cap J}(\Sigma) + h_{I \cup J}(\Sigma) \text{ for any two subsets } I, J \subseteq [d].$$

(1)

This is thanks to what is known as Koteljanskii’s inequalities [15] on the determinants of positive definite matrices, i.e.,

$$\det \Sigma_I \det \Sigma_J \geq \det \Sigma_{I \cap J} \det \Sigma_{I \cup J}.$$  

(2)

In the language of polyhedral geometry, this means that the image of the entropy map

$$H : \text{PD}_d \rightarrow \mathbb{R}^{2^d}, \quad \Sigma \mapsto (h_I(\Sigma))_{I \subseteq [d]},$$

lies inside the supermodular cone $\mathcal{F}_d$ in $\mathbb{R}^{2^d}$. This is the polyhedral cone specified by the inequalities in (1), i.e.,

$$\mathcal{F}_d := \{x = (x_I)_{I \subseteq [d]} \in \mathbb{R}^{2^d} : x_{\emptyset} = 0 \text{ and } x_I + x_J \leq x_{I \cap J} + x_{I \cup J} \text{ for all } I, J \subseteq [d]\}.$$

Since $x_{\emptyset} = 0$ for $x \in \mathcal{F}_d$ we can see $\mathcal{F}_d$ as a cone in $\mathbb{R}^{2^d-1}$.

1.2. **Main results.** In this paper we deal with multivariate Gaussian distributions on local fields, and more generally nonarchimedean valued fields. See Example 2.2 for a discussion. In particular, we shall define an analog to the entropy map and show that it satisfies the same set of inequalities (1). More precisely, we prove the following:

**Theorem 1.3.** The pushforward measure of a multivariate Gaussian measure on a local field by the valuation map is given by a tropical polynomial whose coefficients are given by the entropy map of this measure (see Theorem 3.5). Moreover, these coefficients are supermodular. The entropy map is still well defined on nonarchimedean valued fields in general, and remains supermodular (see Theorem 4.2).
This solves [6, Conjecture 5.1] which roughly states that, given a multivariate Gaussian measure on a local field, its image under the entropy map determines its pushforward under valuation via a tropical polynomial. We shall break down Theorem 1.3 into several pieces. Namely, Theorems 3.5 and 4.2 for the local field case, and the discussion in Section 5 for the general nonarchimedean valued field case.

One motivation for this paper is the search for a suitable definition of tropical Gaussian measures [26]. Tropical stochastics has been an active research area in recent years and has diverse applications from phylogenetics [20; 29] to game theory [2] and economics [4; 27]. One appealing approach to define tropical Gaussians is to tropicalize Gaussian measures on a valued field.

Our text is organized as follows. In Section 3 we show that tropicalizing multivariate Gaussians on local fields yields probability measures on the lattice \( \mathbb{Z}^d \) that are determined by the entropy map via a tropical polynomial. In Section 4 we show the supermodularity of the entropy map and provide a recursive algorithm to compute it. In Section 5 we explain why orthogonality is not a suitable approach to define Gaussian measures when the field \( K \) is not locally compact. Nevertheless, we will see that the entropy map is still well defined and remains supermodular and we explicitly compute its image when \( d = 3 \).

You can find at https://mathrepo.mis.mpg.de/GaussianEntropyMap/index.html implementations, computations and data related to this paper.

**Remark 1.4.** For readers not familiar with local fields, we refer to [14; 23]. Local fields are not commonly used in statistics and probability. However, in recent years there has been a stream of literature addressing probabilistic and statistical questions in the \( p \)-adic setting, starting from the early work of Evans [8; 9] to the more recent developments [5; 10; 17] to mention a few.

### 2. Background on valued fields and Gaussian measures

This section is meant to collect the basic facts and results that we will need in our discussion. Most of these results can be found in the literature on valued fields in number theory [7; 23; 28] and functional analysis [21; 22].

#### 2.1. Valued fields.

Let \( K \) be a field with an additive nonarchimedean valuation \( \text{val} : K \to \mathbb{R} \cup \{+\infty\} \) with valuation group \( \Gamma := \text{val}(K^\times) \). The valuation map \( \text{val} \) defines an equivalence class of exponential valuations or absolute values \( |\cdot| \) on \( K \) via \( |x| := a^{-\text{val}(x)} \) (where \( a \in (1, \infty) \)) and hence also a topology on \( K \). The valuation \( \text{val} \) is called discrete if its valuation group \( \Gamma \) is a discrete subgroup of \( \mathbb{R} \) which, by scaling \( \text{val} \) suitably, we can always assume to be \( \mathbb{Z} \) (we then call \( \text{val} \) a normalized valuation). In the discrete valuation case we fix a uniformizer \( \pi \) of \( K \), i.e., an element \( \pi \in K \) with \( \text{val}(\pi) = 1 \). We denote by \( \mathcal{O} := \{x \in K : \text{val}(x) \geq 0\} \) the valuation ring of \( K \); this is a local ring with unique maximal ideal \( \mathfrak{m} := \{x \in K : \text{val}(x) > 0\} \) and residue field \( k := \mathcal{O}/\mathfrak{m} \). When the valuation is discrete, the ideal \( \mathfrak{m} \) is generated in \( \mathcal{O} \) by \( \pi \), i.e., \( \mathfrak{m} := \pi \mathcal{O} \). We mention typical examples of such fields in Example 2.2.

**Example 2.2.** (1) The field \( \mathbb{F}_q((t)) \) of Laurent series in one variable with coefficients in the finite field \( \mathbb{F}_q \).

(2) The fields \( \mathbb{R}((t)) \) or \( \mathbb{C}((t)) \) of Laurent series with complex or real coefficients. These are fields with an infinite residue field but still in discrete valuation \( \Gamma = \mathbb{Z} \).
(3) The fields $\mathbb{R}\{t\} = \bigcup_{n \geq 1} \mathbb{R}(t^{1/n})$ and $\mathbb{C}\{t\} = \bigcup_{n \geq 1} \mathbb{C}(t^{1/n})$ of Puiseux series in $t$. In this case the valuation group $\Gamma = \mathbb{Q}$ is dense.

(4) The field of generalized Puiseux series $\mathbb{K}$ which has valuation group $\Gamma = \mathbb{R}$. This field consists of formal series $f = \sum_{\alpha \in \mathbb{R}} a_{\alpha} t^{\alpha}$ where $\text{supp}(f) := \{ \alpha \in \mathbb{R} : a_{\alpha} \neq 0 \}$ is either finite or has $+\infty$ as the only accumulation point; see [3].

(5) All the previous fields have the same characteristic as their residue fields. Interesting examples in mixed characteristic are the field of $p$-adic numbers $\mathbb{Q}_p$ where $p$ is prime, its algebraic closure $\overline{\mathbb{Q}}_p$ and the field of $p$-adic complex numbers $\mathbb{C}_p$ (completion of $\overline{\mathbb{Q}}_p$).

2.3. Local fields. These are valued fields that are locally compact. In this section let us assume that $K$ is locally compact. It is then known that $K$ is isomorphic to a finite field extension of $\mathbb{Q}_p$ or $\mathbb{F}_q((t))$ and that its valuation group $\Gamma$ is discrete in $\mathbb{R}$, and its residue field $k$ is finite. In this case, by convention, the absolute value on $K$ is defined as $|x| = q^{-\text{val}(x)}$ (so we choose $a = q$), and there exists a unique Haar measure $\mu$ on $K$ such that $\mu(C) = 1$.

2.4. Lattices. Let $d \geq 1$ an integer. We call a lattice in $K^d$ any $\mathcal{O}$-submodule $\Lambda := \bigoplus_{i=1}^n \mathcal{O} a_i$ generated by a basis $(a_1, \ldots, a_d)$ of $K^d$. The basis $(a_1, \ldots, a_d)$ that generates $\Lambda$ is not unique. We can write $\Lambda = A\mathcal{O}^d$ where $A$ is the matrix with columns $a_1, \ldots, a_d$, which is then called a representative of $\Lambda$. The elements $U$ of the group $\text{GL}_d(K)$ that leave $\mathcal{O}^d$ invariant (i.e., $U\mathcal{O}^d = \mathcal{O}^d$) are exactly the matrices $U \in \text{GL}_d(\mathcal{O})$ with entries in $\mathcal{O}$ whose inverse has all entries in $\mathcal{O}$. The group $\text{GL}_d(\mathcal{O})$ then plays the role of the orthogonal group $O_d(\mathbb{R})$ [11, Theorem 2.4]. Then, like positive definite matrices, lattices are in one-to-one correspondence with left cosets $\text{GL}_d(K) / \text{GL}_d(\mathcal{O})$. In particular, any two representatives of a lattice $\Lambda$ are elements of the same left coset. A lattice $\Lambda$ is called diagonal if it admits a diagonal matrix as a representative. Let us now state a result on lattices over valued fields that will be useful in our discussion.

Lemma 2.5. For any two lattices $\Lambda, \Lambda'$ there exists an element $g \in \text{GL}_d(K)$ such that $g\Lambda$ and $g\Lambda'$ are both diagonal lattices.\(^2\)

Proof. It suffices to show this when $\Lambda$ is the standard lattice $\Lambda = \mathcal{O}^d$. Let $\Lambda \in \text{GL}_d(K)$ be a representative of $\Lambda'$. Thanks to the nonarchimedean singular value decomposition (see [10, Theorem 3.1]), there exists a diagonal matrix $D \in \text{GL}_d(K)$ and $U, V \in \text{GL}_d(\mathcal{O})$ such that $A = UDV$. Hence we deduce that $\Lambda' = UD\mathcal{O}^d$. Picking $g = U^{-1}$ yields $g\Lambda = U^{-1}\mathcal{O}^d = \mathcal{O}^d$ and $g\Lambda' = D\mathcal{O}^d$.\(\square\)

2.6. Gaussian measures. Suppose that $K$ is a local field and $d$ is a positive integer. As shown by Evans [9], one can define multivariate Gaussian measures on $K^d$ using nonarchimedean orthogonality. It turns out that these measures are precisely the uniform distributions on $\mathcal{O}$-submodules of $K^d$. The nondegenerate Gaussians on $K^d$ are then parametrized by full rank submodules of $K^d$, i.e., lattices.

\(^1\)Homothety classes of diagonal lattices form what is called an apartment in the theory of buildings.

\(^2\)This is in fact a property of buildings: any two chambers belong to a common apartment; see [1].
For a lattice \( \Lambda \) in \( K^d \), we denote by \( \mathbb{P}_\Lambda \) the Gaussian measure on \( K^d \) given by \( \Lambda \), i.e., the uniform probability measure on \( \Lambda \). If \( f_\Lambda \) denotes the density (with respect to the Haar measure \( \mu^{\otimes d} \)) of \( \mathbb{P}_\Lambda \), then
\[
f_\Lambda(x) = 1_\Lambda(x)/\mu^{\otimes d}(\Lambda), \quad x \in K^d,
\]
where \( 1_\Lambda \) is the set indicator function of \( \Lambda \).

One can then think of lattices as analogs for the positive definite covariance matrices in the real case since they parametrize nondegenerate multivariate Gaussian measures. In the language of group theorists, one can think of the Bruhat–Tits building for the reductive group \( \text{PGL}_d(K) \) \([1]\) as the parameter space for nondegenerate Gaussians up to scalar multiplication.

### 3. The entropy map of local field Gaussian distributions

In this section we assume that \( K \) is a local field and we fix a positive integer \( d \geq 1 \) and a lattice \( \Lambda \) in \( K^d \). We recall that there is a unique Haar measure \( \mu^{\otimes d} \) on \( K^d \) that is the product measure induced by \( \mu \) on \( K \). Letting \( A \) be a representative of the lattice \( \Lambda \), i.e., \( \Lambda = AC^d \), we can define the entropy \( h(\Lambda) \) of the lattice \( \Lambda \) as
\[
h(\Lambda) = \text{val}(\det A).
\]
This is a well-defined quantity since any other representative of \( \Lambda \) is of the form \( AU \) where \( U \in \text{GL}_d(\mathbb{C}) \) and \( \det U \in \mathbb{C}^\times \) is a unit, so \( \text{val}(\det U) = 0 \). This definition lines up with the definition in the real case because \( \text{val}(x) = -\log_q |x| \) where \( |\cdot| \) is the absolute value on \( K \), so we get
\[
h(\Lambda) = \text{val}(\det A) = -\log_q |\det A|.
\]

The following proposition justifies the nomenclature “entropy” and relates the entropy \( h(\Lambda) \) of a lattice \( \Lambda \) to its measure \( \mu^{\otimes d}(\Lambda) \).

**Proposition 3.1.** We have \( \mu^{\otimes d}(\Lambda) = q^{-h(\Lambda)} \). Moreover, the quantity \( h(\Lambda) \) is the differential entropy of the Gaussian measure \( \mathbb{P}_\Lambda \), i.e.,
\[
h(\Lambda) = \int_{K^d} \log_q (f_\Lambda(x)) \mathbb{P}_\Lambda(dx).
\]

**Proof.** Let \( A \) be a representative of \( \Lambda \). Thanks to the nonarchimedean singular value decomposition (see \([10, \text{Theorem 3.1}]\)), we can write \( A = UDV \), where \( U, V \in \text{GL}_d(\mathbb{C}) \) are two orthogonal matrices and \( D \) is a diagonal matrix. Then we have \( \Lambda = UDC^d \). Since orthogonal linear transformations in \( K^d \) preserve the measure, we have \( \mu^{\otimes d}(\Lambda) = \mu^{\otimes d}(DC^d) \). Let \( \alpha_1, \ldots, \alpha_d \) be the diagonal entries of \( D \). Then we have
\[
\mu^{\otimes d}(\Lambda) = \mu^{\otimes d}(\bigoplus_{i=1}^d \alpha_i C^d) = q^{-\text{val}(\alpha_1)} \cdots -\text{val}(\alpha_d). \quad \text{But } \text{val}(\alpha_1) + \cdots + \text{val}(\alpha_d) = \text{val}(\det A) = h(\Lambda).
\]
The second statement follows from the immediate computation
\[
\int_{K^d} \log_q (f_\Lambda(x)) \mathbb{P}_\Lambda(dx) = \int_{K^d} \log_q (f_\Lambda(x)) f_\Lambda(x) \mu^{\otimes d}(dx) = h(\Lambda). \quad \square
\]

For a subset \( I \) of \([d] := \{1, 2, \ldots, d\} \) we denote by \( \Lambda_I \) the image of \( \Lambda \) under the projection onto the space \( K^{|I|} \) of coordinates indexed by \( I \). This is also a lattice in the space \( K^{|I|} \). So, for any subset \( I \subset [d] \),
we can define the entropy \( h_I(\Lambda) \) of the lattice \( \Lambda_I \). We can then define the entropy map

\[
H : \text{GL}_d(K) / \text{GL}(\mathbb{C}) \to \mathbb{R}^d, \quad \Lambda \mapsto (h_I(\Lambda))_{I \subset [d]},
\]

where \( h_\varnothing(\Sigma) = 0 \) by convention. If \( A \) is a representative of \( \Lambda \) with columns \( a_1, \ldots, a_d \), then the lattice \( \Lambda_I \) is the lattice generated over \( \mathbb{C} \) by the vectors \( a_{iI} \) which are the subvectors of the \( a_i \) with coordinates indexed by \( I \). So we can compute \( h_I(\Lambda) \) from the matrix \( A \) by

\[
h_I(\Lambda) = \min_{J \subset [d], |J| = |I|} \text{val}(\det(A_{I \times J})),
\]

where \( A_{I \times J} \) is the matrix extracted from \( A \) by taking the rows indexed by \( I \) and the columns indexed by \( J \), i.e., \( A_{I \times J} = (A_{i,j})_{i \in I, j \in J} \).

Now let \( X \) be a \( K^d \)-valued random variable with Gaussian distribution \( \mathbb{P}_\Lambda \) given by \( \Lambda \). So for any measurable set \( B \) in the Borel \( \sigma \)-algebra of \( K^d \),

\[
\mathbb{P}_\Lambda(X \in B) = \frac{\mu^\otimes d(\Lambda \cap B)}{\mu^\otimes d(\Lambda)},
\]

and \( V := \text{val}(X) \) its image under coordinate-wise valuation. Notice that, since \( \mathbb{P}_\Lambda(X_i = 0) = 0 \) for any \( i \in \{1, \ldots, d\} \), the vector \( V \) is almost surely in \( \mathbb{Z}^d \). By definition the distribution of \( V \) is the pushforward of the distribution of \( X \) by the map \( \text{val} \). We are interested in the distribution of the valuation vector \( V \), and to determine it we compute its tail distribution function \( Q_\Lambda \), which is defined on \( \mathbb{R}^d \) as

\[
Q_\Lambda(v) := \mathbb{P}_\Lambda(V \geq v) \text{ for any } v \in \mathbb{R}^d,
\]

where \( \geq \) is the coordinate-wise partial order on \( \mathbb{R}^d \). Since \( V \) takes values in \( \mathbb{Z}^d \), this function is completely determined by its values for \( v \in \mathbb{Z}^d \). For a vector \( v = (v_1, \ldots, v_d) \in \mathbb{Z}^d \), let us denote by \( \pi^v \) the \( \mathbb{C} \)-module generated by the basis \( \pi^v e_i \) where \( e_1, \ldots, e_d \) is the standard basis of \( K^d \), i.e.,

\[
\pi^v = \pi^{v_1} \mathbb{C} e_1 \oplus \cdots \oplus \pi^{v_d} \mathbb{C} e_d.
\]

**Definition 3.2.** We define the logarithmic tail distribution function \( \varphi_\Lambda \) as

\[
\varphi_\Lambda : \mathbb{Z}^d \to \mathbb{Z}, \quad v \mapsto -\log_q(Q_\Lambda(v)).
\]

The following lemma relates the tail distribution function \( \varphi_\Lambda \) with the entropy \( h(\Lambda) \) of the lattice \( \Lambda \).

**Lemma 3.3.** We have \( \varphi_\Lambda(v) = h(\Lambda \cap \pi^v) - h(\Lambda) \). Moreover, if \([\Lambda : \Lambda \cap \pi^v] \) denotes the index of \( \Lambda \cap \pi^v \) as a subgroup of \( \Lambda \), then we also have

\[
Q_\Lambda(v) = 1/[\Lambda : \Lambda \cap \pi^v].
\]

**Proof.** By definition we have \( Q_\Lambda(v) = \mathbb{P}_\Lambda(X \in \pi^v) = \mu^\otimes d(\pi^v \cap \Lambda)/\mu^\otimes d(\Lambda) \). So by virtue of Proposition 3.1 we deduce that \( Q_\Lambda(v) = \varphi_\Lambda^{h(\Lambda)}(\Lambda \cap \pi^v) \). The first statement then follows from the definition of \( \varphi_\Lambda \) (Definition 3.2). For the second statement, by definition, \( \Lambda \) can be partitioned into \([\Lambda : \Lambda \cap \pi^v] \) cosets of \( \Lambda \cap \pi^v \). Since the Haar measure \( \mu^\otimes d \) is translation invariant, all of these cosets
have the same measure, i.e., \( \mu^{\otimes d}(\Lambda) = [\Lambda : \Lambda \cap \pi^v]\mu^{\otimes d}(\Lambda \cap \pi^v) \). The result then follows from the fact that \( Q_\Lambda(v) = \mu^{\otimes d}(\pi^v \cap \Lambda)/\mu^{\otimes d}(\Lambda) \) and Definition 3.2.

Next, we introduce a technical tool that we will be using in the proof of our first result.

**Definition 3.4.** For any \( \ell \in \{0, \ldots, d\} \) we define the \( \ell \)-distance \( \phi_\ell(\Lambda, \Lambda') \) of two lattices \( \Lambda, \Lambda' \) as the minimum of \( \text{val}(\det(x_1, \ldots, x_\ell, y_1, \ldots, y_\ell)) \) among all possible choices of \( x_1, \ldots, x_\ell \subseteq \Lambda \) and \( y_1, \ldots, y_\ell \subseteq \Lambda' \) where \( k = d - \ell \).

Since for any \( g \in \text{GL}_d(K) \), and for \( x_1, \ldots, x_\ell \subseteq \Lambda \) and \( y_1, \ldots, y_\ell \subseteq \Lambda' \), we have

\[
\text{val}(\det(gx_1, \ldots, gx_\ell, gy_1, \ldots, gy_\ell)) = \text{val}(\det(x_1, \ldots, x_\ell, y_1, \ldots, y_\ell)) + \text{val}(\det g),
\]

we can see that

\[
\phi_\ell(g\Lambda, g\Lambda') = \phi_\ell(\Lambda, \Lambda') + \text{val}(\det g).
\]

We then deduce that the quantity \( \phi_\ell(\Lambda, \Lambda') - h(\Lambda') \) is invariant under the action \( \text{GL}_d(K) \), i.e., for any \( g \in \text{GL}_d(K) \) we have

\[
\phi_\ell(g\Lambda, g\Lambda') - h(g\Lambda') = \phi_\ell(\Lambda, \Lambda') - h(\Lambda').
\]

When the second lattice \( \Lambda' = \pi^v \) is diagonal and \( \Lambda \) has representative \( A \in \text{GL}_d(K) \), the optimal choice for the vectors \( x_1, \ldots, x_\ell \) and \( y_1, \ldots, y_\ell \) is when the vectors \( x_1, \ldots, x_\ell \) are among the columns \( a_1, \ldots, a_\ell \) of \( A \) and the vectors \( y_1, \ldots, y_\ell \) are among the vectors \( \pi^v e_i \) where \( (e_i)_{1 \leq i \leq d} \) is the standard basis of \( K^d \). So we deduce that \( \phi_\ell(\Lambda, \pi^v) \) can be computed as

\[
\phi_\ell(\Lambda, \pi^v) = \min_{I,J \subseteq [d]} \left( \text{val}(\det A_{I \times J}) + \sum_{j \notin J} v_j \right).
\]

So we also get

\[
\phi_\ell(\Lambda, \pi^v) - h(\pi^v) = \min_{I,J \subseteq [d]} \left( \text{val}(\det A_{I \times J}) - \sum_{j \in J} v_j \right).
\]

In the special case \( \Lambda = \pi^a \), for \( a \in \mathbb{Z}^d \), the determinant of \( A_{I \times J} \) in the above optimization problem is 0 whenever \( J \neq I \), since we can choose \( A \) to be diagonal. So we get

\[
\phi_\ell(\pi^a, \pi^v) - h(\pi^v) = \min_{I \subseteq [d]} \left( \sum_{i \in I} a_i - \sum_{i \in I} v_i \right).
\]

**Theorem 3.5.** The logarithmic tail distribution function \( \varphi_\Lambda \) is a tropical polynomial on \( \mathbb{Z}^d \) given by

\[
\varphi_\Lambda(v) = \max_{I \subseteq [d]} (v_I - h_I(\Lambda)).
\]

**Proof.** First we show this for a diagonal lattice \( \Lambda = \pi^a \) where \( a \in \mathbb{Z}^d \). For any \( v \in \mathbb{Z}^d \), let \( a \vee v \) the vector with coordinates \( \max(a_i, v_i) \). We have \( \pi^a \cap \pi^v = \pi^{a \vee v} \) so we get the entropy \( h(\pi^a) = \sum_{i=1}^d a_i \)

\[
\varphi_\Lambda(v) = \max_{I \subseteq [d]} (v_I - h_I(\Lambda)).
\]
and \( h(\pi^a \cap \pi^v) = h(\pi^{av}) = \sum_{i=1}^d \max(a_i, v_i) \). Hence we have

\[
\varphi_\Lambda(v) = h(\pi^a \cap \pi^v) - h(\pi^a) = \max_{I \subset [d]} \left( \sum_{i \in I} v_i + \sum_{i \not\in I} a_i \right) - \sum_{i=1}^d a_i = \max_{I \subset [d]} (v_I - a_I),
\]

and \( h_I(\pi^a) = a_I \). So the theorem holds for diagonal lattices. To see why it also holds for a general lattice \( \Lambda \), first notice that in the diagonal case \( \Lambda = \pi^a \) we have

\[
\varphi_\Lambda(v) = - \min_{\ell=0, \ldots, d} (\phi_\ell(\Lambda, \pi^v) - h(\pi^v)).
\]

Secondly, notice that the right hand side of the previous equation is invariant under the action of \( \text{GL}_d(K) \). So for \( g \in \text{GL}_d(K) \),

\[
\min_{\ell=0, \ldots, d} (\phi_\ell(g\Lambda, g\pi^v) - h(g\pi^v)) = \min_{\ell=0, \ldots, d} (\phi_\ell(\Lambda, \pi^v) - h(\pi^v)).
\]

By Lemma 3.3, we have \( \varphi_\Lambda(v) = \log_q [\Lambda : \Lambda \cap \pi^v] = \log_q [g\Lambda : g\Lambda \cap g\pi^v] \). Now fix a general lattice \( \Lambda \) and \( v \in \mathbb{Z}^d \). Also, by Lemma 2.5, there exists \( g \in \text{GL}_d(K) \) such that \( g\Lambda \) and \( g\pi^v \) are both diagonal, so

\[
\varphi_\Lambda(v) = \log_q [g\Lambda : g\Lambda \cap g\pi^v] = - \min_{\ell=0, \ldots, d} (\phi_\ell(g\Lambda, g\pi^v) - h(g\pi^v)) = - \min_{\ell=0, \ldots, d} (\phi_\ell(\Lambda, \pi^v) - h(\pi^v)).
\]

Hence, we deduce, thanks to (6), that

\[
\varphi_\Lambda(v) = - \min_{\ell=0, \ldots, d} \left( \min_{I, J \subset [d]} (\text{val}(\det A_{I \times J}) - \sum_{j \in J} v_j) \right).
\]

We can simplify this thanks to (5) to get the desired equation (7).

So the distribution of the random vector of valuations \( V \) is given by a tropical polynomial \( \varphi_\Lambda \) via its tail distribution function \( Q_\Lambda \). The coefficients of this polynomial are exactly the entropies \( h_I(\Lambda) \).

Now we prove a couple of interesting properties of \( \varphi_\Lambda \), namely how the coefficients \( h_I(\Lambda) \) behave under diagonal scaling and permutation of coordinates of the random vector \( X \). To this end, let us denote by \( D_a = \text{diag}(a_1, \ldots, a_n) \) the diagonal matrix with coefficients \( a_i \in K \) and \( P^\sigma \) the permutation matrix corresponding to a permutation \( \sigma \) of \([d]\), i.e., \( P^\sigma_{i,j} = 1 \) when \( j = \sigma(i) \) and 0 otherwise.

**Lemma 3.6.** Let \( \Lambda \) be a lattice in \( K^d \), \( a \in K^d \) and \( \sigma \) a permutation of \([d]\). We have

\[
h_I(D_a \Lambda) = h_I(\Lambda) + \sum_{i \in I} \text{val}(a_i) \quad \text{and} \quad h_I(P^\sigma \Lambda) = h_{\sigma(I)}(\Lambda).
\]

**Proof.** For \( I \subset [d] \), we have \( h_I(D_a \Lambda) = \min_{|J|=|I|} \text{val}(\det(D_a A)_{I \times J}) \), where \( A \) is any representative of \( \Lambda \). Since all the lines of \( D_a A \) are multiples of those of \( A \) by the scalars \( a_i \) we deduce that \( \det(D_a A)_{I \times J} = \det A_{I \times J} \prod_{i \in I} a_i \) and hence we get

\[
h_I(D_a \Lambda) = h_I(\Lambda) + \sum_{i \in I} \text{val}(a_i).
\]

Similarly we can see the effect the permutation of coordinates of \( X \) has on the vector of entropies \( H(\Lambda) = (h_I(\Lambda))_{I \subset [d]} \). 

\[\square\]
4. Supermodularity of the entropy map

As it is the case for real Gaussians, we would like the vector of entropies $H(\Lambda) := (h_1(\Lambda))$ to have values in the supermodular cone $\mathcal{S}_d$ as conjectured in [6]. As a first step towards proving this result, notice that the previous lemma implies that if $\Lambda$ is a lattice such that $H(\Lambda) \in \mathcal{S}_d$, then for any diagonal matrix $D_\sigma$ we still have $H(D_\sigma \Lambda) \in \mathcal{S}_d$ and $H(P^{\sigma} \Lambda) \in \mathcal{S}_d$ for any permutation $\sigma$ of $\{1, \ldots, d\}$.

**Definition 4.1** (Hermite normal form$^3$). Every lattice $\Lambda$ in $K^d$ has a representative $A$ in Hermite normal form, i.e., a matrix $A = (A_{ij})$ in $\text{GL}_d(K)$ satisfying the following conditions:

(i) $A$ is lower triangular, i.e., $A_{ij} = 0$ whenever $i < j$.

(ii) For any $1 \leq j < i \leq d$, we have either $\text{val}(A_{ij}) < \text{val}(A_{jj})$ or $A_{ij} = 0$.

(iii) The diagonal coefficients $A_{ii}$ are of the form $A_{ii} = \pi^a_i$ for some $a_i \in \mathbb{Z}$.

We can now state the second result of this section concerning the supermodularity of the entropy map. However, we first need to give an equivalent definition of the supermodular cone as

$$\mathcal{S}_d = \{(x_i)_{I \subset [d]} \in \mathbb{R}^d : x_i = 0 \text{ and } x_{Ii} + x_{Ij} \leq x_I + x_{Iij} \text{ for any } I \subset [d], i \neq j \in [d] \setminus I\},$$

where we write $Ii$ instead of $I \cup \{i\}$. These are the facet-defining inequalities of the cone $\mathcal{S}_d$ and there are $d(d-1)2^{d-3}$ of them; see [16].

**Theorem 4.2.** The image of the map $H : \Lambda \rightarrow (h_1(\Lambda))_{I \subset [d]}$ lies in the supermodular cone $\mathcal{S}_d$, i.e., for any subset $I \subset [d]$ with $|I| \leq d - 2$ and $i \neq j \in [d] \setminus I$,

$$h_{Ii}(\Lambda) + h_{Ij}(\Lambda) \leq h_I(\Lambda) + h_{Iij}(\Lambda).$$

**Proof.** We prove this by induction on $d$. The result is trivial for $d = 1, 2$. Assume that it holds for lattices in $K^r$ for any $r \leq d$, where $d \geq 3$. Let $\Lambda$ be a lattice in $K^d$ and $A$ its Hermite normal form. For any $I \subset [d]$ of size $|I| < d - 2$ the inequality $h_{Ii}(\Lambda) + h_{Ij}(\Lambda) \leq h_I(\Lambda) + h_{Iij}(\Lambda)$ holds for any $i \neq j$ not in $I$ thanks to the induction hypothesis. This is because, when $|I| \leq d - 2$, we are working on the lattice $\Lambda_{Iij}$ which is a lattice in dimension less than $d$. Then, it suffices to show the inequality when $I$ has size $d - 2$. By Lemma 3.6 we can assume that $I = \{1, \ldots, d - 2\}$ and $i = d - 1$ and $j = d$ (if not, we can just act on $\Lambda$ by a suitable permutation matrix). Let us write down the matrix $A$ as

$$A = \begin{pmatrix}
\pi^{a_1} & 0 & \ldots & 0 & 0 & 0 \\
* & \pi^{a_2} & \ddots & \vdots & \vdots & \\
\vdots & \ddots & \ddots & 0 & 0 & 0 \\
* & \ldots & \ldots & \pi^{a_{d-2}} & 0 & 0 \\
* & \ldots & \ldots & \pi^{a_{d-1}} & 0 \\
* & \ldots & \ldots & \pi^{a_d} & x & \pi^{a_d}
\end{pmatrix}.$$  

---

$^3$The curious reader can see [28, Chapter II] and [6, Proposition 4.2] for details.
Recall that since $A$ is the Hermite form of $\Lambda$ we have $\text{val}(x) < a_d$ or $x = 0$. Now we have

$$
\begin{align*}
    h_{Ii}(\Lambda) &= a_1 + \cdots + a_{d-1}, & h_{Ij}(\Lambda) &= a_1 + \cdots + a_{d-2} + \text{min}(\text{val}(x), a_d), \\
    h_I(\Lambda) &= a_1 + \cdots + a_{d-2} & \text{and} & \quad h_{ij}(\Lambda) &= a_1 + \cdots + a_d.
\end{align*}
$$

The inequality $h_{Ii}(\Lambda) + h_{Ij}(\Lambda) \leq h_I(\Lambda) + h_{ij}(\Lambda)$ then holds simply because $\text{min}(\text{val}(x), a_d) \leq a_d$ and this finishes the proof. \hfill $\square$

This theorem underlines another similarity between the local field Gaussians defined in [9] and classical multivariate Gaussian measures. From Lemma 3.6 we can see that acting on $\Lambda$ by a diagonal matrix just moves the point $H(\Lambda) \in \mathcal{F}_d$ in parallel to the lineality space of the cone $\mathcal{F}_d$, that is, the biggest vector space contained in $\mathcal{F}_d$.

The classical entropy map is tightly related to conditional independence. More precisely, if $\Sigma \in \text{PD}_d$ and $X$ is a Gaussian vector with covariance matrix $\Sigma$, then for any $I \subseteq [d]$ and $i \neq j$ not in $I$, the variables $X_i$ and $X_j$ are independent given the vector $X_I$ if and only if $h_{Ii}(\Sigma) + h_{Ij}(\Sigma) = h_I(\Sigma) + h_{ij}(\Sigma)$ and we write

$$
X_i \perp X_j | X_I \iff h_{Ii}(\Sigma) + h_{Ij}(\Sigma) = h_I(\Sigma) + h_{ij}(\Sigma).
$$

This means that the conditional independence models are exactly the inverse images by $H$ of the faces of $\mathcal{F}_d$ [24, Proposition 4.1]. It turns out that, in the local field setting, the nonarchimedean entropy map $H$ defined in (3) also encodes conditional independence information on the coordinates of the random Gaussian vector $X$ as stated in the following proposition.

**Proposition 4.3.** Assume $d \geq 2$, let $I$ be a subset of $[d]$ and let $i \neq j \in [d] \setminus I$ be two distinct integers. Let $\Lambda$ be a lattice in $K^d$ and $X$ a random Gaussian vector with distribution given by $\Lambda$. Then the conditional independence statement $X_i \perp X_j | X_I$ holds if and only if $h_{Ii}(\Lambda) + h_{Ij}(\Lambda) = h_I(\Lambda) + h_{ij}(\Lambda)$.

**Proof.** Using Lemma 3.6 we reduce to the case $I = [r]$ where $r \leq d - 2$, $i = r + 1$ and $j = i + 1$. Let $A = (a_{i,j})$ be the unique representative in Hermite form of $\Lambda$. We claim that $X_i \perp X_j | X_I$ if and only if $a_{j,i} = 0$. To see why, let $Z = A^{-1}X$ which is a Gaussian vector with uniform distribution on $C^d$. We have $X_i = a_{i,1}Z_1 + \cdots + a_{i,i}Z_i$ and $X_j = a_{j,1}Z_1 + \cdots + a_{j,j}Z_j$. Since $Z_I = A_{I,I}^{-1}X_I$, given $X_I$ we know $Z_I$ and vice versa. Hence $X_{i} \perp X_j | X_I$ holds if and only if $(a_{j,i}Z_i + a_{j,j}Z_j) \perp Z_i$. This happens if and only if the vectors $(1, 0)$ and $(a_{j,i}, a_{j,j})$ in $K^2$ are orthogonal; see [9]. This is equivalent to $\text{val}(a_{j,i}) \leq \text{val}(a_{j,j})$ which means that $a_{j,i} = 0$ since $A$ is in Hermite form. On the other hand, since $A$ is lower triangular, we have

$$
\begin{align*}
    h_I(\Lambda) &= \text{val}(\det A_{I \times I}), & h_{Ii}(\Lambda) &= h_I(\Lambda) + \text{val}(a_{i,i}), \\
    h_{Ij}(\Lambda) &= h_I(\Lambda) + \text{min}(\text{val}(a_{j,i}), \text{val}(a_{j,j})) & \text{and} & \quad h_{ij}(\Lambda) &= h_I(\Lambda) + \text{val}(a_{i,i}) + \text{val}(a_{j,j}).
\end{align*}
$$

So the equality $h_{Ii}(\Lambda) + h_{Ij}(\Lambda) = h_I(\Lambda) + h_{ij}(\Lambda)$ holds if and only if $\text{val}(a_{j,i}) \leq \text{val}(a_{j,j})$. Since $A$ is the Hermite form of $\Lambda$, this happens if and only if $a_{j,i} = 0$. In combination with the calculation above, this finishes the proof. \hfill $\square$
In other terms, the conditional independence statement $X_i \perp \perp X_j | X_I$ holds if and only if the entropy vector $H(\Lambda) = (h_i(\Lambda))$ is on the face of the polyhedral cone $\mathcal{F}_d$ cut by the equation $h_{ii}(\Lambda) + h_{ij}(\Lambda) = h_I(\Lambda) + h_{Ij}(\Lambda)$. This gives an analog of [24, Proposition 4.1].

**Corollary 4.4.** The Gaussian conditional independence models are exactly those subsets of lattices that arise as inverse images of the faces of $\mathcal{F}_d$ under the map $H$.

**Proof.** This follows immediately from the previous proposition. □

This underlines the importance of the map $H$, and also gives reason to think that the suitable analog of the positive definite cone on local fields is the set of lattices, or more precisely the Bruhat–Tits building [1; 6]. A hard question in information theory for classical multivariate Gaussians is to describe the image of the entropy map [24]. This problem turns out to be difficult in this setting as well.

**Problem 4.5.** Characterize the image of the entropy map $H$ and describe how it intersects the faces of $\mathcal{F}_d$. What can you say about the fibers of this map?

**Remark 4.6.** We recall that for any $d \geq 1$, the image $\text{im}(H)$ is invariant under the action of the symmetric group, and by translation in parallel to the lineality space of $\mathcal{F}_d$. This is thanks to Lemma 3.6. We will provide an answer for Problem 4.5 when $d = 2, 3$ at the end of Section 5.

We now provide an algorithm to compute the entropy vector $H(\Lambda)$, i.e., the coefficients of the polynomial $\varphi_\Lambda$. This relies on computing the Hermite form rather than directly solving the optimization problems given by (5).

**Algorithm: Computing $H(\Lambda)$.

- **Input:** A full rank matrix $A = (a_1, \ldots, a_n) \in K^{d \times n}$ with $n \geq d$ generating $\Lambda$.
- **Output:** The entropy vector $H(\Lambda)$.

  for $I \subset [d]$ do
  
  Compute the Hermite form $A_I$ of $\Lambda_I$.
  
  $h_I(\Lambda) \leftarrow \text{val} (\det A_I)$.

  end

  $H(\Lambda) \leftarrow (h_I(\Lambda))_{I \subset [d]}$.

  return $H(\Lambda)$.

Let us now discuss a couple of low-dimensional examples when $K = \mathbb{Q}_p$.

**Example 4.7.** Let $\Lambda$ be the lattice represented by $A = \begin{pmatrix} 1 & 0 \\ p & p^2 \end{pmatrix}$. The coefficients $h_I(\Lambda)$ of the polynomial $\varphi_\Lambda$ can be computed from the representative $A$ using the above algorithm and we have

$$h_\emptyset(\Lambda) = 0, \quad h_1(\Lambda) = 0, \quad h_2(\Lambda) = 1, \quad h_{1,2}(\Lambda) = 2,$$

and then we get

$$\varphi_\Lambda(v_1, v_2) = \max(0, v_1, v_2 - 1, v_1 + v_2 - 2).$$

The independence statement $X_1 \perp \perp X_2$ does not hold since the inequality $h_1(\Lambda) + h_2(\Lambda) \leq h_{12}(\Lambda)$ is strict. The tropical curve of $\varphi_\Lambda$ and its regular triangulation of the square are shown in Figure 1.
Figure 1. Tropical curve of $\varphi_3$ and its regular triangulation of the square for Example 4.7.

**Example 4.8.** Let $\Lambda$ be the lattice represented by

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 1 & \pi^2 & 0 \\ 1 & \pi & \pi^2 \end{pmatrix}.$$

The polynomial $\varphi_3$ can be computed again using the algorithm on page 11 and we get

$$h_\emptyset(3) = 0, \quad h_1(3) = 0, \quad h_2(3) = 0, \quad h_3(3) = 0, \quad h_{1,2}(3) = 2, \quad h_{1,3}(3) = 1, \quad h_{2,3}(3) = 1, \quad h_{1,2,3}(3) = 4.$$ 

So we deduce that

$$\varphi_3(v) = \max(0, v_1, v_2, v_3, v_1 + v_2 - 2, v_1 + v_3 - 1, v_2 + v_3 - 1, v_1 + v_2 + v_3 - 4).$$

We can easily check that the supermodularity inequalities are satisfied. Also, none of the conditional independence statements $X_i \perp X_j | X_k$ are satisfied for $\{i, j, k\} = \{1, 2, 3\}$ since the point $H(3)$ is in the interior of the cone $\mathcal{F}_3$, i.e., all the inequalities $h_{ki}(\Lambda) + h_{kj}(\Lambda) \leq h_i(\Lambda) + h_{ijk}(\Lambda)$ are strict. Figure 2 shows the tropical geometry for the lattice $\Lambda$.

**Remark 4.9.** For any lattice $\Lambda$, there exists a maximal (for inclusion) diagonal lattice inside $\Lambda$ and a minimal diagonal lattice containing $\Lambda$. Let us denote these two lattices by $\pi^a$ and $\pi^b$ respectively, where $a \geq b \in \mathbb{Z}^d$. So, we have the inclusions $\pi^a \subset \Lambda \subset \pi^b$. It is not difficult to see that the region of linearity corresponding to the monomial $v_1 + \cdots + v_d - h(\Lambda)$ in the tropical polynomial $\varphi_\Lambda(v)$ is the orthant $\mathbb{R}^d_{\geq a} := \{x \in \mathbb{R}^d : x \geq a\}$. Similarly, the region of linearity corresponding to the monomial 0 is the orthant $\mathbb{R}^d_{\leq b} := \{x \in \mathbb{R}^d : x \leq b\}$. From this, we can deduce the recursive relation

$$h_{[d]}(\Lambda) = h_{[d-1]}(\Lambda) + a_d.$$ 

This iterative way of computing the entropy map $H(\Lambda)$ is slightly more efficient than the algorithm on page 11 where we have to compute the whole Hermite form of $\Lambda_I$ for every $I \subset [d]$. This iterative algorithm is the one implemented in the website linked to on page 3.
5. The entropy map on nonarchimedean fields

In this section we generalize some of the results in Section 3 to the case where $K$ is a field with a nonarchimedean valuation.

When the residue field $k$ of $K$ is infinite or the valuation group $\Gamma$ is dense in $\mathbb{R}$, the probabilistic framework we had in Section 3 is no longer valid. More precisely, we lose the local compactness and we no longer necessarily have a Haar measure on $K$.

We define the entropy map $H$ of a lattice as in Section 3, i.e., for any $I \subset [d]$,

$$h_I(\Lambda) := \min_{|J| = |I|} \text{val}(\det A_{I \times J}),$$

where $A$ is a representative of $\Lambda$. We can still define a Hermite representative of $\Lambda$.

**Definition 5.1.** Every lattice $\Lambda$ in $K^d$ has a representative $A$ in Hermite normal form, i.e., a matrix $A$ in $\text{GL}_d(K)$ satisfying the following conditions:

(i) $A$ is lower diagonal.

(ii) For any $1 \leq j < i \leq d$, we have either $\text{val}(A_{i,j}) < \text{val}(A_{j,j})$ or $A_{i,j} = 0$.

The same argument used in Theorem 4.2 can be used again to show that the image of $H$ still lies in the supermodular cone $\mathcal{F}_d$. In this setting, however, since the valuation group can be dense in $\mathbb{R}$, the image is not necessarily in $\mathcal{F}_d \cap \mathbb{Z}^{2d-1}$. As in Section 3, the map $H$ fails to be surjective when $d \geq 3$. The algorithm we provide a link to on page 3 computes the map $H$ when $K = \mathbb{Q}[[t]]$ is the field of Puiseux series over $\mathbb{Q}$.

Now we show that the only distribution on the field Laurent series $K = \mathbb{R}((t))$ that satisfies the definition suggested in [9, Definition 4.1] is the Dirac measure at 0. Let $\mathbb{P}$ be such a probability measure. First, we
recall that if $X$ is a random variable with distribution $\mathbb{P}$, then for any $a \in \mathbb{R}^\times$ the random variables $X$ and $aX$ have the same distribution, and we write $X \overset{d}{=} aX$. In particular, for any $a \in \mathbb{R}^\times$ we have $X \overset{d}{=} aX$.

**Proposition 5.2.** The probability distribution $\mathbb{P}$ is the Dirac measure at 0.

**Proof.** We can write the power series expansion of $X$ as $X = X_0 t^V + X_1 t^{V+1} + \cdots$, where $V \in \mathbb{Z}$ is the random valuation of $X$. Hence for $a \in \mathbb{R}^\times$ we have $aX = aX_0 t^V + aX_1 t^{V+1} + \cdots$, and we deduce that $X_k \overset{d}{=} aX_k$ for any $k \geq 0$ and $a \in \mathbb{R}^\times$. We then deduce that $X_k = 0$ almost surely for all $k \geq 0$. Hence $X = 0$ almost surely which finishes the proof. □

Using a variant of this argument, it is not difficult to see that a similar problem would arise when we try to define Gaussian measures by orthogonality for all fields listed in Example 2.2. It is not immediately clear how to fix this problem and find a suitable definition for Gaussian measures on nonarchimedean valued fields.

**Problem 5.3.** Is there a suitable definition for Gaussian measures on the fields listed in Example 2.2?

**Remark 5.4.** We can define a probability measure on $\mathbb{R}^d$ induced by $\Lambda$ via its tail distribution $Q_\Lambda$ as in Section 3. One can see that the support of this distribution is $\text{trop}(\Lambda) := \text{val}(\Lambda \cap (\mathbb{K}^\times)^d)$; the image under valuation of points in $\Lambda$ with no zero coordinates. This is in general a polyhedral complex in $\mathbb{R}^d$ where each edge is parallel to some $e_I := \sum_{i \in I} e_i$. Figure 3 is a drawing of $\text{trop}(\Lambda)$ for a lattice in $K^3$ when $K = \mathbb{K}$ (the field of generalized Puiseux series).

To conclude this section we give a partial answer for Problem 4.5 when $d = 2, 3$ and the valuation group is $\mathbb{R}$.

**Proposition 5.5.** For $d = 2$, the image $\text{im}(H)$ of the entropy map $H$ is exactly $\mathcal{F}_2$.

**Proof.** For $\Lambda$ with representative $\left( \begin{array}{cc} a & 0 \\ b & \rho_{a+b} \end{array} \right)$ with $a, b \in \mathbb{R}$ and $\delta \geq 0$, we have $H(\Lambda) = (a, b, a+b+\delta)$. So $H$ is indeed surjective onto $\mathcal{F}_2$. □

For $d = 3$, the cone $\mathcal{F}_3 \subset \mathbb{R}^7$ has a lineality space $\mathcal{L}_3$ of dimension 3. Since both $\mathcal{F}_3$ and $\text{im}(H)$ are stable under translations in $\mathcal{L}_3$ (see Remark 4.6 and Lemma 3.6 on diagonal scaling of lattices), they are
fully determined by their projection onto a complement of $\mathcal{L}_3$. Let us write vectors $x$ of $\mathbb{R}^7$ in the form
\[
x = (x_1, x_2, x_3; x_{12}, x_{13}, x_{23}, x_{123}),
\]
and let us project $\mathcal{S}_3$ and $\text{im}(H)$ on the linear space $W \subset \mathbb{R}^7$ of vectors of the form
\[
x = (0, x_2, x_3; 0, x_{13}, x_{23}; 0),
\]
which is a complement of $\mathcal{L}_3$ in $\mathbb{R}^7$. We write a vector of $W$ as $(x_2, x_3; x_{13}, x_{23})$ or simply as $(w, x, y, z)$ to simplify notation. Let us denote by $\mathcal{P}$, $\mathcal{C}$ the projections of $\text{im}(H)$ and $\mathcal{S}_3$ respectively onto the space $W$. From Section 4, we clearly have $\mathcal{P} \subset \mathcal{C}$.

The projection $\mathcal{C}$ of $\mathcal{S}_3$ onto $W$ is a polyhedral cone that does not contain any lines. In the language of polyhedral geometry, this is called a pointed cone. Moreover, the dimension of this projection is 4. It is defined in $W$ by the inequalities
\[
\begin{align*}
  w &\leq 0 \quad \text{if } x \leq y, \\
  w + x &\leq z \quad \text{if } y \leq 0, \\
  z &\leq w \quad \text{if } y + z \leq x.
\end{align*}
\]
(8)

This defines $\mathcal{C}$ as a pointed cone over a bipyramid (see Figure 4).

On the other hand, any lattice $\Lambda$ in $\mathbb{K}^3$ can be represented, up to diagonal scaling, by a representative with Hermite form of the shape
\[
\begin{pmatrix}
  1 & 0 & 0 \\
  a & 0 & 0 \\
  b & c & 1
\end{pmatrix},
\]
The entropy vector of a lattice $\Lambda$ with such a Hermite normal form is of the shape
\[
H(\Lambda) = (0, h_2, h_3; 0, h_{13}, h_{23}; 0).
\]
This corresponds to the projection of $\text{im}(H)$ to $W$ parallel to $\mathcal{L}_3$. So the projection $\mathcal{P}$ of $\text{im}(H)$ onto $W$ is the set
\[
\mathcal{P} = \left\{ H(\Lambda) : \Lambda \text{ given by a matrix of the shape } \begin{pmatrix} 1 & 0 & 0 \\ * & 0 & 0 \\ * & * & 1 \end{pmatrix} \text{ in } \mathbb{K}^{3 \times 3} \right\}.
\]
For a lattice $\Lambda$ with representative
\[
A = \begin{pmatrix}
  1 & 0 & 0 \\
  a & 1 & 0 \\
  b & c & 1
\end{pmatrix},
\]
such that $a, b, c \in \mathbb{K}$ with negative or zero valuation (see Definition 4.1), the point $H(\Lambda)$ in $W$ is given by
\[
\begin{align*}
  w &= h_2(\Lambda) = \text{val}(a), \\
  x &= h_3(\Lambda) = \min(\text{val}(b), \text{val}(c)), \\
  y &= h_{13}(\Lambda) = \text{val}(c), \\
  z &= h_{23}(\Lambda) = \min(\text{val}(ac - b), \text{val}(a)).
\end{align*}
\]
Figure 4. The intersection of the affine hyperplane $x + y + z + w + 1 = 0$ with $\mathcal{P}$ (left) and with $\mathcal{C}$ (right). The facets shown on the right are $\perp 2$ (red), $\perp 3$ (blue), $\perp 3$ (green), $\perp 2|3$ (orange), $\perp 3|2$ (yellow) and $2 \perp 3|1$ (gray).

One can check that, for any choice of $a, b, c \in \mathbb{K}$ with negative or zero valuation, the above coordinates satisfy the inequalities in (8). With the constraints on the valuations of $a, b, c$, and from this parametric representation of $\mathcal{P}$, we can see that points of $\mathcal{P}$ have to satisfy the inequalities

$$\begin{align*}
w &\leq 0, \\
x &\leq y, \\
y &\leq 0.
\end{align*}$$

The only part that remains to determine is the inequalities involving the last variable $z$. The ambiguity comes from the fact that cancellations can happen in $ac - b$ which might affect $\text{val}(ac - b)$ and hence also $z$. But, separating the cases where $\text{val}(ac) = \text{val}(b)$ and $\text{val}(ac) \neq \text{val}(b)$, we get the following three sets of inequalities that describe $\mathcal{P}$ as a polyhedral complex:

$$\begin{align*}
\left\{ \begin{array}{l}
w \leq 0, \\
x \leq w + y, \\
y \leq 0, \\
z = x,
\end{array} \right. & \quad \left\{ \begin{array}{l}
w \leq 0, \\
x \leq y, \\
y \leq 0, \\
y + w \leq x, \\
z = y + w,
\end{array} \right. & \quad \left\{ \begin{array}{l}
w \leq 0, \\
x \leq y, \\
y \leq 0, \\
x = y + w, \\
z \leq w, \\
x \leq z.
\end{array} \right.
\end{align*}$$

We can then see that $\mathcal{P}$ is a polyhedral fan of dimension 3 inside $\mathcal{C}$. More precisely, $\mathcal{P}$ is the union of three pointed polyhedral cones of dimension 3 inside $\mathcal{C}$ which is a cone of dimension 4. Figure 4 depicts the intersections of $\mathcal{P}$ and $\mathcal{C}$ with the hyperplane $w + x + y + z + 1 = 0$ (slicing the pointed cones with a hyperplane).

Corollary 5.6. The entropy map $H : \text{GL}_d(\mathbb{K}) / \text{GL}_d(\mathcal{C}_{\mathbb{K}}) \to \mathcal{F}_d$ is not surjective when $d \geq 3$.

We expect this result to hold in every dimension, i.e., the image $\text{im}(H)$ is a polyhedral fan whose facets are polyhedral cones of dimension $\frac{1}{2}d(d + 1)$ inside $\mathcal{F}_d$, which is of dimension $2^d - 1$. 
6. Conclusion

In conclusion, there are many similarities between the classical theory of Gaussian distributions on euclidean spaces and the theory of Gaussian measures on local fields as defined by Evans in [9]. In this paper we have exhibited another similarity in terms of differential entropy. This gives reason to think that the suitable nonarchimedean analog of the positive definite cone is indeed the set of lattices, or more precisely, in the language of group theorists, the Bruhat–Tits building for SL. This analogy can still be carried out for nonarchimedean valued fields in general. However, when the field \( K \) has a dense valuation group or an infinite residue field, we lose the probabilistic interpretation and thus also the notion of entropy.

Acknowledgements

The polyhedral geometry images in Figures 2, 3 and 4 were drawn using Polymake [12]. The author would like to thank the Max Planck Institute for Mathematics in the Sciences for the generous hospitality while working on this project. He would also like to thank Bernd Sturmfels and Ian Le for valuable mathematical discussions. The author is grateful to Avinash Kulkarni for the numerous and valuable exchanges while writing this paper. Many thanks also to Rida Ait El Mansour and Adam Quinn Jaffe for their remarks on early drafts of this manuscript. Finally, the author thanks the referee for valuable comments and remarks.

References


Received 2021-01-03. Revised 2022-01-18. Accepted 2022-03-10.

YASSINE EL MAAZOUZ: yassine.el-maazouz@berkeley.edu
Department of Statistics, University of California, Berkeley, CA, United States