Types of Markov Fields and Tilings

Yuliy Baryshnikov and Jaroslaw Duda and Wojciech Szpankowski, *Fellow, IEEE*

***Abstract*—The method of *types* is one of the most popular techniques in information theory and combinatorics. However, thus far the method has been mostly applied to one-dimensional Markov processes, and it has not been thoroughly studied for general Markov fields. Markov fields over a finite alphabet of size $m \geq 2$ can be viewed as models for multi-dimensional systems with local interactions. The locality of these interactions is represented by a *shape* $\mathcal{S}$ while its marking by symbols of the underlying alphabet is called a *tile*. Two assignments in a Markov field have the same *type* if they have the same empirical distribution, i.e., if they have the same number of tiles of a given type. Our goal is to study the growth of the number of possible Markov field *types* in either a d-dimensional box of lengths $n_1, \ldots, n_d$ or its cyclic counterpart, a d-dimensional torus. We relate this question to the enumeration of nonnegative integer solutions of a large system of Diophantine linear equations called the *conservation laws*. We view a Markov type as a vector in a $D = m^{|\mathcal{S}|}$ dimensional space and count the number of such vectors satisfying the conservation laws, which turns out to be the number of integer points in a certain polytope. For the torus this polytope is of dimension $\mu = D - 1 - \mathrm{rk}(\mathbf{C})$ where $\mathrm{rk}(\mathbf{C})$ is the number of linearly independent conservation laws $\mathbf{C}$. This provides an upper bound on the number of types. Then we construct a matching lower bound leading to the conclusion that the number of types in the torus Markov field is $\Theta(N^\mu)$ where $N = n_1 \cdots n_d$. These results are derived by geometric tools including ideas of discrete and convex multidimensional geometry.**

Index Terms: Markov fields, Markov types, conservation laws, linear Diophantine equations, enumerative combinatorics, analytic and discrete geometry.

I. Introduction

The *method of types* is one of the most popular and useful techniques in information theory and combinatorics. Two sequences of equal length are of the same type if they have identical empirical distributions, thus sequences of the same type are assigned the same probability by all distributions in a given class [6], [25]. The method of types was known for some time in probability and statistical physics. But only in the 1970's Csiszár and his group developed a general method and made it a basic tool of information theory of discrete memoryless sources [5], [6]; see also [4], [8], [11], [17], [25], [31]. The method of types is used in a myriad of applications [6], from the minimax redundancy [11] to simulation of information sources [18]. However, thus far this method has been studied only for *one-dimensional* processes, mostly Markov [11], [12], [31] but also general stationary ergodic processes [25]. Here we investigate types of *Markov fields* (Bayesian networks, Gibbs fields and/or factor graphs) [14], [29] that find applications ranging from sensor networks [22] to images to information retrieval [19]. For example, they are used in [29] to analyze finite covers of factor graphs to estimate the behavior of sum-product algorithms for LDPC decoding, and to approximate the matrix permanent. The focus of [29], however, is on the usage of the methods of types, while we are focusing on the quintessential question of *characterizing the set of possible types*. We develop here a novel approach to Markov field types based on multidimensional discrete and analytic geometry to study this important and intricate problem that has been left open for too long.

The Markov/Gibbs fields are governed by *local interactions*, parameterized by some collections of neighboring sites. We shall call these collections of sites *plaques*; they cover the domain of the Markov fields. If the domain of the underlying Markov field is a subset of Euclidean space (the case in all of our applications), these plaques are in fact *shifts*, or *displacements* of the same *shape* $\mathcal{S}$. For example, for one-dimensional Markov sources of order r discussed below, the shape $\mathcal{S}$ is just the interval $\mathcal{S} = \{0, 1, \ldots, r\}$, and then $\mathfrak{p} = \mathcal{S} + \mathbf{s}$ where $\mathbf{s}$ is a displacement vector (see also [30]). A *marking* of a shape with symbols of the alphabet $\mathcal{A}$ is called a tile $\mathfrak{t}$. In other words, a tile $\mathfrak{t} : \mathcal{S} \to \mathcal{A}$ is an assignment of alphabet letters to all cells of the plaque (one can think of a tile as a labeled plaque).

In this paper, the fields take values in the *finite* alphabet $\mathcal{A} = \{1, 2, \ldots, m\}$. The domains $\mathcal{D}$, where the fields are defined, will be rectangular subsets of a d-dimensional integer *lattice* $\mathbb{Z}^d$, subject to some boundary conditions. We consider here either the free boundary conditions called a *box* and denoted as $\mathcal{I}$ or periodic (in all d dimensions) boundary conditions referred to as a *torus* and denoted as $\mathcal{O}$. The set of all possible configurations on the domain $\mathcal{D}$ (i.e., functions $\mathcal{D} \to \mathcal{A}$) is denoted as $\mathtt{Conf}(\mathcal{D})$. The realizations of a Markov field are written as $\mathbf{x} \equiv \mathbf{x}^{\mathbf{n}} \in \mathtt{Conf}(\mathcal{D})$, where $\mathbf{n} = (n_1, \ldots, n_d)$ is the size of a torus or a box. Finally, the set of all tiles is denoted as $\mathfrak{T} \equiv \mathtt{Conf}(\mathcal{S})$.

A. One-dimensional Fields

In order to gently introduce the questions we address here for Markov fields and their types, we start with a

Parts of this paper were presented at the 2014 IEEE International Symposium on Information Theory, Honolulu.

Y. Baryshnikov is with the Dept. of Mathematics, and Electrical and Computer Engineering, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA (email: ymb@illinois.edu). His work is supported by Grants from ARPA HR0011-07-1-0002 and ONR N000140810668.

J. Duda was with the Center for Science of Information, Purdue University, W. Lafayette, IN 47907, USA. His current address is Institute of Computer Science and Computer Mathematics, Jagiellonian University, Poland (email: dudajar@gmail.com).

W. Szpankowski is with the Department of Computer Science, Purdue University, IN 47907, USA (e-mail: spa@cs.purdue.edu); also with the Faculty of Electronics, Telecommunications and Informatics, Gdańsk University of Technology, Poland. His work was supported by NSF Center for Science of Information (CSoI) Grant CCF-0939370, and in addition by NSF Grants DMS-0800568, CCF-0830140, and CCF-1524312, NIH Grant 1U01CA198941-01, and the NCN grant, grant UMO-2013/09/B/ST6/02258.

one-dimensional Markov chain over a finite alphabet $\mathcal{A} = \{1, 2, \ldots, m\}$. Let us write $\mathbf{x}^n = x_1 \ldots x_n \in \mathcal{A}^n$ for a sequence of length n generated by a Markov source. For Markov sources of order $r = 1$ we have two equivalent representations for the probability $\mathbf{P}(\mathbf{x}^n)$ of $\mathbf{x}^n$:

$$\mathbf{P}(\mathbf{x}^n) = \mathbf{P}(x_1) \prod_{i=2}^{n} \mathbf{P}(x_i|x_{i-1}) =$$

$$= \mathbf{P}(x_1) \prod_{(i,j)\in\mathcal{A}^2} p_{ij}^{\mathbf{T}(ij)}, \qquad \sum_{x_1} \mathbf{P}(x_1) = 1,$$

where p_{ij} is the transition probability from $i \in \mathcal{A}$ to $j \in \mathcal{A}$, and the frequency count $\mathbf{T}(ij)$ is the number of pairs (ij) in the sequence $\mathbf{x}^n$. If we denote $\mathfrak{t} = (ij)$ and $\mathfrak{T} = \mathcal{A}^2$, then the previous equation can be written as

$$\mathbf{P}(\mathbf{x}^n) = \mathbf{P}(x_1) \prod_{\mathfrak{t}\in\mathfrak{T}} p_{\mathfrak{t}}^{\mathbf{T}(\mathfrak{t})}. \tag{1}$$

Similarly, for one-dimensional Markov sources of order r we have

$$\mathbf{P}(\mathbf{x}^n) = \mathbf{P}(x_1^r) \prod_{i=r+1}^{n} \mathbf{P}(x_i|x_{i-1} \ldots x_{i-r}) =$$

$$= \mathbf{P}(x_1^r) \prod_{\mathfrak{t}\in\mathfrak{T}} p_t^{\mathbf{T}(\mathfrak{t})}, \qquad \sum_{x_1^r} \mathbf{P}(x_1^r) = 1, \tag{2}$$

where $\mathfrak{t} = (i_1, \ldots, i_{r+1}) \in \mathfrak{T} := \mathcal{A}^{r+1}$ and $\mathbf{T}(\mathfrak{t})$ counts the appearances of $\mathfrak{t}$ in $\mathbf{x}^n$. In just introduced notation, we have in this case $\mathcal{S} = \{0, 1, \ldots, r\}$, plaque $\mathfrak{p} = \mathcal{S} + \mathbf{s}$, and $|\mathfrak{T}| = m^{r+1}$, where $\mathbf{s}$ is a displacement vector.

B. Gibbs Distributions

It is well known that the distribution $\mathbf{P}(\mathbf{x})$ of a Markov field $\mathbf{x}$ can be rewritten as a Gibbs distribution

$$\mathbf{P}(\mathbf{x}) = Z^{-1} \prod_{\mathfrak{p}} \mathbf{w}(\mathbf{x}|_{\mathfrak{p}}), \tag{3}$$

where $\mathbf{w}$ is a *weight* – a (nonnegative) function (or equivalently vector) on configurations of plaques (shape $\mathcal{S}$ shifted to some position), and $\mathbf{x}|_{\mathfrak{p}}$ is a restriction of $\mathbf{x}$ to $\mathfrak{p}$. The normalization factor Z is the *Gibbs partition function*

$$Z = \sum_{\mathbf{x}} \prod_{\mathfrak{p}} \mathbf{w}(\mathbf{x}|_{\mathfrak{p}}). \tag{4}$$

Observe that (3) is an instance of the Hammersley–Clifford Theorem [20].

In the case when all the plaques have the same shape and the weights are *translation invariant* (i.e., the weight $\mathbf{w}(\mathbf{x}|_{\mathfrak{p}})$ depends only on the restriction $\mathfrak{t} = \mathbf{x}|_{\mathfrak{p}}$, but not on the position of the plaque), the product in (4) can be rewritten as

$$\prod_{\mathfrak{p}} \mathbf{w}(\mathbf{x}|_{\mathfrak{p}}) = \mathbf{w}^{\mathbf{T}}, \tag{5}$$

where $\mathbf{w} = (w_{\mathbf{x}_1}, \ldots, w_{\mathbf{x}_K})$ for some K is the vector of weights, with the convention that $\mathbf{a}^{\mathbf{b}} = \prod_k a_k^{b_k}$ for two vectors $\mathbf{a} = (a_1, \ldots, a_K), \mathbf{b} = (b_1, \ldots, b_K)$. Here, the vector $\mathbf{T} = \{\mathbf{T}(\mathfrak{t})\}_{\mathfrak{t}\in\mathfrak{T}}$ is the *type* of the configuration $\mathbf{x}$, i.e., the number of the plaques $\mathfrak{p}$ (obtained by shifting the shape $\mathcal{S}$) for which the restriction $\mathbf{x}|_{\mathfrak{p}}$ is the same as $\mathfrak{t}$.

Observe that the partition function (4) can be rewritten as

$$Z = \sum_{\mathbf{T}} M(\mathbf{T})\mathbf{w}^{\mathbf{T}}, \tag{6}$$

where the summation is over all *types* $\mathbf{T}$, that is, the vector of numbers $\{\mathbf{T}(\mathfrak{t})\}_{\mathfrak{t}\in\mathfrak{T}}$ of plaques $\mathfrak{p} = \mathbf{s} + \mathcal{S}$ such that the restriction of the configuration $\mathbf{x}$ to $\mathfrak{p}$ has the labeling $\mathfrak{t}$, and $M(\mathbf{T})$ is the number of configurations $\mathbf{x} \in \mathtt{Conf}(\mathcal{O})$ (or in $\mathtt{Conf}(\mathcal{I})$) having the type $\mathbf{T}$. This reformulation (6) allows one to decouple the effects of the weights $\mathbf{w}$ and combinatorics of the model encoded by the shape of the domain $\mathcal{O}$ or $\mathcal{I}$, and the shape of the plaques $\mathcal{S}$.

C. Counting the Types

The preceding discussion leads naturally to the following two questions:

1) For a given type, i.e., for the collection of counts $\{\mathbf{T}(\mathfrak{t})\}_{\mathfrak{t}\in\mathfrak{T}}$, how many fields $\mathbf{x}^n$ realize it (i.e., what is $M(\mathbf{T})$ for a given type $\mathbf{T}$)?
2) How many *distinct types* are there, that is, what is the size of the set of types $\mathbf{T}$ for which $M(\mathbf{T}) > 0$?

In the language of Markov sources, the types $\{\mathbf{T}(\mathfrak{t})\}_{\mathfrak{t}\in\mathfrak{T}}$, as we define them, specialize to the familiar *Markov types*: for a one-dimensional Markov source, the type just encodes the number of times the transition $i \to j$ ($i, j \in \mathcal{A}$) is observed in a sequence. We should point out that the condition that a given type can be observed in a Markov trajectory is equivalent to a certain multigraph representing it to be *Eulerian*, as discussed in [12]. The question, for a given Eulerian type $\mathbf{T}$ to determine the number of trajectories having this type is, in our notation, the question of finding $M(\mathbf{T})$.

These two problems (i.e., number of sequences of a given type and number of types) were studied quite extensively in the past for one-dimensional Markov sources. The number of sequences of a given Markov type was first addressed by Whittle [31] and then re-established by analytic method in [11]. A precise evaluation of the number of Markov types was recently presented in [12] (see also [17] for tree models).

In this paper, we address a more general, and much harder problem: the *enumeration of Markov field types* (i.e., the number of distinct empirical distributions of tile counts), that can be realized by a "trajectory" $\mathbf{x} \in \mathtt{Conf}(\mathcal{O})$. Let us be more precise. For dimension $d \in \mathbb{N}$, let $\mathbf{n} = (n_1, n_2, \ldots, n_d)$ and $N := n_1 \cdot n_2 \cdot \ldots \cdot n_d$. Define the box $\mathcal{I}_{\mathbf{n}} = I_{n_1} \times I_{n_2} \times \ldots \times I_{n_d} \subset \mathbb{Z}^d$ with $I_n := \{0, 1, \ldots, n-1\}$ on which the underlying Markov field is defined. We mostly work here with the *rectangular torus*, i.e., the fields on $\mathbb{Z}^d$ subject to $\mathbf{n}$-periodic boundary conditions (see Figure 1). Our results remain valid for the general case, where the periodicity lattice is not necessarily rectangular. In the 1-dimensional (1D) case analyzed in [12] these periodic conditions translate into the *cyclic sequences* of symbols in $\mathcal{A}$, that is, sequences $\mathbf{x}^n$ in which $x_{k+n} = x_k$.

Now, the *tile count*, or *type* of the field $\mathbf{x}$ is a function $\mathbf{T} : \mathfrak{T} \to \mathbb{N}$ counting how often each tile occurs in the field

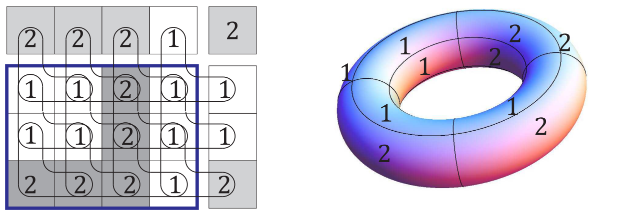

Figure 1. Illustration of cyclic fields: here $\mathbf{x}$ is defined on the torus $\mathcal{O}_{4,3}$ that is constructed from the box on the left by gluing the left and the right as well as the top and the bottom edges.

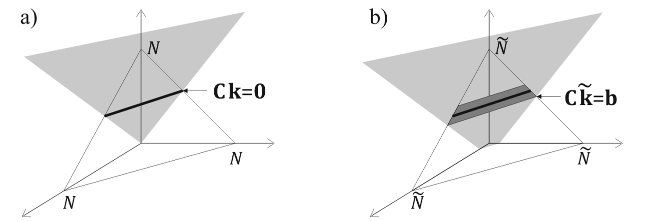

Figure 2. Geometry of type vectors for (a) torus $\{\mathbf{T}(\mathfrak{t})\}$ and (b) box $\{\tilde{\mathbf{T}}(\mathfrak{t})\}$. Here the gray area denotes the cone $\mathcal{C}$ of non-negative type vectors satisfying the conservation laws while the intersection with either the simplex $\{\sum_{\mathfrak{t}} \mathbf{T}(\mathfrak{t}) = N\}$ or the simplex $\{\sum_{\mathfrak{t}} \mathbf{T}(\mathfrak{t}) = \tilde{N}\}$ representing a polyhedron is displayed in bold.

$\mathbf{x}$, that is,

$$\mathbf{T}(\mathfrak{t}) = |\{\mathbf{s} \in \mathcal{D} \ : \ \mathcal{S} + \mathbf{s} \subset \mathcal{D}, \text{and } \mathbf{x}|_{\mathcal{S}+\mathbf{s}} = \mathfrak{t}\}| \quad (7)$$

where $f|_B$ denotes the restriction of a function f to a smaller domain B, and $\mathbf{s}$ is a displacement vector.

While tilings and their asymptotic counting are discussed in many references [1], [13], [15], [21], our problem is distinctly different: these references are concerned with (asymptotic) evaluation of what we call $M(\mathbf{T})$. Here we address the issue of the *support* of the function M, especially of its size:

$$\mathcal{P}_{\mathbf{n}} = \mathcal{P}_{\mathbf{n}}(\mathcal{A}, \mathcal{S}) =$$

$$\{\mathbf{T} \in \mathbb{Z}^{\mathfrak{T}} : \text{exists } \mathbf{x} \in \mathtt{Conf}(\mathcal{D}), \text{ such that } \mathbf{x} \text{ is of type } \ \mathbf{T}\}. \quad (8)$$

The cardinality of $\mathcal{P}_{\mathbf{n}}$, i.e., the number of realizable *types*, is our main concern in this paper. While the question of understanding the structure of the set of types for multi-dimensional fields (lattice) is very natural and important one, we could not find any relevant literature (however, see [29]), beyond the 1D lattice situation.

D. Overview of the Results

We shall view the types $\{\mathbf{T}(\mathfrak{t})\}_{\mathfrak{t}\in\mathfrak{T}}$ as a vector of dimension $D := |\mathfrak{T}| = m^{|\mathcal{S}|}$ equal to the number of possible tiles $\mathfrak{t} \in \mathfrak{T}$. Clearly, $\mathbf{T}(\mathfrak{t}) \geq 0$ for all $\mathfrak{t} \in \mathfrak{T}$. However, this vector satisfies a number of *equality* constraints that have a major impact on the cardinality of $\mathcal{P}_{\mathbf{n}}$. First of all, one has the normalization condition

$$\sum_{\mathfrak{t}} \mathbf{T}(\mathfrak{t}) = \mathtt{I}(\mathcal{D}), \quad (9)$$

where $\mathtt{I}(\mathcal{D})$ is the number of different plaques $\mathfrak{p} = \mathbf{s} + \mathcal{S}$ in $\mathcal{D}$. It is quite obvious for the torus that $\mathtt{I} = N = n_1 \cdot \ldots \cdot n_d$. Further, in order to tile a torus the number of tiles "ending" with a subtile $\mathfrak{t}' : \mathcal{S}' \to \mathcal{A}$ for some subshape $\mathcal{S}' \subset \mathcal{S}$ must be equal to the number of tiles that "begin" with $\mathfrak{t}'$ (see Figure 1). This leads, as in the 1D case, to what we call the *conservation laws* (discussed in depth in Section II):

$$\sum_{\mathfrak{t}:\mathfrak{t}|_{\mathcal{S}_1'} = \mathfrak{t}'} \mathbf{T}(\mathfrak{t}) = \sum_{\mathfrak{t}:\mathfrak{t}|_{\mathcal{S}_2'} = \mathfrak{t}'} \mathbf{T}(\mathfrak{t}) \quad (10)$$

for all pairs of subshapes $\mathcal{S}_1', \mathcal{S}_2' \subset \mathcal{S}$ such that $\mathcal{S}_2' = \mathcal{S}_1' + \mathbf{s}$ for some $\mathbf{s}$, and $\mathfrak{t}' : \mathcal{S}' \to \mathcal{A}$.

The system of equations (9)–(10) constitutes a *linear system of Diophantine equations* in $\mathbb{Z}^D$. We denote by $\mathcal{F}_{\mathbf{n}} :=$ $\mathcal{F}_{\mathbf{n}}(\mathcal{A}, \mathcal{S})$ the set of nonnegative integer solutions to (9)–(10). Clearly, $|\mathcal{P}_{\mathbf{n}}| \leq |\mathcal{F}_{\mathbf{n}}|$ since all types in $\mathcal{P}_{\mathbf{n}}$ satisfy the conservation laws, and thus lie in $\mathcal{F}_{\mathbf{n}}$. However, we will see that unlike the 1D situation, these sets are very different.

As we discussed, little is known about the set of realizable types in higher dimensions. Let us briefly survey the available 1D results where we set $N = n_1 := n$. In [12] an analytic approach was used to enumerate precisely $\mathcal{F}_n$ for $d = 1$. Another analytic approach is suggested in Stanley [27], however, it allows only to find the order of growth. We remark here that extending analytic techniques of [12] to estimate asymptotically $|\mathcal{F}_{\mathbf{n}}|$ is in general quite complicated, however, in Section II we discuss it in some details. Furthermore, for the $d = 1$ case $|\mathcal{P}_n| \sim |\mathcal{F}_n|$ as $n \to \infty$ (meaning: $\lim_{n\to\infty} |\mathcal{P}_n|/|\mathcal{F}_n| = 1$). This seems not to hold any longer for the multidimensional case where the set of types $\mathcal{P}_{\mathbf{n}}$ is asymptotically smaller: $\lim_{n\to\infty} |\mathcal{P}_{\mathbf{n}}|/|\mathcal{F}_{\mathbf{n}}| < 1$. Thus we can only establish an upper bound on the size of the set of types through $\mathcal{F}_{\mathbf{n}}$, and we propose another approach to find a lower bound.

To analyze the cardinality of $\mathcal{F}_{\mathbf{n}}$ and, ultimately, $\mathcal{P}_{\mathbf{n}}$ we need to understand the geometry of D-dimensional count vector $\{\mathbf{T}(\mathfrak{t})\}_{\mathfrak{t}\in\mathfrak{T}}$. In particular, we must estimate the dimension of the affine subspace spanned by $\mathcal{F}_{\mathbf{n}}$. To accomplish it we shall write the conservation law (10) as $\mathbf{C} \cdot \mathbf{T} = \mathbf{0}$ where $\mathbf{C}$ is a matrix describing a system of conservation laws (10), or, perhaps, its submatrix of the same rank. This allows us to define the cone $\mathcal{C}$ (recall that a set $\mathcal{C}$ is a cone if $\mathbf{T} \in \mathcal{C}$ implies $\lambda\mathbf{T} \in \mathcal{C}$ for all $\lambda \geq 0$):

$$\mathcal{C} \equiv \mathcal{C}(\mathcal{A}, \mathcal{S}) = \{\mathbf{T} \in \mathbb{R}^D_{\geq 0} \ : \ \mathbf{C} \cdot \mathbf{T} = \mathbf{0}\},$$

and the corresponding commutative monoid (a "lattice analogue of a cone")

$$\mathcal{C}_{\mathbb{Z}} := \mathcal{C} \cap \mathbb{Z}^D.$$

Then

$$\mathcal{F}_{\mathbf{n}} \equiv \mathcal{F}_{\mathbf{n}}(\mathcal{A}, \mathcal{S}) = \{\mathbf{T} \in \mathcal{C}_{\mathbb{Z}} : \sum_{\mathfrak{t}} \mathbf{T}(\mathfrak{t}) = N\}. \quad (11)$$

The dimensionality (of the affine span of) $\mathcal{F}_{\mathbf{n}}$ depends on D and the set of constraints represented by the matrix $\mathbf{C}$. We shall show that $\mathcal{F}_{\mathbf{n}}$ lies in an affine subspace of dimension $\mu = D - 1 - \mathrm{rk}(\mathbf{C})$ where $\mathrm{rk}(\mathbf{C})$ is the rank of $\mathbf{C}$. This is illustrated in Figure 2(a). In our first main result Theorem 3 we present a precise characterization of $\mathrm{rk}(\mathbf{C})$.

Our ultimate goal, however, is to estimate the cardinality of the number of types $\mathcal{P}_{\mathbf{n}}$, that is, the number of *realizable* tiling types, or the number of distinct count vectors $\mathbf{T}$. We shall see that the Hausdorff distance between the *normalized* set $\hat{\mathcal{P}}_{\mathbf{n}} := \mathcal{P}_{\mathbf{n}}/N$ is close to $\hat{\mathcal{F}}_{\mathbf{n}} := \mathcal{F}_{\mathbf{n}}/N$ leading to our main Theorem 7 in which we establish that $|\mathcal{P}_{\mathbf{n}}| = \Theta(N^{\mu})$ where $\mu = D - 1 - \mathrm{rk}(\mathbf{C})$. However, unlike $d = 1$, where we proved $|\mathcal{P}_n| \sim |\mathcal{F}_n|$, in the multidimensional case $|\mathcal{F}_{\mathbf{n}}|$ seems not to be asymptotically equivalent to $|\mathcal{P}_{\mathbf{n}}|$ even if the growth of both is the same. Finally, we briefly discuss the non-cyclic Markov field types and provide an upper bound on the number of types in a box. In this case lack of cyclic boundary conditions introduces some imbalance in the conservation laws replacing $\mathbf{C} \cdot \mathbf{T} = \mathbf{0}$ by $\mathbf{C} \cdot \mathbf{T} = \mathbf{b}$ for some vector $\mathbf{b}$ as illustrated in Figure 2(b). This leads to an upper bound $O(N^{D-1}/(n)^{\mathrm{rk}(\mathbf{C})})$ on the number of types in the box $\mathcal{I}_{\mathbf{n}}$. However, whether this is the right growth for the number of types in the box case, remains an open question.

In summary, we prove that $|\mathcal{P}_{\mathbf{n}}| = \Theta(N^{\mu})$ for a torus. But this is only a starting point to study other interesting questions; for example, regarding the redundancy of a (universal) code [7] for Markov fields. To solve the redundancy problem, we first would require to generalize Rissanen's lower bound [23] to Markov fields, a quest that has been wanting for some time. Second, we would need to construct a code achieving this lower bound. By answering the second question first, we hope to actually formulate precisely a Rissanen-like lower bound for Markov fields. To accomplish it we need a conjecture regarding the number of fields of a given type (that is, about the size of a typical $M(\mathbf{T})$). In Conjecture 1 of Section II we propose that the number of "typical" fields of a given type is asymptotically $AN^{-\mu/2}2^{NH}$ where H is the entropy and A is a constant. Then, the results of this paper would imply that the average redundancy is

$$R_N = \frac{\mu}{2}\log N + o(\log N)$$

where μ is the dimension of the underlying affine space defined above. Thus the number of free parameters is exactly the dimension μ computed in this paper, which seems to be unknown till now.

Regarding methodology used to establish our results. As mentioned before, in [12] for $d = 1$ we applied an analytic approach through multidimensional Cauchy's integral. We still can use this approach for some simple shapes in the multidimensional case, as discussed in Section II. However, in the general case we have to switch to tools of discrete, convex, and analytic multidimensional geometry that somewhat resembles the method discussed in [27]. In particular, we need to understand how to count the number of lattice points in a polytope [2], [9]. This will allow us to find the number of nonnegative integer solutions of a linear system of Diophantine equations (i.e., conservation laws) that leads to the enumeration of the Markov field types.

The paper is organized as follows. In the next section we present our main results and some consequences. Most of the proofs are delayed till the last section, i.e., Section III.

II. Main Results

In this main section of the paper, we formulate the conservation laws and present our main results regarding the enumeration of types for Markov fields and tilings. We sprinkle this section with many examples to illustrate our definitions and results.

A. Basic Definitions and Examples

We start with some examples illustrating our definitions discussed in the introduction.

Example 1: *1D Markov chain.*
The first example we consider is the 1D case. For $d = 1$ the torus becomes a length $N = n_1$ cycle, fields are length N sequences with cyclic boundary condition: $x_{N+i} = x_i$. We are usually interested in the distribution of pairs, so the shape is $\mathcal{S} = \{0, 1\}$; for r-order Markov the shape is $\mathcal{S} = \{0, \ldots, r\}$. For example, when $m = 2$ and $\mathcal{O}_{10} = \{0, \ldots, 9\}$, let us consider a 1D (cyclic) sequence $\mathbf{x} = (1122111212)$. Clearly, $\mathbf{T}(21) = 3$ because pattern "21" (i.e., $\mathfrak{t}(0) = 2, \mathfrak{t}(1) = 1$) appears in $\mathbf{x}$ for 3 different shifts: $\mathbf{s} \in \{3, 7, 9\}$. Similarly, $\mathbf{T}(11) = 3$, $\mathbf{T}(12) = 3$, $\mathbf{T}(22) = 1$. Also, $\sum_{(ij)\in\{1,2\}} T(ij) = 10$. Note that we can view $\mathbf{T}$ as a $D = m^2 = 4$ dimensional vector.

Example 2: *2D Markov Field with the* $\mathtt{L}$ *Shape.*
Let $d = 2$. The torus is an $n_1 \times n_2$ rectangle with cyclic boundary conditions: $\mathbf{x}(i, j) = \mathbf{x}(i + n_1, j) = \mathbf{x}(i, j + n_2)$. Let us take the 4×3 torus $\mathcal{O}_{4,3} = \{0, 1, 2, 3\}\times\{0, 1, 2\}$. Fields assign an element from the alphabet $\mathcal{A} = \{1, 2\}$ to each point of this torus. For example, for the field

$$\mathbf{x} = \begin{bmatrix} 1121 \\ 1121 \\ 2221 \end{bmatrix}$$

we have $\mathbf{x}(0, 0) = 2,\ \mathbf{x}(0, 1) = 1,\ \mathbf{x}(1, 0) = 2$, but also $\mathbf{x}(4, 0) = \mathbf{x}(0, 3) = \mathbf{x}(4, 3) = \mathbf{x}(0, 0)$, where we use north-east coordinates with $\mathbf{x}(0, 0)$ in the lower left corner. The first shape we consider here is the simplest nontrivial L-shape: $\mathcal{S} = \{(0, 0), (0, 1), (1, 0)\}$. We find

$$\mathbf{T}\begin{pmatrix} 1 \\ 12 \end{pmatrix} = 2$$

because this pattern appears in $\mathbf{s} \in \{(3, 0), (1, 1)\}$ positions.

Example 3: *2D Markov Field with the Square* $\square$ *Shape.*
The second 2D shape we consider is a 2×2 square shape $\mathcal{S} = \{0, 1\} \times \{0, 1\}$. For the same torus $\mathcal{O}_{4,3}$ and field $\mathbf{x}^{\mathbf{n}}$ as in the previous example, we find

$$\mathbf{T}\begin{pmatrix} 11 \\ 11 \end{pmatrix} = 2$$

because this pattern appears in $\mathbf{s} \in \{(0, 1), (3, 1)\}$ positions. $\square$

B. Conservation Laws

Conservation laws are associated with the different ways we can embed a smaller shape $\mathcal{S}'$ into a larger shape $\mathcal{S}$. Recall that shapes are subsets of $\mathbb{Z}^d$, and thus our embeddings are just displacements by a vector in $\mathbb{Z}^d$. For example, the subshape $\mathcal{S}' = \{0,1\} \times \{0\}$ has six embeddings into $\mathcal{S} = \{0,1,2\} \times \{0,1,2\}$, that can be identified with $\mathbf{s} \in \{0,1\} \times \{0,1,2\}$ shifts: $\mathcal{S}' + \mathbf{s} \subset \mathcal{S}$.

Let $\boldsymbol{\varepsilon} : \mathcal{S}' \to \mathcal{S}$ be an embedding. A tile on $\mathcal{S}$ is a mapping $\mathfrak{t} : \mathcal{S} \to \mathcal{A}$, and composing it with $\boldsymbol{\varepsilon}$ we obtain a (sub)tile $\mathfrak{t}'$ on the smaller shape: restriction of $\mathfrak{t}$ to $\boldsymbol{\varepsilon}(\mathcal{S}')$ we denote as $\mathfrak{t}' = \boldsymbol{\varepsilon}^*(\mathfrak{t}) : \mathcal{S}' \to \mathcal{A}$. Further, recall that a type $\mathbf{T} : \mathcal{A}^{\mathcal{S}} \to \mathbb{N}$ is a vector with components indexed by tiles on $\mathcal{S}$. The mapping $\mathfrak{t} \mapsto \boldsymbol{\varepsilon}^*(\mathfrak{t})$ defines a mapping $\hat{\varepsilon} : \mathcal{T}_{\mathcal{S}} \to \mathcal{T}_{\mathcal{S}'}$, where $\mathcal{T}_{\mathcal{S}}$ is set of types for shape $\mathcal{S}$, taking a type $\mathbf{T}$ (on shape $\mathcal{S}$) into a type $\mathbf{T}' = \hat{\varepsilon}\mathbf{T}$ defined on $\mathcal{S}'$. Clearly,

$$(\hat{\varepsilon}\mathbf{T})(\mathfrak{t}') = \mathbf{T}'(\mathfrak{t}') := \sum_{\mathfrak{t}:\ \boldsymbol{\varepsilon}^*(\mathfrak{t}) = \mathfrak{t}'} \mathbf{T}(\mathfrak{t}) \qquad (12)$$

is just the sum of the counts $\mathbf{T}(\mathfrak{t})$ over all tiles $\mathfrak{t}$ such that their restriction to $\boldsymbol{\varepsilon}(\mathcal{S}')$ coincides with $\mathfrak{t}'$. Now, if there are two different embeddings $\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2 : \mathcal{S}' \to \mathcal{S}$, one obtains two types on $\mathcal{S}'$, namely $\hat{\varepsilon}_1\mathbf{T}$ and $\hat{\varepsilon}_2\mathbf{T}$ having the same subtile $\mathfrak{t}'$. The next lemma introduces a conservation law.

Lemma 1. *If the type* $\mathbf{T}$ *is the count vector for a configuration* $\mathbf{x}$ *on a torus* $\mathcal{O}_{\mathbf{n}}$*, then*

$$(\hat{\varepsilon}_1\mathbf{T})(\mathfrak{t}') - (\hat{\varepsilon}_2\mathbf{T})(\mathfrak{t}') = 0 \qquad (13)$$

where $\hat{\varepsilon}_1$ *and* $\hat{\varepsilon}_2$ *are mappings with the same corresponding subtile* $\mathfrak{t}'$ *(i.e.,* $\boldsymbol{\varepsilon}_1^*(\mathfrak{t}) = \boldsymbol{\varepsilon}_2^*(\mathfrak{t}) = \mathfrak{t}'$*) satisfying (12).*

This obvious lemma again generalizes the Eulerian condition (that every vertex has the same number of incoming and outgoing edges) in a multigraph describing types in the 1D situation, and is the starting point of our study. It is also equivalent to (10) discussed in the introduction.

Example 4: *Continuation of Example 1.*
Returning to the 1D case of Example 1, we have $\mathcal{S} = \{0,1\}$, $\mathcal{S}' = \{0\}$, with $\boldsymbol{\varepsilon}_1$ placing the node (subshape) 0 at 0, and $\boldsymbol{\varepsilon}_2$ at 1. In this case,

$$(\hat{\varepsilon}_1\mathbf{T})(1) = \mathbf{T}(11) + \mathbf{T}(12) =: \mathbf{T}(1*)$$

$$(\hat{\varepsilon}_1\mathbf{T})(2) = \mathbf{T}(21) + \mathbf{T}(22) =: \mathbf{T}(2*)$$

and

$$(\hat{\varepsilon}_2\mathbf{T})(1) = \mathbf{T}(11) + \mathbf{T}(21) =: \mathbf{T}(*1)$$

$$(\hat{\varepsilon}_2\mathbf{T})(2) = \mathbf{T}(12) + \mathbf{T}(22) =: \mathbf{T}(*2).$$

(We use the mnemonic $\mathbf{T}(a*)$ and like to denote the summation over the don't-care variable.) Observe also that the conservation law $(\hat{\varepsilon}_1\mathbf{T})(1) = (\hat{\varepsilon}_2\mathbf{T})(1)$ simply means that the number of edges (in the corresponding multigraph description) *entering* 1 is the same as the number of edges *leaving* 1, that is, the Euler condition. Furthermore, using the vector count $(\mathbf{T}(11), \mathbf{T}(12), \mathbf{T}(21), \mathbf{T}(22))$ in the space of types $\mathbb{R}^4$, we can re-write this conservation law as

$$\begin{aligned} 0 &= (\hat{\varepsilon}_1\mathbf{T})(1) - (\hat{\varepsilon}_2\mathbf{T})(1) = \mathbf{T}(12) - \mathbf{T}(21) \\ &= (0,1,-1,0) \cdot (\mathbf{T}(11), \mathbf{T}(12), \mathbf{T}(21), \mathbf{T}(22))^t, \end{aligned}$$

where in the last line we use the matrix $\mathbf{C} = (0,1,-1,0)$, and the superscript t means transposed. In fact, this suggests that we can re-phrase our discussion in terms of linear functionals and dual spaces as discussed in details below. In particular, in this example, we can use the following linear function (functional):

$$\boldsymbol{v}_{(\mathcal{S}', \mathfrak{t}' = \text{"}1\text{"}, \boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2)}(\mathbf{T}) = (0,1,-1,0) \cdot \mathbf{T},$$

which is formally a covector in the dual space (space of all covectors). □

We now re-formulate our conservation laws in the language of *linear functionals* and *dual spaces* [16]. This formalism allows us to rigorously prove our statements.

Definition 2. *Consider the* vector space of types*,* $\mathscr{V} := \mathbb{R}^D$*. Let* $\mathcal{S}'$ *be a shape,* $\boldsymbol{\varepsilon}_i : \mathcal{S}' \to \mathcal{S}$*,* $i \in \{1,2\}$*, be different embeddings of* $\mathcal{S}'$ *into* $\mathcal{S}$*, and* $\mathfrak{t}'$ *be a tile on* $\mathcal{S}'$*. The linear function (*covector*)* $\boldsymbol{v}_{(\mathcal{S}', \mathfrak{t}', \boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2)} \in \mathscr{V}^*$ *defined by*

$$\boldsymbol{v}_{(\mathcal{S}', \mathfrak{t}', \boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2)}\mathbf{T} \mapsto (\hat{\varepsilon}_1\mathbf{T})(\mathfrak{t}') - (\hat{\varepsilon}_2\mathbf{T})(\mathfrak{t}')$$

is called the conservation law *corresponding to the tuple* $(\mathcal{S}', \mathfrak{t}', \boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2)$*.*

In the standard basis of $\mathscr{V}$, the mapping $\hat{\varepsilon}$ is a linear combination of $\mathbf{T}$ coordinates with 0 or 1 coefficients, $\boldsymbol{v}$ is the difference of two of them, so all nonzero coefficients of $\boldsymbol{v}_{(\mathcal{S}', \mathfrak{t}', \boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2)}$ are ± 1. In fact, all conservation laws for all possible $\mathcal{S}', \boldsymbol{\varepsilon}, \mathfrak{t}'$ form a (huge) matrix $\mathbf{C}$ with coefficients in $\{-1,0,1\}$. In Example 4, $\mathfrak{t}' =$ "1" leads to $(0,1,-1,0)$ vector, forming the matrix

$$\mathbf{C} = \begin{pmatrix} 0 & 1 & -1 & 0 \\ 0 & -1 & 1 & 0 \end{pmatrix}$$

with linearly dependent rows (redundant conservation laws). Generally the number of independent rows (i.e., rank of $\mathbf{C}$) is much smaller than the number of all possible conditions obtained this way.

We aim at finding a matrix $\mathbf{C}_m$ with independent rows. There are several sources of such dependencies among the rows of $\mathbf{C}$:

1) The normalization equation $\sum_{\mathfrak{t}' \in \mathtt{Conf}(\mathcal{S}')} (\hat{\varepsilon}\mathbf{T})(\mathfrak{t}') = \mathbf{T}(*) = N$ implies that

$$\sum_{\mathfrak{t}' \in \mathtt{Conf}(\mathcal{S}')} \boldsymbol{v}_{(\mathcal{S}', \mathfrak{t}', \boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2)} = 0$$

for any two embeddings $\boldsymbol{\varepsilon}_i : \mathcal{S}' \to \mathcal{S}$. This eliminates for every pair of embeddings $\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2 : \mathcal{S}' \to \mathcal{S}$ one equation (reducing the number of rows from $m^{|\mathcal{S}'|}$ to $m^{|\mathcal{S}'|} - 1$ equations), since summing over all $\mathfrak{t}'$ we obtain the trivial equation $N = N$.

2) Clearly, the functional $\boldsymbol{v}_{(\mathcal{S}', \mathfrak{t}', \boldsymbol{\varepsilon}_2, \boldsymbol{\varepsilon}_3)}$ can be represented as

$$\boldsymbol{v}_{(\mathcal{S}', \mathfrak{t}', \boldsymbol{\varepsilon}_2, \boldsymbol{\varepsilon}_3)} = \boldsymbol{v}_{(\mathcal{S}', \mathfrak{t}', \boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_3)} - \boldsymbol{v}_{(\mathcal{S}', \mathfrak{t}', \boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2)}.$$

Hence, for any $\mathcal{S}'$ (admitting at least two different embeddings into $\mathcal{S}$), we can fix one of the embeddings

$$\varepsilon_{\mathcal{S}'} : \mathcal{S}' \to \mathcal{S}$$

as the canonical one, and restrict our attention only to the conservation laws

$$\boldsymbol{v}_{(\mathcal{S}', \mathfrak{t}', \varepsilon_{\mathcal{S}'}, \varepsilon)},$$

where ε runs over all embeddings $\varepsilon : \mathcal{S}' \to \mathcal{S}$ different from $\varepsilon_{\mathcal{S}'}$: there are $(|\{\varepsilon : \mathcal{S}' \to \mathcal{S}\}| - 1)$ such choices. (We will discuss a way to produce such a choice consistently later on.) We remark that the number of conservation laws after these first two restrictions are:

$$(m^{|\mathcal{S}'|} - 1)(|\{\varepsilon : \mathcal{S}' \to \mathcal{S}\}| - 1).$$

3) These two reductions are sufficient for small shapes S. However, for larger shapes such as the 2×2 squares, there are further relations. Specifically, let us choose one symbol $m \in \mathcal{A}$. Observe that counts using this symbol can be expressed without it, for example $T(m1) = T(*1) - \sum_{i=1}^{m-1} T(i1)$. In this case, we can express conservation laws over the whole alphabet $\mathcal{A}$ without using one symbol, say the m-th one. It is sufficient to focus on laws for the reduced alphabet $\mathcal{A}' = \mathcal{A} \setminus \{m\}$: there are $(m-1)^{|\mathcal{S}'|}$ of them for a given embedding.

We are now in the position to formulate our first result regarding the rank of the matrix $\mathbf{C}$. We shall do it in the formalism we have just established. In particular, we use the notion of the *kernel* or *null space* of the underlying linear functionals defined in our case as $\{\mathbf{T} : \ \mathbf{C} \cdot \mathbf{T} = \mathbf{0}\}$. To pick up further information on linear algebra and linear functionals the reader is referred to [16]. In Section III we establish the following result.

Theorem 3. (i) *The submatrix $\mathbf{C}_m$ of $\mathbf{C}$ formed by the rows corresponding to the functionals*

$$\boldsymbol{v}_{(\mathcal{S}', \mathfrak{t}', \varepsilon_{\mathcal{S}'}, \varepsilon)} \tag{14}$$

with $\mathfrak{t}'$ over $\mathcal{A}' = \mathcal{A} \setminus \{m\}$ has the same rank as the full matrix $\mathbf{C}$, and therefore defines the same kernel (i.e., $\{\mathbf{T} : \mathbf{C}_m \cdot \mathbf{T} = \mathbf{0}\} = \{\mathbf{T} : \mathbf{C} \cdot \mathbf{T} = \mathbf{0}\}$). Here $\varepsilon_{\mathcal{S}'}$ is a canonical *embedding of a shape $\mathcal{S}'$ embeddable into $\mathcal{S}$.*
(ii) *The matrix $\mathbf{C}$ has the corank (the dimension of its kernel: $\{\mathbf{T} : \mathbf{C} \cdot \mathbf{T} = \mathbf{0}\}$) equal to the number of tilings of the reduced alphabet $\mathcal{A}' = \mathcal{A} - \{m\}$ of all subshapes $\mathcal{S}'$ (including the empty one) embeddable into $\mathcal{S}$, i.e.,*

$$\mu + 1 = \sum_{\mathcal{S}' : |\{\varepsilon : \mathcal{S}' \to \mathcal{S}\}| \geq 1} (m-1)^{|\mathcal{S}'|}. \tag{15}$$

The rank of the matrix $\mathbf{C}$ is given by

$$\mathrm{rk}(\mathbf{C}) = D - \mu - 1 =$$

$$= \sum_{\mathcal{S}' : |\{\varepsilon : \mathcal{S}' \to \mathcal{S}\}| \geq 1} (|\{\varepsilon : \mathcal{S}' \to \mathcal{S}\}| - 1)(m-1)^{|\mathcal{S}'|}, \tag{16}$$

where the summation is again over all shapes $\mathcal{S}'$ embeddable into $\mathcal{S}$.

For the box shapes, formula (16) (requiring an enumeration of all subshapes fitting into $\mathcal{S}$) can be significantly simplified:

Corollary 4. *If $\mathcal{S} = I_{l_1} \times I_{l_2} \times \ldots \times I_{l_d}$ (recall that $I_l = \{0, 1, \ldots, l-1\}$), one has*

$$\mu = D - 1 - \mathrm{rk}(\mathbf{C}) = \sum_{\mathbf{s} \in \{0,1\}^d} m^{\prod_i (l_i - s_i)} \cdot (-1)^{\sum_i s_i} \tag{17}$$

where $\mathbf{l} = (l_1, \ldots, l_d) \in \mathbb{N}^d$.

C. More Examples

We now discuss a few examples illustrating the dimensionality reductions associated with the conservation laws and Theorem 3. We already observed in Example 4 that there is a single independent conservation law corresponding to $(0, 1, -1, 0) \cdot \mathbf{T} = 0$ in the $D = 4$ dimensional space of types leading to $\mu = 4 - 1 - 1 = 2$.

Example 5: *2D Markov Field with the L-Shape – Continuation.*
For the $\mathtt{L}$-shape $\mathcal{S} = \{(0,0), (0,1), (1,0)\}$ in 2D and $m = 2$, the frequency vector $\mathbf{T}$ has $D = m^3 = 8$ coordinates

$$\left(\begin{pmatrix} 1 \\ 11 \end{pmatrix}, \begin{pmatrix} 1 \\ 21 \end{pmatrix}, \begin{pmatrix} 1 \\ 12 \end{pmatrix}, \begin{pmatrix} 1 \\ 22 \end{pmatrix}, \right.$$

$$\left. \begin{pmatrix} 2 \\ 11 \end{pmatrix}, \begin{pmatrix} 2 \\ 21 \end{pmatrix}, \begin{pmatrix} 2 \\ 12 \end{pmatrix}, \begin{pmatrix} 2 \\ 22 \end{pmatrix} \right);$$

however, only five of them are independent. The subshape of a single point $\mathcal{S}' = \{(0,0)\}$, which can be embedded in all three positions: $\varepsilon_1((0,0)) = (0,0)$, $\varepsilon_2((0,0)) = (1,0)$, $\varepsilon_3((0,0)) = (0,1)$ leading to the following conservation laws:

$$0 = \mathbf{T}\begin{pmatrix} * \\ 1* \end{pmatrix} - \mathbf{T}\begin{pmatrix} * \\ *1 \end{pmatrix} =$$

$$= \mathbf{T}\begin{pmatrix} 1 \\ 12 \end{pmatrix} + \mathbf{T}\begin{pmatrix} 2 \\ 12 \end{pmatrix} - \mathbf{T}\begin{pmatrix} 1 \\ 21 \end{pmatrix} - \mathbf{T}\begin{pmatrix} 2 \\ 21 \end{pmatrix},$$

$$0 = \mathbf{T}\begin{pmatrix} * \\ 1* \end{pmatrix} - \mathbf{T}\begin{pmatrix} 1 \\ ** \end{pmatrix} =$$

$$= \mathbf{T}\begin{pmatrix} 2 \\ 11 \end{pmatrix} + \mathbf{T}\begin{pmatrix} 2 \\ 12 \end{pmatrix} - \mathbf{T}\begin{pmatrix} 1 \\ 21 \end{pmatrix} - \mathbf{T}\begin{pmatrix} 1 \\ 22 \end{pmatrix},$$

$$0 = \mathbf{T}\begin{pmatrix} * \\ *1 \end{pmatrix} - \mathbf{T}\begin{pmatrix} 1 \\ ** \end{pmatrix} =$$

$$= \mathbf{T}\begin{pmatrix} 2 \\ 11 \end{pmatrix} + \mathbf{T}\begin{pmatrix} 2 \\ 21 \end{pmatrix} - \mathbf{T}\begin{pmatrix} 1 \\ 12 \end{pmatrix} - \mathbf{T}\begin{pmatrix} 1 \\ 22 \end{pmatrix}.$$

These equations define the functionals $\boldsymbol{v}_{(\mathcal{S}', \mathfrak{t}', \varepsilon_1, \varepsilon_2)}$, $\boldsymbol{v}_{(\mathcal{S}', \mathfrak{t}', \varepsilon_1, \varepsilon_3)}$ and $\boldsymbol{v}_{(\mathcal{S}', \mathfrak{t}', \varepsilon_2, \varepsilon_3)}$ with $\mathfrak{t}' = 1$. Obviously one of these equations is redundant – choosing the lower left position as the canonical embedding $\varepsilon_{\mathcal{S}'} := \varepsilon_1$, there remain only the first two of the above equations. In the basis above, they can be written as:

$$\mathbf{C}\mathbf{T} = \begin{pmatrix} 0 & -1 & 1 & 0 & 0 & -1 & 1 & 0 \\ 0 & -1 & 0 & -1 & 1 & 0 & 1 & 0 \end{pmatrix} \cdot \mathbf{T} = \mathbf{0}.$$

These two independent conservation laws restrict the space of $\mathbf{T}$ to a $\mu+1$ =6–dimensional cone, and the normalization equation further restricts it to a $\mu = 5$ dimensional polytope.

Example 6: *2D Markov Field with Square Shape $\square$ – Continuation*.
For the 2×2 square shape and $m = 2$ the frequency vector $\mathbf{T}$ is in $D = m^4 = 16$–dimensional space. As $\mathcal{A}' = \mathcal{A}\backslash\{m\} = \{1\}$, the ultimate set of independent conservation laws (16) are

$$\mathbf{T}\begin{pmatrix} ** \\ 1* \end{pmatrix} = \mathbf{T}\begin{pmatrix} 1* \\ ** \end{pmatrix} = \mathbf{T}\begin{pmatrix} *1 \\ ** \end{pmatrix} = \mathbf{T}\begin{pmatrix} ** \\ *1 \end{pmatrix},$$

$$\mathbf{T}\begin{pmatrix} ** \\ 11 \end{pmatrix} = \mathbf{T}\begin{pmatrix} 11 \\ ** \end{pmatrix}, \quad \mathbf{T}\begin{pmatrix} 1* \\ 1* \end{pmatrix} = \mathbf{T}\begin{pmatrix} *1 \\ *1 \end{pmatrix}.$$

The first line contains three equations for a single point shape. The second line contains the remaining two single conditions for $\mathcal{S}' = \{(0,0),(1,0)\}$ and $\mathcal{S}' = \{(0,0),(0,1)\}$, respectively, and both their embeddings. By combining these five equations we can obtain the remaining ones. For example,

$$\mathbf{T}\begin{pmatrix} 1* \\ ** \end{pmatrix} = \mathbf{T}\begin{pmatrix} 1* \\ 1* \end{pmatrix} + \mathbf{T}\begin{pmatrix} 1* \\ 2* \end{pmatrix}$$

implies

$$\mathbf{T}\begin{pmatrix} 1* \\ 2* \end{pmatrix} = \mathbf{T}\begin{pmatrix} 1* \\ ** \end{pmatrix} - \mathbf{T}\begin{pmatrix} 1* \\ 1* \end{pmatrix} =$$

$$= \mathbf{T}\begin{pmatrix} *1 \\ ** \end{pmatrix} - \mathbf{T}\begin{pmatrix} *1 \\ *1 \end{pmatrix} = \mathbf{T}\begin{pmatrix} *1 \\ *2 \end{pmatrix}.$$

Thus, $\mathbf{T}$ in $D = m^4 = 16$-dimensional space has $\mu+1 = 11$ components by the above five independent conservation laws. The normalization restricts it further to $\mu = 10$–dimensional polytope. In Figure 3 we show all 21 vertices of this polytope and the corresponding tiling. Observe that some vertices (vectors) are not realizable by a Markov field. $\square$

Finally, we illustrate (17) of Corollary 4 for the box shape.

Example 7. *The Box Shape*.
Let us now consider a general box shape $\mathcal{I}_{l_1} \times \ldots \times \mathcal{I}_{l_d}$. Observe that:

- For the $d = 1$ dimensional shape $\mathcal{S} = \{(0),(1)\}$ we have $\mu = m^2 - m$, as known already from [12]. For $\mathcal{S} = \{(0),(1),(2)\}$ we find $\mu = m^3 - m^2$, while for $\mathcal{S} = \{(0),(1),(2),(3)\}$ we have $\mu = m^4 - m^3$.
- For $d = 2$ the 2×2 square shape ($l_1 = l_2 = 2$) leads to $\mu = m^4 - 2m^2 + m$, the 3×2 rectangular shape gives $\mu = m^6 - m^4 - m^3 + m^2$ and the 3×3 square ends up with $\mu = m^9 - 2m^6 + m^4$.
- For $d = 3$ the $2\times2\times2$ cube leads to $\mu = m^8 - 3m^4 + 3m^2 - m$, while the $2\times3\times4$ box gives $\mu = m^{24} - m^{18} - m^{16} - m^{12} + m^{12} + m^9 + m^8 - m^6$.
- For $d = 4$ we have $\mu = m^{16} - 4m^8 + 6m^4 - 4m^2 + m$ for the $2\times2\times2\times2$ cube.
- Finally, in $d = 5$ space the $2\times2\times2\times2\times2$ cube leads to $\mu = m^{32} - 5m^{16} + 10m^8 - 10m^4 + 5m^2 - m$.

D. Geometry and Enumeration

We explore now the geometry of the vector counts $\mathbf{T} = \{\mathbf{T}(\mathfrak{t})\}_{\mathfrak{t}\in\mathfrak{T}}$ in the D-dimensional space. As discussed, the conservation laws (which we write as a linear system $\mathbf{CT} = \mathbf{0}$) together with $\mathbf{T} \geq 0$ restrict the vectors $\mathbf{T}$ to a $D - \mathrm{rk}(\mathbf{C}) = \mu + 1$ dimensional cone $\mathcal{C}$ and the normalization equation $\sum_{\mathfrak{t}} \mathbf{T}(\mathfrak{t}) = N$ (for torus) further restricts $\mathbf{T}$ to the polytope $\mathcal{F}_{\mathbf{n}}$. Formally, let us define

$$\mathcal{C} = \{\mathbf{T} \in \mathbb{R}^D_{\geq 0} \;:\; \mathbf{C}\cdot\mathbf{T} = \mathbf{0}\}, \tag{18}$$

$$\mathcal{F}_{\mathbf{n}} = \mathcal{F}_N \equiv \mathcal{F}_{\mathbf{n}}(\mathcal{A},\mathcal{S}) =$$
$$= \{\mathbf{T} \in \mathcal{C}\cap\mathbb{N}^D : \sum_{\mathfrak{t}} \mathbf{T}(\mathfrak{t}) = N\}. \tag{19}$$

We also define the *normalized* polytope $\hat{\mathcal{F}}$ of frequency vectors $\hat{\mathbf{T}}$ as

$$\hat{\mathcal{F}} \equiv \hat{\mathcal{F}}(\mathcal{A},\mathcal{S}) = \{\hat{\mathbf{T}} \in \mathbb{R}^D_{\geq 0} \;: \mathbf{C}\cdot\hat{\mathbf{T}} = \mathbf{0};\; \sum_{\mathfrak{t}} \mathbf{T}(\mathfrak{t}) = 1\} \tag{20}$$

and let

$$\mathcal{F}_N = N\hat{\mathcal{F}} \cap \mathbb{N}^D. \tag{21}$$

Finally, the *normalized* (rescaled) set of realizable count vectors (*types*) is

$$\hat{\mathcal{P}} \equiv \hat{\mathcal{P}}(\mathcal{A},\mathcal{S}) = \bigcup_{\mathbf{n}} \mathcal{P}_{\mathbf{n}}(\mathcal{A},\mathcal{S})/N(\mathbf{n}). \tag{22}$$

Where $N(\mathbf{n}) = \prod_i n_i$ is the size of the torus. Obviously, $\hat{\mathcal{P}} \subset \hat{\mathcal{F}}$.

Observe that $\hat{\mathcal{F}}$ is a compact polyhedron, hence (from basic convex analysis [24]) a polytope, i.e., the convex hull of its *extremal points*. These extremal points are the intersections of the linear subspace $\{\mathbf{T} : \mathbf{CT} = \mathbf{0},\; \sum_{\mathfrak{t}} \mathbf{T}(\mathfrak{t}) = 1\}$ with some μ of D conditions of type $T(\mathfrak{t}) = 0$. The number of the extremal points obtained this way is finite and at most $\binom{D}{\mu}$.

Example 8. *Polytopes in the 2D Case*.
For the $\mathtt{L}$-shape in the 2D case with $m = 2$, we have a $\mu = 5$ dimensional polytope in $D = 8$ dimensional space. Among $\binom{8}{5} = 56$ possible ways to choose zero coordinates, it turns out that there are only 7 vertices with all nonnegative coordinates. These seven vertices have the following coordinates:

$$\{(0,0,0,0,0,0,0,1),\; (0,0,0,\tfrac{1}{3},0,\tfrac{1}{3},\tfrac{1}{3},0),$$
$$(0,0,0,\tfrac{1}{2},\tfrac{1}{2},0,0,0),\; (0,0,\tfrac{1}{2},0,0,\tfrac{1}{2},0,0),$$
$$(0,\tfrac{1}{2},0,0,0,0,\tfrac{1}{2},0),\; (0,\tfrac{1}{3},\tfrac{1}{3},0,\tfrac{1}{3},0,0,0),$$
$$(1,0,0,0,0,0,0,0)\}.$$

All of these points correspond to periodic tilings (the same periodic tilings as for cases 1 to 7 in Figure 3). On the other hand, for the 2×2 square shape and $m = 2$, we have 21 vertices of a $\mu = 10$ dimensional polytope in $D = 16$ dimensional space as shown in Figure 3. Surprisingly, now some of the vertices do *not* correspond to periodic tilings, so in general not all points in $\hat{\mathcal{F}}$ lead to a realizable tiling and therefore a point in $\hat{\mathcal{P}}$ (see Figure 4). $\square$

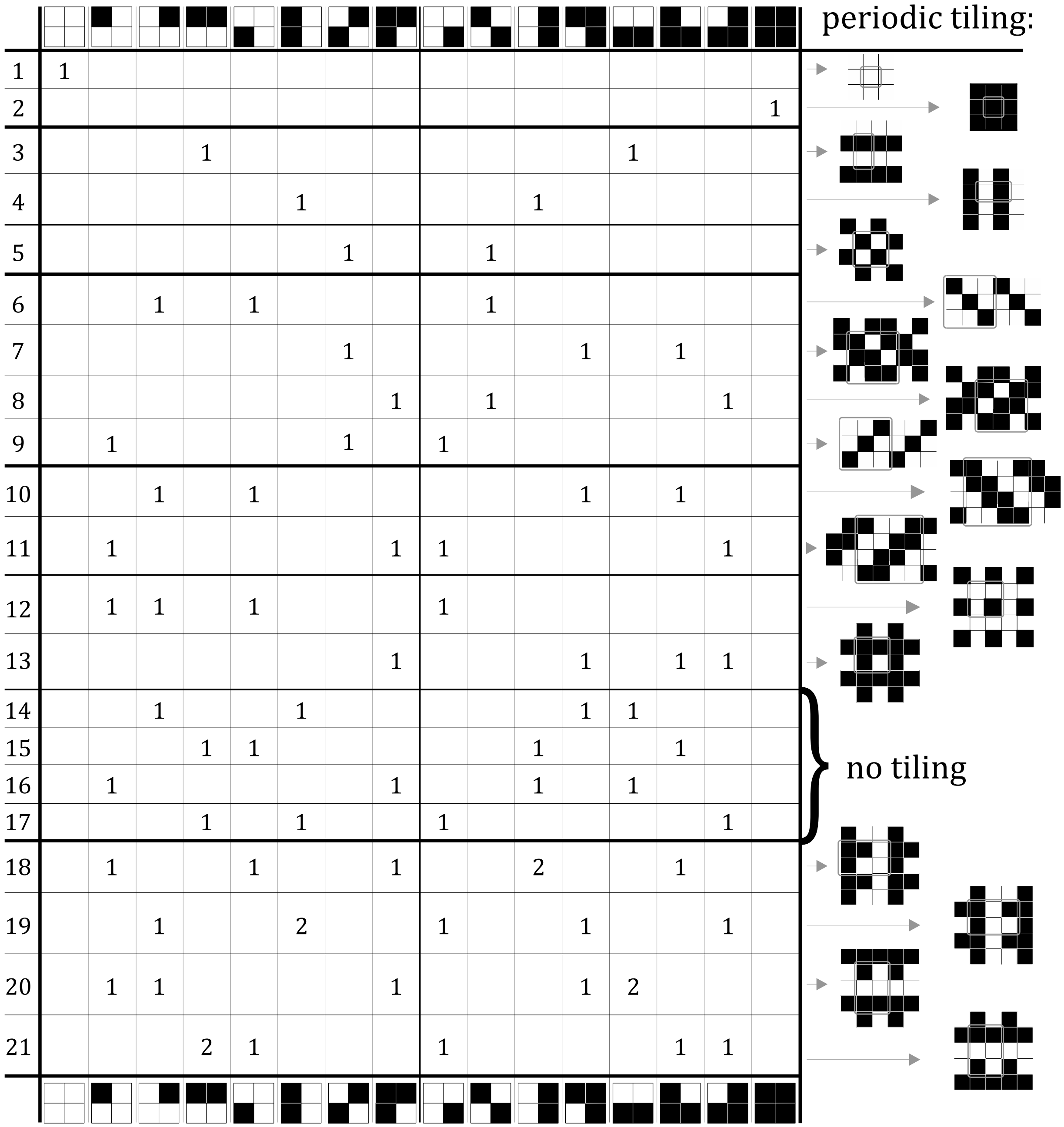

Figure 3. For $m = 2$ and the 2×2 square shape $\square$, we show all 21 vertices (only nonzero coordinates are displayed) of the $\mu = 10$ dimensional polytope in a $D = 16$ dimensional space. On the right-hand side, we also show the corresponding realizable tilings: Four of them (14-17) cannot be realized. Periodic tilings 1-7 correspond to all 7 vertices for the L like shape.

Interestingly enough, we can prove that the topological closure of $\hat{\mathcal{P}}$ is still a convex subset of $\hat{\mathcal{F}}$. This is illustrated in Figure 4 and proved below.

Lemma 5. *The closure* $\mathrm{cl}(\hat{\mathcal{P}})$ *of* $\hat{\mathcal{P}}$ *in a torus is a convex subset of* $\hat{\mathcal{F}}$.

Proof: To prove convexity of a closed set, it is enough to show that for any two points in this set, the point in the center between them is also in the underlying set. For every point $\hat{\mathbf{T}} \in \mathrm{cl}(\hat{\mathcal{P}})$ one can find a sequence of periodic tilings, whose (rescaled) frequency vectors converge to this points.

Consider two sequences of fields: $\mathbf{x}'_i$ and $\mathbf{x}''_i$ such that their frequency vectors converge to $\hat{\mathbf{T}}'_i \to \hat{\mathbf{T}}'$ and $\hat{\mathbf{T}}''_i \to \hat{\mathbf{T}}''$ as $i \to \infty$. We need to construct a sequence of fields $\mathbf{x}_i$ with frequency vectors $\hat{\mathbf{T}}_i \to \hat{\mathbf{T}} = (\hat{\mathbf{T}}' + \hat{\mathbf{T}}'')/2$. For this purpose, having $\mathbf{x}'_i$ and $\mathbf{x}''_i$ with correspondingly $\hat{\mathbf{T}}'_i$ and $\hat{\mathbf{T}}''_i$ frequency vectors, we shall construct $\mathbf{x}_i$ field with $\hat{\mathbf{T}}_i$ frequency vector in at most $\epsilon_i > 0$ distance from $(\hat{\mathbf{T}}'_i + \hat{\mathbf{T}}''_i)/2$, where $\epsilon_i \to 0$ is some arbitrary sequence. To accomplish it we cover one half of a large torus with $\mathbf{x}'_i$ tiling and the second with $\mathbf{x}''_i$. If the size of such a torus grows to infinity, the obtained frequency tends to $(\hat{\mathbf{T}}'_i + \hat{\mathbf{T}}''_i)/2$, as desired. ■

The set $\mathcal{F}_N$ consists of all integer points inside the polytope: $N\hat{\mathcal{F}} \cap \mathbb{N}^D$. The volume of $N\hat{\mathcal{F}}$ is proportional to N^μ, and

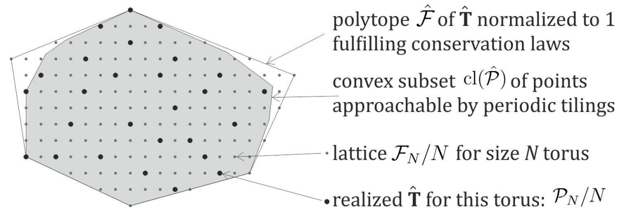

Figure 4. An illustration of the polytope $\hat{\mathcal{F}}$: $\hat{\mathbf{T}}$ vectors realized by periodic tilings create some irregular subset of the lattice and while $N \to \infty$ they densely cover some convex subset of $\hat{\mathcal{F}}$.

we expect the number of integer points inside also grows asymptotically as N^μ. This is indeed the case by Ehrhart's theorem [9]:

Theorem 6 (Ehrhart, 1967)**.** *If $\hat{\mathcal{F}}_N$ is the rational polytope*[1] *given by*

$$\begin{array}{rclc} \mathbf{Cv} & = & N\mathbf{b}, & \mathbf{b} \in \mathbb{Q}^D, \quad \mathbf{C} \in \mathbb{Q}^D \times \mathbb{Q}^D, \\ \mathbf{v} & \geq & 0, & \mathbf{v} \in \mathbb{R}^D, \end{array}$$

then there exist a period $p \in \mathbb{N}$ and real coefficients $c_{i,j}$ such that $c_{\mu,j} \neq 0$ for some j and

$$|\mathcal{F}_N| = a_{\mu,j}N^\mu + a_{\mu-1,j}N^{\mu-1} + \ldots + a_{0,j}$$
$$\text{if} \quad N \equiv j \pmod{p}$$

where $\mathcal{F}_N$ is a set of integers points inside $\hat{\mathcal{F}}_N$, and μ is the dimension of $\hat{\mathcal{F}}$, i.e., $D - \mathrm{rk}(\mathbf{C})$.

Indeed, by the construction the vertices of $\hat{\mathcal{F}}_N$ are solutions of a system of linear equations with integer coefficients (actually, ± 1), making it a rational polytope. Since $\mathcal{F}_\mathbf{n}$ upper bounds $\mathcal{P}_\mathbf{n}$ (as a set), the volume N^μ of $\mathcal{F}_\mathbf{n}$ provides an upper bound for the number of types. We need now a matching lower bound to prove that $|\mathcal{P}_\mathbf{n}(\mathcal{A}, \mathcal{S})| = \Theta(N^\mu)$. In Section III we construct such a bound, leading to the main result of this paper.

Theorem 7. *Consider the torus $\mathcal{O}_\mathbf{n}$. There exist constants $0 < c^- \leq c^+$ such that for n_i large enough we have*

$$c^- N^\mu \leq |\mathcal{P}_\mathbf{n}(\mathcal{A}, \mathcal{S})| \leq c^+ N^\mu \tag{23}$$

where, we recall, $N = n_1 \cdots n_d$.

We should point out that in [12] for $d = 1$ it was proved that $|\mathcal{F}_N|$ is asymptotically equivalent to $|\mathcal{P}_N|$, that is, $|\mathcal{F}_N|/|\mathcal{P}_N| \to 1$ as $N \to \infty$: the set of realizable types in dimension 1 is essentially given by the conservation laws. Remarkably, this seems not to be true in general in higher dimensions. However, in some special cases we can say more about $|\mathcal{F}_N|$ but not necessary about $|\mathcal{P}_N|$. This is discussed in the next subsection.

We end this section with a conjecture related to the average redundancy of a code for a Markov field source. Consider a two-stage code for Markov fields:

$$(\text{type of } \mathbf{x^n}, \ \text{position within the type}).$$

If we denote by $S(\mathbf{x^n})$ the number of fields of the same type as $\mathbf{x^n}$, then the code length $L(\mathbf{x^n})$ is $L(\mathbf{x^n}) = \log(|\mathcal{P}_\mathbf{n}|) + \log(S(\mathbf{x^n}))$. We proved in Theorem 7 that $\log(|\mathcal{P}_\mathbf{n}|) \sim \mu \log N$, so to calculate the code length we need to estimate the size of *typical* types, that is, the number of fields of a typical type. But this is hard (see [11] for $d = 1$ case). Nevertheless, we shall put forward the following conjecture.

[1] A polytope with vertices in $\mathbb{Q}^D$ is called a *rational polytope*.

Conjecture 1. *For a typical $\mathbf{x^n}$ the number of fields of the same type is*

$$S(\mathbf{x^n}) \sim \frac{A}{N^{\mu/2}} 2^{NH}$$

where H is the entropy of the underlying Markov field and A is a constant.

Provided the conjecture is true, the average redundancy becomes

$$\begin{aligned} R_N &= \mathbf{E}[L(\mathbf{x^n})] - NH \\ &= \log O(N^\mu) + \log\left(\frac{A}{N^{\mu/2}} 2^{NH}\right) - NH \\ &= \frac{\mu}{2}\log N + o(\log N). \end{aligned} \tag{24}$$

Proving (24) may be very challenging. But we have already made the first step by providing a precise formula for the number of free parameters μ (the coefficient at $\log N$), that is, actually formulating precisely a Rissanen-like lower bound for Markov fields.

E. Analytic Approach

In [12] an analytic approach was used to enumerate $\mathcal{F}_N$ (here $N = n_1$, the length of the underlying sequence) for $d = 1$. We first recall some results from [12] and then extend them to any dimension and shapes. We should point out, however, that through this approach we will only get better asymptotics for $|\mathcal{F}_N|$ but not for $|\mathcal{P}_N|$. This is actually of interest on its own since it allows us to enumerate precisely the number of nonnegative solutions of a multidimensional system of linear Diophantine equations; not an easy task, as argued in [27].

Let us recall some facts from [12]. We first assume $d = 1$ and enumerate $\mathcal{F}_N$. We accomplished it by finding the following generating function

$$F_m^*(z) = \sum_{N \geq 0} |\mathcal{F}_N(m)| z^N$$

and then taking the coefficient at z^N which is written as $[z^N]F_m^*(z) = |\mathcal{F}_N(m, \mathcal{S})|$.

Let $\mathbf{z} = \{z_\mathfrak{t}\}_{\mathfrak{t} \in \mathfrak{T}}$, where in our case $\mathfrak{t} = (\alpha, \beta) \in \mathcal{A}^2$ is a pair of symbols or in other words the shape is $\mathcal{S} = \{0, 1\}$. We also write $\mathbf{z^T} = \prod_{\alpha\beta} z_{\alpha\beta}^{\mathbf{T}(\alpha\beta)}$. We introduce a multidimensional generating function $\sum_\mathbf{T} \mathbf{z^T}$ that we estimate in two different ways for $z_{\alpha\beta} = z\frac{y_\alpha}{y_\beta}$ for some $(y_\alpha)_{\alpha \in \mathcal{A}}$ vector:

$$\sum_\mathbf{T} \mathbf{z^T} = \prod_{\alpha\beta}\left(\sum_{\mathbf{T}(\alpha\beta)}\left(z\frac{y_\alpha}{y_\beta}\right)^{\mathbf{T}(\alpha\beta)}\right) =$$

$$= \prod_{\alpha\beta}\left(1 - z\frac{y_\alpha}{y_\beta}\right)^{-1},$$

$$\sum_{\mathbf{T}} \mathbf{z}^{\mathbf{T}} = \sum_{\mathbf{T}} z^{\sum_{\alpha,\beta} \mathbf{T}(\alpha\beta)} \prod_{\alpha} y_\alpha^{\sum_\beta \mathbf{T}(\alpha\beta) - \sum_\beta \mathbf{T}(\beta\alpha)}.$$

Now if $\mathbf{T} \in \mathcal{F}_N$, that is,

$$\sum_{\alpha,\beta} \mathbf{T}(\alpha,\beta) = N, \qquad \sum_{\beta} \mathbf{T}(\alpha\beta) - \sum_{\beta} \mathbf{T}(\beta\alpha) = 0,$$

then

$$F_m^*(z) = \sum_{N\geq 0} |\mathcal{F}_N(m,\mathcal{S})| z^N =$$

$$= [y_1^0 y_2^0 \ldots y_m^0] \prod_{\alpha,\beta=1}^{m} \left(1 - z\frac{y_\alpha}{y_\beta}\right)^{-1} \tag{25}$$

where $[\mathbf{y}^0]F(\mathbf{y}) := [y_1^0 y_2^0 \ldots y_m^0]F(\mathbf{y})$ denotes the zeroth power coefficient of $F(\mathbf{y})$.

There is a simple interpretation of formula (25): Its right hand side can be seen as a product of m^2 geometric series, while α, β terms correspond to "$\alpha\beta$" pattern (pair) in our sequence. The auxiliary $\mathbf{y}$ variables are used to restrict $\mathbf{T}$ to those satisfying the conservation laws: each symbol should appear the same number of times on the left and on the right position of $\mathcal{S}$. Thanks to the y_α/y_β term, the power of y_α increases by 1 every time α appears in left position of "$\alpha\beta$", and decreases by 1 when it appears in the right position. However, in addition we have the normalization equation which allows us to eliminate one of the variables, for example by setting $y_m = 1$.

Let us now move to the multidimensional case $d > 1$. Each auxiliary variable corresponds to a single equation of the conservation laws. We can reduce the set of equations by considering only independent variables, as discussed in Theorem 3.

Let us start with some examples. For the L shape as in Example 2 we have $\mathcal{S} = \{(0,0),(0,1),(1,0)\}$, and

$$F_m^*(z) =$$

$$= [x_1^0 x_2^0 \ldots x_m^0 y_1^0 y_2^0 \ldots y_m^0] \prod_{\alpha,\beta,\gamma} \left(1 - z\frac{x_\alpha}{x_\beta}\frac{y_\alpha}{y_\gamma}\right)^{-1},$$

where the auxiliary variables $\mathbf{x}$ now guard the conservation law in one direction, $\mathbf{y}$ in the other direction. In other words, the L shape tile $\mathbf{t}$ is marked as follows

$$\begin{pmatrix} \gamma & \\ \alpha & \beta \end{pmatrix}$$

and then the conservation laws are

$$\sum_{\beta\gamma} \mathbf{T}\begin{pmatrix} \gamma & \\ \alpha & \beta \end{pmatrix} = \sum_{\beta\gamma} \mathbf{T}\begin{pmatrix} \gamma & \\ \beta & \alpha \end{pmatrix},$$

$$\sum_{\beta\gamma} \mathbf{T}\begin{pmatrix} \gamma & \\ \alpha & \beta \end{pmatrix} = \sum_{\beta\gamma} \mathbf{T}\begin{pmatrix} \alpha & \\ \gamma & \beta \end{pmatrix}.$$

We can choose $x_m = y_m = 1$. Interestingly enough, in this case $\mathrm{cl}(\hat{\mathcal{P}}) = \hat{\mathcal{F}}$ since both are spanned on 7 vertices in $D = 8$ dimensional space, as shown in Figure 2. This, however, does not imply that $|\mathcal{F}_N| \sim |\mathcal{P}_N|$.

For the analogous 3D L shape $\mathcal{S} = \{(0,0,0),(0,0,1),(0,1,0),(1,0,0)\}$, we find

$$F_m^* = [x_1^0 x_2^0 \ldots x_m^0 y_1^0 y_2^0 \ldots y_m^0 u_1^0 u_2^0 \ldots u_m^0]$$
$$\prod_{\alpha,\beta,\gamma,\delta} \left(1 - z\frac{x_\alpha}{x_\beta}\frac{y_\alpha}{y_\gamma}\frac{u_\alpha}{u_\delta}\right)^{-1}$$

and we can set $x_m = y_m = u_m = 1$. Using partial fraction expansions, as in [12], we can obtain asymptotic expressions for $|\mathcal{F}_N|$, as illustrated below.

Example 8. For $m = 2$ in the 2D case and the L shape, we have

$$F_2^*(z) = \frac{1 - z + z^2}{(z-1)^6(z+1)^2(z^2+z+1)}.$$

Using the partial fraction expression we find after some algebra

$$|\mathcal{F}_N(2,L)| = \frac{N^5}{12 \cdot 5!} + O(N^4).$$

For the analogous 3D L shape when $m = 2$ we arrive at

$$F_2^*(z) = \frac{Q(z)}{D(z)}$$

where $D(z) = (z-1)^{13}(z+1)^5(z^2+1)(z^2+z+1)^3(z^4+z^3+z^2+z+1)$ and $Q(z) = 1 + 2z + 22z^2 + 50z^3 + 94z^4 + 138z^5 + 175z^6 + 184z^7 + 163z^8 + 120z^9 + 76z^{10} + 38z^{11} + 16z^{12} + 2z^{13} + z^{14}$. Using the partial fraction decomposition and Cauchy's integral formula we find

$$|\mathcal{F}_N(2,L)| = \frac{541}{4320}\frac{N^{12}}{12!} + O(N^{11})$$

for large N. □

Let us now look at a situation with a more subtle dependence between the conservation laws, for example for the 2×2 square shape in 2D discussed in Example 3. The first approach could be:

$$F_m^*(z) = [x_{11}^0 x_{12}^0 \ldots x_{mm}^0 y_{11}^0 y_{12}^0 \ldots y_{mm}^0]$$
$$\prod_{\alpha,\beta,\gamma,\delta} \left(1 - z\frac{x_{\alpha\beta}}{x_{\gamma\delta}}\frac{y_{\alpha\gamma}}{y_{\beta\delta}}\right)^{-1}$$

where we can initially set $x_{mm} = y_{mm} = 1$. This formula corresponds to $\mathcal{S}'$ and s being selected as $\mathcal{S}' = \{(0,0),(1,0)\}$, $\mathbf{s} = (0,1)$ and $\mathcal{S}' = \{(0,0),(0,1)\}$, $\mathbf{s} = (1,0)$ or the following marking of the square tile

$$\begin{pmatrix} \gamma & \delta \\ \alpha & \beta \end{pmatrix}.$$

It leads to a complete set of conservation laws, but with some linear dependencies, as the conservation law for $\mathcal{S}' = \{(0,0)\}$, $\mathbf{s} = (1,1)$ can be induced in two ways. To ensure using only independent variables/conservation laws, we use Theorem 3 to deduce the set of independent conservation laws. This leads to

$$F_m^* = [u_1^0..u_m^0 v_1^0..v_m^0 w_1^0..w_m^0 x_{11}^0..x_{mm}^0 y_{11}^0..y_{mm}^0]$$
$$\prod_{\alpha,\beta,\gamma,\delta} \left(1 - z\frac{u_\alpha}{u_\beta}\frac{v_\alpha}{v_\gamma}\frac{w_\alpha}{w_\delta}\frac{x_{\alpha\beta}}{x_{\gamma\delta}}\frac{y_{\alpha\gamma}}{y_{\beta\delta}}\right)^{-1}$$

where $u_m = v_m = w_m = 1$ and $x_{im} = x_{mi} = y_{im} = y_{mi} = 1$ for any $i \in \mathcal{A}$.

Example 9. Consider $m = 2$. Then both approaches lead to

$$F_2^*(z) = \frac{Q_1(z)}{(z-1)^{11}(z+1)^7(z^2+1)^3(z^2-z+1)(z^2+z+1)^3}$$

where $Q_1(z) = 1 + 2z^3 + 5z^4 + 2z^6 + 6z^7 + 8z^8 + 6z^9 + 2z^{10} + 5z^{12} + 2z^{13} + z^{16}$, from which we find

$$|\mathcal{F}_{\text{Size}}(2, \square)| = \frac{5}{3456}\frac{N^{10}}{10!} + O(N^9)$$

for large N. □

For a general shape we consider the conservation laws (14), attach a variable $\mathbf{y}$ to each of them, and choose a fraction of some of these variables in the product of $m^{|\mathcal{S}|}$ geometric series to enforce the conservation laws by zeroing the power of these variables. This allows us to find a general expression for the underlying generating function, that is,

$$F_m^*(z) =$$

$$= [\mathbf{y}^0] \prod_{\mathfrak{t}:\mathcal{S}\to\mathcal{A}} \left(1 - z \prod_{\mathcal{S}' \text{ embeddable in } \mathcal{S},\ \varepsilon \neq \varepsilon_{\mathcal{S}'}} \frac{y_{\varepsilon^*(\mathfrak{t})}}{y_{\varepsilon^*_{\mathcal{S}'}(\mathfrak{t})}}\right)^{-1} \quad (26)$$

where $\varepsilon_{\mathcal{S}'}$ is the canonical embedding, $[\mathbf{y}^0]$ denotes taking zeroth power of all used y_i.

F. Number of Types in a Box – An Upper Bound

Finally, we comment on the number of types $\tilde{\mathcal{P}}_{\mathbf{n}}(m, S)$ in the box $\mathcal{I}_{\mathbf{n}} = I_{n_1} \times I_{n_2} \times \ldots \times I_{n_d} \subset \mathbb{Z}^d$. We only discuss an upper bound, leaving establishing the proper growth to a forthcoming paper. Let $\mathbf{x} = \mathbf{x}^{\mathbf{n}}$ be a configuration in the box $\mathcal{I}_{\mathbf{n}}$. Its type in the box is now defined by shifts (embeddings) that fit into the box, that is,

$$\tilde{\mathbf{T}}(\mathfrak{t}) = |\{\mathbf{s} \in \tilde{\mathcal{I}}_{\mathbf{n}} : \mathbf{x}|_{\mathcal{S}+\mathbf{s}} = \mathfrak{t}\}|$$
$$\text{where} \qquad \tilde{\mathcal{I}}_{\mathbf{n}} = \{\mathbf{s} : \mathcal{S} + \mathbf{s} \subset \mathcal{I}_{\mathbf{n}}\}.$$

We assume that $\mathbf{0} \in \mathcal{S}$ and $\tilde{\mathcal{I}}_{\mathbf{n}} \subset \mathcal{I}_{\mathbf{n}}$. We know that $\mathbf{T}$ in the torus satisfies the conservation laws $\mathbf{C} \cdot \mathbf{T} = \mathbf{0}$. For the box, however, we must re-define the type $\tilde{\mathbf{T}}$ by taking into account the boundary effect on $\mathbf{T}$, that is,

$$\tilde{\mathbf{T}}(\mathfrak{t}) = \mathbf{T}(\mathfrak{t}) - |\{\mathbf{s} \in \mathcal{I}'_{\mathbf{n}} : \mathbf{x}|_{\mathcal{S}+\mathbf{s}} = \mathfrak{t}\}|$$
$$\text{where} \qquad \mathcal{I}'_{\mathbf{n}} = \mathcal{I}_{\mathbf{n}} \setminus \tilde{\mathcal{I}}_{\mathbf{n}}, \quad (27)$$

that is, we need to eliminate that shifts that drive $\mathcal{S}$ outside the box. Multiplying (27) by $\mathbf{C}$ and using $\mathbf{C} \cdot \mathbf{T} = \mathbf{0}$ we find the following conservation laws for the box:

$$\mathbf{C}\tilde{\mathbf{T}} = \mathbf{b}$$

$$\text{for} \quad \mathbf{b} = \mathbf{C} \cdot \mathbf{b}', \quad \mathbf{b}' = (-|\{\mathbf{s} \in \mathcal{I}'_{\mathbf{n}} : \mathbf{x}|_{\mathcal{S}+\mathbf{s}} = \mathfrak{t}\}|)_{\mathfrak{t}\in\mathfrak{T}}.$$

Notice that the norm of the $\mathbf{b}'$ vector is bounded by the size of the boundary, that is, $\sum_{\mathfrak{t}} |\mathbf{b}'(\mathfrak{t})| \leq |\mathcal{I}'_{\mathbf{n}}|$, which is of order $\Theta(N/\min_i n_i)$. Furthermore, the matrix $\mathbf{C}$ does not depend on the size of the boundary (only on $\mathcal{S}$ and m), therefore the number of $\mathbf{b}$ is bounded by $\Theta(N/\min_i n_i)$. Finally, by the normalization

$$\sum_{\mathfrak{t}} \tilde{\mathbf{T}}(\mathfrak{t}) = |\tilde{\mathcal{I}}_{\mathbf{n}}| =: \tilde{N},$$

types $\tilde{\mathbf{T}}$ in the box are in a $D - 1$ dimensional linear subspace. For every $\mathbf{b}$, the conservation laws $\mathbf{C}\tilde{\mathbf{T}} = \mathbf{b}$ have $O(N^\mu)$ nonnegative solutions inside a polygon. The freedom of choosing $\mathbf{b} \in \mathbb{N}^{\text{rk}(\mathbf{C})}$ allows us to shift this polygon in the remaining $\text{rk}(\mathbf{C}) = D - 1 - \mu$ dimensions by at most $\Theta(N/\min_i n_i)$, so that $\mathbf{b}$ is inside the ball of $O((N/\min_i n_i)^{D-1-\mu})$ integer points. This leads to the upper bound $O(N^{D-1}/(\min_i\ n_i)^{\text{rk}(\mathbf{C})})$ on the number of types. However, whether it is the right order of growth in this case remains an open question.

III. Analysis

In this section we provide the proofs of Theorems 3 and 7.

A. Proof of Theorem 3

In Theorem 3 we present a complete set of independent conservation laws. Specifically, we take every subshape $\mathcal{S}'$ embeddable in $\mathcal{S}$ and select one of them as the canonical one. Then we consider all $\mathfrak{t}' : \mathcal{S}' \to \mathcal{A}'$ where $\mathcal{A}' = \mathcal{A} \setminus \{m\}$ and the corresponding conservation laws $(\hat{\varepsilon}_{\mathcal{S}'}\mathbf{T})(\mathfrak{t}') - (\hat{\varepsilon}\mathbf{T})(\mathfrak{t}') = 0$ for all other $\hat{\varepsilon}$ embeddings of this subshape. Observe that the dropped symbol $m \in \mathcal{A}$ is automatically included since the following holds:

$$\mathbf{T}''(\mathfrak{t}' \cup o_m) = \mathbf{T}'(\mathfrak{t}') - \sum_{i\in\mathcal{A}'} \mathbf{T}''(\mathfrak{t}' \cup o_i)$$

where o_i is a single point/position outside $\mathfrak{t}'$ that takes value $i \in \mathcal{A}$ there. Clearly, $\mathbf{T}' = \hat{\varepsilon}'(\mathbf{T})$ and $\mathbf{T}'' = \hat{\varepsilon}''(\mathbf{T})$, where ε', ε'' are embeddings corresponding to $\mathfrak{t}'$ and $\mathfrak{t}' \cup o_i$.

To prove independence of the conservation laws, we have to define some order among them and show that $\mathbf{C}$ becomes triangular. We order the conservation laws by the size of $\mathcal{S}'$ (referred as height). We illustrate it in Figure 5, where the ordering of the columns is shown in gray leading to a triangular form of $\mathbf{C}$.

To make this more formal, let us introduce a certain basis in the space $\mathbb{R}^D$ of functions on the configurations on $\mathcal{S}$. Fix the standard basis $\{\mathbf{e}(\mathfrak{t}), \mathfrak{t} \in \mathfrak{T}(\mathcal{A}, \mathcal{S})\}$. For each tiling $\mathfrak{t}$, we can split out the *inessential part*, the cells $\boxed{b} \in \mathcal{S}$ where $\mathfrak{t}(b) = m$, and the *support* of $\mathfrak{t}$, i.e., the collection of boxes where $\mathfrak{t}(b) \neq m$. Alternatively we can enumerate the tilings $\mathfrak{t}$ of $\mathcal{S}$ by the shape $\mathcal{S}'$ of their support, by the embedding ε of this support into $\mathcal{S}$, and by the tiling of $\mathcal{S}'$ over the reduced alphabet $\mathcal{A} \setminus \{m\} =: \mathcal{A}'$. Such a basis vector we will denote as

$$\mathbf{e}(\mathcal{S}', \varepsilon, \mathfrak{t}).$$

We will call the *height* of a basis vector $\mathbf{e}(\mathcal{S}', \varepsilon, \mathfrak{t})$ the size of its support, that is

$$\mathtt{H}(\mathbf{e}(\mathcal{S}', \varepsilon, \mathfrak{t})) = |\mathcal{S}'|.$$

Further, we assign to each basis vector its *weight*, defined as follows. We first denote by $\#(b)$ the number of cells $\boxed{b}$.

	$\mathcal{S}'$	s																
1		↑		1		1	-1		-1			1		1	-1		-1	
2		↗			1	1	-1	-1					1	1	-1	-1		
3		→					-1	-1	-1	-1	1	1	1	1				
4		↑		1	1	1	-1				-1				-1			
5		→		-1	1		-1	-1			1		1					

Figure 5. The nonzero coordinates for all 5 conservation laws discussed in Theorem 3 for a 2×2 box shape and $m = 2$: the upper row shows all $D = 16$ squares corresponding to all tiles $\mathfrak{t}$. $\mathcal{S}'$ denotes the canonical embedding and $\mathbf{s}$ denotes shift for the second embedding in the $(\hat{\boldsymbol{\varepsilon}}_{\mathcal{S}'}\mathbf{T})(\mathfrak{t}') - (\hat{\boldsymbol{\varepsilon}}\mathbf{T})(\mathfrak{t}') = 0$ conservation law. Reduced alphabet is $\mathcal{A}' = \{1\}$, so we need to consider only constant $\mathfrak{t}' = 1$.

Then we number all the cells of $\mathcal{S}$ and we assign the weight of b to be $\epsilon^{\#(b)}$ for some small $\epsilon > 0$. The weight of a basis vector $\mathbf{e}(\mathcal{S}', \boldsymbol{\varepsilon}, \mathfrak{t})$ is the sum

$$\sum_{b \in \boldsymbol{\varepsilon}(\mathcal{S}')} \epsilon^{\#(b)}.$$

There is nothing very specific about this choice of the weights; the only property we will use is that for small enough ϵ, the weights corresponding to different embeddings of the same subshape $\mathcal{S}'$ are all different (which is easy to verify). In particular, for any embeddable $\mathcal{S}'$, there exists a unique embedding $\boldsymbol{\varepsilon}_{\mathcal{S}'}$ of $\mathcal{S}'$ having *maximal weight* among all such embeddings $\boldsymbol{\varepsilon} : \mathcal{S}' \to \mathcal{S}$. We will be calling this basis vector the *anchor* of the embeddable shape $\mathcal{S}'$ and its embedding as *canonical*.

We group the basis vectors $\mathbf{e}(\mathcal{S}', \boldsymbol{\varepsilon}, \mathfrak{t})$ according to their height (increasing left to right), and within each height by the support shape $\mathcal{S}'$ and within each group corresponding to a support shape $\mathcal{S}'$ by the tiling $\mathfrak{t}$ of $\mathcal{S}'$ over the reduced alphabet $\mathcal{A}'$. Finally, within each such group (corresponding to a given subshape $\mathcal{S}'$ and its tiling $\mathfrak{t}$), we order the basis vectors $\mathbf{e}(\mathcal{S}', \boldsymbol{\varepsilon}, \mathfrak{t})$ by the *weight of the embeddings* $\boldsymbol{\varepsilon}$. In particular, the anchor within each group is the rightmost element. This defines a complete ordering on the basis vectors.

Now we are ready to prove Theorem 3. We will be using the basis consisting of the standard vectors $\mathbf{e}(\mathcal{S}', \boldsymbol{\varepsilon}, \mathfrak{t})$ ordered as described above, left to right. The rows of the (sub)matrix $\mathbf{C}_m$ (defined in Theorem 3) correspond to the covectors

$$\boldsymbol{v}_{(\mathcal{S}', \mathfrak{t}, \boldsymbol{\varepsilon}_{\mathcal{S}'}, \boldsymbol{\varepsilon})}.$$

Each such covector has exactly two components,

$$\mathbf{e}(\mathcal{S}', \boldsymbol{\varepsilon}_{\mathcal{S}'}, \mathfrak{t}) - \mathbf{e}(\mathcal{S}', \boldsymbol{\varepsilon}, \mathfrak{t})$$

in the group of height $\mathtt{H}(\mathbf{e}(\mathcal{S}', \boldsymbol{\varepsilon}_{\mathcal{S}'}, \mathfrak{t})) = \mathtt{H}(\mathbf{e}(\mathcal{S}', \boldsymbol{\varepsilon}, \mathfrak{t}))$; all other components have *higher height*.

It follows that, if one augments $\mathbf{C}_m$ by the rows with basis vectors $\mathbf{e}(\mathcal{S}', \boldsymbol{\varepsilon}_{\mathcal{S}'}, \mathfrak{t})$, running over all embeddable subshapes $\mathcal{S}'$, and their tilings $\mathfrak{t}$, then the leftmost vectors in the rows will be all different. Finally, sorting the rows according to these leftmost elements will result in the upper-triangular matrix.

This, in turn, implies that the basis elements

$$\mathbf{e}(\mathcal{S}', \boldsymbol{\varepsilon}_{\mathcal{S}'}, \mathfrak{t})$$

span a complement to the kernel of $\mathbf{C}_m$, and therefore the kernel of $\mathbf{C}$ has dimension at most the number of tilings by symbols of $\mathcal{A}'$ of embeddable shapes $\mathcal{S}'$.

Denote the subspace of $\mathscr{V} \cong \mathbb{R}^D$ spanned by the basis vectors as

$$L_{\mathcal{S}} := \mathbf{e}(\mathcal{S}', \boldsymbol{\varepsilon}_{\mathcal{S}'}, \mathfrak{t}), \mathfrak{t} \in (\mathcal{A}')^{\mathcal{S}'}, \mathcal{S}' \text{embeddable into } \mathcal{S}.$$

To prove that $\operatorname{Ker} \mathbf{C} = \operatorname{Ker} \mathbf{C}_m$, we will produce for any torus $\mathcal{O}_{\mathbf{n}}$ of sufficiently large $\mathbf{n}$, a collection of tilings, of size $\dim(L_{\mathcal{S}})$, such that their frequency vectors, paired with the basis vectors spanning $L_{\mathcal{S}}$, result in a triangular matrix with $\pm$ on the diagonal. This implies that $\operatorname{Ker} \mathbf{C} = \operatorname{Ker} \mathbf{C}_m$.

Let $\mathbf{n}(\mathcal{S})$ be the smallest vector $\mathbf{n}$ such that the box $\mathcal{I}_{\mathbf{n}}$ contains $\mathcal{S} + \mathcal{S} = \{a + b : a, b \in \mathcal{S}\}$ (understood as the Minkowski sum). We will be always assuming that $\mathcal{S}$ is embedded into this interval (denoted as $\mathcal{I}_{\mathcal{S}}$) in a fixed way.

Let $\mathfrak{t} \in \mathfrak{T}(\mathcal{A}, \mathcal{S})$ be a tiling of the shape $\mathcal{S}$, and $\mathcal{S}'$ its support (i.e., the set of cells $\boxed{b}$ where $\mathfrak{t}(b) \neq m$), and $\mathfrak{t}'$ the corresponding tiling of $\mathcal{S}'$ by symbols of the reduced alphabet $\mathcal{A}'$. For any large enough torus $\mathcal{O}_{\mathbf{n}}$, place a single copy of $\mathfrak{t}$, in an arbitrary way in the torus, extending it to the rest of the torus by the symbol m. Denote the corresponding frequency vector $\mathbf{T}_{\mathcal{S}', \mathfrak{t}'}$.

Lemma 8. *Consider the matrix of scalar products* $\mathbf{T}_{\mathcal{S}', \mathfrak{t}'}$ *and* $\mathbf{e}(\mathcal{S}'', \boldsymbol{\varepsilon}_{\mathcal{S}''}, \mathfrak{t}'')$ *where both* $\mathcal{S}', \mathfrak{t}'$ *and* $\mathcal{S}'', \mathfrak{t}''$ *run over all embeddable subshapes and their tilings by the reduced alphabet. Then, if the subshapes are ordered by their heights, the matrix is upper triangular, with* $1/N$ *on the diagonal.*

Proof. The proof is straightforward: the type vector $\mathbf{T}_{\mathcal{S}', \mathfrak{t}'}$ is produced by scanning through the torus by the shifts of $\mathcal{S}$. There is a unique position where the support lands on the anchor of $\mathcal{S}'$, and all other positions are either not anchored (thus yielding zero products with the basis vectors $\mathbf{e}(\mathcal{S}'', \boldsymbol{\varepsilon}_{\mathcal{S}''}, \mathfrak{t}'')$), or have lower height. ■

We remark that one can modify the tiling of the torus: in lieu of a single copy of the interval containing a copy of $\mathcal{S}$, one could tile $\mathcal{O}_{\mathbf{n}}$ with $\Theta(N)$ copies of the shape $\mathcal{S}'$, supporting $\mathfrak{t}'$. In this case, the rescaled frequency vector $\mathbf{T}/N$ would converge, as $\mathbf{n}$ increases, to some vector $\mathbf{T}_{\mathcal{S}', \mathfrak{t}'} \in \hat{\mathcal{P}}$.

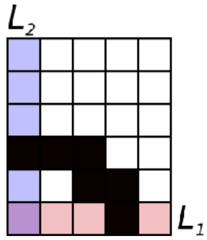

Figure 6. Example of a 6-cell shape $\mathcal{S}'$ in its lowest position.

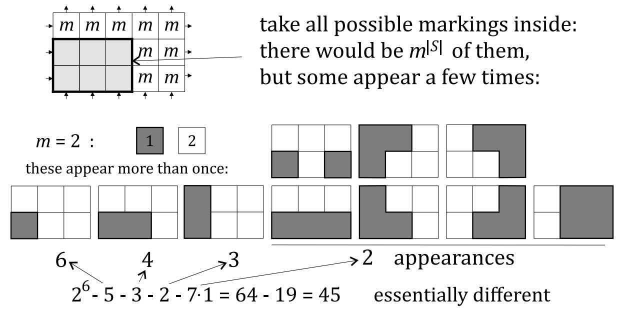

Figure 7. An illustration to the construction in Lemma 9: We place a tile with $m^{|\mathcal{S}|}$ possible patterns in the torus with all remaining positions filled with symbol m. The number of such fields is equal to $D - \mathrm{rk}(\mathbf{C}) = 2^6 - 19 = 45$, as desired.

B. Proof of Corollary 4

By Theorem 3, we need to sum $(m-1)^{|\mathcal{S}'|}$ over all shapes embeddable into $\mathcal{S}$ as in (16). Among all possible embeddings, there is a unique one that is (lexicographically) minimal. One can think about a gravity force pointing along the vector $(-1, \dots, -1)$ and forcing $\mathcal{S}'$ to slide inside the box $\mathcal{S}$ to its lowest position. Clearly, this lowest position is characterized by the condition that $\mathcal{S}'$ has non-empty intersection with the lowest k-th coordinate layer

$$\mathsf{L}_k = I_{l_1} \times \ldots \times \{1\} \times \ldots \times I_{l_d}$$

(here $\{1\}$ stands in k-th place in the product), for each $k \in \{1, \dots, d\}$; see Figure 6.

Alternatively, the sum we need to evaluate is the total number of all tilings of the box shape $\mathcal{S}$ with each of the layers $\mathsf{L}_k, k = 1, \dots, d$, containing at least one cell marked with a symbol of the reduced alphabet $\mathcal{A}'$.

The set of tilings with at least one cell in L_k marked by an element of $\mathcal{A}'$ is, clearly, the complement of the set of all tilings with the tilings having *all cells in* L_k *marked with* m. Denote the latter set of such tilings by M_k. The size of the set of tilings we are interested in is therefore,

$$|\mathfrak{T}(\mathcal{S})| - \sum_k |M_k|.$$

By inclusion-exclusion, this is equal to

$$\sum_{J \subset \{1,\dots,d\}} (-1)^{|J|} |\cap_{j \in J} M_j|,$$

where the summation is over all subsets of $\{1, \dots, d\}$ (for the empty subset, we take the summand to be $|\mathfrak{T}(\mathcal{S})|$).

The cardinality of the set $\cap_{j \in J} M_j$ is, clearly, the number of all tilings which have cells in $\cup_{j \in J} \mathsf{L}_j$ equal to m, which is, obviously, the number of all tilings of $\mathcal{S} - \cup_{j \in J} \mathsf{L}_j$. Put together, these formulae imply the corollary. ■

C. Proof of Theorem 7

Since $\mathcal{P}_\mathbf{n} \subset \mathcal{F}_\mathbf{n}$ and $\hat{\mathcal{F}}$ is a convex polytope, we conclude the upper bound on $|\mathcal{P}_\mathbf{n}|$ from the Ehrhart's Theorem 6 applied to $\mathcal{F}_\mathbf{n}$. Therefore, we can now focus on establishing a lower bound. We will accomplish it by constructing a family of tilings with a set of frequency vectors growing as N^μ. Specifically, we will first construct building blocks: $\mu + 1$ small linearly independent tilings. Then we will construct large tilings by concatenating these small ones, obtaining a regular lattice of frequency vectors in the μ dimensional simplex on these $\mu + 1$ vertices.

Let us first observe that if the torus is not large enough, there are some additional constraints due to the cyclical boundary condition. For example, in 1D case for $\mathcal{S} = \{0, 1, 2\}$ and torus/cycle $\mathcal{O} = \{0, 1, 2, 3\}$, the tile "111" automatically enforces the tile having "1*1", where "*" is any letter on the remaining position. These additional constraints can reduce the dimension of realizable frequency vectors. For example for 3×2 rectangular shape and $m = 2$, there are only 21 linearly independent possible tilings of a 3×3 torus. For 4×3 torus this number grows to 42, and finally saturates at the promised value $\mu + 1 = 45$ for a 5×3 torus.

In the next lemma we construct $\mu + 1$ linearly independent frequency vectors. To formulate it, we need to define the *width* of the shape $\mathcal{S}$ as the smallest $(w_1, \dots, w_d) \in \mathbb{N}^d$ such that for some shift $S \subset I_{w_1} \times \ldots \times I_{w_d}$.

Lemma 9. *If $n_i \geq 2w_i - 1$ for all $i = 1, \dots, d$, then there exist $\mu + 1$ tilings of $\mathcal{O}_\mathbf{n}$ with linearly independent frequency vectors.*

Proof: We will construct these tilings as

$$\texttt{Conf}^0(\mathcal{D}) :=$$
$$= \{\mathbf{x} : \mathcal{O}_\mathbf{n} \to \mathcal{A} \ : \ \mathbf{x}(\mathbf{a}) = m \quad \text{for all} \quad \mathbf{a} \in \mathcal{O}_\mathbf{n} \setminus \mathcal{S}\},$$

that is, the torus is filled with $m \in \mathcal{A}$ outside the $\mathcal{S}$ shape. The remaining $|\mathcal{S}|$ values $\mathbf{x}|_\mathcal{S}$ can be selected in $m^{|\mathcal{S}|} = D$ ways, which is more than $\mu + 1 = D - \mathrm{rk}(\mathbf{C})$. However, this set contains tilings differing only by a shift and hence having identical frequency vectors. For a given field $\mathbf{x} \in \texttt{Conf}^0(\mathcal{D})$, let $\mathcal{S}' \subset \mathcal{S}$ be a subset on which the corresponding tiling $\mathfrak{t}'$ has values different than m, that is,

$$\mathcal{S}' = \{\mathbf{a} : \mathbf{x}(\mathbf{a}) \in \{1, \dots, m-1\}\} \subset \mathcal{S}. \qquad (28)$$

Observe that if $\mathcal{S}' + \mathbf{s} \subset \mathcal{S}$ for some $\mathbf{s} \in \mathbb{Z}^d$, then there exists an element of $\texttt{Conf}^0(\mathcal{D})$ differing from $\mathbf{x}$ only by the shift $\mathbf{s}$. There are $|\{\mathbf{s} : \mathcal{S}' + \mathbf{s} \subset \mathcal{S}\}| - 1$ such elements of $\texttt{Conf}^0(\mathcal{D})$ having identical frequency vector. For given $\mathcal{S}'$ there are $(m-1)^{|\mathcal{S}'|}$ such situations ($\mathbf{x}|_{\mathcal{S}'}$), thus from the initial $|\texttt{Conf}^0(\mathcal{D})| = m^{|\mathcal{S}|}$ tilings we need to subtract

$$\sum_{\mathcal{S}' \subset \mathcal{S}} (m-1)^{|\mathcal{S}'|} \left(|\{\mathbf{s} : \mathcal{S}' + \mathbf{s} \subset \mathcal{S}\}| - 1\right)$$

redundant shifted copies. This gives exactly $\mathrm{rk}(\mathbf{C})$ as in (16). Finally, if we count only once all elements of $\texttt{Conf}^0(\mathcal{D})$

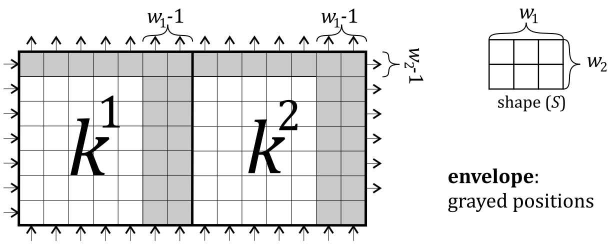

Figure 8. Illustration to Lemma 10.

differing by a shift, then their number is exactly $m^{|S|} - \mathrm{rk}(\mathbf{C}) = \mu + 1$ as desired. This is illustrated in Figure 7 for $m = 2$.

It remains to show that this set of $\mu + 1$ frequency vectors is linearly independent. This is in essence already shown in the proof of the Theorem 3: we exhibited a collection of $\mu + 1$ covectors such that the pairing matrix of these covectors with the type vectors we constructed is upper triangular, in the ordering constructed there. The result follows immediately. ∎

To continue our construction, we now concatenate the just constructed tilings of size $2\mathbf{w} - 1$ designing a set of tilings of size growing as desired N^μ. Generally, such concatenation can lead to some new tiles near the boundary, but this problem disappears if concatenated tilings are identical on the *envelope* defined as

$$E = (I_{n_1} \times \ldots \times I_{n_d}) \setminus (I_{n_1 - w_1 + 1} \times \ldots \times I_{n_d - w_d + 1})$$

which we assume to hold. To complete the proof, we need another simple lemma.

Lemma 10. *If $\mathbf{x}^1,\ \mathbf{x}^2$ tilings of size $\mathbf{n}$ and frequency vectors $\hat{\mathbf{T}}^1$, $\hat{\mathbf{T}}^2$ are identical on the envelope E, that is, $\mathbf{x}^1|_E = \mathbf{x}^2|_E$, then the frequency vector of a tiling constructed by concatenating them is $\hat{\mathbf{T}}^{12} = (\hat{\mathbf{T}}^1 + \hat{\mathbf{T}}^2)/2$.*

Proof: Observe that the resulting tile appears in all positions from both original tilings. But the size of the torus is twice as big leading to the average frequency vector $\hat{\mathbf{T}}^{12} = (\hat{\mathbf{T}}^1 + \hat{\mathbf{T}}^2)/2$. This is illustrated in Figure 8. ∎

We are now in the position to complete the **proof of Theorem 7**. For $n_i \geq 2w_i - 1$, we construct a family of periodic tilings for which the set of frequency vectors grows like N^μ. We take $\mu + 1$ tilings with linearly independent size $2\mathbf{w} - 1$ frequency vectors $\hat{\mathbf{T}}^1, \ldots, \hat{\mathbf{T}}^{\mu+1}$ as discussed in Lemma 9. We can concatenate them into larger tori and then the resulting frequency vector corresponds to a convex combination. Thus the resulting frequency vectors are

$$\left\{ \frac{a_1 \hat{\mathbf{T}}^1 + \ldots + a_{\mu+1} \hat{\mathbf{T}}^{\mu+1}}{a_1 + \cdots + a_{\mu+1}} : \forall_i\ a_i \in \mathbb{N},\ \sum_i a_i = \frac{N}{N'} \right\}$$

where $N' = \prod_i (2w_i - 1)$ and a_i is the number of tiles with the frequency vector $\hat{\mathbf{T}}_i$. Observe now that the size of this discrete simplex is determined by the number of integer solutions of

$$\sum_i a_i = \frac{N}{N'}$$

which is

$$\binom{N/N' + \mu - 1}{\mu} = O(N^\mu).$$

This implies the existence of a lower bound when n_i are integer multiples of $2w_i - 1$. In the general case we can fill the remaining positions with m. This completes the construction of a lower bound, and the proof of Theorem 7.

References

[1] F. Ardila and R. Stanley, Tilings, Clay Public Lecture at the IAS/Park City Mathematics Institute, July, 2004; see also the Clay Institute webasite: http://www.claymath.org/library/senior-scholars/stanley-ardila-tilings.pdf

[2] M. Beck and S. Robins, *Computing the Continuous Discretely*, Undergraduate Texts in Mathematics. New York: Springer, 2007.

[3] P. Billingsley, Statistical methods in Markov chains, *Ann. Math. Statistics*, 32, 12-40, 1961.

[4] T. Cover and J.A. Thomas, *Elements of Information Theory*, John Wiley & Sons, New York, 1991.

[5] I. Csziszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*, Academic Press, New York, 1981.

[6] I. Csziszár, The method of types, *IEEE Trans. Information Theory*, 44, 2505-2523, 1998.

[7] L. Davisson, Universal noiseless coding, *IEEE Trans. Inf. Theory*, 19, 783–795, 1973.

[8] M. Drmota and W. Szpankowski, Precise minimax redundancy and regret, *IEEE Trans. Information Theory*, 50, 2686-2707, 2004.

[9] E. Ehrhart, Sur une probleme de geomtrie diophantine lineaire, *J. reine angew. Math.*, 227, 1-29, 1967.

[10] M. Feder, N. Merhav, and M. Gutman, Universal prediction of individual sequences, *IEEE Trans. Information Theory*, 38, 1258–1270, 1992.

[11] P. Jacquet and W. Szpankowski, Markov types and minimax redundancy for Markov sources, *IEEE Trans. Information Theory*, 50, 1393-1402, 2004.

[12] P. Jacquet, C. Knessl, W. Szpankowski, Counting Markov types, balanced matrices, and Eulerian graphs. *IEEE Transactions on Information Theory* 58(7), 4261-4272, 2012.

[13] P. Kasteleyn, The statistics of dimers on a lattice. I. The number of dimer arrangements on a quadratic lattice, *Physica*, 27 (12), 1961.

[14] R. Kindermann and J. Snell, *Markov Random Fields and Their Applications*, American Mathematical Society, 1980.

[15] R. Kenyon, The planar dimer model with boundary: a survey, Directions in mathematical quasicrystals, CRM Monogr. Ser. 13, Providence, R.I.: American Mathematical Society, 307-328, 2000.

[16] S. Lang, *Linear Algebra*, Springer, New York, 1976.

[17] A. Martín, G. Seroussi, and M. J. Weinberger, Type classes of tree models, *IEEE Trans. Inf. Theory*, 58, 4077-4093, 2012. *Proc. ISIT 2007*, Nice, France, 2007.

[18] N. Merhav and M. J. Weinberger, On universal simulation of information sources using training data, *IEEE Trans. Inf. Theory*, 50, 1, 5-20, 2004.

[19] D. Metzler and W. Croft, A Markov random field model for term dependencies, *SIGIR'05*, Salvador, Brazil, 2005.

[20] Mézard and A. Montanari, *Information, Physics, and Computation*, Oxford University Press, 2009.

[21] I. Pak, Tile invariants: New survey, *Theoretical Computer Science*, 303, 303-331, 2003.

[22] Y. Rachlin, R. Negi and P. Khosla, Sensing capacity for Markov random fields, *Proc. IEEE Int. Symp. Inf. Theory*, Adelaide, 2005.

[23] J. Rissanen, Fisher information and stochastic complexity, *IEEE Trans. Information Theory*, 42, 40-47, 1996.

[24] R. Rockafellar, *Convex Analysis*, Princeton University Press, 1996.

[25] G. Seroussi, On universal types, *IEEE Trans. Information Theory*, 52, 171-189, 2006.

[26] P. Shields, Universal redundancy rates do not exist, *IEEE Trans. Information Theory*, 39, 520-524, 1993.

[27] R. Stanley, *Enumerative Combinatorics*, Vol. II, Cambridge University Press, Cambridge, 1999.

[28] W. Szpankowski, *Average Case Analysis of Algorithms on Sequences*, Wiley, New York, 2001.

[29] P.O. Vontobel, Counting in graph covers: A combinatorial characterization of the Bethe entropy function, *IEEE Trans. Information Theory*, 59, 6018-6048, 2013.

[30] J.S. Yedidia, W.T. Freeman, and Y. Weiss, Constructing free-energy approximations and generalized belief propagation algorithms, *IEEE Trans. Information Theory*, 51, 2282 - 2312, 2005.
[31] P. Whittle, Some distribution and moment formulæ for Markov chain, *J. Roy. Stat. Soc.,* Ser. B., 17, 235-242, 1955.

Yuliy Barsyhnikov

Jarosław (Jarek) Duda received the S.M. degrees in computer science, mathematics and physics in 2004, 2005 and 2006 respectively, then Ph.D. degrees in computer science and physics in 2010 and 2012 respectively, all in Jagiellonian University, Krakow, Poland. In 2013-2014 he was a postdoctoral researcher in the Center of Science of Information at Purdue University. He is currently an assistant professor at Jagiellonian University, Krakow, Poland. His main interest is information theory.

Wojciech Szpankowski is Saul Rosen Distinguished Professor of Computer Science and (by courtesy) Electrical and Computer Engineering at Purdue University where he teaches and conducts research in analysis of algorithms, information theory, bioinformatics, analytic combinatorics, random structures, and stability problems of distributed systems. He received his M.S. and Ph.D. degrees in Electrical and Computer Engineering from Gdansk University of Technology. He held several Visiting Professor/Scholar positions, including McGill University, INRIA, France, Stanford, Hewlett-Packard Labs, Universite de Versailles, University of Canterbury, New Zealand, Ecole Polytechnique, France, the Newton Institute, Cambridge, UK, ETH, Zurich, and Gdańsk University of Technology, Poland. He is a Fellow of IEEE, and the Erskine Fellow. In 2010 he received the Humboldt Research Award. He published two books: ”Average Case Analysis of Algorithms on Sequences”, John Wiley & Sons, 2001, and ”Analytic Pattern Matching: From DNA to Twitter”, Cambridge, 2015. He has been a guest editor and an editor of technical journals, including THEORETICAL COMPUTER SCIENCE, the ACM TRANSACTION ON ALGORITHMS, the IEEE TRANSACTIONS ON INFORMATION THEORY, FOUNDATION AND TRENDS IN COMMUNICATIONS AND INFORMATION THEORY, COMBINATORICS, PROBABILITY, AND COMPUTING, and ALGORITHMICA. In 2008 he launched the interdisciplinary Institute for Science of Information, and in 2010 he became the Director of the newly established NSF Science and Technology Center for Science of Information.