J Stat Phys (2017) 166:765–782
DOI 10.1007/s10955-016-1595-8

Translation Invariant Extensions of Finite Volume Measures

S. Goldstein[1] · T. Kuna[2] · J. L. Lebowitz[1,3] · E. R. Speer[1]

Received: 14 June 2016 / Accepted: 30 July 2016 / Published online: 11 August 2016
© Springer Science+Business Media New York 2016

Abstract We investigate the following questions: Given a measure μ_Λ on configurations on a subset Λ of a lattice $\mathbb{L}$, where a configuration is an element of Ω^Λ for some fixed set Ω, does there exist a measure μ on configurations on all of $\mathbb{L}$, invariant under some specified symmetry group of $\mathbb{L}$, such that μ_Λ is its marginal on configurations on Λ? When the answer is yes, what are the properties, e.g., the entropies, of such measures? Our primary focus is the case in which $\mathbb{L} = \mathbb{Z}^d$ and the symmetries are the translations. For the case in which Λ is an interval in $\mathbb{Z}$ we give a simple necessary and sufficient condition, *local translation invariance* (*LTI*), for extendibility. For LTI measures we construct extensions having maximal entropy, which we show are Gibbs measures; this construction extends to the case in which $\mathbb{L}$ is the Bethe lattice. On $\mathbb{Z}$ we also consider extensions supported on periodic configurations, which are analyzed using de Bruijn graphs and which include the extensions with minimal entropy. When $\Lambda \subset \mathbb{Z}$ is not an interval, or when $\Lambda \subset \mathbb{Z}^d$ with $d > 1$, the LTI condition is necessary but not sufficient for extendibility. For $\mathbb{Z}^d$ with $d > 1$, extendibility is in some sense undecidable.

Keywords Local translation invariance · Translation invariant extensions · Maximal entropy extensions · de Bruijn graphs · Bethe lattice

Dedicated to David Ruelle and Yasha Sinai on the occasion of their eightieth birthdays.

✉ E. R. Speer
speer@math.rutgers.edu

1 Department of Mathematics, Rutgers University, New Brunswick, NJ 08903, USA

2 Department of Mathematics and Statistics, University of Reading, Whiteknights, PO Box 220, Reading RG6 6AX, UK

3 Department of Physics, Rutgers University, New Brunswick, NJ 08903, USA

1 Introduction

We consider an extension problem for measures describing configurations on a lattice $\mathbb{L}$, in which the value η_i of the configuration η at site (or equivalently vertex) i of $\mathbb{L}$ belongs to some fixed set Ω. Let $X = \Omega^{\mathbb{L}}$ be the set of all configurations on $\mathbb{L}$ and, for $\Lambda \subset \mathbb{L}$, let $X_\Lambda = \Omega^\Lambda$ be the set of configurations on Λ. Then we address the following extension problem: for a fixed $\Lambda \subset \mathbb{L}$ and probability measure μ_Λ on X_Λ, does there exist a probability measure μ on X such that (i) μ is invariant under some specified group of symmetries of $\mathbb{L}$ and (ii) the marginal measure of μ on X_Λ is μ_Λ? When μ exists we speak of it as an *invariant extension* of μ_Λ or, since we consider only invariant extensions, simply as an *extension*. In this case we say that μ_Λ is *extendible*.

For most of this paper the lattice $\mathbb{L}$ will be taken to be $\mathbb{Z}^d$ for some $d \geq 1$, and the symmetries of interest will be translations or, occasionally, translations plus reflections about a coordinate axis. In this case we will refer to an invariant extension μ as a *TI extension* or, as above, simply as an *extension*. In Sect. 3 we discuss the case in which $\mathbb{L}$ is the Bethe lattice.

Schlijper [27–30] studied the problem of TI extensions on $\mathbb{Z}^d$ in the 1980's, in connection with its application to the *cluster-variation method* [22], used to obtain approximations to the equilibrium thermodynamic properties of lattice systems. The problem was also studied by Pivato [24] in 2001 and by Chazottes et al. [6] in 2012. For completeness we will summarize or rederive some of the results of these papers in the course of describing our own work; we discuss this in more detail at the end of this section.

For $\Lambda \subset \mathbb{Z}^d$ we say that measure μ_Λ on X_Λ is *locally translation invariant* (LTI) if for any subsets $A, A' \subset \Lambda$, with A' a translate of A, the marginal of μ_Λ on $X_{A'}$ is the translate of the marginal on X_A. The LTI property is clearly necessary for the existence of a TI extension μ of μ_Λ and one question we investigate is whether or not it is also sufficient.

From now on we will assume that, unless otherwise stated, Λ is finite. Our primary focus will be the case in which Ω is finite. Our results then do not depend significantly on the cardinality of Ω; for simplicity of presentation we will often take Ω to be $\{0, 1\}$, so that a configuration $\eta \in X$ may be thought of as describing a particle system or point process on the lattice. We will occasionally point out instances in which our results extend to certain infinite Ω.

For finite Ω we have a good understanding of the situation when $d = 1$ and Λ is an interval. In particular, the LTI condition is in this case sufficient for the existence of the TI extension [24,29], as we verify in Sect. 2 by an explicit construction of such an extension; the resulting μ is in fact the unique TI extension which maximizes the entropy of its marginal on $X_{\Lambda'}$ for any interval $\Lambda' \supset \Lambda$. Moreover, if Λ contains $k+1$ sites then μ is the shift-invariant measure on the sample paths of a k-step Markov process, and as such is a Gibbs measure with TI interactions which have range at most k and involve at most $k+1$ sites. We obtain an explicit expression for the Gibbs–Shannon entropy of $\mu_{\Lambda'}$ as an affine function of the size of Λ'. This extension method applies also when Λ is an infinite strip in $\mathbb{Z}^2$ or a higher-dimensional analogue of such a strip, or is one of a certain class of subgraphs of the Bethe lattice which are analogous to intervals in $\mathbb{Z}$.

In the one-dimensional case, with Λ an interval, Ω finite, and μ_Λ LTI, there exist also TI extensions of μ_Λ which are supported on a finite number of periodic configurations [6]. We discuss this and related results in Sects. 4 and 5 . Every TI measure of this type has finite entropy, in a sense which will be made precise later, and we show that every extension with minimal entropy is of this type (there is in general no unique minimal entropy extension). However, for general μ_Λ we are not able to give an explicit construction of a minimal entropy

extension or even to determine the minimal entropy, although we do give an *a priori* bound for the latter. A key element of our analysis here is the identification of a measure on X_Λ with an edge-weighting of a de Bruijn graph; the LTI condition on the measure then corresponds to current conservation at the vertices of the graph.

When one passes to the extension problem for Λ not an interval of $\mathbb{Z}$ the situation is more complex. In Sect. 6 we give simple examples of LTI measures μ_Λ which are not extendible; for example, with $\Lambda = \{i, i+1, i+3\} \subset \mathbb{Z}$ and Λ a unit square in $\mathbb{Z}^2$ (with four lattice sites). For the latter example it is relevant to remark that when $\Lambda \subset \mathbb{Z}^2$ is a product of intervals, and μ_Λ is LTI, the maximal entropy extension procedure mentioned above can be used to extend μ_Λ in one direction, i.e., to an infinite strip, and the resulting extension is TI in the direction of extension. In this process, however, the LTI property in the second direction may be lost; only if it is preserved can one repeat the process to obtain an extension to all of $\mathbb{Z}^2$. Similar remarks apply in higher dimensions.

As noted above, when $\Omega = \{0, 1\}$ we are studying point processes on $\mathbb{Z}^d$. In this context the current paper may be viewed as continuing work on a long-standing problem: the existence of point processes when one is given partial information about the desired process (see, for example [8,23,31] and further references given in [17]). In our own previous work [4,7,16,17] we considered the case in which the information given was all k-point correlation functions of the process up to some order k_0; this is in essence a *truncated moment problem*. Extending the work of [1] we obtained a sufficient condition, essentially one of low density, for the existence of the process. Complete necessary and sufficient conditions were found in [17]; these are worked out in detail for the case $k_0 = 2$, but the method is general. See also [18].

We finally note that, for fixed $\Lambda \subset \mathbb{Z}^d$, the set of measures μ_Λ on X_Λ which can be extended to a TI measure on $\mathbb{Z}^d$ is closed in the weak topology. For if $(\mu_{\Lambda,n})_{n=1}^\infty$ is a sequence of measures on X_Λ converging weakly to a measure μ_Λ, and $\hat{\mu}_n$ is a TI extension of $\mu_{\Lambda,n}$ to X, then some subsequence of $(\hat{\mu}_n)_{n=1}^\infty$ will converge, by compactness, and the limit will be a TI extension of μ_Λ.

The outline of the rest of the paper is as follows. In Sect. 2 we show that the LTI condition is sufficient, when Λ is an interval in $\mathbb{Z}$, for the existence of a TI extension; we do so by explicit construction of the maximal entropy, Markovian, extension. We also point out that this extension is Gibbsian. In Sect. 3 we take up the extension problem on the Bethe lattice; there also we construct maximal entropy extensions which are Markovian and Gibbsian. In Sect. 4 we return the the case in which Λ is an interval in $\mathbb{Z}$ and discuss a different class of extensions, those supported on periodic configurations, and in Sect. 5 show that these include the extensions of minimal possible entropy. In Sect. 6 we give several examples to show that, in $\mathbb{Z}^d$, the LTI condition is not sufficient for extendibility when Λ is not an interval in $\mathbb{Z}$; we give also a corresponding example for the Bethe lattice. Section 7 takes up the issue of the decidability of the extension problem.

We finally discuss briefly the overlap of this paper with other work. As noted above, the Markovian extension of Sects. 2 and 2.1 has been discussed before, see in particular [24,29], although its Gibbsian nature has not been noted. Periodic extensions are discussed [6,24], and our results in Sect. 4 partially overlap with this work. Our discussion of the Bethe lattice in Sect. 3 is, we believe, new, as is the discussion of minimal entropy measures in Sect. 5. Examples 16 and 17 of Sect. 6 are new. We first learned of the existence of undecidability questions for the extension problem in higher dimensions from Ruelle [26]; aspects of this problem have also been discussed in [6,24,30], but we give a formulation that we feel has something new to offer.

2 Maximal Entropy Extensions in One Dimension

We will adapt the general notation of the introduction to the one-dimensional case considered in this and in Sects. 4 and 5 by letting Λ_k, for $k \geq 0$, be the subset $\{0, 1, \ldots, k\}$ of $\mathbb{Z}$, $X_k = \{0, 1\}^{\Lambda_k}$ be the set of particle configurations on Λ_k, and $\pi_k \mu$, for μ a measure on X or on X_j with $j > k$, be the marginal of μ on X_k. We suppose that we are given an LTI measure μ_k on X_k and construct explicitly a (TI) measure μ on $X = \{0, 1\}^{\mathbb{Z}}$ which extends μ_k, i.e., satisfies $\pi_k \mu = \mu_k$.

We will usually invoke the condition that μ_k is LTI as the requirement that the marginal $\mu_{k-1} = \pi_{k-1}\mu_k$ of μ_k on X_{k-1} be the translation of the marginal on configurations on the sites $\{1, 2, \ldots, k\}$:

$$\begin{aligned} \mu_{k-1}(\eta_0, \ldots, \eta_{k-1}) &= \sum_{\sigma=0,1} \mu_k(\eta_0, \ldots, \eta_{k-1}, \sigma), \\ &= \sum_{\sigma=0,1} \mu_k(\sigma, \eta_0, \ldots, \eta_{k-1}) \end{aligned} \tag{1}$$

We say that μ_k is *symmetric* if it is invariant under reflections about the center of Λ_k. It is easy to verify that if μ_k is LTI and $j < k$ then the marginal of μ_k on X_j is also LTI, and if μ_k is also symmetric then so is this marginal. In fact, a symmetric measure μ_k is LTI if and only if its marginal on X_{k-1} is symmetric.

We begin the construction of μ by extending μ_k to an LTI measure μ_{k+1} on X_{k+1}, defining $\mu_{k+1}(\eta_0, \eta_1, \ldots, \eta_{k+1}) = 0$ if $\mu_{k-1}(\eta_1, \ldots, \eta_k) = 0$, and

$$\mu_{k+1}(\eta_0, \eta_1, \ldots, \eta_{k+1}) = \frac{\mu_k(\eta_0, \eta_1, \ldots, \eta_k)\mu_k(\eta_1, \eta_2, \ldots, \eta_{k+1})}{\mu_{k-1}(\eta_1, \ldots, \eta_k)}, \tag{2}$$

otherwise. One verifies easily that μ_{k+1} is LTI and has marginal μ_k on X_k. Moreover, if μ_k is symmetric, so is μ_{k+1}. This leads immediately to the main result of this section.

Theorem 1 *If μ_k is an LTI measure on X_k then there exists a TI measure μ on X which extends μ_k. Moreover, if μ_k is symmetric then μ may be taken to be symmetric under any reflection.*

Proof Suppose $\Lambda = \{-l, -l + 1, \ldots, j\}$ with $l \geq 0$ and $j \geq k$. By the repeated application of (2) (and a translation) one may find, for any l and j, an LTI measure μ_Λ on $\{0, 1\}^\Lambda$ whose marginal on X_k is μ_k. The result then follows from the Kolmogorov existence theorem. The preservation of symmetry is immediate. □

There is another approach to this extension procedure. Equation (2) can be written as

$$\mu_{k+1}(\eta_0, \eta_1, \ldots, \eta_{k+1}) = \mu_k(\eta_0, \eta_1, \ldots, \eta_k)\mu_k(\eta_{k+1} \mid \eta_1, \ldots \eta_k), \tag{3}$$

where the conditional probability $\mu_k(\zeta_k \mid \zeta_0, \ldots \zeta_{k-1})$ is defined to be zero whenever $\mu_{k-1}(\zeta_0, \ldots \zeta_{k-1}) = 0$, and then one sees that the extension procedure described in the proof of Theorem 1 is equivalent to

$$\mu_j(\eta_0, \eta_1, \ldots, \eta_j) = \mu_k(\eta_0, \ldots, \eta_k) \prod_{i=1}^{j-k} \mu_k(\eta_{i+k} \mid \eta_i, \ldots, \eta_{i+k-1}). \tag{4}$$

Equation (4) says that we may regard the extension procedure as defining a Markov chain with state space $\{0, 1\}$ having memory, for which the transition probabilities depend on states

at the previous $k-1$ time steps. The TI extension μ of μ_k given in Theorem 1 is then just the invariant measure on sample paths for this chain. Equivalently one may think of a 1-step Markov chain with state space X_k.

Remark 2 It is sometimes convenient to use spin notation rather than particle (i.e., lattice gas) notation in describing configurations in X_k or X. As usual if η is a particle variable taking values in $\{0, 1\}$ we introduce a corresponding spin variable $\sigma = 2\eta - 1$ taking values in $\{-1, +1\}$. This is convenient in particular because it permits us to write a measure μ_k directly in terms of the corresponding spin correlations:

$$\mu_k(\sigma_0, \ldots, \sigma_k) = 2^{-(k+1)}\left(1 + \sum_{A \subset \Lambda_k, A \neq \emptyset} \langle \sigma_A \rangle \, \sigma_A\right), \tag{5}$$

where as usual $\sigma_A = \prod_{i \in A} \sigma_i$ and the spin correlation $\langle \sigma_A \rangle$ denotes the expectation $\mu_k(\sigma_A)$ of σ_A in the measure μ_k. One may think of the expectations $\langle \sigma_A \rangle$ in (5) as parameters which determine the measure; the condition that μ_k be LTI is simply that $\langle \sigma_A \rangle = \langle \sigma_B \rangle$ whenever B is a translate of A. Note, however, that in using the form (5) to construct a measure with certain given $\langle \sigma_A \rangle$ one must check that it assigns nonnegative probability to each configuration.

Remark 3 (a) If Ω is any finite set then everything said above extends immediately to measures defined on Ω^{Λ_k}. As an example, take $\Omega = \{0, 1\}^{\Lambda_m}$ so that Ω^{Λ_k} may be identified with the space $\{0, 1\}^{\Lambda_k \times \Lambda_m}$ of particle configurations on a $k \times m$ rectangle. If such a measure is LTI under translations in the horizontal direction then it may be extended by an obvious generalization of (4) to a measure on configurations on $\Lambda \times \Lambda_m$, where Λ is any interval containing Λ_k, and hence to a measure on configurations on $\mathbb{Z} \times \Lambda_m$ as in the proof of Theorem 1. However, even if the original measure is also LTI in the vertical direction, this extension procedure need not maintain the LTI property for translations in the vertical direction. Example 16 (see Sect. 6) is of this nature: an extension from the original 2×2 square to a 3×2 rectangle destroys the vertical LTI property.

(b) The extension method in the Markovian form (4) is well defined for a large class of spaces Ω; it suffices that the $\mu_k(\eta_k \mid \eta_0, \ldots \eta_{k-1})$ be well defined (regular) conditional probabilities. This will be true, for example, if Ω is a complete separable metric space [2] or, essentially equivalently, a *standard Borel space* [10]. Consider, for example, an LTI measure for a strip in $\mathbb{Z}^2$, that is, an LTI measure on $\{0, 1\}^{\mathbb{Z} \times \Lambda_m} = \left(\{0, 1\}^{\mathbb{Z}}\right)^{\Lambda_m}$; since $\{0, 1\}^{\mathbb{Z}}$ is a standard Borel space we may then use (4) to extend vertically to a measure on configurations on $\{0, 1\}^{\mathbb{Z}^2}$. In this context the original measure might be obtained by beginning with an LTI measure on a rectangle $\{0, 1\}^{\Lambda_k \times \Lambda_m}$ and making the maximal entropy extension in the horizontal direction. As noted above, in general the LTI property in the vertical direction will not be preserved by this extension, but if it is, the above remarks apply.

2.1 Entropy

Let us denote the Gibbs–Shannon entropy of a probability measure μ on a discrete set Q (typically for us a set of configurations) by $S(\mu)$:

$$S(\mu) = -\sum_{q \in Q} \mu(q) \log \mu(q). \tag{6}$$

The entropy of the extension μ_j given in (3) or (4) is easily found to be

$$S(\mu_j) = S(\mu_k) + (j-k)[S(\mu_k) - S(\mu_{k-1})], \quad j \geq k. \tag{7}$$

Now recall the strong subadditivity inequality for entropy [20]: if μ is a probability measure on a product $Q = Q_1 \times Q_2 \times Q_3$ then the marginal measures μ_{12}, μ_{23}, and μ_2 on $Q_1 \times Q_2$, $Q_2 \times Q_3$, and Q_2, respectively, satisfy

$$S(\mu) \leq S(\mu_{12}) + S(\mu_{23}) - S(\mu_2). \tag{8}$$

This implies that if ν_j is any measure on X_j for which the marginal $\nu_j^{(i)}$ on $\{0, 1\}^{\{i,i+1,\ldots,i+k\}}$ satisfies

$$\nu_j^{(i)}(\eta_i, \ldots, \eta_{i+k}) = \mu_k(\eta_i, \ldots, \eta_{i+k}), \quad i = 0, 1, \ldots j - k, \tag{9}$$

then $S(\nu_j)$ is at most equal to the right hand side of (7). Thus μ_j is in fact the maximal entropy extension of μ_k to X_j under the constraint (9); there is a unique such extension by the strict concavity of the entropy. Note that satisfaction of (9) is necessary for ν_j to be LTI but is not sufficient, since it does not impose translation invariance for sets of diameter greater than $k + 1$. Nevertheless, we know from Theorem 1 that the measure μ_j, which maximizes the entropy under only the constraint (9), is LTI, so that the same measure would be obtained by maximizing the entropy over all LTI measures.

2.2 Gibbs Measures

Now suppose that μ_k does not assign zero probability to any configuration. Then it is easy to see that the Markovian extension μ on X provided by the proof of Theorem 1 is an infinite-volume Gibbs state (see [26] for definitions and results) for the translation invariant interaction energy U defined formally by

$$U(\eta) = \sum_{i \in \mathbb{Z}} \Phi(\eta_i, \ldots, \eta_{i+k}), \quad \eta \in X, \tag{10}$$

where the interaction potential Φ is given by

$$\Phi(\eta_0, \ldots, \eta_k) = -\ln \mu_k(\eta_k \mid \eta_0, \ldots, \eta_{k-1}). \tag{11}$$

If μ_k assigns zero measure to some configurations then the same arguments and conclusions apply. We may either simply allow for the interaction Φ to sometimes be $+\infty$ and consider the corresponding Gibbs states, or we may follow [26] and restrict the allowable configurations to those whose restriction to any interval of $k + 1$ sites has nonzero probability under μ_k.

Conversely, suppose that one starts with a TI measure ν on X which is a Gibbs state for some interaction potential with interactions of range at most k; again we may allow the interaction potential to take the value $+\infty$ or alternatively may restrict the allowable configurations by local constraints of range k. The marginal ν_k of ν on X_k will then be LTI. Since ν is Gibbs, it is Markovian of range k [26] and hence agrees with the Markovian extension of ν_k to X provided by the proof of Theorem 1. Note that the original interaction in terms of which ν is defined need not agree with the interaction (11) naturally associated the Markovian extension procedure applied to ν_k; different interactions will give the same Gibbs measure, and thus be essentially equivalent, if they give the same answer, up to a constant, when they are used to compute the energy in a finite volume with boundary condition (the constant may depend on the boundary condition) [11, Sect. 2.4].

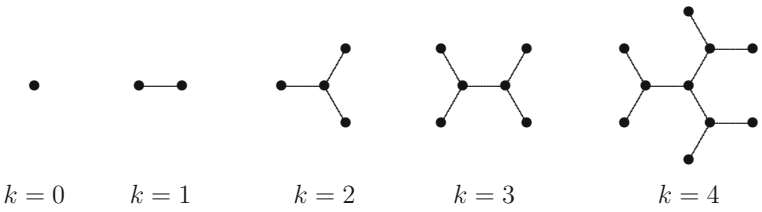


Fig. 1 Full subgraphs of diameter k, $k = 0, \ldots, 4$, when $q = 2$

3 Extension of Measures on the Bethe Lattice

As discussed in Remark 3a above, the Markovian extension technique of Sect. 2 does not in general yield extensions of measures in higher dimensions. It does do so, however, on the Bethe lattice. To discuss this we fix $q \geq 1$ and let $\mathbb{T}$ denote the (infinite) Bethe lattice in which every vertex has degree $q+1$. For a subgraph H of $\mathbb{T}$ we may then consider a measure μ_H on $X_H = \Omega^H$ and ask for an extension to $X = \Omega^{\mathbb{T}}$ which is invariant under the group of all isomorphisms of $\mathbb{T}$. Again there is an obviously necessary *local invariance (LI)* condition for the existence of such an extension: that if H_1 and H_2 are connected isomorphic subgraphs of H, with $\phi : H_1 \to H_2$ an isomorphism, and ν_j is the marginal of μ_H on X_{H_j}, then the induced mapping ϕ_* of measures satisfies $\phi_* \nu_2 = \nu_1$.

A special role in what follows will be played by *full subgraphs* of $\mathbb{T}$: the maximal subgraphs of $\mathbb{T}$ with fixed diameter k. Any two such are isomorphic (see Fig. 1). When k is even, such a full subgraph has $(q+1)$-fold symmetry about a central vertex; when k is odd, it has 2-fold symmetry across a central edge. We fix some vertex v_0 in $\mathbb{T}$ and some edge e_0 incident on it, and for $k = 0, 1, \ldots$ let T_k denote the particular full subgraph which has central vertex v_0 if k is even and central edge e_0 if k is odd, so that in particular $T_0 \subset T_1 \subset \cdots$. We abbreviate X_{T_k} as X_k. If $q = 1$ then the Bethe lattice $\mathbb{T}$ is just $\mathbb{Z}$ and T_k is the interval Λ_k of $k+1$ sites. For general q the graphs T_k play a role for extensions to $\mathbb{T}$ similar to that of the Λ_k for extensions to $\mathbb{Z}$: when one begins with an LI measure on X_k there is no obstruction to extension.

Theorem 4 *Let μ_k be an LI measure on T_k, $k = 0, 1, \ldots$. Then μ_k has an extension to an invariant measure μ on X.*

Note that, because the invariance we consider for $\mathbb{T}$ is invariance under all isomorphisms, the $q = 1$ case of Theorem 4 corresponds to the case in Theorem 1 of symmetric extensions of symmetric measures.

Proof of Theorem 4 We first extend μ_k to a measure μ_{k+1} on T_{k+1}; in doing so we will need only two consequences of the local invariance of μ_k: (i) μ_k is invariant under automorphisms of T_k, and (ii) the marginal of μ_k on X_{k-1} is invariant under automorphisms of T_{k-1}. (In fact, (i) and (ii) imply local invariance, but we do not need this.) T_{k+1} contains r_k copies $T_k^1, \ldots, T_k^{r_k}$ of T_k (one of which is T_k itself), where $r_k = 2$ if k is even and $r_k = q+1$ if k is odd, and these intersect in T_{k-1}. From (i) it follows that we may speak without ambiguity of the copy μ_k^i of μ_k on T_k^i, and from (ii) that the marginal of any μ_k^i on T_{k-1} is independent of i. Then the extension formula is

$$\mu_{k+1}(\eta) = \mu_k(\eta_{T_{k-1}})^{-(r_k-1)} \prod_{i=1}^{r_k} \mu_k^i \left(\eta_{T_k^i}\right) \tag{12}$$

$$= \mu_k(\eta_{T_{k-1}}) \prod_{i=1}^{r_k} \mu_k^i \left(\eta_{T_k^i \setminus T_{k-1}} \mid \eta_{T_{k-1}}\right), \tag{13}$$

where for $\eta \in X_{k+1}$ we have written η_H for the restriction of η to any subgraph H of T_{k+1}. Note that (12) is a direct generalization of (2), while (13) is a symmetrized version of (3).

It is easy to see that μ_{k+1} is indeed an extension of μ_k. Moreover, it again satisfies properties (i) and (ii); (i) follows from the fact that any automorphism of T_{k+1} will permute the T_k^i and preserve T_{k-1}, and because μ_{k+1} extends μ_k, (ii) is an immediate consequence of property (i) for μ_k. Thus we may apply the above construction repeatedly to extend μ_k to any X_j with $j > k$, maintaining properties (i) and (ii) at each step, and thus obtain the extension μ on X as a limit. μ is invariant under the group G of isomorphisms of $\mathbb{T}$, because, as is easily seen, G is generated by the subgroup G_{v_0} of G of isomorphisms that fix the central vertex v_0, together with the subgroup G_{e_0} of isomorphisms that fix the central edge e_0, and invariance of μ under these subgroups follows from properties (i) and (ii) of the finite extensions. □

Remark 5 There is an alternative but equivalent version of the last step of this proof, that is, of the extension beyond X_{k+1}. The procedure is of course again inductive; we consider the extension from X_j to X_{j+1}, $j \geq k+1$. One obtains T_{j+1} from T_j by adding q new edges and vertices to each vertex v in a certain subset V_j of the vertices of T_j of degree one. Each such v, together with the new edges and vertices attached to it, is part of a unique copy $\widehat{T}_k^v$ of T_k in T_{j+1}, and the extension is

$$\mu_{j+1}(\eta) = \mu_j(\eta_{T_j}) \prod_{v \in V_j} \mu_k \left(\eta_{\widehat{T}_k^v \setminus H^v} \mid \eta_{H^v}\right), \tag{14}$$

where $H^v = \widehat{T}_k^v \cap T_j$. Thus the distribution determined by μ_{j+1} on the new vertices attached to v depends only on the configuration in H^v. From this it follows that the distribution on the complement of any T_j which arises from conditioning the extension μ on the configuration in T_j, actually depends only on the configuration on vertices of T_j a distance at most k from the complement. In this sense the extension constructed here is maximally Markovian.

It is straightforward to calculate the entropy of these extensions, through the recurrence

$$S(\mu_{j+1}) = S(\mu_j) + (r_j - 1)[S(\mu_j) - S(\mu_{j-1})], \tag{15}$$

where μ_{k-1} is the marginal of μ_k on X_{k-1}. For $q = 1$ the entropy grows linearly with j; see Sect. 2.1. For $q \geq 2$, $S(\mu_j)$ grows as $q^{j/2}$ for large j; more specifically, with $\Delta S = [S(\mu_k) - S(\mu_{k-1})]$ we find that for $j \geq k$ and $p = \lfloor (j - k + 1 - (k \bmod 2))/2 \rfloor$,

$$S(\mu_j) = S(\mu_k) + \frac{r_{j+1} q^p - r_{k+1}}{q-1} \, q^{k \bmod 2} \, \Delta S. \tag{16}$$

A simple calculation proves the existence of the thermodynamic limit of the entropy per site, and an argument as in Sect. 2.1 implies that (16) is the maximum possible entropy of any extension.

The considerations of Sect. 2.2 apply here essentially unchanged; in particular, the measure on $\mathbb{T}$ constructed here is Gibbsian with interaction potential

$$\Phi(\eta_{T_k}) = -\ln \mu_k \left(\eta_{T_k \setminus T_{k-1}} \mid \eta_{T_{k-1}}\right). \tag{17}$$

4 Periodic Extensions in One Dimension

We now turn again to the study of TI extensions to $X = \Omega^{\mathbb{Z}}$ of an LTI measure μ_k defined on the space X_k of configurations on the interval Λ_k. We consider in particular extensions which are of a different character from the maximal entropy extensions considered in Sect. 2: those supported on periodic configurations in X. We will call such a measure a *periodic configuration* (PC) measure. Within the class of PC measures we single out the *basic periodic configuration* (BPC) measures: a BPC measure is obtained from a particular p-periodic configuration by giving weight $1/p$ to each of its p translates. The set of PC measures is clearly convex and its extreme points are the BPC measures; since the set of BPC measures is countable, every PC measure ν has a representation

$$\nu = \sum c_\alpha \nu_\alpha, \quad c_\alpha \geq 0, \quad \sum c_\alpha = 1, \tag{18}$$

where the sums run over the set $\{\nu_\alpha\}$ of all BPC measures.

In this section and the next we will for simplicity consider only the case of particle configurations, i.e., $\Omega = \{0, 1\}$; the extension of the results to any finite Ω is straightforward. We will write a configuration $\eta \in X_k$ as a string $\eta_0\eta_1 \dots \eta_k$ (rather than as a $(k+1)$-tuple $(\eta_0, \dots, \eta_k)$) and will similarly write $\eta \in X$ as a doubly infinite string; since we do not use any products of the η_i no confusion should arise.

4.1 BPC and Extremal LTI Measures on X_k

Let M_k denote the set of all LTI probability measures on X_k; M_k is manifestly a convex polytope. Since the LTI condition (1) and the normalization $\sum_{\eta \in X_k} \mu_k(\eta) = 1$ represent 2^k independent constraints on the 2^{k+1} variables $\mu_k(\eta)$, $\eta \in X_k$, M_k has dimension 2^k. We wish to characterize the extreme points of this polytope. We will say that a periodic configuration $\eta \in X$ and the corresponding BPC measure ν supported on the translates of η are *k-primitive* if no segment of length k is repeated within one (minimal) period p of η, that is, if $\eta_i\eta_{i+1} \dots \eta_{i+k-1} = \eta_j\eta_{j+1} \dots \eta_{j+k-1}$ only when $j - i$ is a multiple of p. Note that then $p \leq 2^k$. We will prove shortly, using the concept of a *de Bruijn graph*, that the extreme points of M_k are precisely the marginals $\pi_k\nu$ with ν a k-primitive BPC measure.

The (binary) *De Bruijn graph of order* $k \geq 1$ (or sometimes *dimension* k) [9,13], G_k, is the directed graph with 2^k vertices and 2^{k+1} edges, labeled respectively by all binary strings of length k and length $k+1$, in which for any binary digits a and b and binary string θ of length $k-1$ the edge $a\theta b$ runs from vertex $a\theta$ to vertex θb (see Fig. 2). Note in particular that G_k contains two loops, labeled respectively by $00\cdots0$ and $11\cdots1$, but no multiple edges. Since the edges of G_k are labeled by the elements of X_k, it is clear that any probability measure μ_k on X_k corresponds to an assignment of a nonnegative *current* j_η to each edge η of G_k in such a way that $\sum_\eta j_\eta = 1$; the correspondence is of course via $j_\eta = \mu_k(\{\eta\})$. The terminology "current" is appropriate because one checks easily that μ_k is LTI if and only if the current is conserved at each vertex of G_k, that is, if for each vertex ξ the sum of the currents on the two edges entering ξ is equal to the sum of the currents on the two edges leaving ξ. (Such a current assignment is called a *circulation* in the theory of graphs.)

Suppose that $\mathcal{P}$ is a closed path in G_k, that is, a sequence of $|\mathcal{P}|$ edges $\eta^{(1)}, \dots, \eta^{(|\mathcal{P}|)}$ in that order (possibly with repetitions), and that $\mathcal{P}$ is *minimal* in the sense that there is no shorter closed path $\mathcal{P}'$ such that $\mathcal{P}$ is obtained by tracing P' several times. Each $\eta^{(j)}$ is a string of $k+1$ symbols (see Fig. 2), and we may obtain from $\mathcal{P}$ a periodic configuration by tracing $\mathcal{P}$ repeatedly, recording the first symbol on each edge encountered: with $\eta^{(j)} = \eta_0^{(j)} \dots \eta_k^{(j)}$

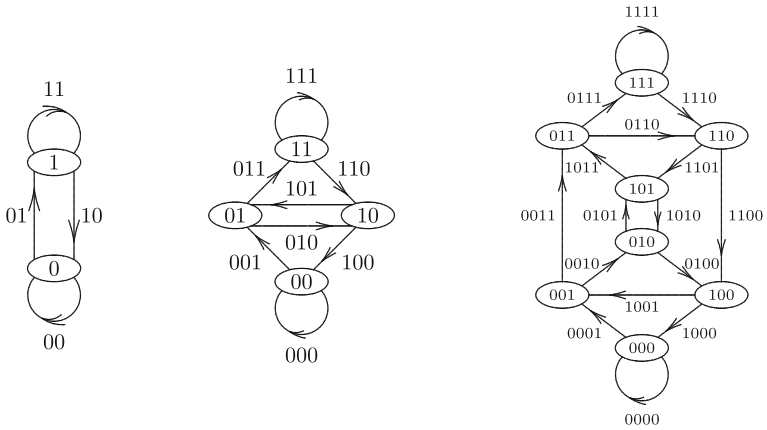


Fig. 2 The De Bruijn graphs G_1, G_2, and G_3

this configuration is $\ldots \eta_0^{(|\mathcal{P}|)} \eta_0^{(1)} \ldots \eta_0^{(|\mathcal{P}|)} \eta_0^{(1)} \ldots \in X$. Let $\nu_{\mathcal{P}}$ denote the BPC measure associated with this configuration; the map $\mathcal{P} \mapsto \nu_{\mathcal{P}}$ establishes a bijective correspondence between minimal closed paths and BPC measures. As usual a *cycle* $\mathcal{C}$ in G_k is a closed path with no repeated vertices; a loop is a cycle with one edge and one vertex.

Theorem 6 *The mapping $\mathcal{C} \to \nu_{\mathcal{C}}$ gives a bijective correspondence between the cycles in G_k and the k-primitive BPC measures on X. The extreme points of M_k are precisely the marginals $\pi_k \nu_{\mathcal{C}}$, with $\mathcal{C}$ a cycle in G_k; equivalently, every LTI measure μ_k on X_k has the form $\pi_k \nu$, where ν is a (finite) convex combination of the k-primitive BPC measures $\nu_{\mathcal{C}}$. In particular, every such μ_k has an extension to a PC measure.*

Proof The condition that a cycle $\mathcal{C}$ contain no repeated vertices is precisely the condition of k-primitivity on $\nu_{\mathcal{C}}$, and the bijective nature of the correspondence $\mathcal{C} \leftrightarrow \nu_{\mathcal{C}}$ is clear. If $|\mathcal{C}| = p$ then the measure $\pi_k \nu_{\mathcal{C}}$ gives weight $1/p$ to each edge of $\mathcal{C}$ and weight 0 to all other edges, and is thus clearly an extreme point of M_k. Conversely, given an element $\mu_k \in M_k$, the set of edges η of G_k for which $\mu_k(\eta) > 0$ must by current conservation contain a cycle $\mathcal{C}$, say of length p, and if α is the smallest current assigned by μ_k to any edge of $\mathcal{C}$ then we may write

$$\mu_k = (p\alpha)\pi_k \nu_{\mathcal{C}} + (1 - p\alpha)\mu_k^*, \quad \text{where} \quad \mu_k^* = \frac{1}{1 - p\alpha}(\mu_k - p\alpha \pi_k \nu_{\mathcal{C}}), \tag{19}$$

thus representing μ_k as a convex combination of the LTI probability measures $\nu_{\mathcal{C}}$ and μ_k^*. We may apply the same construction to μ_k^* and by repeating this process obtain a representation of μ_k as a convex combination

$$\mu_k = \sum_{\mathcal{C}} \beta_{\mathcal{C}} \pi_k \nu_{\mathcal{C}}, \quad \sum_{\mathcal{C}} \beta_{\mathcal{C}} = 1, \tag{20}$$

as required; the process must terminate because μ_k^* as defined by (19) gives zero probability to at least one more configuration than does μ. □

Remark 7 (a) It is known [12,21] that G_k contains cycles with every length p satisfying $1 \le p \le 2^k$. In particular, G_k contains two cycles of length 1, the loops mentioned above, which correspond to the two period-one configurations in X, and $2^{(2^{k-1}-k)}$ cycles of length 2^k [9].

(b) Although it follows from (a) that M_k contains at least $2^{(2^{k-1}-k)}$ extreme points, a result of Carathéodory [5,14] implies that, because M_k is of dimension 2^k, every LTI measure μ_k can be written as a convex combination of at most $2^k + 1$ measures $\pi_k \nu_{\mathcal{C}}$.

4.2 Extensions of LTI Measures to a Ring

Let $\mathbb{Z}_L$ denote the ring (one-dimensional discrete torus) with L sites: $\mathbb{Z}_L = \{0, \ldots, L-1\}$, with translation defined via periodic boundary conditions. Rather than considering extensions of μ_k to TI measures on X one may instead ask whether there exists an integer $L \geq k+1$ such that μ_k has an extension to a TI measure on $Y_L = \{0,1\}^{\mathbb{Z}_L}$, the set of configurations on $\mathbb{Z}_L$, and, if so, for what values of L this is true. The first question is clearly equivalent to that of whether there is an extension to a measure ν on $\mathbb{Z}$ which is a convex combination of a finite number of BPC measures; if so, L may then be taken to be the least common multiple of their periods. Theorem 6 implies that the answer is affirmative; here we discuss a few questions about the values of L.

Theorem 8 *The smallest integer $L \geq k+1$ such that every LTI measure on X_k may be extended to a TI measure on Y_L is*

$$L_0(k) = \operatorname{lcm}\left\{1, 2, 3, \ldots, 2^k - 1, 2^k\right\}. \tag{21}$$

Proof Suppose that $L \geq k+1$. It is easy to see that the measure $\nu_{\mathcal{C}}$, with $\mathcal{C}$ a cycle in G_k, can be extended to a TI measure on Y_L if and only if $|\mathcal{C}|$ divides L, and from the representation (20) it follows that every LTI measure on X_k can be extended if and only if L is divisible by $|\mathcal{C}|$ for every cycle $\mathcal{C}$. The result (21) now follows from Remark 7a. □

Theorem 8 does not address the natural question of the minimal value of L, or of all possible values of L, for the extension of a given measure μ_k. We will not try to give a general discussion of this matter, but rather content ourselves with discussing two examples.

Example 9 Suppose that we begin with a measure μ_1 defined on two sites. The De Bruijn graph G_1 has three cycles $\mathcal{C}_0, \mathcal{C}_1, \mathcal{C}_2$ (see Fig. 2), where $\mathcal{C}_0$ and $\mathcal{C}_1$ denote the loops 00 and 11 respectively, and $\mathcal{C}_2$ the cycle with edges 01 and 10, so that $\mu_1 = \sum_{i=0}^{2} c_i \pi_1 \nu_{\mathcal{C}_i}$ with coefficients in the simplex $\Sigma = \{(c_0, c_1, c_2) \mid c_i \geq 0,\ \sum_{i=0}^{2} c_i = 1\}$. μ_1 has a TI extension to Y_L for any even L, since the measures corresponding to the three extreme points $(1,0,0)$, $(0,1,0)$, and $(0,0,1)$ of Σ have extensions which give respectively probability one to 0^L, probability one to 1^L, and probability one half to each of $(01)^{L/2}$ and $(10)^{L/2}$ (where for any finite binary string η, η^n denotes the concatenation of n copies of η). If L is odd, say $L = 2\ell + 1$, then μ_1 has an extension to Y_L if and only if $c_2 \leq 1 - 1/L$. For the measures corresponding to the four extreme points $(1,0,0)$, $(0,1,0)$, $(1/L, 0, 1-1/L)$, and $(0, 1/L, 1-1/L)$ of $\Sigma \cap \{(c_0, c_1, c_2) \mid c_2 \leq 1 - 1/L\}$ all have extensions to Y_L; the first two are as above and the remaining two give probability $1/L$ to each of the L translates of $0(01)^\ell$ and $1(01)^\ell$, respectively. Conversely, if L is odd and μ_1 has a TI extension to Y_L then the fact that every configuration in Y_L contains either two consecutive 1's or two consecutive 0's implies that $c_0 + c_1 \geq 1/L$.

Example 10 If μ_k is supported on a union of vertex disjoint cycles (and gives positive probability to each) then an analysis as in the proof of Theorem 8 shows that μ_k has an extension to Y_L if and only if L is divisible by the least common multiple of the lengths of these cycles. The maximal length of a cycle is 2^k (see Remark 7a), but this least common multiple may

Table 1 Maximum values of $\mathrm{lcm}\{|\mathcal{C}_1|, \ldots, |\mathcal{C}_n|\}$ for $\mathcal{C}_1, \ldots, \mathcal{C}_n$ vertex disjoint cycles in G_k

k	L_k
1	$2 = 2$
2	$4 = 4$
3	$12 = 3 \times 4$
4	$140 = 4 \times 5 \times 7$
5	$5460 = 3 \times 4 \times 5 \times 7 \times 13$
6	$2{,}042{,}040 = 3 \times 5 \times 7 \times 8 \times 11 \times 13 \times 17$
7	$7{,}216{,}569{,}360 = 5 \times 7 \times 9 \times 11 \times 13 \times 16 \times 17 \times 19 \times 31$

be much larger. Using Maple, we find that for $k = 1, \ldots, 7$ the maximum value L_k of this least common multiple, taken over all collections of vertex disjoint cycles, is as given in Table 1; there the factorization corresponds to the lengths of the maximizing set of cycles. In all of these cases except $k = 3$, $L_k = g(2^k)$, where g is *Landau's function* [19]: $g(n)$ is the maximum, over all partitions $\{n_1, \ldots, n_q\}$ of n, of $\mathrm{lcm}\{n_i, \ldots, n_q\}$. Landau showed that $\log g(n) \sim \sqrt{n \log n}$ for large n. We do not know whether $L_k = g(2^k)$ for some or all $k \geq 8$; the question is whether or not there exists a set of vertex-disjoint cycles in G_k whose lengths are given by the maximizing partition of 2^k.

5 Minimal Entropy Extensions

In this section we consider TI extensions μ of μ_k with finite total entropy, where the *total entropy* of any TI measure μ on X is defined by

$$S(\mu) = \lim_{j \to \infty} S(\pi_j \mu), \tag{22}$$

with the entropy $S(\pi_j \mu)$ given by (6). The TI extensions described in Theorem 6 clearly have finite entropy. We begin with a partial converse.

Lemma 11 *Every TI measure μ on X with finite total entropy is a PC measure.*

Proof Let μ be a TI measure on X. Since any configuration η with $\mu(\eta) > 0$ must be periodic, μ is atomic if and only if it is a PC measure. Suppose then that μ is not atomic, so that $\mu = t\mu_a + (1-t)\mu_b$ with $0 \leq t < 1$, μ_a an atomic probability measure, and μ_b a nonatomic probability measure. Because μ_b is nonatomic there must exist, for any $n \geq 1$, a partition $X = \bigcup_{i=1}^n A_i$ with the A_i measurable sets satisfying $\mu_b(A_i) = 1/n$ [15, p. 174]. Then one can approximate the A_i arbitrarily closely by disjoint cylinder sets A_i', where for some large N, $A_i' = \pi_{[-N,N]}^{-1}(B_i)$ with $B_i \subset X_{[-N,N]}$. In particular, one can make this approximation sufficiently close that

$$\begin{aligned} S(\pi_{[-N,N]}\mu_b) &\geq -\sum_{i=1}^n (\pi_{[-N,N]}\mu_b)(B_i) \log (\pi_{[-N,N]}\mu_b)(B_i) \\ &\geq -\frac{1}{2} \sum_{i=1}^n \mu_b(A_i) \log \mu_b(A_i) \\ &= \frac{1}{2} \log n. \end{aligned} \tag{23}$$

Thus μ_b and hence also μ does not have finite (total) entropy. □

Now let μ_k be an LTI measure on X_k; from Lemma 11 and the discussion in Sect. 4.1 we know that every finite entropy extension ν of μ_k has the form

$$\nu = \sum_{\mathcal{P}} c_{\mathcal{P}} \nu_{\mathcal{P}}, \quad c_{\mathcal{P}} \geq 0, \quad \sum_{\mathcal{P}} c_{\mathcal{P}} = 1, \tag{24}$$

where the sum is over some collection of minimal closed paths in G_k. Note that the BPC measure $\nu_{\mathcal{P}}$ has entropy $S(\nu_{\mathcal{P}}) = \log |\mathcal{P}|$ and that the measure ν of (24) has entropy

$$S(\nu) = \sum c_{\mathcal{P}} (\log |\mathcal{P}| - \log c_{\mathcal{P}}). \tag{25}$$

Theorem 12 *Let μ_k be an LTI measure on X_k. Then μ_k has at least one (TI) extension of minimal entropy of the form*

$$\hat{\mu} = \sum_{i=0}^{2^k} c_i \nu_{\mathcal{C}_i}, \quad c_i \geq 0, \quad \sum_{i=0}^{2^k} c_k = 1, \tag{26}$$

for some cycles $\mathcal{C}_0, \ldots, \mathcal{C}_{2^k}$ in G_k.

Proof We will first show that an entropy-minimizing extension exists, then that there is such a measure with the form (24) having $c_{\mathcal{P}} > 0$ only if $\mathcal{P}$ is a cycle, and finally that there is such a minimizing measure which is a convex combination of at most $2^k + 1$ measures $\nu_{\mathcal{C}}$, with $\mathcal{C}$ a cycle. Let M be the space of all TI probability measures on X which extend μ_k. Then, in the weak topology, M is compact and moreover (22) exhibits S as the increasing limit of functions continuous on M; thus S is lower semi-continuous on M and so achieves its minimum there. Each minimizing measure will have the form (24).

Now we show that there is a minimizing measure in the form (24) for which each closed path appearing with nonzero coefficient is a cycle. To do so we note that among all the minimizing measures we can choose one, say

$$\tilde{\mu} = \sum_{\mathcal{P}} \tilde{c}_{\mathcal{P}} \nu_{\mathcal{P}}, \quad \tilde{c}_{\mathcal{P}} \geq 0, \quad \sum_{\mathcal{P}} \tilde{c}_{\mathcal{P}} = 1, \tag{27}$$

which contains, among the (minimal) closed paths $\mathcal{P}$ with $\tilde{c}_{\mathcal{P}} > 0$, the maximal number of cycles. Then in fact all the $\mathcal{P}$ in (27) for which $\tilde{c}_{\mathcal{P}} > 0$ must be cycles. For if there is some such path $\tilde{P}$ which is not a cycle it may be decomposed into a cycle $\tilde{C}$ and some remaining path $\mathcal{P}'$ edge-disjoint from $\tilde{C}$. If now in (27) we make the replacement

$$\tilde{c}_{\tilde{P}} \nu_{\tilde{P}} \quad \longrightarrow \quad \frac{|\tilde{C}|}{|\tilde{P}|} \tilde{c}_{\tilde{P}} \nu_{\tilde{C}} + \frac{|\mathcal{P}'|}{|\tilde{P}|} \tilde{c}_{\tilde{P}} \nu_{\mathcal{P}'}, \tag{28}$$

then $\pi_k \tilde{\mu}$ is unchanged and a simple computation shows that $S(\tilde{\mu})$ does not increase. By choice of $\tilde{\mu}$, then, $\tilde{C}$ must appear in (27) with positive coefficient. But then the same computation shows that $S(\tilde{\mu})$ must strictly decrease under (28), contradicting its minimality.

To complete the proof we suppose that N is the minimum number of cycles needed to express an entropy-minimizing extension of μ_k in the form $\tilde{\mu} = \sum_{i=0}^{N} a_i \nu_{\mathcal{C}_i}, a_i > 0$, assume that $N > 2^k + 1$, and derive a contradiction. Now certainly $\mu_k = \sum_{i=0}^{N} a_i \pi_k \nu_{\mathcal{C}_i}$, and it then follows from the proof in [14] of the Carathéodory result mentioned in Remark 7(b) that

(possibly after renumbering the cycles) we may also write $\mu_k = \sum_{i=0}^{N-1} a_i^* \pi_k \nu_{\mathcal{C}_i}$ for some nonnegative coefficients a_i^*. Set $a_N^* = 0$ and consider the measure

$$\tilde{\mu}(t) = \sum_{i=0}^{N} a_i(t) \nu_{\mathcal{C}_i}, \quad a_i(t) = (1-t)a_i + t a_i^*, \tag{29}$$

and let $[u, 1]$, with $u < 0$, be the maximal interval such that $a_i(t) \geq 0$ for $i = 0, \ldots, N$ if $t \in [u, 1]$. Note then that $\pi_k \tilde{\mu}(t) = \mu_k$ for $t \in [u, 1]$ and that the entropy $S(\tilde{\mu}(t))$, which is concave in t, achieves its minimum at $t = 0$, and hence is constant. Thus $S(\tilde{\mu}(1)) = S(\tilde{\mu})$, contradicting the minimality of N. □

Remark 13 (a) The argument of the second step in the proof of Theorem 12 may be used to strengthen the conclusion expressed in (24) by showing that in fact every entropy-minimizing extension of μ_k must be a *finite* convex combination of BPC measures.

(b) From (25) and (26) it follows easily that the entropy of a minimal entropy extension of μ_k can be at most $(2^k + 1)[\log 2^k + \log(2^k + 1)]$.

Theorem 12 reduces the search for a minimal entropy extension of a given μ_k to a finite problem: for each set of $2^k + 1$ cycles in G_k one tests whether μ_k is a convex combination of the measures $\pi_k \nu_{\mathcal{C}}$ for cycles $\mathcal{C}$ in that set; if it is, the coefficients are uniquely determined and one may compute the entropy of the corresponding combination of the $\nu_{\mathcal{C}}$. One then chooses the minimum entropy among the latter. However, this is not practical for k very large, and we have no better approach at the moment.

There is, however, one case in which the minimal entropy extension of μ_k may be obtained explicitly. Suppose that μ_k is such that there is a bijective measure preserving mapping $\phi : X_k \to X_k$ which for each vertex ξ of G_k, $\xi \in X_{k-1}$, maps the set of edges entering ξ to the set of edges leaving ξ, that is, $\phi(\{0\xi, 1\xi\}) = \{\xi 0, \xi 1\}$. Then ϕ determines a set of pairwise edge-disjoint paths $\mathcal{P}_1, \ldots, \mathcal{P}_n$ by the condition that if $\eta^{(1)}, \ldots, \eta^{(|\mathcal{P}_i|)}$ are the edges of $\mathcal{P}_i$ in order then $\eta^{(j+1)} = \phi(\eta^{(j)})$, $j = 1, \ldots, |\mathcal{P}_i| - 1$, and $\phi(\eta^{(|\mathcal{P}_i|)}) = \eta^{(1)}$. Now because ϕ is measure preserving, μ_k gives equal weight to the edges of each $\mathcal{P}_i$, and thus the measure $\sum_{i=1}^{n} w_i |\mathcal{P}_i| \nu_{\mathcal{P}_i}$, where $w_i = \mu_k(\eta)$ for any edge η of $\mathcal{P}_i$, is an extension of μ_k with the same entropy as μ_k and is thus the minimal entropy extension. This is precisely the case in which μ_k has a TI extension that is $(k+1)$-deterministic, in the sense that is is supported on configurations that are determined by any segment of length $k + 1$.

A special case of the above occurs when μ_k is supported on a set of pairwise vertex-disjoint cycles $\mathcal{C}_1, \ldots, \mathcal{C}_n$; this means that for each vertex of G_k at most one incoming and one outgoing edge is assigned positive probability by μ_k. Then μ_k has the *unique* TI extension $\sum_{i=1}^{n} w_i |\mathcal{C}_i| \nu_{\mathcal{C}_i}$, with w_i as above. In particular, the maximal entropy extension of Sect. 2 and the minimal entropy extension of Theorem 12 are the same. The total entropy of the extension is the entropy of μ_k, so that one sees from (7) (or directly) that this is also the entropy of the restriction μ_{k-1} of μ_k to X_{k-1}. The TI extension in this case is k-deterministic.

Remark 14 One might try to obtain a minimum entropy extension of μ_k by a step-by-step procedure, extending first to a measure μ_{k+1} on X_{k+1}, then to μ_{k+2} on X_{k+2}, etc., choosing at each step the minimum entropy extension for that step. In fact, there is a simple algorithm for each single step, but the procedure may not produce the extension with the minimal total entropy, even for the case $k = 2$. We omit details.

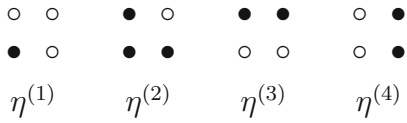
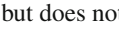
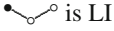


Fig. 3 Configurations given nonzero probability in Example 16. Filled circles represent occupied sites

6 Nonextendible LTI Measures

In this section we give examples to show that the local invariance condition (LTI or, for the Bethe lattice, LI) is in general not sufficient for extension to an invariant measure, unless Λ is an interval in $\mathbb{Z}$, a strip in $\mathbb{Z}^2$, a higher-dimensional analogue of the latter, or a full subgraph of the Bethe lattice.

Example 15 (a) Let $\Lambda = \{0, 1, 3\} \subset \mathbb{Z}$ and $\Omega = \{\circ, \bullet\}$, and suppose that μ_Λ assigns probability $1/2$ to each of the two configurations $(\bullet, \bullet, \cdot, \circ)$ and $(\circ, \circ, \cdot, \bullet)$. Then μ_Λ is LTI but does not even have an LTI extension to the interval $\{0, 1, 2, 3\}$, and hence not a TI extension to $\mathbb{Z}$.

(b) We may give a similar example for the Bethe lattice, again with $\Omega = \{\circ, \bullet\}$. In the notation of Sect. 3 we take $q = 2$ and let H be one of the connected subgraphs of T_2 with three vertices. Then the measure which assigns probability $1/4$ to each of the configurations  , and  is LI but does not extend to an LI measure on X_2.

Note that when $\Lambda \subset \mathbb{Z}$ and Λ_k is the minimal interval containing Λ, the extendibility of an LTI measure from X_Λ to X_k, which is necessary and, by Theorem 1, sufficient for extendibility to X, may be determined by solving a standard linear programming feasibility problem. If H is a finite subgraph of $\mathbb{T}$ then a similar remark applies to the extension of an LI measure from X_H to X_T, where T is the minimal full subgraph of $\mathbb{T}$ containing H. Thus the problem of extension in one dimension, or on the Bethe lattice, is decidable; this is in contrast to the situation in higher dimension, as discussed in Sect. 7.

Example 16 Let Λ be a 2×2 square in $\mathbb{Z}^2$, let $\Omega = \{\circ, \bullet\}$, and let μ be the measure which assigns probability $1/4$ to each of the configurations $\eta^{(i)}$, $i = 1, \ldots, 4$, shown in Fig. 3. μ is LTI, since the marginal on the top, bottom, left, or right two sites of Λ is Bernoulli with parameter $1/2$, but μ cannot be extended to a measure μ' on configurations on a 3×3 square Λ' containing Λ. For suppose that μ' is such an extension; μ' can give positive probability to a configuration η' only if the restriction of η' to each of the four 2×2 subsquares of Λ' is one of the $\eta^{(i)}$. Moreover, there must be an η' with $\mu'(\eta') > 0$ having restriction $\eta^{(1)}$ to the lower left subsquare, and η' must then have restrictions $\eta^{(3)}$ and $\eta^{(4)}$ to the upper left and lower right subsquares, respectively. But then the restriction of η' to upper right subsquare will not be one of the $\eta^{(i)}$, whether or not the upper right hand site of Λ' is occupied in η'.

The measure μ in Example 16 is not invariant under any of the symmetry operations of the square; one could hope that the imposition of some symmetry might give extensibility. The next example shows that this is not the case.

Example 17 Let Λ be a 3×3 square in $\mathbb{Z}^2$ and let μ be the measure which assigns probability $1/7$ to each of the configurations $\eta^{(i)}$, $i = 1, \ldots, 7$, shown in Fig. 4. μ is easily seen to be LTI and is clearly invariant under any symmetry of the square. However, μ cannot be extended to any 4×4 square Λ' containing Λ. The proof is similar to the proof of nonextendibility given in Example 16: one checks, for example, that there is no configuration η' in Λ' which reduces to $\eta^{(2)}$ in the lower left hand 3×3 subsquare and to one of the $\eta^{(i)}$ in all the other 3×3 subsquares.

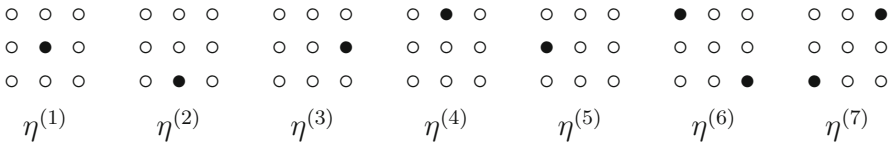


Fig. 4 Configurations given nonzero probability in Example 17

The measures in the above examples appear rather special, in that they assign zero probability to most configurations. Recall from Sect. 1, however, that the set of nonextendible measures (for fixed Λ) is open; thus small perturbations of these measures will also be nonextendible. Such perturbations may be taken to assign nonzero probability to every configuration in Λ, to be be again LTI, and, for Example 17, to have all the symmetries of the square.

7 Undecidability

As discussed above, the situation in $\mathbb{Z}^d$ for $d \geq 2$ is quite different from that for $d = 1$ or for the Bethe lattice. In particular, there is no formula similar to (4) or (13) for an extension, and the LTI condition is not sufficient for extendibility. More surprisingly, the extendibility problem for $\mathbb{Z}^d$ with $d \geq 2$ is in several senses undecidable. To discuss this we introduce the idea of *tiling*. It suffices to fix Λ to be a 2×2 box in $\mathbb{Z}^2$, and we will do so throughout this section. Let Y be a finite subset of $X_\Lambda = \Omega^\Lambda$ constructed with some Ω; we call the elements of Y *tiles* and say that Y *tiles the plane* if there is a configuration in X whose restriction to every 2×2 box is an element of Y. Here we do not consider Ω as fixed, so there is an infinite number of possible sets Y of tiles; given Y we will consider only tilings by Y, so that for all practical purposes Y determines Ω, and we will omit mention of the latter in what follows. Berger [3] proved:

Theorem 18 *The tiling problem is undecidable: there is no decision procedure which, given a set of tiles Y, determines whether or not Y tiles the plane.*

The statement of the theorem means that there is no algorithm (or Turing machine) to which we can submit a tile set Y and which will then yield the result "tiles" when Y in fact tiles and "does not tile" when it does not. Note that it is crucial here that the set of possible tile sets Y be infinite, since otherwise there would exist such a decision procedure, though we might not know what it is: if there were N tile sets $Y_1, \ldots, Y_N$ then it would be one of the 2^N procedures which answer the N questions "Does Y_k tile?" in all 2^N possible ways.

We should remark that Berger worked with a different sort of tile, originally introduced by Wang [32], but it is easy to see [30] that his result implies the above theorem. A simplified version of the proof of Berger's theorem is due to Robinson [25].

There is an alternative formulation of undecidability which follows from Berger's result.

Corollary 19 *There exists a set of tiles Y_* such that Y_* tiles the plane but cannot be proved to do so.*

Proof It is known [25] that if Y does not tile the plane then there is some finite rectangle which it does not tile, so that the fact that Y does not tile is provable. Thus if the conclusion of the theorem were false an algorithm generating all provable theorems would provide a decision procedure for the tiling problem. □

It should be noted that Corollary 19 is not really stated precisely; whether or not Y can be proved to tile depends on the axiom system on which the proof is to be based. In fact, however, a tile set Y_* satisfying the theorem will always exist, but will in general depend on the (consistent and sound) axiom system chosen.

Schlijper [30] observed a connection between the tiling problem and the extendibility of measures:

Theorem 20 *(Schlijper) A set of tiles Y tiles the plane if and only if there exists an extendible measure ν_Y on X_Λ with support Y.*

Then an undecidability result follows:

Corollary 21 *There is no decision procedure which, given a set of tiles Y, determines whether or not there exists an extendible measure ν_Λ on X_Λ with support Y. Equivalently, there is no decision procedure which, given an LTI measure μ_Λ on X_Λ, determines whether or not there exists an extendible measure ν_Λ on X_Λ with the same support.*

This result certainly suggests that there should be no decision procedure for our original problem of determining whether or not a given LTI measure μ_Λ is extendible, that is, that this problem is undecidable. However, since most of the measures μ_Λ are presumably not expressible—not computable—most cannot in fact be "given" to the decision procedure to begin with.

Nevertheless, Corollary 21 does lead to a conclusion which furnishes a strong sense of undecidability for the original problem. Suppose there were a condition C for extendibility which was "useful", that is, simpler and more transparent than extendibility itself—perhaps something in the nature of the LTI condition (although of course we know that this is not a such a condition for $d \geq 2$). No matter how simple C were, however, there could exist no decision procedure which, given a set Y of tiles, would determine whether or not there exists a measure ν_Λ on X_Λ with support Y which satisfies C.

Another formulation of undecidability for the extension problem, alternative to Corollary 21 in the sense that Corollary 19 is alternative to Theorem 18, follows immediately from Corollary 19 and Theorem 20:

Corollary 22 *There exists a set of tiles Y_* such that (1) there is an extendible measure ν_Λ on X_Λ with support Y_*, but (2) this fact cannot be proven.*

Here, as for Corollary 19, Y_* will in general depend on the axiom system to be used in the proof. It of course follows from Corollary 22 that if we specify a measure on X_Λ with support Y_* to arbitrary but not perfect precision—that is, specify that it lie in some arbitrarily small set of measures—then for some such specification there exists an extendible measure satisfying it whose existence can't be proven. This suggests that there exists an extendible measure ν_Λ on X_Λ for which the proposition that it is extendible can't be proven. However this is not so clear, since the relevant measure may not be suitably expressible and thus the desired proposition may not exist.

As with Corollary 21, however, Corollary 22 does have a consequence which strongly suggests undecidability for the original problem: if there were a condition C for extendibility, "useful" in the sense discussed above, then the statement that there exists a measure with support Y_* which satisfies C would be true but unprovable.

Acknowledgements The work J.L.L. was supported in part by NSF Grant DMR 1104500 and AFOSR Grant FA9550-16-1-0037. We thank A. C. D. van Enter and M. Hochman for bringing to our attention previous work on this problem, M. Hochman, M. Saks, S. Thomas, and A. C. D. van Enter for helpful discussions, and D. Avis for making the computer program *lrs* available to the public and for helpful advice on its use.

References

1. Ambartzumian, R.V., Sukiasian, H.S.: Inclusion-exclusion and point processes. Acta Appl. Math. **22**, 15–31 (1991)
2. Ash, R.B.: Real Analysis and Probability. Academic Press, New York (1972)
3. Berger, R.: The Undecidability of the Domino Problem. Memoirs of the American Mathematical Society, vol. 66. The American Mathematical Society, Providence (1966)
4. Caglioti, E., Kuna, T., Lebowitz, J.L., Speer, E.: Point processes with specified low order correlations. Markov Process. Relat. Fields **12**, 257–272 (2006)
5. Carathéodory, C.: Über den Variabilitätsbereich der Fourierschen Konstanten von positiven harmonischen Funktionen. Rend. Circ. Mat. Palermo **32**, 193–217 (1911)
6. Chazottes, J.-R., Gambaudo, J.-M., Hochman, M., Ugalde, E.: On the finite-dimensional marginals of shift-invariant measures. Ergod. Theory Dyn. Syst. **32**, 1485–1500 (2012)
7. Costin, O., Kuna, T., Lebowitz, J.L., Speer, E.: On the realizability of point processes with specified one and two particle densities. In: Landim, C., Olla, S., Spohn, H. (eds.) Large Scale Stochastic Dynamics, vol. 43. Mathematische Forschungsinstitut Oberwolfach, Oberwolfach (2004)
8. Crawford, J., Torquato, S., Stillinger, F.H.: Aspects of correlation function realizability. J. Chem. Phys. **119**, 7065–7074 (2003)
9. de Bruijn, N.G.: A combinatorial problem. K. Ned. Akad. van Wet. **A49**, 758–764 (1946)
10. Dudley, R.M.: Real Analysis and Probability. Cambridge University Press, New York (2002)
11. Georgii, H.-O.: Gibbs Measures and Phase Transitions, 2nd edn. DeGruyter, Berlin (2011)
12. Golomb, S.W.: Shift Register Sequences, Revised edn. Aegean Park Press, Laguna Hills (1982)
13. Good, I.J.: Normal recurring decimals. J. Lond. Math. Soc. **21**, 169–172 (1946)
14. Grünbaum, V.: Convex Polytopes, 2nd edn. Springer, New York (2003)
15. Halmos, P.R.: Measure Theory. D. van Nostrand, Princeton (1950)
16. Kuna, T., Lebowitz, J.L., Speer, E.: Realizability of point processes. J. Stat. Phys. **129**, 417–440 (2007)
17. Kuna, T., Lebowitz, J.L., Speer, E.: Necessary and sufficient conditions for realizability of point processes. Ann. Appl. Probab. **21**, 1253–1281 (2011)
18. Lachieze-Rey, R., Molchanov, I.: Regularity conditions in the realisability problem with applications to point processes and random closed sets. Ann. Appl. Probab. **25**, 116–149 (2015)
19. Landau, E.: Über die Maximalordnung der Permutationen gegebenen Grades. Arch. Math. Phys. Ser. 3 **5**, 92–103 (1903)
20. Lieb, E.: Some convexity and subadditivity properties of entropy. Bull. Am. Math. Soc. **81**, 1–13 (1975)
21. Maurer, U.M.: Asymptotically-tight bounds on the number of cycles in generalized de Bruijn graphs. Discret. Appl. Math. **37/38**, 421–436 (1992)
22. Pelizzola, A.: Cluster variation method in statistical physics and probabilistic graphical models. J. Phys. A **38**, R309–R339 (2005)
23. Percus, J.K.: The pair distribution function in classical statistical mechanics. In: Frisch, H.L., Lebowitz, J.L. (eds.) The Equilibrium Theory of Classical Fluids. Benjamin, New York (1964)
24. Pivato, M.: Building a stationary stochastic process from a finite-dimensional marginal. Can. J. Math. **53**, 382–413 (2001)
25. Robinson, R.M.: Undecidability and nonperiodicity for tilings of the plane. Invent. Math. **12**, 177–209 (1971)
26. Ruelle, D.: Thermodynamic Formalism: The Mathematical Structures of Classical Equilibrium Statistical Mechanics. Addison Wesley, Reading (1978) (second edition Cambridge University Press, 2004)
27. Schlijper, A.G.: Convergence of the cluster-variation method in the thermodynamic limit. Phys. Rev. B **27**, 6841–6848 (1983)
28. Schlijper, A.G.: Exact variational methods and cluster-variation approximations. J. Stat. Phys. **35**, 285–301 (1984)
29. Schlijper, A.G.: On some variational approximations in two-dimensional classical lattice systems. J. Stat. Phys. **40**, 1–27 (1985)
30. Schlijper, A.G.: Tiling problems and undecidability in the cluster-variation method. J. Stat. Phys. **50**, 689–714 (1988)
31. Stillinger, F.H., Torquato, S.: Pair correlation function realizability: lattice model implications. J. Phys. Chem. B **108**, 19589 (2004)
32. Wang, H.: Proving theorems by pattern recognition-II. Bell Syst. Tech. J. **40**, 1–41 (1961)