

FREEZING PHASE TRANSITIONS FOR LATTICE SYSTEMS AND HIGHER-DIMENSIONAL SUBSHIFTS

J.-R. CHAZOTTES, T. KUCHERENKO, A. QUAS

ABSTRACT. Let $X = \mathcal{A}^{\mathbb{Z}^d}$, where $d \geq 1$ and \mathcal{A} is a finite set, equipped with the action of the shift map. For a given continuous potential $\varphi : \mathcal{A}^{\mathbb{Z}^d} \rightarrow \mathbb{R}$ and $\beta > 0$ (“inverse temperature”), there exists a (nonempty) set of equilibrium states $\text{ES}(\beta\varphi)$. The potential φ is said to exhibit a “freezing phase transition” if $\text{ES}(\beta\varphi) = \text{ES}(\beta'\varphi)$ for all $\beta, \beta' > \beta_c$, while $\text{ES}(\beta\varphi) \neq \text{ES}(\beta'\varphi)$ for any $\beta < \beta_c < \beta'$, where $\beta_c \in (0, \infty)$ is a critical inverse temperature depending on φ .

In this paper, given any proper subshift X_0 of X , we explicitly construct a continuous potential $\varphi : X \rightarrow \mathbb{R}$ for which there exists $\beta_c \in (0, \infty)$ such that $\text{ES}(\beta\varphi)$ coincides with the set of measures of maximal entropy on X_0 for all $\beta > \beta_c$, whereas for all $\beta < \beta_c$, $\mu(X_0) = 0$ for all $\mu \in \text{ES}(\beta\varphi)$. This phenomenon was previously studied only for $d = 1$ in the context of dynamical systems and for restricted classes of subshifts, with significant motivation stemming from quasicrystal models. Additionally, we prove that under a natural summability condition – satisfied, for instance, by finite-range potentials or exponentially decaying potentials – freezing phase transitions are impossible.

Keywords and phrases: equilibrium states, Gibbs states, measures of maximal entropy, maximizing measures, ground states, phase transitions, multidimensional symbolic dynamics, substitution subshifts, subshifts of finite type, aperiodic tilings, models of quasicrystals.

CONTENTS

1. Introduction	2
1.1. Equilibrium states, temperature going to 0 and maximizing measures	3
1.2. Subshifts and quasicrystals	5
1.3. Freezing at non-zero temperature.	5
1.4. Statistical physics	6
2. Main results	7
2.1. Freezing on any proper subshift and a no-go result	7

Date: April 30, 2025.

The authors warmly thank Aernout van Enter for his thoughtful email exchanges. We did our best to draw a connection with statistical mechanics, and our conversations with him played a meaningful role in guiding that process. AQ and TK acknowledge the financial support from CNRS and the hospitality of the Centre de Physique Théorique (CPHT) at École Polytechnique, where part of this work was conducted.

2.2. Some remarks about Theorem A	9
2.3. Intermezzo on statistical physics	9
2.4. Previously known results	10
2.5. Some illustrative examples	12
3. Freezing on any proper subshift	13
3.1. Notation and definitions	13
3.2. Dyadic tiling lemma	14
3.3. Potentials that freeze on a given subshift	16
4. Obstruction to freezing phase transitions	22
4.1. Proof of Theorem B	22
4.2. Proof of Theorem B'	28
5. The one-dimensional case	29
6. Proofs of some Folklore results	35
6.1. Properties of accumulation points of $(\mu_{\beta_n\varphi})_n$ as $\beta_n \rightarrow \infty$	35
6.2. When the pressure function is affine on a segment	37
6.3. Proof of Proposition 2.3	38
6.4. Gibbs states have full support	38
References	39

1. INTRODUCTION

The connections between statistical physics and dynamical systems have long been fruitful, with ideas initially flowing from physics into the study of dynamical systems. The foundational developments in ergodic theory and chaotic dynamics were deeply influenced by statistical mechanics, particularly through the notion of Gibbs states. These measures, introduced by the pioneering work of Boltzmann and Gibbs, became a crucial tool in understanding long-term statistical behavior in dynamical systems, notably in the study of hyperbolic systems and Anosov flows. The development of thermodynamic formalism, initiated by Sinai, Ruelle, and Bowen, brought a statistical physics perspective to dynamical systems, using concepts such as entropy, pressure, Gibbs and equilibrium states to describe the statistical properties of chaotic systems [12, 14].

However, in recent years, this flow of ideas has started to reverse in certain contexts. Techniques originating in ergodic theory and symbolic dynamics – historically developed for understanding dynamical systems – are now being applied back to lattice models in statistical physics. Lattice models have played a pivotal role in the study of phase transitions, providing a mathematical framework to understand how systems change between distinct states of matter (e.g., from a disordered phase to an ordered phase). These models are defined on a discrete lattice (e.g., \mathbb{Z}^d) and describe interactions between variables (e.g., spins, particles, or other degrees of freedom) located at the lattice sites; see for instance [23, 27].

Recent studies on low-temperature limits and ground states of lattice models have increasingly leveraged techniques from symbolic dynamics [3, 10, 11, 13, 15, 26]. In this paper, we focus on freezing phase transitions, where the set of equilibrium states can collapse at low temperatures – not just at zero temperature. While our results on freezing phase transitions are abstract, they were motivated by understanding how equilibrium states collapse in a way that they become concentrated on subshifts without any periodic configuration, which serve as models for quasicrystals in the context of lattice systems.

Despite the close connections between statistical physics and dynamical systems through thermodynamic formalism [46], these two fields remain largely distinct communities. This separation arises not only from differences in the types of questions they address but also from fundamental differences in their core objects and frameworks. In statistical physics, the primary object is an interaction, sometimes called an “interaction potential”, which is a family of functions indexed by finite subsets of \mathbb{Z}^d . This object describes how components of a system interact locally and provides a natural, flexible foundation for building models from the ground up. In contrast, in dynamical systems, the central object is a function on the phase space – often called a potential as well.

Schematically (we will revisit this point in more detail later), interactions are used to define Gibbs states as well as equilibrium states¹. The construction of these states relies on an averaged version of the interactions over finite subsets of \mathbb{Z}^d containing the origin. These averaged functions correspond to the potentials used in dynamical systems, where the natural objects are equilibrium states rather than Gibbs states (notably because the former can be defined in a very general setting). The complication arises from the fact that the mapping that transforms an interaction into a potential lacks good properties. As a result, it is generally not possible to automatically translate results obtained in the context of dynamical systems to apply them directly to interactions. This effect will show up in this article: although our results are stated in the spirit of dynamical systems, we will show how to adapt and translate them into the framework of statistical physics.

1.1. Equilibrium states, temperature going to 0 and maximizing measures. Consider the lattice \mathbb{Z}^d and a space of configurations X of the form $\mathcal{A}^{\mathbb{Z}^d}$ where \mathcal{A} is a finite set (often called an alphabet in the context of symbolic dynamics). On X we consider the action of the shift map. A potential φ is a real-valued function on X . Given $\beta \in \mathbb{R}$, the equilibrium states for $\beta\varphi$ are the shift-invariant probability measures which maximize

¹ In the context of statistical physics, one uses the term “state” to mean a probability measure, especially for Gibbs and equilibrium measures which describe the states of the system.

the functional

$$\nu \mapsto h(\nu) + \beta \int \varphi d\nu,$$

where $h(\nu)$ is the entropy of ν . This is the variational principle. We also define the so-called pressure function

$$(1) \quad \beta \mapsto p_\varphi(\beta) := \sup_\nu \left\{ h(\nu) + \beta \int \varphi d\nu \right\},$$

where the supremum is taken over shift-invariant (probability)² measures. The pressure function is convex and in particular continuous. (For this material, we refer to [33].)

The parameter β is interpreted as the inverse temperature in the context of statistical physics so, physically, $\beta \geq 0$. For $\beta = 0$ (infinite temperature), there is only one equilibrium state which is the measure of maximal entropy on $\mathcal{A}^{\mathbb{Z}^d}$, that is, the product measure whose entropy is equal $\log |\mathcal{A}|$. This is the most disorderly phase possible, where φ plays no role.

The other extreme is the limit $\beta \rightarrow +\infty$ (the zero-temperature limit) which is associated with ground states and is considerably more subtle. It has only begun to be the subject of systematic research in the last fifteen years, and revealed a number of surprises, see [3, 10, 11, 13, 15, 26] and references therein. A maximizing measure for φ is a shift-invariant measure μ such that $\int \varphi d\mu \geq \int \varphi d\nu$ for all shift-invariant measures ν . The term “maximizing measure” originates from the field of ergodic optimization. In statistical physics, one would instead use the term “ground state”. However, due to the lack of consensus on the precise definition of ground states, we adopt the terminology from dynamical systems here. For a comprehensive survey on ergodic optimization in the context of a continuous map acting on a compact metric space – corresponding to a \mathbb{Z} -action when the map is invertible – we refer the reader to [32]. While some general results extend naturally to \mathbb{Z}^d -actions, others do not. If φ is not cohomologous to a constant and satisfies a natural uniform summability condition, the support of any of its maximizing measures is contained in a proper subshift of $\mathcal{A}^{\mathbb{Z}^d}$, that is, a closed and shift-invariant subset of $\mathcal{A}^{\mathbb{Z}^d}$. This subshift can be characterized in terms of maximizing configurations which are configurations whose energy cannot be increased by finite modifications [25], so we call it the maximizing subshift for φ .

Now, if we take an arbitrary sequence $(\beta_n)_{n \geq 1}$ of inverse temperatures such that $\beta_n \rightarrow +\infty$, and consider a sequence $(\mu_n)_{n \geq 1}$ where, for each n , μ_n is an equilibrium state for $\beta_n \varphi$, then any accumulation point of $(\mu_n)_{n \geq 1}$ is a maximizing measure and it maximizes entropy among the possible maximizing measures. There are two special cases where the zero-temperature limit does exist, namely if the maximizing subshift is uniquely ergodic (i.e., carries a single shift-invariant measure), or if it has a unique measure of maximal entropy. The existence of the zero-temperature limit turns out to

² Unless otherwise specified, all measures in this paper are probability measures.

depend both on the regularity of φ and d (the dimension of the lattice). For $d = 1$, the limit always exists for locally constant potentials (corresponding to finite-range interactions), but may fail to exist for Lipschitz potentials (i.e., exponentially decaying pair interactions). For $d \geq 2$, the limit generally fails to exist even for locally constant potentials.

The pressure function, in any dimension and for any continuous φ , exhibits a slant asymptote, that is,

$$(2) \quad \lim_{\beta \rightarrow +\infty} (p_\varphi(\beta) - (s_\varphi \beta + h_\varphi)) = 0,$$

where

$$(3) \quad s_\varphi = \sup_\nu \int \varphi d\nu \quad \text{and} \quad h_\varphi = \sup \left\{ h(\eta) : \int \varphi d\eta = s_\varphi \right\},$$

and the supremum defining s_φ is taken over shift-invariant measures. (For completeness, we provide a proof of this fact in Section 6.)

1.2. Subshifts and quasicrystals. We can approach the problem in reverse: starting with a subshift that models either a crystal or a quasicrystal, we can investigate how to construct an interaction for which this subshift will serve as the maximizing one. Here a crystal is a subshift consisting of finitely many configurations, each being a shift of another. A quasicrystal, by contrast, is a uniquely ergodic, zero-entropy subshift with infinitely many configurations, and thus no periodic configurations. We emphasize that there is no universally accepted definition of a quasicrystal, even within the framework of lattice models. One could for instance refine the previous definition by additionally requiring the subshift to be minimal in the ergodic-theoretic sense, which informally means that all configurations share the same patterns. On the other hand, one could simply require a subshift that contains uncountably many configurations, with none of them being periodic.

There is a striking difference between $d = 1$ and $d \geq 2$. To model a quasicrystal for $d \geq 2$, one can consider subshifts of finite type without periodic configurations, which correspond to aperiodic Wang tilings (we will mention examples and references below). For $d = 1$, however, subshifts of finite type are either reduced to finitely many periodic configurations (crystal), or are of positive topological entropy which is interpreted as chaos/disorder. Consequently, in dimension 1, natural zero-entropy uniquely ergodic subshifts such as the Thue-Morse substitution have been commonly used, and infinite-range interactions are required [24]. For further discussions and background, we refer the reader to [17, 20].

1.3. Freezing at non-zero temperature. For $0 < \beta < \infty$, the variational principle described above indicates that equilibrium states exhibit a balance between “order” and “disorder”, with the potential φ acting as the driving force behind ordering. However, phase transitions can still occur. From this perspective, we can reconsider the previous discussion by examining the behavior at nonzero temperature and addressing the following questions:

Given a subshift modeling, for example, a quasicrystal, can we construct a function φ whose equilibrium states coincide, for all $\beta > \beta_c$, with the unique invariant measure supported by this subshift? Is it possible that for $\beta < \beta_c$, none of the equilibrium states for φ are supported on the given subshift?

For a more general subshift, one might wish for the equilibrium states to correspond, for all $\beta > \beta_c$, to the measures of maximal entropy associated with this subshift.

If this scenario occurs, the phase transition corresponds to an abrupt contraction of the support of the equilibrium states, which become concentrated on a significantly smaller set of maximizing configurations for the given potential. This type of phase transition differs from the classical phase transitions typically encountered in statistical physics [23, 27].

In this article, given any proper subshift X_0 of $X = \mathcal{A}^{\mathbb{Z}^d}$, we are able to construct a continuous potential, with controlled modulus of continuity, that “freezes” on X_0 . Specifically, for $\beta > \beta_c$, the equilibrium states coincide with the measures of maximal entropy supported on X_0 , whereas for $\beta < \beta_c$, none of the equilibrium states for the given potential are supported on X_0 . Additionally, we establish a no-go theorem, demonstrating that potentials satisfying a classical summability condition cannot exhibit such phase transitions.

1.4. Statistical physics. Recall that $X = \mathcal{A}^{\mathbb{Z}^d}$. The primary objects here are interactions. An interaction Φ is a family of functions $(\Phi_\Lambda)_{\Lambda \in \mathbb{Z}^d}$, where for each (finite) subset Λ of \mathbb{Z}^d , $\Phi_\Lambda : X \rightarrow \mathbb{R}$ depends only on the configurations restricted to Λ . Interactions are the building blocks of Gibbs states that we will not describe here (we refer to [23, 27, 46]). For a shift-invariant interaction, one can construct a function $\varphi : \mathcal{A}^{\mathbb{Z}^d} \rightarrow \mathbb{R}$ given by

$$\varphi = \sum_{\substack{\Lambda \in \mathbb{Z}^d \\ 0 \in \Lambda}} \frac{\Phi_\Lambda}{|\Lambda|},$$

where $A \subseteq B$ means A is a finite subset of B . This function appears in the definition of equilibrium states (variational principle) [27, 46]. It is commonly referred to as a potential in dynamical systems theory, and is continuous if $\sum_{\Lambda \in \mathbb{Z}^d, 0 \in \Lambda} |\Lambda|^{-1} \|\Phi_\Lambda\|_\infty < +\infty$.

A fundamental question is to understand the structure of the set of Gibbs and equilibrium states for the interaction $\beta\Phi$ where $\beta \geq 0$. Let us emphasize that equilibrium states coincide with shift-invariant Gibbs states for sufficiently regular interactions, the natural condition being absolute summability, that is, $\sum_{\Lambda \in \mathbb{Z}^d, 0 \in \Lambda} \|\Phi_\Lambda\|_\infty < +\infty$; see [27, Theorem 15.30, p. 322] or [46, Theorem 4.2, p. 58]. In the absence of uniqueness, there exist examples of non-shift-invariant Gibbs states, which therefore cannot be equilibrium states, as well as equilibrium states that are not Gibbs states. However,

when the Gibbs state is unique, it must be shift-invariant and coincides with the unique equilibrium state.

Notice that sign conventions used in dynamical systems are different from those used in statistical physics, basically because physicists prefer to minimize energy.

2. MAIN RESULTS

2.1. Freezing on any proper subshift and a no-go result. We first define what we precisely mean by a *freezing phase transition*. (We will mention several concepts that will be precisely defined in the next section.) Let $X = \mathcal{A}^{\mathbb{Z}^d}$ be a full d -dimensional shift on a finite alphabet \mathcal{A} . For $\beta \in \mathbb{R}$ and a continuous potential $\varphi : X \rightarrow \mathbb{R}$, we define the following three subsets of shift-invariant measures. The set of equilibrium states for $\beta\varphi$ is

$$\text{ES}(\beta\varphi) = \left\{ \mu \text{ shift-invariant} : h(\mu) + \beta \int \varphi d\mu = p_\varphi(\beta) \right\},$$

where $p_\varphi(\beta)$ is defined in (1). Then, the set of maximizing measures for φ is

$$\text{MM}(\varphi) = \left\{ \mu \text{ shift-invariant} : \int \varphi d\mu = s_\varphi \right\},$$

where s_φ is defined in (3). Finally, define the set of zero-temperature accumulation points as

$$\text{ZTA}(\varphi) = \left\{ \mu : \exists \beta_n \rightarrow \infty, \mu_n \in \text{ES}(\beta_n\varphi) \text{ with } \mu_n \rightsquigarrow \mu \right\},$$

where $\mu_n \rightsquigarrow \mu$ means that μ_n converges to μ for the vague topology. A well-known folk theorem states that $\text{ZTA}(\varphi) \subseteq \text{MM}(\varphi)$, with the measures in $\text{ZTA}(\varphi)$ being among those that maximize entropy among all maximizing measures (see Theorem 6.2 in Section 6).

Definition 2.1 (Freezing phase transition for a potential). Let $\varphi : X \rightarrow \mathbb{R}$ be a continuous function. We will say that φ has a freezing phase transition at β_c if $\text{ES}(\beta\varphi) = \text{ES}(\beta'\varphi)$ for all $\beta, \beta' > \beta_c$, while $\text{ES}(\beta\varphi) \neq \text{ES}(\beta'\varphi)$ for any $\beta < \beta_c < \beta'$.

Remark 2.2. By the following proposition, which characterizes freezing phase transitions, it follows that imposing $\text{ES}(\beta\varphi) \neq \text{ES}(\beta'\varphi)$ for any $\beta < \beta_c < \beta'$ necessarily implies that $\text{ES}(\beta\varphi) \cap \text{ES}(\beta'\varphi) = \emptyset$ for all $\beta < \beta_c < \beta'$. Also, if $\text{ES}(\beta\varphi) = \text{ES}(\beta'\varphi)$ for all $\beta, \beta' > \beta_c$, then $\text{ES}(\beta\varphi) = \text{ZTA}(\varphi)$ for all $\beta > \beta_c$. However, we prefer to keep the definition as it is.

The previous definition can be rephrased in terms of the pressure function's affine behavior to the right of β_c , as follows (see (1) for the definition). The proof is given in Section 6.

Proposition 2.3. Let $X = \mathcal{A}^{\mathbb{Z}^d}$ be a full d -dimensional shift on a finite alphabet \mathcal{A} , $\varphi : X \rightarrow \mathbb{R}$ be a continuous function, and $\beta_c \in \mathbb{R}$. Then, φ has a freezing phase transition at β_c if and only if the pressure function is

affine on $[\beta_c, \infty)$ and not on any larger interval. Moreover, in this case $p_\varphi(\beta) = s_\varphi\beta + h_\varphi$, where s_φ, h_φ are defined in (3).

According to the above proposition, if φ has a freezing phase transition, the pressure function $p_\varphi(\beta)$ would exhibit a graph similar to the one shown in Figure 1.

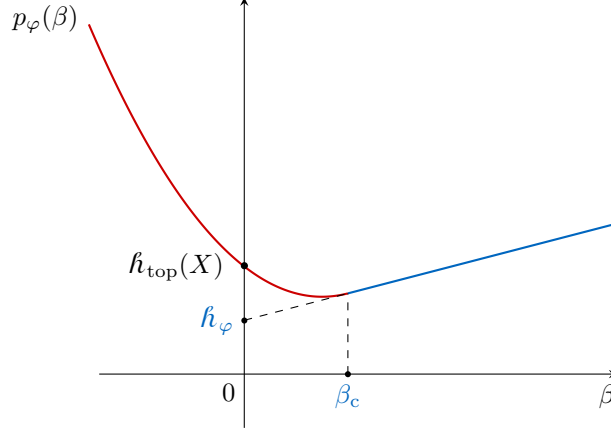


FIGURE 1. Freezing phase transition at β_c .

We now address our main question: How can we construct a continuous function φ that *freezes on a given subshift* X_0 of X ? We make the following definition.

Definition 2.4 (Freezing on a subshift). Let X_0 be a proper subshift of X . A continuous potential $\varphi: X \rightarrow \mathbb{R}$ is said to *freeze on* X_0 if it has a freezing phase transition at some parameter β_c and X_0 is the smallest subshift (with respect to the inclusion) which contains the supports of all measures in $\text{ZTA}(\varphi)$.

Our main result is the following.

Theorem A. *For any proper subshift X_0 of a full d -dimensional shift X on a finite alphabet, there is an explicit potential φ with a freezing phase transition at some $\beta_c > 0$ such that for $\beta > \beta_c$, $\text{ES}(\beta\varphi)$ is the set of measures of maximal entropy on X_0 (and coincides with $\text{ZTA}(\varphi)$), while for $\beta < \beta_c$, $\mu(X_0) = 0$ for all $\mu \in \text{ES}(\beta\varphi)$. In particular, if the smallest subshift containing the supports of all measures of maximal entropy on X_0 is X_0 itself, then φ freezes on X_0 at β_c .*

A more detailed statement is given in Theorem 3.6 below. In particular, the potential takes countably-many values and its variation is bounded above by a term arising from the rate of convergence in the definition of topological entropy (4) and an error term of order $\log \log n/n^d$. In particular, when the

subshift X_0 is a full shift with alphabet $\mathcal{A}' \subsetneq \mathcal{A}$, $\text{var}_n \varphi = O\left(\frac{\log \log n}{n^d}\right)$. Another particular case is when X_0 is a constant-box substitution shift, in which case $\text{var}_n \varphi = O\left(\frac{\log n}{n^d}\right)$ (see Corollary 3.8).

However, as our next theorem shows, a freezing phase transition is impossible for potentials φ satisfying $\text{var}_n \varphi \leq \frac{1}{n^{d+\epsilon}}$ for some $\epsilon > 0$ and all $n \geq 1$.

Theorem B. *Let X be a full d -dimensional shift on a finite alphabet and $\varphi : X \rightarrow \mathbb{R}$ be a continuous potential. If $\sum_{n \geq 1} n^{d-1} \text{var}_n \varphi < \infty$ then φ cannot have a freezing phase transition.*

2.2. Some remarks about Theorem A. One may wonder about what happens for $\beta = \beta_c$. As demonstrated already by the classical example of Hofbauer (see below), $\text{ES}(\beta_c \varphi)$ may or may not coincide with $\text{ZTA}(\varphi)$, and the pressure function may or may not be differentiable at β_c .

Mathematically, one can consider negative inverse temperatures. Associated with the limit $\beta \rightarrow -\infty$, one can consider minimizing measures. The corresponding freezing question relates to whether it is possible to reach the associated slant asymptote, which would occur to the left of a certain inverse temperature β'_c (which can be $\neq \beta_c$). Clearly, this is equivalent to analyzing the problem for $\beta'(-\varphi)$ with $\beta' \rightarrow \infty$ since $\beta\varphi = (-\beta)(-\varphi)$.

The last remark is related to the measure of maximal entropy on X . If $\varphi = u \circ \sigma - u + c$ where $u : X \rightarrow \mathbb{R}$ is continuous, $c \in \mathbb{R}$, and $\sigma : X \rightarrow X$ is the shift map, then $p_\varphi(\beta) = c\beta + \log |\mathcal{A}|$. Moreover, these potentials correspond to the same equilibrium state which is the measure of maximal entropy of X whose value is $\log |\mathcal{A}| = h_{\text{top}}(X)$, the topological entropy of X (see next section for more details). All shift-invariant measures are maximizing measures for this φ .

2.3. Intermezzo on statistical physics. One may naturally ask what happens when starting from an interaction $\Phi = (\Phi_\Lambda)_{\Lambda \in \mathbb{Z}^d}$. Unfortunately, the correspondence between Φ and φ is rather intricate in general. For instance, the mapping $\Phi \mapsto \varphi$ is not injective, and there is no general method to construct preimages of φ ; see [46, pp. 34-35]). Hence, some information is inevitably lost when comparing the regularity of Φ with that of φ .

For finite-range interactions – which encompass most models in dimension $d \geq 2$ – the situation is straightforward, as this implies that $\text{var}_n \varphi$ vanishes for all sufficiently large n . Hence, Theorem B implies that finite-range interactions cannot have a freezing phase transition. But of course the condition $\sum_{n \geq 1} n^{d-1} \text{var}_n \varphi < \infty$ can be satisfied by long-range interactions, provided they decay fast enough. (If there exists a finite $R > 0$ such that $\Phi_\Lambda = 0$ for all sets $\Lambda \in \mathbb{Z}^d$ whose diameter is larger than R , the potential is said to have finite range. Otherwise, it has infinite range.) The issue is that one cannot directly derive the natural assumption on Φ from that on φ , which separates interactions based on whether they can induce a freezing phase transition or not. Therefore, we state a theorem specifically for interactions.

Theorem B'. *Let X be a full d -dimensional shift on a finite alphabet and $\Phi = (\Phi_\Lambda)_{\Lambda \in \mathbb{Z}^d}$ be an interaction. If $\sum_{\Lambda \in \mathbb{Z}^d, \Lambda \ni 0} \|\Phi_\Lambda\|_\infty < \infty$ then Φ cannot have a freezing phase transition.*

The summability condition in the previous statement is satisfactory because it holds precisely for interactions belonging to the classical space of interactions. We provide more information in Section 4.2. We do not claim that Theorem B' is entirely new, but we were unable to find a suitable reference.

We end this section with the following remark regarding the analog of Theorem A in terms of interactions.

Remark 2.5. Theorem 3.6 below provides a more precise formulation of Theorem A by explicitly expressing the potentials that exhibit a freezing phase transition in terms of the target subshift X_0 . For each such potential, we derive a formula for the corresponding interaction, which, as expected, is not summable in the sense of Theorem B' (see Remark 3.7 for further details).

2.4. Previously known results. Previous studies on freezing phase transitions have focused exclusively on the case $d = 1$ and specific examples. The constructions were carried out in one-sided settings, i.e., $X = \mathcal{A}^{\mathbb{N}}$, and the advantage is two-fold. First, Ruelle's Perron-Frobenius operator can be used to evaluate the pressure. Second, a potential defined on a one-sided shift can be readily extended to the corresponding two-sided shift without altering the pressure function. Hence, the results concerning phase transitions immediately carry over to the two-sided settings. We note that the converse is not true. Walters [47] provides necessary and sufficient conditions for a continuous two-sided potential to be cohomologous to a one-sided potential, and these conditions cannot be satisfied by any potential with a freezing phase transition.

In the context of shift dynamical systems, the first example of a freezing phase transition is due to Hofbauer [30], though his motivation was to exhibit a continuous potential φ with two distinct (ergodic) equilibrium states (there is no inverse temperature in his paper). He works with the one-sided full shift $\{0, 1\}^{\mathbb{N}}$ and defines the potential φ by $\varphi(0^\infty) = 0$ and $\varphi(x) = a_n$ if $\text{dist}(x, 0^\infty) = 2^{-n}$, where $(a_n)_{n \geq 0}$ is a sequence of negative real numbers such that $\sum_{n \geq 0} e^{a_0 + \dots + a_n} = 1$. If $\sum_{n \geq 0} (n+1) e^{a_0 + \dots + a_n} < \infty$, there are exactly two ergodic equilibrium states. Conversely, if $\sum_{n \geq 0} (n+1) e^{a_0 + \dots + a_n} = \infty$, there is a unique equilibrium state. Reformulated in our context, he proves that this potential has a freezing phase transition at $\beta_c = 1$ with $X_0 = \{0^\infty\}$, hence $s_\varphi = h_\varphi = 0$ and $\text{ES}(\beta\varphi) = \text{MM}(\varphi) = \text{ZTA}(\varphi) = \{\delta_{0^\infty}\}$ for all $\beta > 1$. In these examples, prior to the freezing phase transition, the unique equilibrium state is fully supported, whereas after the phase transition, the unique equilibrium state is supported on a single point. When the previous series is finite, the pressure function is not differentiable at β_c , while it

is differentiable when it is infinite. Hofbauer was inspired by the Fisher-Felderhof droplet model [22], of which it is a crude simplification. We refer to [40] for a shorter, alternative proof of the nonuniqueness part of Hofbauer's results. A systematic study of a family of Hofbauer potentials exhibiting a phase transition at some value β_c was done by Lopes [42].

Another example in the same spirit as Hofbauer's example is given in [44] to exhibit non-analyticity of multifractal spectra for pointwise dimension. The author considers $X = \{0, 1, 2\}^{\mathbb{N}}$ and, reformulated in our context, he has $X_0 = \{0, 1\}^{\mathbb{N}}$ and $\beta_c = 1$. A general case where $X = \mathcal{A}^{\mathbb{N}}$ for a finite alphabet \mathcal{A} and $X_0 \subset X$ is any subshift of finite type is studied in [36], where an explicit construction of a potential with a freezing phase transition on X_0 is carried out.

In the context of one-dimensional, one-sided shifts, with potentials of the same form as the ones used in this paper, references [5, 6] assert that they have established a freezing phase transition with X_0 defined by the subshift generated by either the Thue-Morse or Fibonacci substitutions. Unfortunately, the proofs presented in these articles are flawed. Work in progress [4] establishes a freezing phase transition in the Thue-Morse subshift, building on the techniques of [5], but requiring significant combinatorial refinements. In particular, since these papers use Ruelle's operator, this approach cannot be extended to higher dimensions.

Recent work [7] broadly extends [5, 6, 4] by proving existence of freezing phase transitions for all zero-entropy (one-sided and two-sided) subshifts, instead of very specific ones. However, the proof is not constructive, since the potentials are ultimately obtained through the Hahn-Banach theorem. Recently, leveraging the same underlying techniques – which are inherently non-constructive – it was demonstrated in [29], within a context encompassing ours, that for any nonempty collection of ergodic measures with a constant entropy function, there exists a continuous φ such that, for all $\beta \geq 1$, its closed convex hull coincides with $\text{ES}(\beta\varphi)$.

There is an abstract result in [37] that guarantees the existence of φ leading to a pressure function as in Figure 1. Precisely, given a compact topological dynamical system (X, T) with positive entropy and upper semi-continuous entropy map, and any closed invariant subset $X_0 \subset X$ with positive topological entropy, there exists a continuous potential $\varphi : X \rightarrow \mathbb{R}$ satisfying (1), (2) and (3) (see Lemma 3.5 below). The proof is based on the fact that equilibrium states can be viewed as tangent functionals to the pressure and relies on a theorem by Israel about approximations of shift-invariant measures by equilibrium states; see [31] or [27, Proposition 17.7, p. 341]. It does not provide any information on the type of potentials which would generate a freezing phase transition.

To conclude this section, we observe that in Hofbauer's example, as well as in [4], there exists a unique equilibrium state with full support for all $0 \leq \beta < \beta_c$. In the construction provided in the proof of Theorem A (detailed

in Theorem 3.6), it remains unclear whether fully supported measures $\mu \in \text{ES}(\beta\varphi)$ exist for all $0 \leq \beta < \beta_c$. We leave this as an open question.

2.5. Some illustrative examples. Theorem A is applicable to any subshift; however, we highlight specific examples of particular interest, especially in light of the motivations outlined in the introduction.

We note that our results have interesting implications already in one dimension ($d = 1$). For example, the objective of [5] and [6] was to determine for which sequences $(a_n)_{n \geq 1}$ the Hofbauer-type potential $\varphi(x) = a_n$ if $\text{dist}(x, X_0) = 2^{-n}$ has a freezing phase transition given that X_0 is the Thue-Morse shift or the Fibonacci subshift respectively. Their conclusion is that φ freezes when $a_n \approx -\frac{1}{n^\alpha}$, $\alpha \in (0, 1)$ for the former and when $a_n \approx -\frac{1}{n}$ for the latter. Their techniques rely heavily on the specific properties of the Thue-Morse subshift and Fibonacci shift, which are very well understood, and the use of the transfer operator on a certain induced system. The proofs of [5, 6] appear to be incomplete, and the ongoing effort [4] aims to justify them in the Thue-Morse case as well as establish freezing when $a_n \approx -\frac{1}{n}$. Meanwhile, a direct application of Theorem 3.6 (for the two-sided Thue-Morse and Fibonacci subshifts) gives that φ freezes when $a_n \approx -\frac{\log n}{n}$. Also, Theorem B shows that φ does not freeze for $a_n \leq -\frac{1}{n^{1+\epsilon}}$ for all n and for some $\epsilon > 0$. We discuss further freezing phase transitions for one-dimensional subshifts in Section 5. In particular, up to a logarithmic factor, we recover the result of [4] on the Thue-Morse subshift using a completely different approach. Moreover, Theorem 5.3 applies to any subshift.

In [19], the authors consider Sturmian subshifts (which can be used for instance to code irrational rotations on the unit circle). Here $\mathcal{A} = \{0, 1\}$ and these subshifts are minimal, uniquely ergodic, and have zero topological entropy. They construct an infinite-range potential, decaying as fast as one wants, whose maximizing subshift is a Sturmian subshift. Using Theorem A (see Theorem 3.6 for a more precise statement), we know how to construct infinite-range potentials, decaying slowly enough, such that a freezing phase transition occurs. However, using Theorem B, we know that the potential in [19] cannot have a freezing phase transition.

We now turn to higher-dimensional subshifts, focusing on subshifts of finite type (SFTs) which exhibit significantly greater diversity compared to their one-dimensional counterparts [41], and are relevant to model quasicrystals, as mentioned in the introduction. Informally, $X_0 \subset \mathcal{A}^{\mathbb{Z}^d}$ is an SFT if no configuration in X_0 contains a pattern from a prescribed collection of patterns F , where $F \subset \mathcal{A}^\Lambda$ and $\Lambda \subset \mathbb{Z}^d$ is finite.

The first example we mention is taken from [9, Example 1.5]. Let $M \in \mathbb{N}$ and define $\mathcal{A} = \{-M, -M+1, \dots, -2, -1, 1, 2, \dots, M-1, M\}$. Then, $x \in X_0$ if and only if, whenever $\|\mathbf{i} - \mathbf{j}\|_1 = 1$, we have $(x_{\mathbf{i}}, x_{\mathbf{j}}) \notin \{(a, a') \in \mathcal{A} \times \mathcal{A} : aa' \leq -2\}$. In other words, the only way for a negative number to have a positive neighbor is through the pair ± 1 . This defines a (strongly) irreducible subshift of finite type (SFT), which consequently has positive

topological entropy. Theorem 1.17 in [9] then asserts that if $M > 4e28^d$, there are exactly two extremal (ergodic) measures of maximal entropy.

In dimension $d \geq 2$, there exist subshifts of finite type (SFTs) that have uncountably many configurations but no periodic configurations – a phenomenon that is impossible for $d = 1$. Let us focus on the case $d = 2$. Such SFTs are typically constructed using aperiodic Wang tilings. Given a finite collection of 1×1 square tiles with colored edges, copies of the tiles are placed (without rotation or reflection) at the points of \mathbb{Z}^2 , arranged so that adjacent edges match in color. For further details, see [41]. A notable example is the Kari-Culik shift, which is constructed in a fundamentally different way from classical examples and, moreover, has positive topological entropy [16]. According to a folk theorem, this SFT cannot be uniquely ergodic. Recently, a family of minimal, uniquely ergodic, aperiodic Wang shifts was constructed in [38, 39]. These shifts are self-similar in the sense that they can also be obtained using two-dimensional substitutions, where the inflation factors are the n -th metallic means. This provides an elegant family of quasicrystal models for which we have explicit constructions of potentials that freeze on them.

3. FREEZING ON ANY PROPER SUBSHIFT

3.1. Notation and definitions. Let $X = \mathcal{A}^{\mathbb{Z}^d}$ be the d -dimensional full shift on a finite alphabet \mathcal{A} , *i.e.*, X is the space of all maps $x : \mathbb{Z}^d \rightarrow \mathcal{A}$. We will call them configurations. We put the discrete topology on \mathcal{A} and the product topology on X (which is thus compact). For $\Lambda \subset \mathbb{Z}^d$, $x|_{\Lambda}$ is the projection of x onto Λ (or the restriction of x to Λ as a map). For $\mathbf{n} \in \mathbb{Z}^d$ we define the shift map $\sigma^{\mathbf{n}}$ of X by $(\sigma^{\mathbf{n}}x)_{\mathbf{m}} = x_{\mathbf{n}+\mathbf{m}}$ for $x = (x_{\mathbf{m}})_{\mathbf{m} \in \mathbb{Z}^d} \in X$. The shift map is the natural action of \mathbb{Z}^d on X by translations. We take $\text{dist}(x, y) = 2^{-n}$, where $n = n(x, y) = \min\{\|\mathbf{n}\|_{\infty} : x_{\mathbf{n}} \neq y_{\mathbf{n}}\}$, as the distance between two configurations x, y (compatible with the product topology).

The space of shift-invariant probability measures is equipped with the vague topology, under which it is compact.

A subshift X_0 of X is a closed subset $X_0 \subset X$ which is $\sigma^{\mathbf{n}}$ -invariant for every $\mathbf{n} \in \mathbb{Z}^d$. Given $\Lambda \Subset \mathbb{Z}^d$ (that is, Λ is a finite subset of \mathbb{Z}^d), a pattern $w \in \mathcal{A}^{\Lambda}$ appears in a configuration x if there exists $\mathbf{m}_0 \in \mathbb{Z}^d$ such that $w_{\mathbf{m}} = x_{\mathbf{m}+\mathbf{m}_0}$ for all $\mathbf{m} \in \Lambda$. We recall the definition of the language of X_0 . For $\Lambda \Subset \mathbb{Z}^d$, define $\mathcal{L}_{\Lambda}(X_0)$ as the set of patterns supported on Λ that appear in some configuration $x \in X_0$. The language of X_0 is defined as $\mathcal{L}(X_0) = \bigcup_{\Lambda \Subset \mathbb{Z}^d} \mathcal{L}_{\Lambda}(X_0)$.

Throughout the paper, we define $\Lambda_n^+ = [0, n-1]^d \cap \mathbb{Z}^d$ and $\Lambda_n = [-n, n]^d \cap \mathbb{Z}^d$; these appear more frequently in ergodic theory and statistical physics respectively. For $\Lambda \subset \mathbb{Z}^d$, we denote by $|\Lambda|$ its volume, that is, the number of sites in Λ . The topological entropy of X_0 is

$$(4) \quad h_{\text{top}}(X_0) = \lim_{n \rightarrow \infty} \frac{1}{|\Lambda_n^+|} \log |\mathcal{L}_{\Lambda_n^+}(X_0)|.$$

A function $\varphi : X \rightarrow \mathbb{R}$ is continuous if and only if $\text{var}_n(\varphi) := \sup\{|\varphi(x) - \varphi(y)| : x|_{\Lambda_n} = y|_{\Lambda_n}\}$ goes to 0. (Note that $\text{var}_n(\varphi) = \sup\{|\varphi(x) - \varphi(y)| : x, y \in X, \text{dist}(x, y) \leq 2^{-n}\}$, the modulus of continuity of φ .) Given $\varphi : X \rightarrow \mathbb{R}$ continuous, recall that $\mu \in \text{ES}(\varphi)$ means that

$$h(\nu) + \int \varphi d\mu = \sup \left\{ h(\nu) + \int \varphi d\nu : \nu \text{ shift-invariant} \right\},$$

where $h(\nu)$ is the (measure-theoretic) entropy of ν , that is,

$$h(\nu) = \lim_{n \rightarrow \infty} -\frac{1}{|\Lambda_n^+|} \sum_{w \in \mathcal{A}^{\Lambda_n^+}} \nu([w]) \log \nu([w])$$

with $[w] = \{x \in X : x|_{\Lambda_n^+} = w\}$. By compactness in vague topology, there exists at least one equilibrium state for φ , that is, at least one μ maximizing the $h(\nu) + \int \varphi d\nu$, as a function of ν ; this is the variational principle. By definition, the above supremum is the topological pressure of φ , denoted $P(\varphi)$. If $\varphi \equiv 0$ then $P(\varphi) = P(0) = h_{\text{top}}(X) = \log |\mathcal{A}|$. (More generally, this is true if $\varphi = u \circ \sigma - u + c$ where $u : X \rightarrow \mathbb{R}$ is continuous, $c \in \mathbb{R}$, see [46, p. 61].) Given $\beta \in \mathbb{R}$, we can define, as done in the introduction, the equilibrium states of $\beta\varphi$ in the obvious way, as well as the pressure function $\beta \mapsto p_\varphi(\beta) := P(\beta\varphi)$ (which is convex and Lipschitz continuous). Given a subshift X_0 of X , its measures of maximal entropy are the measures whose measure-theoretic entropy is the same as $h_{\text{top}}(X_0)$.

3.2. Dyadic tiling lemma. The following lemma plays a key role in our construction of potentials that freeze on a subshift. To show that equilibrium states are supported on the given subshift, we will show that ergodic measures not supported on the subshift do not achieve the supremum in the variational principle. Our approach to this is based on efficiently partitioning the orbit of a point lying outside the subshift into blocks that do lie inside the subshift.

Definition 3.1. A *tiling* t of \mathbb{Z}^d is a partition of \mathbb{Z}^d into a disjoint union of finite non-empty subsets: $\mathbb{Z}^d = \bigcup_{i \in I} S_i$. Given $\mathbf{j} \in \mathbb{Z}^d$, we let $t(\mathbf{j})$ denote the set S_i containing \mathbf{j} .

We let $\mathcal{T}(\mathbb{Z}^d)$ denote the collection of all tilings of \mathbb{Z}^d . If $t = \{S_i : i \in I\}$ is a tiling and $\mathbf{j} \in \mathbb{Z}^d$, we define the *translation of t by \mathbf{j}* to be $\text{Tr}_{\mathbf{j}}(t) = \{S_i - \mathbf{j} : i \in I\}$, where $S_i - \mathbf{j}$ denotes the translate of S_i by the vector \mathbf{j} , that is, $S_i - \mathbf{j} = \{\mathbf{k} - \mathbf{j} : \mathbf{k} \in S_i\}$.

Definition 3.2. Let T be a measurable \mathbb{Z}^d -action on a measure space (Z, ρ) . An *equivariant tiling* is a map τ from Z to $\mathcal{T}(\mathbb{Z}^d)$ with the property that $\tau(T^{\mathbf{j}}z) = \text{Tr}_{\mathbf{j}}(\tau(z))$.

Rather than writing $\tau(z)$ for the tiling, we will write τ^z , so that $\tau^z(\mathbf{j})$ is the tile in $\tau(z)$ containing \mathbf{j} .

We describe the standard dyadic odometer as the space $\Sigma = \{0, 1\}^{\mathbb{N}_0}$, where we write a point $y \in \Sigma$ as a left-infinite sequence $\dots y_2 y_1 y_0$ and define the homeomorphism ι of Σ by

$$\iota(y) = \begin{cases} \dots y_{n+1} 1 0 \dots 0 & \text{if } y_n \dots y_0 = 0 1 \dots 1 \text{ for some } n \geq 0; \\ \dots 0 0 0 0 & \text{if } y = \dots 1 1 1 1. \end{cases}$$

That is, in our presentation of the standard odometer, the map ι is binary addition of 1 with carry to the left. Given $y \in \Sigma$, we define a family of equivalence relations on \mathbb{Z} by $i \sim_{n,y} j$ if $\iota^i(y)$ and $\iota^j(y)$ agree in all coordinates at level n and above. That is, the $\sim_{0,y}$ -equivalence classes corresponding to y partition \mathbb{Z} into singletons; the $\sim_{1,y}$ -equivalence classes corresponding to y partition \mathbb{Z} into intervals of length 2; and the $\sim_{n,y}$ -equivalence classes corresponding to y partition \mathbb{Z} into intervals of length 2^n . Further, each $\sim_{n,y}$ -equivalence class is the union of exactly two $\sim_{n-1,y}$ -equivalence classes. If τ_n is the map sending y to the collection of $\sim_{n,y}$ -equivalence classes, we see that each τ_n is a \mathbb{Z} -equivariant tiling.

Definition 3.3. The \mathbb{Z}^d -dyadic odometer is the space

$$Y = \{(y_1, \dots, y_d) : y_i \in \Sigma\},$$

with the \mathbb{Z}^d -action on Y given by $T_Y^j(y) = (\iota^{j_1}(y_1), \dots, \iota^{j_d}(y_d))$.

The \mathbb{Z}^d -dyadic odometer has a family of natural equivariant tilings given by

$$\tau_n^y = \{S_n(j, y) : j \in \mathbb{Z}^d\},$$

where

$$S_n(j, y) = \{k \in \mathbb{Z}^d : j_i \sim_{n, y_i} k_i \text{ for each } 1 \leq i \leq d\}.$$

That is, τ_n^y is a decomposition of \mathbb{Z}^d into a grid of $2^n \times \dots \times 2^n$ boxes, and each τ_n^y is the union of $2^d \tau_{n-1}^y$ boxes.

The following *Dyadic Tiling Lemma* is closely related to a result that appears in the work of Aaronson and Weiss [2]; see also notes of Meyerovitch [43] for related work.

Lemma 3.4 (Dyadic Tiling Lemma). *Let X_0 be a proper subshift of $X = \mathcal{A}^{\mathbb{Z}^d}$ and let μ be an ergodic shift-invariant measure supported on $X \setminus X_0$. Let Y be the dyadic \mathbb{Z}^d -odometer with unique invariant measure ν and let $\bar{\mu}$ be an ergodic joining of μ and ν , where \mathbb{Z}^d acts by the product action, $T^j(x, y) = (\sigma^j(x), T_Y^j(y))$.*

Then there exists an equivariant tiling $\tau : X \times Y \rightarrow \mathcal{T}(\mathbb{Z}^d)$ with the following properties:

- (1) *For $\bar{\mu}$ -a.e. (x, y) and each $j \in \mathbb{Z}^d$, there exists $n \in \mathbb{N}_0$ such that $\tau^{(x,y)}(j) = \tau_n^y(j)$. That is, all tiles appearing in $\tau(x, y)$ are dyadic tiles defined by y (at various scales).*

- (2) For $\bar{\mu}$ -a.e. (x, y) and each $\mathbf{j} \in \mathbb{Z}^d$, if $\tau^{(x,y)}(\mathbf{j}) = \tau_n^y(\mathbf{j})$ and $n > 0$, then $x|_{\tau^{(x,y)}(\mathbf{j})} \in \mathcal{L}(X_0)$. That is, the restriction of x to any $\tau^{(x,y)}$ -tile that is larger than $1 \times \cdots \times 1$ belongs to the language of X_0 ;
- (3) For $\bar{\mu}$ -a.e. (x, y) and each $\mathbf{j} \in \mathbb{Z}^d$, if $\tau^{(x,y)}(\mathbf{j}) = \tau_n^y(\mathbf{j})$, then $x|_{\tau_{n+1}^y(\mathbf{j})} \notin \mathcal{L}(X_0)$. That is, the restriction of x to any dyadic tile larger than the $\tau^{(x,y)}$ -tile does not belong to the language of X_0 .

Taken as a whole, the lemma states that the tiling $\tau(x, y)$ is dyadic and each tile larger than the 0th level is the maximal dyadic tile for which the corresponding x word belongs to the language of X_0 .

Proof. Let μ and ν be as in the statement and let $\bar{\mu}$ be an ergodic joining of them. Define

$$(5) \quad n^{(x,y)}(\mathbf{j}) = \begin{cases} 0 & \text{if } x_{\mathbf{j}} \notin \mathcal{L}(X_0); \\ \max\{n \in \mathbb{N}_0 : x|_{\tau_n^y(\mathbf{j})} \in \mathcal{L}(X_0)\} & \text{otherwise.} \end{cases}$$

Note that since μ is ergodic and μ is not supported on X_0 , $n^{(x,y)}$ is almost surely finite. We then define the tiling τ by

$$\tau^{(x,y)}(\mathbf{j}) = \tau_{n^{(x,y)}(\mathbf{j})}^y(\mathbf{j}) \text{ for all } (x, y) \text{ and } \mathbf{j}.$$

That is, $\tau^{(x,y)}(\mathbf{j})$ is the largest dyadic tile (as determined by y) for which the corresponding x -block belongs to $\mathcal{L}(X_0)$. To see that this is a tiling, let $(x, y) \in X \times Y$, let $\mathbf{j} \in \mathbb{Z}^d$ and let $n = n^{(x,y)}(\mathbf{j})$. First notice that $\mathbf{j} \in \tau^{(x,y)}(\mathbf{j})$ for each $\mathbf{j} \in \mathbb{Z}^d$, so that the tiles cover \mathbb{Z}^d . Next note that if $\mathbf{k} \in \tau^{(x,y)}(\mathbf{j})$, then $\tau_n^y(\mathbf{j}) = \tau_n^y(\mathbf{k})$, so that $\tau_m^y(\mathbf{j}) = \tau_m^y(\mathbf{k})$ for all $m \geq n$. From the definition of $n^{(x,y)}$, it follows that $n^{(x,y)}(\mathbf{k}) = n^{(x,y)}(\mathbf{j})$, so that $\tau^{(x,y)}(\mathbf{k}) = \tau^{(x,y)}(\mathbf{j})$. This ensures that distinct tiles do not overlap. The three properties then follow by construction of $\tau^{(x,y)}$. \square

3.3. Potentials that freeze on a given subshift. We will prove Theorem A, which, as stated above, follows as a consequence of a more detailed theorem. The following proposition provides sufficient conditions for a potential to freeze on a given subshift.

Proposition 3.5. *Let X_0 be a proper subshift of X and $\varphi : X \rightarrow \mathbb{R}$ a continuous potential such that*

- (1) $\varphi|_{X_0} \equiv 0$ and $\varphi(x) < 0$ for all $x \notin X_0$;
- (2) $p_\varphi(1) = h_{\text{top}}(X_0)$;
- (3) *The set of equilibrium states for φ is exactly the set of measures of maximal entropy for X_0 .*

Then there exists $\beta_c \in (0, 1]$ such that φ has a freezing phase transition at β_c . For $\beta > \beta_c$, $\text{ES}(\beta\varphi)$ is the set of measures of maximal entropy on X_0 ; while for $\beta < \beta_c$, $\mu(X_0) = 0$ for all $\mu \in \text{ES}(\beta\varphi)$. In particular, if X_0 is the smallest subshift containing the support of all measures of maximal entropy on X_0 , then φ freezes on X_0 at β_c .

Proof. Since $\varphi \leq 0$, $p_\varphi(\beta)$ is non-increasing, so for $\beta \geq 1$ $p_\varphi(\beta) \leq p_\varphi(1) = h_{\text{top}}(X_0)$. If μ is any measure of maximal entropy on X_0 , then since $\varphi|_{X_0} \equiv 0$, we have $h(\mu) + \beta \int \varphi d\mu = h_{\text{top}}(X_0)$, so that $p_\varphi(\beta) = h_{\text{top}}(X_0)$ for any $\beta \geq 1$.

On the other hand, for $\beta = 0$ we have $p_\varphi(\beta\varphi) = p_\varphi(0) = h_{\text{top}}(X) > h_{\text{top}}(X_0)$, since X_0 is a proper subshift of X . Let $\beta_c = \inf\{\beta: p_\varphi(\beta) = h_{\text{top}}(X_0)\}$, so that $0 < \beta_c \leq 1$. By Proposition 2.3, φ has a freezing phase transition at β_c . By Remark 2.2, $\text{ES}(\beta\varphi)$ is the set of measures of maximal entropy on X_0 for $\beta > \beta_c$.

Hence, in the case that X_0 is the smallest subshift containing the supports of all measures of maximal entropy on X_0 , we have shown φ freezes on X_0 at β_c .

Finally, if $0 < \beta < \beta_c$, then $p_\varphi(\beta) > h_{\text{top}}(X_0)$ so $\text{ES}(\beta\varphi)$ does not contain any measures supported on X_0 . By ergodicity, for $\mu \in \text{ES}(\beta\varphi)$, $\mu(X_0) = 0$. \square

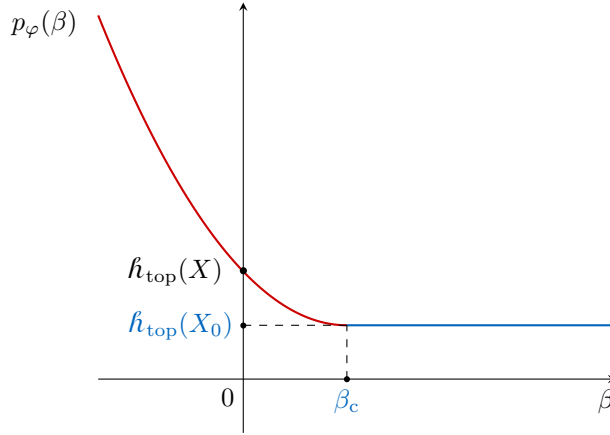


FIGURE 2. Freezing on a proper subshift X_0 .

As the next theorem shows, it is indeed possible to construct potentials satisfying the assumptions of the previous proposition, with a rather precise control on their variation.

Theorem 3.6. *Suppose $X = \mathcal{A}^{\mathbb{Z}^d}$ is a d -dimensional full shift on a finite alphabet \mathcal{A} and X_0 is a proper subshift of X . Define*

$$(6) \quad \varphi(x) = \begin{cases} 0 & \text{if } x \in X_0, \\ -a_j & \text{if } \text{dist}(x, X_0) = 2^{-j}, \end{cases}$$

where (a_j) is a decreasing sequence of positive numbers, converging to zero and satisfying the condition

$$a_{2^{i+1}} \geq \left(\frac{\log n_i}{2^{id}} - h_{\text{top}}(X_0) \right) + \frac{2 \log^+ i + 3}{2^{id}} \text{ for all } i \in \mathbb{N}_0,$$

where n_i denotes the number of admissible $2^i \times \cdots \times 2^i$ blocks in X_0 and $\log^+(\cdot) = \max(\log x, 0)$. Then φ satisfies the conditions of Proposition 3.5.

Remark 3.7. An interaction Φ can be associated with the φ given in (6), namely

$$\Phi_{\Lambda_n}(x) = \begin{cases} a_{n+1} - a_n & \text{if } x|_{\Lambda_n} \notin \mathcal{L}_{\Lambda_n}(X_0); \\ 0 & \text{otherwise.} \end{cases}$$

This interaction lies in the space \mathcal{B} (see (15) for the definition), but clearly not in the space \mathcal{S} of absolutely summable interactions (see (14)). Indeed, as we demonstrate, absolute summability precludes the possibility of a freezing phase transition (see Theorem B' above).

Notice that $\frac{\log n_i}{2^{id}} - \hat{h}_{\text{top}}(X_0)$ is non-increasing and converges to 0 by sub-multiplicativity of the pattern-counting function and the definition of topological entropy; and also $(2 \log^+ i + 3)/2^{id}$ decreases to 0. If we define (a_j) by

$$a_j = \left(\frac{\log n_i}{2^{id}} - \hat{h}_{\text{top}}(X_0) \right) + \frac{2 \log^+ i + 3}{2^{id}} \quad \text{when } 2^i < j \leq 2^{i+1}$$

as above, then we note that $\text{var}_j \varphi = a_j$. In particular, we obtain

$$\text{var}_j \varphi = \kappa_{\lfloor \log_2 j \rfloor} + O\left(\frac{\log \log j}{j^d}\right),$$

where $\kappa_i = \frac{1}{2^{id}} \log n_i - \hat{h}_{\text{top}}(X_0)$. That is, κ_i measures the rate of convergence of $|\Lambda_{2^i}^+|^{-1} \log |\mathcal{L}_{\Lambda_{2^i}^+}(X_0)|$ to $\hat{h}_{\text{top}}(X_0)$.

For subshifts generated by substitutions, we have the following estimate.

Corollary 3.8. *For a constant-box \mathbb{Z}^d -substitution shift X_0 there is a potential φ of the form of (6) with $\text{var}_j \varphi = O(\frac{\log j}{j^d})$ such that φ has a freezing phase transition on X_0 .*

Observe that for any d , if $\text{var}_j \varphi$ is as in the corollary, then $\sum_j j^{d-1} \text{var}_j \varphi = \infty$ (compare with Theorem B).

Proof. It is known [45, Theorem 7.17] that a \mathbb{Z}^d -substitution dynamical system obtained from a substitution rule where each symbol is mapped to a $m_1 \times m_2 \times \cdots \times m_d$ box has a word complexity function given by $|\mathcal{L}_{\Lambda_n^+}(X_0)| = O(n^{\log(m_1 \cdots m_d)/\log \min m_j})$. Since [45] does not include a proof but refers to an unpublished thesis, we include a proof of this below. Given the estimate, we deduce

$$\kappa_i = \frac{1}{2^{id}} \log O(2^{i(\log(m_1 \cdots m_d)/\log \min m_j)}) = O(i/2^{id}).$$

It follows that $a_n = O(\log n/n^d)$ as claimed.

The bound for the number of words may be obtained as follows. Let $m = \min(m_1, \dots, m_d)$ and write $m_j = m^{a_j}$ for each j so that $a_j = \log m_j / \log m$ and let $Q = a_1 + \cdots + a_d$. Now given n , let k be the smallest natural

number such that $m^k \geq n$, that is $k = \lceil \log n / \log m \rceil$. If x is any point in X_0 and $\mathbf{q} \in \mathbb{Z}^d$, then $x_{\mathbf{q}}$ lies in the k -fold substitution of some symbol a . The word $x_{\mathbf{q} + \Lambda_n^+}$ may intersect up to $2^d k$ -fold substitution boxes, since each such box has dimensions at least n in each direction. If \mathbf{o} is the origin of the box containing $x_{\mathbf{q}}$, the 2^d boxes that may be intersected are based at $\mathbf{o} + \epsilon_1 m_1^k \mathbf{e}_1 + \dots + \epsilon_d m_d^k \mathbf{e}_d$ where $\epsilon_i \in \{0, 1\}$ for each i and \mathbf{e}_i is the i th standard basis vector. Thus $x_{\mathbf{q} + \Lambda_n^+}$ is fully determined by the choice of the symbols that were k -times substituted in the 2^d boxes and the difference $\mathbf{q} - \mathbf{o}$. Hence we obtain the bound $|\mathcal{L}_{\Lambda_n^+}(X_0)| \leq |A|^{2^d} (m_1 \dots m_d)^k = |A|^{2^d} (m^Q)^k = |A|^{2^d} (m^k)^Q$. Since $m^k < mn$, we deduce $|\mathcal{L}_{\Lambda_n^+}(X_0)| \leq |A|^{2^d} m^Q n^Q$, so that $|\mathcal{L}_{\Lambda_n^+}(X_0)| = O(n^Q)$ as claimed. \square

Proof of Theorem 3.6. Let φ be a potential on X as described in the statement of the theorem. Since $\varphi|_{X_0} = 0$, if μ_0 is a measure of maximal entropy on X_0 , then $\hat{h}(\mu_0) + \int \varphi d\mu_0 = \hat{h}_{\text{top}}(X_0)$. To complete the proof, we show that if μ is any ergodic measure not supported on X_0 , then $\hat{h}(\mu) + \int \varphi d\mu < \hat{h}_{\text{top}}(X_0)$.

Let μ be an ergodic measure on X such that the support of μ does not lie in X_0 (so that by ergodicity, $\mu(X_0) = 0$). We let $\bar{\mu}$ be an ergodic joining of μ and ν where ν is the unique invariant measure on the dyadic \mathbb{Z}^d odometer. By Lemma 3.4, there exists a $\bar{\mu}$ -a.e. defined equivariant tiling $\tau^{(x,y)}$ of \mathbb{Z}^d satisfying the properties given in that lemma.

Let $E = \{(x, y) : \text{lex min } \tau^{(x,y)}(\mathbf{0}) = \mathbf{0}\}$. For each $i \in \mathbb{N}_0$, let $E_i = \{(x, y) \in E : n^{(x,y)}(\mathbf{0}) = i\}$ with $n^{(x,y)}(\mathbf{0})$ as defined in (5), let $n_i = |\mathcal{L}_{\Lambda_{2^i}^+}(X_0)|$ denote the number of legal $2^i \times \dots \times 2^i$ -blocks in $\mathcal{L}(X_0)$, and let $W_1^i, \dots, W_{n_i}^i$ be an enumeration of these blocks. We let $E_{i,j} = \{(x, y) \in E_i : x|_{\Lambda_{2^i}^+} = W_j^i\}$. We now define three partitions:

$$\begin{aligned} \mathcal{P}_0 &= \{E, E^c\}; \\ \mathcal{P}_1 &= \{E^c, E_0, E_1, \dots\}; \text{ and} \\ \mathcal{P}_2 &= \{E^c, E_{i,j} : i \in \mathbb{N}_0, 1 \leq j \leq n_i\}. \end{aligned}$$

Let $p_i = \bar{\mu}(E_i)/\bar{\mu}(E)$. Clearly, $\sum_{i=0}^{\infty} p_i = 1$. We claim that additionally $\bar{\mu}(E) \sum_i 2^{di} p_i = 1$. To see this second equality, notice that since $\tau^{(x,y)}$ is almost surely a tiling of \mathbb{Z}^d , the sets $\{T^{\mathbf{j}}(E_i) : \mathbf{j} \in \Lambda_{2^i}^+\}$ form a partition (up to sets of measure 0) of $X \times Y$. Summing the measures of the sets, we obtain $1 = \sum_{i=0}^{\infty} 2^{di} \bar{\mu}(E_i) = \bar{\mu}(E) \sum_{i=0}^{\infty} 2^{di} p_i$. In the next two lemmas, we estimate $\hat{h}(\mu)$ and $\int \varphi d\mu$ in terms of the quantities we have introduced.

Lemma 3.9. *Let μ be an ergodic invariant measure on X that is not supported on X_0 . With the notation defined above,*

$$\hat{h}(\mu) < \bar{\mu}(E) \sum_{i=0}^{\infty} p_i (3 + 2 \log^+ i + \log n_i).$$

Lemma 3.10. *Let μ be an ergodic invariant measure on X that is not supported on X_0 . With the notation defined above,*

$$\int \varphi d\mu \leq -\bar{\mu}(E) \sum_{i=0}^{\infty} p_i 2^{id} a_{2^{i+1}}.$$

We now complete the proof of Theorem 3.6, deferring the proof of the lemmas until afterwards.

Summing the two contributions from the lemmas, we see

$$\begin{aligned} h(\mu) + \int \varphi d\mu &< \bar{\mu}(E) \sum_{i=0}^{\infty} p_i (3 + 2 \log^+ i + \log n_i - 2^{id} a_{2^{i+1}}) \\ &= h_{\text{top}}(X_0) + \bar{\mu}(E) \sum_{i=0}^{\infty} p_i (3 + 2 \log^+ i + \log n_i - 2^{id} h_{\text{top}}(X_0) - 2^{id} a_{2^{i+1}}). \end{aligned}$$

By hypothesis, all the terms in the sum are non-positive, so that $h(\mu) + \int \varphi d\mu < h_{\text{top}}(X_0)$ for any ergodic μ that is not supported on X_0 . \square

Proof of Lemma 3.9. Notice that $\mathcal{P}_0 \prec \mathcal{P}_1 \prec \mathcal{P}_2$. Let $\mathcal{Q}_X = \{[a] \times Y : a \in \mathcal{A}\}$ be the partition according to the symbol x_0 and let $\mathcal{B}_X = \bigvee_{j \in \mathbb{Z}^d} T^j \mathcal{Q}_X$. Since $\bar{\mu}$ is a joining of μ and another measure, we have $h(\bar{\mu}, \mathcal{Q}_X) = h(\mu)$. We claim that the elements of \mathcal{Q}_X (and hence of $T^j \mathcal{Q}_X$ and hence of \mathcal{B}_X) agree with elements of $\bigvee_{j \in \mathbb{Z}^d} T^j \mathcal{P}_2$ up to measure 0 sets. For $\bar{\mu}$ -a.e. (x, y) , $\bigvee_{j \in \mathbb{Z}^d} T^j \mathcal{P}_0$ determines the tiling $\tau^{(x, y)}$; and $\bigvee_{j \in \mathbb{Z}^d} T^j \mathcal{P}_2$ determines x_0 . We define $\mathcal{B}^- = \bigvee_{j <_{\text{lex}} 0} T^j \mathcal{P}_2$, where $<_{\text{lex}}$ refers to the standard lexicographic (total) ordering on \mathbb{Z}^d .

By the above, we see $h(\mu) \leq h(\bar{\mu}, \mathcal{P}_2) \leq h(\bar{\mu}) \leq h(\mu)$, where the last inequality comes from that fact that Y has zero topological entropy. Hence we see $h(\mu) = h(\bar{\mu}) = H_{\bar{\mu}}(\mathcal{P}_2 | \mathcal{B}^-)$.

We now estimate

$$\begin{aligned} (7) \quad h(\mu) &= H_{\bar{\mu}}(\mathcal{P}_2 | \mathcal{B}^-) \\ &= H_{\bar{\mu}}(\mathcal{P}_0 | \mathcal{B}^-) + H_{\bar{\mu}}(\mathcal{P}_1 | \mathcal{B}^- \vee \mathcal{P}_0) + H_{\bar{\mu}}(\mathcal{P}_2 | \mathcal{B}^- \vee \mathcal{P}_0 \vee \mathcal{P}_1) \\ &\leq H_{\bar{\mu}}(\mathcal{P}_0 | \mathcal{B}^-) + H_{\bar{\mu}}(\mathcal{P}_1 | \mathcal{P}_0) + H_{\bar{\mu}}(\mathcal{P}_2 | \mathcal{P}_1). \end{aligned}$$

We observe that \mathcal{P}_0 is \mathcal{B}^- measurable, as the locations and sizes of the tiles in $\tau^{(x, y)}$ starting in the lexicographic past determine whether or not a new tile is to start at the origin. In particular, we deduce

$$H_{\bar{\mu}}(\mathcal{P}_0 | \mathcal{B}^-) = 0.$$

A standard estimate shows

$$H_{\bar{\mu}}(\mathcal{P}_1 | \mathcal{P}_0) = -\bar{\mu}(E) \sum_{i=0}^{\infty} p_i \log p_i,$$

as the partition \mathcal{P}_1 is obtained from \mathcal{P}_0 by splitting a set of measure $\bar{\mu}(E)$ into pieces that have conditional measures p_0, p_1, \dots . For $i = 0, 1$, we have

$-p_i \log p_i \leq \frac{1}{e}$, the maximum value of $-x \log x$. For $i \geq 2$, we partition the terms into those i 's for which $p_i < \frac{1}{i^2}$ and those for which $p_i \geq \frac{1}{i^2}$. In the first case, since $-x \log x$ is increasing on $[0, \frac{1}{e}]$, $-p_i \log p_i \leq -\frac{1}{i^2} \log \frac{1}{i^2} = \frac{2}{i^2} \log i$. In the second case, we have $-\log p_i \leq 2 \log i$. Hence we have

$$\begin{aligned} H_{\bar{\mu}}(\mathcal{P}_1|\mathcal{P}_0) &\leq \bar{\mu}(E) \left(\frac{2}{e} + \sum_{i \geq 2, p_i < 1/i^2} \frac{2 \log i}{i^2} + 2 \sum_{i \geq 2, p_i \geq 1/i^2} p_i \log i \right) \\ &\leq \bar{\mu}(E) \left(\frac{2}{e} + \sum_{i \geq 2} \frac{2 \log i}{i^2} + 2 \sum_{i \geq 2} p_i \log i \right) \\ &< \bar{\mu}(E) \left(3 + 2 \sum_{i \geq 2} p_i \log i \right) \\ &= \bar{\mu}(E) \sum_{i \geq 0} p_i (3 + 2 \log^+ i). \end{aligned}$$

Similarly, we have

$$H_{\bar{\mu}}(\mathcal{P}_2|\mathcal{P}_1) \leq \sum_{i=0}^{\infty} \bar{\mu}(E_i) \log n_i = \bar{\mu}(E) \sum_{i=0}^{\infty} p_i \log n_i,$$

since \mathcal{P}_2 is obtained by splitting each E_i , a set of measure $\bar{\mu}(E_i) = p_i \bar{\mu}(E)$, into n_i pieces.

Substituting the bounds in (7), we obtain

$$h(\mu) < \bar{\mu}(E) \sum_{i=0}^{\infty} p_i (3 + 2 \log^+ i + \log n_i),$$

which completes the proof of the lemma. \square

Proof of Lemma 3.10. Define a function $\bar{\varphi}$ on $X \times Y$ by $\bar{\varphi}(x, y) = \varphi(x)$. As observed above, $\{T^{\mathbf{j}} E_i : i \in \mathbb{N}_0, \mathbf{j} \in \Lambda_{2^i}^+\}$ forms a countable partition of $X \times Y$. If $(x, y) \in T^{\mathbf{j}} E_i$ for some $\mathbf{j} \in \Lambda_{2^i}^+$ then $\mathbf{j} \in \tau^{(x, y)}(\mathbf{0}) = \tau_i^y(\mathbf{0})$. By property (3), $x|_{\tau_{i+1}^y(\mathbf{0})} \notin \mathcal{L}(X_0)$. It follows that $\text{dist}(x, X_0) \geq 2^{-2^{i+1}}$ so that $\bar{\varphi}(x, y) = \varphi(x) \leq -a_{2^{i+1}}$. Hence, we see

$$\begin{aligned} \int \varphi d\mu &= \int \bar{\varphi} d\bar{\mu} \\ &\leq - \sum_i \sum_{\mathbf{j} \in \Lambda_{2^i}^+} a_{2^{i+1}} \bar{\mu}(T^{\mathbf{j}} E_i) \\ &= -\bar{\mu}(E) \sum_i p_i 2^{id} a_{2^{i+1}}. \end{aligned}$$

This completes the proof. \square

4. OBSTRUCTION TO FREEZING PHASE TRANSITIONS

4.1. Proof of Theorem B. In this section we show that the equilibrium states for potentials with summable variation are fully supported, and then use that to prove Theorem B.

Theorem 4.1. *Let X be the full shift $\mathcal{A}^{\mathbb{Z}^d}$ and let $\varphi: X \rightarrow \mathbb{R}$ be a continuous potential satisfying*

$$(8) \quad \sum_{n=1}^{\infty} n^{d-1} \text{var}_n \varphi < \infty.$$

Then any equilibrium state for φ is fully supported (every cylinder set has a strictly positive measure).

For a sketch of the proof, if μ is a shift-invariant probability measure on X which is not fully supported, we use μ to construct another measure $\tilde{\mu}$ on X by occasionally inserting blocks that lie outside the support of μ to ensure that $h(\tilde{\mu}) + \int \varphi d\tilde{\mu} > h(\mu) + \int \varphi d\mu$. This implies that μ is not an equilibrium state of the potential φ .

Proof. Let μ be an ergodic invariant measure that is not fully supported. We let W be a block of symbols such that $\mu([W]) = 0$, where $[W]$ denotes the cylinder set $\{x: x|_{\Lambda} = W\}$ and Λ is the set of coordinates of symbols appearing in W . By extending W if necessary, we may assume $\Lambda = \Lambda_n^+$ for some n . We show that μ is not an equilibrium state for φ by constructing a shift-invariant measure $\tilde{\mu}$ such that

$$h(\tilde{\mu}) + \int \varphi d\tilde{\mu} > h(\mu) + \int \varphi d\mu.$$

Informally, $\tilde{\mu}$ will be constructed by starting from a realization, x , of μ and overwriting blocks in x with copies of the pattern W at random and with very low frequency to obtain a new point z . Since W 's do not occur in x , any W that occurs in z occurs as a result of this overwriting process. This allows one to recover from z the original point x (up to regions that were overwritten with W 's), as well as the locations where the overwriting took place. There may be some ambiguity in identifying the overwriting locations because when a W is inserted it is possible that some of the symbols of the inserted W together with some of the existing symbols in x may make additional copies of W . We show that the entropy of $\tilde{\mu}$ is bounded below by the entropy of μ plus the entropy of the overwriting process minus the entropy of the parts of x that are overwritten and with a correction for the ambiguity in identifying the overwrite locations. If the overwriting process has frequency δ , the entropy of $\tilde{\mu}$ will be $h(\mu) + \delta|\log \delta| - O(\delta)$. Similarly we will show $\int \varphi d\tilde{\mu} \geq \int \varphi d\mu - O(\delta)$. Combining these, we will see that for small $\delta > 0$, the desired inequality holds.

More precisely, let $\delta > 0$ be a parameter and let ν_0 be the Bernoulli measure on $Y_0 = \{0, 1\}^{\mathbb{Z}^d}$ where 1's occur with frequency δ . We define a

continuous shift-commuting map $\Psi: Y_0 \rightarrow Y_0$ by

$$\Psi(y)_v = \begin{cases} 1 & \text{if } y_v = 1; \text{ and } y_u = 0 \text{ for all } \mathbf{u} \text{ with } 0 < \|\mathbf{u} - \mathbf{v}\|_\infty < n \\ 0 & \text{otherwise.} \end{cases}$$

That is, Ψ removes 1's that are closer than n apart, this distance being chosen to ensure that overwritten W 's don't overlap. We then define ν to be $\Psi_*(\nu_0)$. Since Ψ is shift-commuting, ν is another invariant measure on Y_0 , and is still mixing as it is a factor of a mixing measure. We claim that $\bar{d}(\nu, \nu_0) \leq (2n)^d \delta^2$, which is intuitively unsurprising since this is an upper bound for the probability that Ψ changes a 1 to a 0.

Recall that the Ornstein's \bar{d} -distance between measures ν and ν_0 is given by $\bar{d}(\nu, \nu_0) = \inf_{\bar{\nu}} \int \mathbb{1}_{\{y_0 \neq y'_0\}} d\bar{\nu}(y, y')$, where $\bar{\nu}$ runs over all joinings of ν and ν_0 ; see *e.g.* [28]. To estimate $\bar{d}(\nu, \nu_0)$, we define a joining on $Y_0 \times Y_0$ by introducing the map $\bar{\Psi}: Y_0 \rightarrow Y_0 \times Y_0$ given by $\bar{\Psi}(y) = (y, \Psi(y))$ and letting $\bar{\nu} = \bar{\Psi}_*(\nu_0)$. This measure is then a joining of ν_0 in the first coordinate and ν in the second. Further, one can see that y_0 and $\Psi(y)_0$ (i.e., the zeroth coordinates of the two points in $\bar{\Psi}(y)$) only differ if $y_0 = 1$ and there exists \mathbf{u} with $0 < \|\mathbf{u}\|_\infty < n$ such that $y_{\mathbf{u}} = 1$. The probability of this is less than $(2n)^d \delta^2$ as claimed. This shows that $\bar{d}(\nu, \nu_0) \leq (2n)^d \delta^2$. Standard estimates show $|\hat{h}(\nu) - \hat{h}(\nu_0)| \leq -\bar{d}(\nu, \nu_0) \log \bar{d}(\nu, \nu_0) - (1 - \bar{d}(\nu, \nu_0)) \log(1 - \bar{d}(\nu, \nu_0))$. This gives $\hat{h}(\nu) \geq \hat{h}(\nu_0) - O(\delta^2 |\log \delta|)$. We then obtain the bound

$$\hat{h}(\nu) \geq \delta |\log \delta| + O(\delta).$$

Although we don't need the more refined estimate, we note that in fact $\hat{h}(\nu) \geq \delta |\log \delta| + \delta - O(\delta^2 |\log \delta|)$.

We let Y denote the range of Ψ : the set of $y \in \{0, 1\}^{\mathbb{Z}^d}$ such that if $y_{\mathbf{u}} = y_{\mathbf{v}} = 1$ and $\mathbf{u} \neq \mathbf{v}$, then $\|\mathbf{u} - \mathbf{v}\|_\infty \geq n$. We now define a shift-commuting map $\Xi: X \times Y \rightarrow X$ (that is, $\Xi(\sigma^{\mathbf{v}} x, \sigma^{\mathbf{v}} y) = \sigma^{\mathbf{v}} \Xi(x, y)$ for all $\mathbf{v} \in \mathbb{Z}^d$) by

$$\Xi(x, y)_v = \begin{cases} W_{\mathbf{j}} & \text{if there exists } \mathbf{j} \in \Lambda_n^+ \text{ such that } y_{\mathbf{v}-\mathbf{j}} = 1; \\ x_v & \text{otherwise.} \end{cases}$$

That is, $\Xi(x, y)$ replaces all blocks in x starting at positions where $y_{\mathbf{u}} = 1$ by copies of W . The fact that $y \in Y$ ensures that $\Xi(x, y)$ is well-defined as there cannot be two distinct \mathbf{j} 's in Λ_n^+ for which $y_{\mathbf{v}-\mathbf{j}} = 1$. Since ν is mixing, $\mu \times \nu$ is an ergodic invariant measure with respect to the action $\sigma^{\mathbf{v}}(x, y) := (\sigma^{\mathbf{v}} x, \sigma^{\mathbf{v}} y)$. We let $\tilde{\mu} = \Xi_*(\mu \times \nu)$, the invariant measure obtained by editing realizations from μ using the "instructions" from an independent realization of ν .

We complete the proof by giving lower bounds for $\hat{h}(\tilde{\mu})$ and $\int \varphi d\tilde{\mu}$. Let $\bar{X} = X \times Y \times X$ and let $\bar{\Xi}: X \times Y \rightarrow X \times Y \times X$ be given by $\bar{\Xi}(x, y) = (x, y, \Xi(x, y))$. The push-forward, $\bar{\mu} = \bar{\Xi}_*(\mu \times \nu)$ is a joining of μ , ν and $\tilde{\mu}$ where the first two coordinates are independent. We introduce a family of partitions of \bar{X} : $\mathcal{P}_X = \{[a] \times Y \times X : a \in \mathcal{A}\}$, $\mathcal{P}_Y = \{X \times [0] \times X, X \times [1] \times X\}$

and $\mathcal{P}_Z = \{X \times Y \times [a] : a \in \mathcal{A}\}$. We also define $\mathcal{P}_{XYZ} = \mathcal{P}_X \vee \mathcal{P}_Y \vee \mathcal{P}_Z$ and $\mathcal{P}_{XY} = \mathcal{P}_X \vee \mathcal{P}_Y$. The \mathcal{P}_{XYZ} is a generating partition, so that $\hat{h}(\bar{\mu}) = \hat{h}(\bar{\mu}, \mathcal{P}_{XYZ})$. Since Ξ is an isomorphism (its inverse is given by forgetting the third coordinate), we see that $\hat{h}(\bar{\mu}) = \hat{h}(\mu \times \nu) = \hat{h}(\mu) + \hat{h}(\nu)$. By properties of joinings

$$\begin{aligned}\hat{h}(\bar{\mu}, \mathcal{P}_Z) &= \hat{h}(\tilde{\mu}) \\ \hat{h}(\bar{\mu}, \mathcal{P}_X) &= \hat{h}(\mu) \\ \hat{h}(\bar{\mu}, \mathcal{P}_{XY}) &= \hat{h}(\mu) + \hat{h}(\nu),\end{aligned}$$

where we used independence in the third equality. (For background on entropy, we refer to [33].) We let $\mathcal{F}_X, \mathcal{F}_Y, \mathcal{F}_Z, \mathcal{F}_{YZ}$ and so on be the σ -algebras generated by $\mathcal{P}_X, \mathcal{P}_Y, \mathcal{P}_Z, \mathcal{P}_{YZ}$ etc. We then use conditional entropy to write

$$\hat{h}(\bar{\mu}) = \hat{h}(\bar{\mu}, \mathcal{P}_Z) + \hat{h}(\bar{\mu}, \mathcal{P}_Y | \mathcal{F}_Z) + \hat{h}(\bar{\mu}, \mathcal{P}_X | \mathcal{F}_{YZ}).$$

This yields

$$\begin{aligned}(9) \quad \hat{h}(\tilde{\mu}) &= \hat{h}(\bar{\mu}, \mathcal{P}_Z) \\ &= \hat{h}(\mu) + \hat{h}(\nu) - \hat{h}(\bar{\mu}, \mathcal{P}_Y | \mathcal{F}_Z) - \hat{h}(\bar{\mu}, \mathcal{P}_X | \mathcal{F}_{YZ}) \\ &\geq \hat{h}(\mu) + \delta |\log \delta| + O(\delta) - \hat{h}(\bar{\mu}, \mathcal{P}_Y | \mathcal{F}_Z) - \hat{h}(\bar{\mu}, \mathcal{P}_X | \mathcal{F}_{YZ}).\end{aligned}$$

We have $\hat{h}(\bar{\mu}, \mathcal{P}_Y | \mathcal{F}_Z) \leq H_{\bar{\mu}}(\mathcal{P}_Y | \mathcal{F}_Z)$. Note that for $\bar{\mu}$ -a.e. (x, y, z) , if $z|_{\Lambda_n^+} \neq W$, then $y_0 = 0$, so that $I_{\bar{\mu}}(\mathcal{P}_Y | \mathcal{F}_Z) = 0$ on the \mathcal{F}_Z -measurable set $[W]^c$. Since $\#\mathcal{P}_Y = 2$, it follows that $H_{\bar{\mu}}(\mathcal{P}_Y | \mathcal{F}_Z) \leq \tilde{\mu}([W]) \log 2$. Since each 1 in the Y -layer can create no more than $(2n-1)^d$ W 's in the Z -layer (each W created must overlap with the W that is rewritten), we see $\tilde{\mu}([W]) \leq (2n-1)^d \nu([1]) < (2n-1)^d \delta$. Hence

$$\hat{h}(\bar{\mu}, \mathcal{P}_Y | \mathcal{F}_Z) \leq (2n-1)^d \delta \log 2 = O(\delta).$$

Similarly, $\hat{h}(\bar{\mu}, \mathcal{P}_X | \mathcal{F}_{YZ}) \leq H_{\bar{\mu}}(\mathcal{P}_X | \mathcal{F}_{YZ})$. Let $A = \bigcup_{j \in \Lambda_n^+} \sigma^j(X \times [1] \times X)$, that is A is the \mathcal{F}_Y -measurable event that (x, y, z) contains a 1 in the $(-\Lambda_n^+)$ -box, which is also the event that z_0 is part of a block that was overwritten from x . We have that for $\bar{\mu}$ -a.e. (x, y, z) , if $(x, y, z) \in A^c$, then $z_0 = x_0$, and x_0 is \mathcal{F}_{YZ} -measurable, so that $I_{\bar{\mu}}(\mathcal{P}_X | \mathcal{F}_{YZ}) = 0$ on A^c . Meanwhile, $I_{\bar{\mu}}(\mathcal{P}_X | \mathcal{F}_{YZ}) \leq \log \#\mathcal{A}$ on A so that $H_{\bar{\mu}}(\mathcal{P}_X | \mathcal{F}_{YZ}) \leq \bar{\mu}(A) \log \#\mathcal{A}$. Since $\bar{\mu}(A) = n^d \nu([1]) < n^d \delta$, we obtain

$$\hat{h}(\bar{\mu}, \mathcal{P}_X | \mathcal{F}_Y \vee \mathcal{F}_Z) \leq n^d \delta \log \#\mathcal{A} = O(\delta).$$

Substituting in (9), we obtain

$$(10) \quad \hat{h}(\tilde{\mu}) \geq \hat{h}(\mu) + \delta |\log \delta| - O(\delta).$$

To bound $\int \varphi d\tilde{\mu}$ from below, we estimate $|\int \varphi d\tilde{\mu} - \int \varphi d\mu|$. Since $\bar{\mu}$ is a joining of μ , ν and $\tilde{\mu}$, we have

$$\begin{aligned} \left| \int \varphi d\tilde{\mu} - \int \varphi d\mu \right| &= \left| \int (\varphi \circ \pi_Z - \varphi \circ \pi_X) d\bar{\mu} \right| \\ &\leq \int |\varphi \circ \pi_Z - \varphi \circ \pi_X| d\bar{\mu}. \end{aligned}$$

Recall that for $\bar{\mu}$ -a.e. $(x, y, z) \in \bar{X}$, the only differences between x and z occur in places where y has a 1, when blocks of x are overwritten by W 's. Letting $r(y)$ be $\min\{\|\mathbf{j}\|_\infty : y_{\mathbf{j}} = 1\} - (n-1)$ or 0 if that quantity is negative, we have $\text{dist}(x, z) \leq 2^{-r(y)}$ for $\bar{\mu}$ -a.e. (x, y, z) . Hence

$$\begin{aligned} \left| \int \varphi d\tilde{\mu} - \int \varphi d\mu \right| &\leq \int \text{var}_{r(y)} \varphi d\bar{\mu} \\ (11) \quad &= \sum_{k=0}^{\infty} \bar{\mu}\{(x, y, z) : r(y) = k\} \text{var}_k \varphi \\ &= \sum_{k=0}^{\infty} \nu\{y : r(y) = k\} \text{var}_k \varphi. \end{aligned}$$

We observe that

$$\begin{aligned} \{y : r(y) = 0\} &= \bigcup_{\mathbf{j} \in [1-n, n-1]^d} \sigma^{-\mathbf{j}}[1] \quad \text{and} \\ \{y : r(y) = k\} &\subseteq \bigcup_{\|\mathbf{j}\|_\infty = k+(n-1)} \sigma^{-\mathbf{j}}[1] \quad \text{for } k > 0. \end{aligned}$$

Hence

$$\begin{aligned} \nu\{y : r(y) = 0\} &\leq (2n-3)^d \nu([1]) \quad \text{and} \\ \nu\{y : r(y) = k\} &\leq ((2k+2n-3)^d - (2k+2n-5)^d) \nu([1]) \\ &< 2d(2k+2n-3)^{d-1} \nu([1]) \quad \text{for } k > 0. \end{aligned}$$

In the last inequality, we bounded the difference between the two cubes as the sum of the faces. Substituting these bounds in (11), we obtain

$$\left| \int \varphi d\tilde{\mu} - \int \varphi d\mu \right| \leq C \nu([1]),$$

where

$$C = (2n-3)^d \text{var}_0 \varphi + \sum_{k=1}^{\infty} 2d(2k+2n-3)^{d-1} \text{var}_k \varphi.$$

Since the sequence $2d(2k+2n-3)^{d-1}/k^{d-1}$ converges to $2^d d$ as $k \rightarrow \infty$, there exists a constant $L > 0$ such that $2d(2k+2n-3)^{d-1} \leq Lk^{d-1}$ for each

$k \geq 1$. It then follows that $C \leq 2(2n-3)^d \|\varphi\| + L \sum_{k=1}^{\infty} k^{d-1} \text{var}_k \varphi$, and so $C < \infty$ by the hypothesis. Hence

$$\left| \int \varphi d\tilde{\mu} - \int \varphi d\mu \right| \leq C\delta,$$

so that

$$\int \varphi d\tilde{\mu} \geq \int \varphi d\mu - O(\delta).$$

Combining this with (10), we see

$$h(\tilde{\mu}) + \int \varphi d\tilde{\mu} \geq h(\mu) + \int \varphi d\mu + \delta |\log \delta| - O(\delta).$$

In particular, for sufficiently small $\delta > 0$, one sees $h(\tilde{\mu}) + \int \varphi d\tilde{\mu} > h(\mu) + \int \varphi d\mu$, so that μ is not an equilibrium state. \square

Proposition 4.2. *Let X be the full shift $\mathcal{A}^{\mathbb{Z}^d}$ and let φ be a continuous potential satisfying*

$$\sum_{n=1}^{\infty} n^{d-1} \text{var}_n \varphi < \infty.$$

If there exist shift-invariant probability measures μ, ν such that $\int \varphi d\mu \neq \int \varphi d\nu$, then there exists a cylinder set C such that $\eta(C) = 0$ for any $\eta \in \text{MM}(\varphi)$.

Proof. Let μ_1, μ_2 be shift-invariant probability measures such that $\int \varphi d\mu_1 < \int \varphi d\mu_2$. Let $A = \sum n^{d-1} \text{var}_n \varphi$. We prove two simple claims. Firstly, for any $\ell \in \mathbb{N}$, if $x, x' \in X$ are such that $x|_{\Lambda_\ell^+} = x'|_{\Lambda_\ell^+}$, we claim

$$(12) \quad \sum_{j \in \Lambda_\ell^+} |\varphi(\sigma^j x) - \varphi(\sigma^j x')| \leq 2\ell^{d-1} dA.$$

To see this, notice there are less than $2d\ell^{d-1}$ j 's in Λ_ℓ^+ that are at any given distance $1 \leq r \leq \lfloor \frac{\ell}{2} \rfloor$ from $(\Lambda_\ell^+)^c$, so that the left side is bounded above by $2d\ell^{d-1}(\text{var}_1 \varphi + \dots + \text{var}_{\ell/2} \varphi) < 2d\ell^{d-1}A$.

Secondly, for any $\ell \in \mathbb{N}$, if $x, x' \in X$ are such that $x|_{\Lambda_\ell^{+c}} = x'|_{\Lambda_\ell^{+c}}$, and S is any finite set that is disjoint from Λ_ℓ^+ , we claim

$$(13) \quad \sum_{j \in S} |\varphi(\sigma^j x) - \varphi(\sigma^j x')| \leq 2^d \ell^{d-1} dA.$$

The argument is similar: the number of j in S that are at a distance r from Λ_ℓ^+ is at most $(\ell + 2r)^d - (\ell + 2(r-1))^d < 2d(\ell + 2r)^{d-1} < 2d(2\ell r)^{d-1} = 2^d \ell^{d-1} d r^{d-1}$.

Let $\int \varphi d\mu_2 - \int \varphi d\mu_1 = \delta > 0$ and let ℓ be such that $\delta\ell^d - (2^d + 4)\ell^{d-1}dA \geq 1$. Since $\int \sum_{j \in \Lambda_\ell^+} \varphi d\mu_2 - \int \sum_{j \in \Lambda_\ell^+} \varphi d\mu_1 = \delta\ell^d$, there exist $u, u' \in X$ with $\sum_{j \in \Lambda_\ell^+} (\varphi(\sigma^j u') - \varphi(\sigma^j u)) \geq \delta\ell^d$. Let W be the pattern $u|_{\Lambda_\ell^+}$ and $W' = u'|_{\Lambda_\ell^+}$.

We will show that $\mu([W]) = 0$ for any φ -maximizing measure. The idea is simple: we show that replacing copies of W with copies of W' increases the

sum over the orbit. If μ is any invariant measure with $\mu([W]) > 0$, then by systematically replacing copies of W with copies of W' , one obtains a measure μ' with $\int \varphi d\mu' > \int \varphi d\mu$, so that μ could not have been a maximizing measure.

Let x, x' be such that x and x' agree outside Λ_ℓ^+ ; $x_{\Lambda_\ell^+} = W$ and $x'_{\Lambda_\ell^+} = W'$. Let S be any finite set such that $S \supseteq \Lambda_\ell^+$. Then we have

$$\begin{aligned} & \sum_{j \in S} (\varphi(\sigma^j x') - \varphi(\sigma^j x)) \\ & \geq \sum_{j \in \Lambda_\ell^+} (\varphi(\sigma^j x') - \varphi(\sigma^j x)) - \sum_{j \in S \setminus \Lambda_\ell^+} |\varphi(\sigma^j x') - \varphi(\sigma^j x)|. \end{aligned}$$

The subtracted term on the right is at most $2^d \ell^{d-1} dA$ by (13). For the first term, we have

$$\begin{aligned} & \sum_{j \in \Lambda_\ell^+} (\varphi(\sigma^j x') - \varphi(\sigma^j x)) \\ & \geq \sum_{j \in \Lambda_\ell^+} (\varphi(\sigma^j u') - \varphi(\sigma^j u)) \\ & \quad - \sum_{j \in \Lambda_\ell^+} (|\varphi(\sigma^j u') - \varphi(\sigma^j x')| + |\varphi(\sigma^j u) - \varphi(\sigma^j x)|) \\ & \geq \delta \ell^d - 4d \ell^{d-1} A. \end{aligned}$$

Hence we see

$$\sum_{j \in S} (\varphi(\sigma^j x') - \varphi(\sigma^j x)) \geq \delta \ell^d - (2^d + 4d) \ell^{d-1} A \geq 1.$$

Suppose that μ is an ergodic invariant measure for which $\mu([W]) > 0$. Let x be a generic point for μ so that $\frac{1}{N^d} \sum_{j \in \Lambda_N^+} \varphi(\sigma^j x) \rightarrow \int \varphi d\mu$ and $\frac{1}{N^d} \#\{j \in \Lambda_{N-\ell+1}^+ : \sigma^j x \in [W]\} \rightarrow \mu([W])$. For each N , we inductively modify x inside Λ_N^+ replacing W 's with W' 's one at a time until there are no W 's remaining. Let the resulting point be y^N . By the previous estimate, each replacement of a W by a W' causes the sum of φ over Λ_N^+ to increase by at least 1, so that

$$\frac{1}{N^d} \sum_{j \in \Lambda_N^+} \varphi(\sigma^j y^N) \geq \frac{1}{N^d} \sum_{j \in \Lambda_N^+} \varphi(\sigma^j x) + \frac{\#(\text{replacements})}{N^d}.$$

Since replacing one W with a W' can get rid of at most $(3\ell)^d$ W 's, one has $\#(\text{replacements}) \geq \frac{1}{3\ell^d} \#\{j \in \Lambda_{N-\ell+1}^+ : \sigma^j x \in [W]\}$. Hence

$$\limsup_{N \rightarrow \infty} \frac{1}{N^d} \sum_{j \in \Lambda_N^+} \varphi(\sigma^j y^N) \geq \int \varphi d\mu + \frac{\mu([W])}{(3\ell)^d}.$$

Taking a subsequential limit of the empirical measures arising from translating the points y^N by the elements of Λ_N^+ , we obtain a shift-invariant

measure μ' such that $\int \varphi d\mu' \geq \int \varphi d\mu + \mu([W])/(3\ell)^d$, hence μ cannot be a maximizing measure for φ . \square

Proof of Theorem B. We consider two cases. First, if there exist invariant measures μ and ν for which $\int \varphi d\mu \neq \int \varphi d\nu$, then Proposition 4.2 implies that maximizing measures are not fully supported, while Theorem 4.1 implies that equilibrium measures for $\beta\varphi$ are fully supported for all $\beta \in \mathbb{R}$. This implies that there is no freezing phase transition.

Secondly, if $\int \varphi d\mu = \int \varphi d\nu$ for all invariant measures, then the unique equilibrium state is the uniform Bernoulli measure on X for all β , so again there is no freezing phase transition. \square

We remark that by a result of Bousch [8], in the case $d = 1$, if φ satisfies $\sum \text{var}_n \varphi < \infty$, then $\int \varphi d\mu$ is equal for all invariant measures if and only if φ is cohomologous to a constant function.

Remark 4.3. We established that if φ satisfies (8), then the corresponding equilibrium states have full support, meaning that every cylinder set has strictly positive measure. An alternative proof exists, though it relies on the concept of Gibbs states. This follows from Keller's book [33, Chapter 5], as we now outline. Notice that our proof takes a completely different approach, relying on ergodic theoretic techniques and notably avoiding any use of Gibbs states.

Keller first defines Gibbs states associated with a potential φ . (It is worth noting that the classical definition of a Gibbs state is based on an interaction Φ , as in [46] and [27].) He then shows that these Gibbs states – which are not necessarily shift-invariant – assign a strictly positive measure to all cylinder sets. This result follows from a precise estimate involving the Birkhoff sum of φ and its pressure. Finally, he demonstrates that under condition (8), the shift-invariant Gibbs states coincide with the equilibrium states – the key point being to establish that the equilibrium states of φ are indeed Gibbs states. Hence, the fact that equilibrium states for $\beta\varphi$ have full support follows at once for any β .

4.2. Proof of Theorem B'. The proof scheme is the same as that of Theorem B.

One first proves that if $\sum_{\Lambda \in \mathbb{Z}^d, \Lambda \neq \emptyset} \|\Phi_\Lambda\|_\infty < \infty$, then the corresponding Gibbs states for $\beta\Phi$ have full support for any β (see Section 6.4 for a proof). For this class of interactions, shift-invariant Gibbs states for $\beta\Phi$ coincide with equilibrium states for $\beta\varphi$ [46, Theorem 4.2], and consequently, the latter also have full support.

Next, [25, Theorem 2.4 (v)] establishes that if $\mu \in \text{MM}(\varphi)$, then $\text{supp}(\mu) \subseteq X_\Phi$, where $X_\Phi \subseteq X$ is the subshift of maximizing configurations for Φ (that we will not define here)³. However, Theorem 6.2 tells us that $\text{ZTA}(\varphi) \subseteq$

³ Note that [25] does not explicitly state that X_Φ is a subshift.

$\text{MM}(\varphi)$. Therefore, if $X_\Phi \subsetneq X$, a freezing phase transition cannot occur since there exist at least one cylinder set in $X \setminus X_\Phi$.

This leaves only the degenerate case where $X = X_\Phi$. By [25, Theorem 2.4 (iv)], if ν is any shift-invariant measure such that $\text{supp}(\nu) \subseteq X_\Phi$, then $\nu \in \text{MM}(\varphi)$. This implies that $\int \varphi d\nu = s_\varphi$ for any shift-invariant measure ν , and by the variational principle, for any $\beta \in \mathbb{R}$, $p(\beta\varphi) = \sup_\nu \{h(\nu) + \beta s_\varphi\} = \log |\mathcal{A}| + \beta s_\varphi$. Since the pressure function $\beta \mapsto p(\beta\varphi)$ is affine on \mathbb{R} , no freezing phase transition can occur in this case either.⁴

Remark 4.4. Two commonly-studied (Banach) spaces of interactions are \mathcal{S} and \mathcal{B} . The properties of interactions in these two spaces are reminiscent of the properties of potentials with summable variation and that are continuous respectively. The first one is

$$(14) \quad \mathcal{S} := \left\{ \Phi : \|\Phi\|_{\mathcal{S}} := \sum_{\substack{\Lambda \in \mathbb{Z}^d \\ \Lambda \ni 0}} \|\Phi_\Lambda\|_\infty < \infty \right\}.$$

Unfortunately, assuming that $\Phi \in \mathcal{S}$ does not imply that φ satisfies (8). A bigger (Banach) space is

$$(15) \quad \mathcal{B} := \left\{ \Phi : \|\Phi\|_{\mathcal{B}} := \sum_{\substack{\Lambda \in \mathbb{Z}^d \\ \Lambda \ni 0}} |\Lambda|^{-1} \|\Phi_\Lambda\|_\infty < \infty \right\}.$$

If $\Phi \in \mathcal{B}$ then φ is continuous. The space \mathcal{B} is “pathological” in many respects [18].

5. THE ONE-DIMENSIONAL CASE

Theorem A gives an explicit construction of a potential φ . An interesting question concerns the optimality of the potential in (6) in a sense that defining φ using a sequence which converges to zero slower than a_j generates a freezing phase transition, but using a faster converging sequence does not. For a subshift $X_0 \subsetneq \mathcal{A}^{\mathbb{Z}}$ we know that

$$\varphi(x) \approx \left(h_{\text{top}}(X_0) - \frac{\log(\#j\text{-words in } X_0)}{j} \right) - \frac{\log \log j}{j} \text{ if } \text{dist}(x, X_0) = \frac{1}{2j}$$

has a freezing phase transition and the zero-temperature accumulation measures for φ are the measures of maximal entropy of X_0 . It follows from Theorem 5.1 below that when X_0 is a subshift of finite type, the term $\frac{\log \log j}{j}$ can be replaced by $\frac{1}{j}$. Moreover, we establish this fact in the one-sided settings.

⁴ In fact, if $\int \varphi d\nu = s_\varphi$ for all shift-invariant measures ν , it follows from [18, Prop. 2.34, p. 46] that $\varphi - s_\varphi$ belongs to the closed linear span of the family of functions $\{f - f \circ \sigma^i : f : X \rightarrow \mathbb{R} \text{ continuous, } i \in \mathbb{Z}^d\}$.

Theorem 5.1 (One-sided SFT's). *Suppose $X^+ = \mathcal{A}^{\mathbb{N}}$ is a full shift and $X_0^+ \subset X^+$ is a one-step subshift of finite type. Then the potential $\varphi^+ : X^+ \rightarrow \mathbb{R}$ defined by*

$$(16) \quad \varphi^+(x) = \begin{cases} 0 & \text{if } x \in X_0^+, \\ -a_j & \text{if } \text{dist}(x, X_0^+) = 2^{-j}, \end{cases}$$

has a freezing phase transition provided (a_j) is a decreasing sequence converging to 0, such that $a_j \geq \frac{1}{j}$ for all $j \in \mathbb{N}_0$. In this case, $\text{ZTA}(\varphi)$ coincides with the set of maximal entropy measures of X_0^+ . In particular, if all transitive components of X_0^+ have the same entropy then φ freezes on X_0^+ .

Proof. Let μ be an ergodic measure which is not supported on X_0^+ . Then μ assigns positive measure to the set $E = \{x : x_0x_1 \notin \mathcal{L}(X_0^+)\}$. We induce on E . Denote by $r(x)$ the first return time to E , i.e., $r(x) = \min\{j \geq 1 : \sigma^j(x) \in E\}$. We obtain the ergodic system (E, μ_E, σ_E) with corresponding potential φ_E where

$$\mu_E = \frac{\mu|_E}{\mu(E)}, \quad \sigma_E(x, z) = \sigma^{r(x)}(x), \quad \varphi_E(x) = \sum_{i=0}^{r(x)-1} \varphi^+(\sigma^i(x)).$$

Then Abramov's formula and Kac's theorem imply that

$$h(\mu) + \int \varphi^+ d\mu = \mu(E) \left(h(\mu_E) + \int \varphi_E d\mu_E \right).$$

Let $\mathcal{Q} = \{Q_j\}_{j \in \mathbb{N}}$ be the first return time partition of E , i.e., $Q_j = \{x \in E : r(x) = j\}$, and let \mathcal{P} be its further refinement into cylinder sets. Denote by $q_j = \mu_E(Q_j)$ and by n_j the number of words of length j which are admissible in X_0 . We have

$$\sum_{j=1}^{\infty} q_j = 1, \quad \sum_{j=1}^{\infty} j q_j = \frac{1}{\mu(E)}, \quad \text{card}\{A \in \mathcal{P} : A \subset Q_j\} \leq n_j.$$

We estimate the entropy of measure μ_E by the μ_E -entropy of the partition \mathcal{P} :

$$h(\mu_E) \leq H(\mathcal{P}) = H(\mathcal{Q}) + H(\mathcal{P}|\mathcal{Q}) \leq \sum_{j=1}^{\infty} -q_j \log q_j + \sum_{j=1}^{\infty} q_j \log n_j,$$

where in the last step we used the fact that each \mathcal{Q}_j contains at most n_j elements of \mathcal{P} . Note that when $q_j > \frac{1}{j^2}$ we have $-\log q_j < 2 \log j$. Therefore,

$$\begin{aligned}
 \sum_{j=1}^{\infty} -q_j \log q_j &= \sum_{q_j > \frac{1}{j^2}} -q_j \log q_j + \sum_{q_j \leq \frac{1}{j^2}} -q_j \log q_j \\
 &\leq \sum_{q_j > \frac{1}{j^2}} 2q_j \log j + \sum_{q_j \leq \frac{1}{j^2}} -q_j \log q_j \\
 (17) \quad &= \sum_{j=1}^{\infty} 2q_j \log j + \sum_{q_j \leq \frac{1}{j^2}} (-q_j 2 \log j - q_j \log q_j) \\
 &= \sum_{j=1}^{\infty} 2q_j \log j + \sum_{q_j \leq \frac{1}{j^2}} \frac{1}{j^2} \left(j^2 q_j \log \frac{1}{q_j j^2} \right).
 \end{aligned}$$

Since the function $x \log \frac{1}{x}$ is bounded above by $\frac{1}{e}$ and $\sum_{j=1}^{\infty} \frac{1}{j^2} = \frac{\pi^2}{6}$, we get

$$(18) \quad h(\mu_E) < \sum_{j=1}^{\infty} q_j (2 \log j + \log n_j) + 1.$$

We now turn our attention to $\int \varphi_E d\mu_E$. Recall that for $x \in Q_j$ we have $\varphi_E(x) = \sum_{i=0}^{j-1} \varphi^+(\sigma^i x)$. We know that $x_1 \dots x_j$ is admissible in X_0^+ , but $x_1 \dots x_{j+1}$ is not admissible in X_0^+ since $x_j x_{j-1} \notin \mathcal{L}(X_0^+)$. Therefore, $\text{dist}(x, X_0^+) = 2^{-1}$ and $\text{dist}(\sigma^i p(x), X_0^+) \geq 2^{-j+i}$ for $0 < i < j$. Hence $\sum_{i=0}^{j-1} \varphi(\sigma^i x) \leq -a_1 - a_{j-1} - \dots - a_2 = -\sum_{i=1}^j a_i$. This gives

$$(19) \quad \int \varphi_E d\mu_E = \sum_{j=1}^{\infty} \int_{Q_j} \varphi_E d\mu_E \leq -\sum_{j=1}^{\infty} \left(q_j \sum_{i=1}^j a_i \right).$$

Combining the last inequality with the entropy estimate (18), and using that $\sum_{j=1}^{\infty} q_j = 1$, we obtain

$$\begin{aligned}
 h(\mu_E) + \int \varphi_E d\mu_E &< \sum_{j=1}^{\infty} q_j \left(2 \log j + \log n_j + 1 - \sum_{i=1}^j a_i \right) \\
 &= \sum_{j=1}^{\infty} j q_j \left(\frac{2 \log j}{j} + \frac{\log n_j}{j} + \frac{1}{j} - \frac{1}{j} \sum_{i=1}^j a_i \right).
 \end{aligned}$$

Since n_j is the number of words of length j in X_0^+ , it follows from Perron-Frobenius theorem that $h_{\text{top}}(X_0^+) \leq \frac{\log n_j}{j} \leq h_{\text{top}}(X_0^+) + \frac{c}{j}$ for some constant $c > 0$. Hence, taking $a_i = \frac{2(c+2)}{i}$ and using that $\sum_{i=1}^j \frac{1}{i} > \frac{1}{2} + \log j$, we

obtain

$$\begin{aligned}
& h(\mu_E) + \int \varphi_E d\mu_E \\
& < \sum_{j=1}^{\infty} jq_j \left(\frac{2 \log j}{j} + h_{\text{top}}(X_0^+) + \frac{c+1}{j} - \frac{2(c+2)}{j} \left(\frac{1}{2} + \log j \right) \right) \\
& = \sum_{j=1}^{\infty} jq_j \left(h_{\text{top}}(X_0^+) - \frac{2(c+1) \log j + 1}{j} \right) \\
& < \sum_{j=1}^{\infty} jq_j h_{\text{top}}(X_0^+) = \frac{h_{\text{top}}(X_0^+)}{\mu(E)},
\end{aligned}$$

where equality comes from Kac's Theorem. \square

Remark 5.2. These methods do not carry over directly to more general one-sided shifts. The reason is that in the case where X_0^+ is a one-sided shift of finite type, if $x \in X^+ \setminus X_0^+$, this is a result of localized defects where $x_j x_{j+1} \notin \mathcal{L}(X_0^+)$ for some j . In particular, if $\text{dist}(x, X_0^+) = 2^{-n}$ for some $n \geq 1$, then $\text{dist}(\sigma x, X_0^+) = 2^{-(n-1)}$. However, for a general one-sided subshift $\text{dist}(x, X_0^+) = 2^{-n}$ implies $x_0 \cdots x_{n-1} \notin \mathcal{L}(X_0^+)$, but this does not imply that $x_1 \cdots x_{n-1} \notin \mathcal{L}(X_0^+)$.

For a general subshift X_0^+ of a one-sided full shift X^+ we have the following result.

Theorem 5.3 (One-sided subshifts). *Suppose $X^+ = \mathcal{A}^{\mathbb{N}_0}$ is a one-sided full shift and X_0^+ is a proper subshift of X^+ . Then the potential $\varphi^+ : X^+ \rightarrow \mathbb{R}$ defined by (6) has a freezing phase transition with $\text{ZTA}(\varphi)$ consisting of measures of maximal entropy of X_0^+ provided the decreasing to zero sequence (a_j) satisfies*

$$a_{3j} \geq \frac{2 \log j}{j} + \frac{1}{j} \sum_{i=0}^{\lfloor \log_2 j \rfloor} 2^i \kappa_i + \frac{1}{j},$$

where $\kappa_i = \frac{\log |\mathcal{L}_{2^i}^{(X_0^+)}|}{2^i} - h_{\text{top}}(X_0^+)$.

Remark 5.4. We note that the sequence which bounds a_{3j} from below tends to zero as $j \rightarrow \infty$. To verify this we use the fact that κ_i monotonically decreases to zero and make a crude estimate on the summation above replacing κ_i by $\log |\mathcal{A}|$ in the first half of the terms and by $\kappa_{\lfloor \log_2 \sqrt{j} \rfloor}$ in the second half of the terms. We obtain

$$\begin{aligned}
\frac{1}{j} \sum_{i=0}^{\lfloor \log_2 j \rfloor} 2^i \kappa_i & \leq \frac{1}{j} \left(2^{\log_2 \sqrt{j}} \log |\mathcal{A}| + \kappa_{\lfloor \log_2 \sqrt{j} \rfloor} 2^{1+\log_2 j} \right) \\
& \leq \frac{\log |\mathcal{A}|}{\sqrt{j}} + 2 \left(\frac{\log |\mathcal{L}_{\lfloor \sqrt{j} \rfloor}^{(X_0^+)}|}{\lfloor \sqrt{j} \rfloor} - h_{\text{top}}(X_0^+) \right) \xrightarrow{j \rightarrow \infty} 0.
\end{aligned}$$

Proof. Consider $X = \mathcal{A}^{\mathbb{Z}}$ and let $X_0 \subset X$ be the two-sided subshift generated by the language of X_0^+ . We define the natural projection map

$$(20) \quad p : X \rightarrow X^+; \quad p((x_i)_{i \in \mathbb{Z}}) = (x_i)_{i \in \mathbb{N}}.$$

Then $p(X_0) = X_0^+$. Any shift-invariant measure μ^+ on X^+ can be written as $\mu^+ = p_*\mu$, where μ is a shift-invariant probability measure on X . In this case we have $h(\mu) = h(\mu^+)$. For a given continuous one-sided potential $\varphi^+ : X^+ \rightarrow \mathbb{R}$, the corresponding two-sided potential $\varphi : X \rightarrow \mathbb{R}$ is given by $\varphi = \varphi^+ \circ p$. Clearly, $\int \varphi^+ d\mu^+ = \int \varphi d\mu$, and hence the pressure functions of φ and φ^+ coincide. We show that $h(\mu) + \int \varphi d\mu < 0$ for any ergodic measure μ which is not supported on X_0 .

We consider the product space $X \times Z$ with map $\bar{\sigma}(x, z) = (\sigma x, \sigma z)$ where $Z = \{0, 1\}^{\mathbb{Z}}$. Let \bar{X} be the subshift of $X \times Z$ consisting of pairs (x, z) with the following properties:

- If $i < j$, and $z_k = 0$ for $i < k \leq j$ (possibly with $z_i = 1$) then $x_i \cdots x_j$ is admissible in X_0 ;
- If $i < j$, $z_i = 1$, $z_j = 1$ then $j - i = 2^k$ for some $k \in \mathbb{N}_0$ and $x_i \cdots x_{2j-i-1}$ is not admissible in X_0 .

Hence, we partition each $x \in X$ into maximal dyadic blocks which are admissible in X_0 . We note that the map $\pi : \bar{X} \rightarrow X$ defined by $\pi(x, z) = x$ is a countable-to-one factor map: for any x and for any n such that $z_n = 1$, $(z_m)_{m \geq n}$ is uniquely determined by the condition that $(x, z) \in \bar{X}$. As a consequence, $h(\bar{\mu}) = h(\mu)$.

We refer to the space \bar{X} as the pin-sequence space and to the 1's as pins. We say that a particular pin is a superpin if the dyadic interval which directly follows the pin is not larger than the preceding one. We induce on the set of points (x, z) such that z_0 is a superpin, i.e.

$$E = \{(x, z) \in \bar{X} : z_0 = 1 \text{ and } \min\{i > 0 : z_i = 1\} \leq \min\{i > 0 : z_{-i} = 1\}\}.$$

As in the proof of Theorem 5.1, we obtain the induced system $(E, \bar{\mu}_E, \bar{\sigma}_E)$ and the induced potential $\bar{\varphi}_E$. We denote the first return time partition of E by $\mathcal{Q} = \{Q_j\}$ where $Q_j = \{(x, z) \in E : r(x, z) = j\}$. Let $q_j = \bar{\mu}_E(Q_j)$ as before.

We estimate $\int \bar{\varphi}_E d\bar{\mu}_E$ first. Suppose $(x, z) \in Q_j$. Since the superpin is at z_j , we know that the admissible dyadic block which starts at position $j + 1$ cannot be longer than the previous dyadic block, and, in particular, cannot be longer than j . Hence, the distance to the next pin is at most j . Recall that the pins partition x into maximal dyadic blocks admissible in X_0 , so the block of length twice the distance to this pin which starts at $j + 1$ cannot be admissible. We conclude that $x_j \cdots x_{3j} \notin \mathcal{L}(X_0)$. Therefore, for $i = 0, \dots, j - 1$ we have $\text{dist}(p(\sigma^i x), X_0^+) \geq \frac{1}{2^{3j}}$, where p is the projection map given by (20). We see that $\varphi(\sigma^i x) \leq -a_{3j}$ for $i = 0, \dots, j - 1$, whence

$\bar{\varphi}_E(x, z) = \sum_{i=0}^{j-1} \varphi(\sigma^i)(x) \leq -ja_{3j}$. We obtain

$$(21) \quad \int \bar{\varphi}_E d\bar{\mu}_E \leq - \sum_{j=1}^{\infty} ja_{3j}q_j.$$

We estimate $h(\bar{\mu})$ by the entropy of the partition \mathcal{P} which is a subpartition of \mathcal{Q} , in which each Q_j is partitioned into length j cylinder sets. We write $H(\mathcal{P}) \leq H(\mathcal{Q}) + H(\mathcal{P}|\mathcal{Q})$ and note that $H(\mathcal{Q}) = \sum_{j=1}^{\infty} -q_j \log q_j \leq \sum_{j=1}^{\infty} 2q_j \log j + 1$ as was shown in (17). It remains to estimate $H(\mathcal{P}|\mathcal{Q})$.

Since the pins between two adjacent superpins at z_0 and z_j are separated by gaps that are powers of 2 of increasing order, they are uniquely determined by the binary expansion of j . Suppose there are $m(j) - 1$ pins between z_0 and z_j with gaps of length $2^{i_1}, \dots, 2^{i_{m(j)-1}}$ so that $j = 2^{i_1} + \dots + 2^{i_{m(j)-1}}$. Clearly, $m(j) - 1 \leq \lfloor \log_2 j \rfloor$. Then the number of different cylinder sets of Q_j is at most $n_{i_1} \times \dots \times n_{i_{m(j)-1}}$ where $n_i = |\mathcal{L}_{2^i}(X_0)|$. Denote $\kappa_i = \frac{\log n_i}{2^i} - h_{\text{top}}(X_0)$. Then,

$$\begin{aligned} h(\bar{\mu}_E) &\leq \sum_{j=1}^{\infty} 2q_j \log j + 1 + \sum_{j=1}^{\infty} q_j \log(n_{i_1} \times \dots \times n_{i_{m(j)-1}}) \\ &\leq \sum_{j=1}^{\infty} q_j \left(2 \log j + \sum_{k=1}^{m(j)} \log n_{i_k} + 1 \right) \\ &\leq \sum_{j=1}^{\infty} q_j \left(2 \log j + \sum_{k=1}^{m(j)} 2^{i_k} \kappa_{i_k} + j h_{\text{top}}(X_0) + 1 \right) \\ &\leq \sum_{j=1}^{\infty} jq_j \left(\frac{2 \log j}{j} + \frac{1}{j} \sum_{i=0}^{\lfloor \log_2 j \rfloor + 1} 2^i \kappa_i + \frac{1}{j} + h_{\text{top}}(X_0) \right). \end{aligned}$$

Combining the last inequality with (21) we obtain

$$h(\bar{\mu}_E) + \int \bar{\varphi}_E d\bar{\mu}_E \leq \sum_{j=1}^{\infty} jq_j \left(\frac{2 \log j}{j} + \frac{1}{j} \sum_{i=0}^{\lfloor \log_2 j \rfloor} 2^i \kappa_i + \frac{1}{j} + h_{\text{top}}(X_0) - a_{3j} \right).$$

Taking (a_j) to be a monotone sequence such that

$$a_{3j} \geq \frac{2 \log j}{j} + \frac{1}{j} \sum_{i=0}^{\lfloor \log_2 j \rfloor} 2^i \kappa_i + \frac{1}{j}$$

gives $h(\bar{\mu}_E) + \int \bar{\varphi}_E d\bar{\mu}_E \leq h_{\text{top}}(X_0) \sum_{j=1}^{\infty} jq_j = \frac{h_{\text{top}}(X_0)}{\bar{\mu}_E(E)}$, so that $\int \varphi d\mu + h(\mu) \leq h_{\text{top}}(X_0)$. \square

Corollary 5.5. *Suppose a subshift $X_0^+ \subset \mathcal{A}^{\mathbb{N}}$ has a word complexity function that grows at most polynomially and is the smallest subshift which supports*

all its invariant measures. Then the potential φ defined by (6) freezes on X_0^+ provided (a_j) is a decreasing sequence satisfying $a_j \geq \frac{\log^2 j}{j}$, $j \in \mathbb{N}$.

Proof. Suppose $|\mathcal{L}_j(X_0)| \leq p(j)$ for all j where $p(j)$ is a polynomial of degree $q \in \mathbb{N}$. Then $h(X_0^+) = 0$ and $2^i \kappa_i \leq cq \log i$ for some constant $c > 0$. We then estimate

$$\frac{1}{j} \sum_{i=0}^{\lfloor \log_2 j \rfloor + 1} 2^i \kappa_i \leq \frac{cq \log j (\lfloor \log_2 j \rfloor + 2)}{j} \leq \frac{2cq \log^2 j}{j}.$$

□

Bruin and Leplaideur [5] consider the Thue-Morse one-sided subshift X_0^+ and $\varphi(x) = -\frac{1}{j^a} + o(\frac{1}{j^a})$ if $\text{dist}(x, X_0^+) = 2^{-j}$. (For this subshift $\mathcal{A} = \{0, 1\}$ and the generating substitution is given by the rules $0 \rightarrow 01, 1 \rightarrow 10$.) They assert that φ exhibits a freezing phase transition for $a \in (0, 1)$ but undergoes no phase transition for $a > 1$. They further posed the question of whether φ undergoes a freezing phase transition at $a = 1$. The proofs presented in [5] were incomplete but were later corrected in [4]. Additionally, the question of freezing for $\varphi(x) = -\frac{1}{j}$ was affirmatively resolved. We observe that the Thue-Morse subshift has linear complexity. Consequently, Corollary 5.5 ensures freezing when $\varphi(x) \approx -\frac{\log^2 j}{j}$, provided that $\text{dist}(x, X_0) = 2^{-j}$. It also follows from Theorem 4.1 that freezing phase transition does not occur for $\varphi \approx -\frac{1}{n^a}$, $a > 1$.

6. PROOFS OF SOME FOLKLORE RESULTS

For the reader's convenience, we provide proofs of the following facts mentioned in the introduction and Section 2, as we could not find a readily available reference.

Remark 6.1. It can be readily checked that Theorem 6.2 and Proposition 6.3 below are valid for any general expansive continuous \mathbb{Z}^d -action on a compact metric space.

6.1. Properties of accumulation points of $(\mu_{\beta_n \varphi})_n$ as $\beta_n \rightarrow \infty$. Let $\varphi : \mathcal{A}^{\mathbb{Z}^d} \rightarrow \mathbb{R}$ be a continuous function. Recall that a shift-invariant measure μ is maximizing for φ if $\int \varphi d\mu \geq \int \varphi d\eta$ for all shift-invariant measures η . By weak*-compactness of the set of shift invariant measures and continuity of the mapping $\eta \mapsto \int \varphi d\eta$, the maximum is attained, so that $\text{MM}(\varphi)$, the set of maximizing measures, is non-empty. Let

$$s_\varphi = \sup_\nu \int \varphi d\nu \quad \text{and} \quad h_\varphi = \sup \{ h(\eta) : \eta \in \text{MM}(\varphi) \}.$$

In statistical physics, h_φ , known as the residual entropy, can also be determined as the exponential growth rate of the number of ground configurations in increasingly large boxes. [1].

Recall also that the entropy function $\eta \mapsto h(\eta)$, where η runs over the set of shift-invariant measures, is bounded between 0 and $\log |\mathcal{A}|$, and is upper semi-continuous [33, Theorem 4.5.6].

Using the variational principle, we define the pressure function $p_\varphi : \mathbb{R}_+ \rightarrow \mathbb{R}$ as

$$\beta \mapsto p_\varphi(\beta) := \sup_\nu \left\{ h(\nu) + \beta \int \varphi d\nu \right\},$$

where the supremum is taken over all shift-invariant measures. Since the functional $\nu \mapsto h(\nu) + \beta \int \varphi d\nu$ is upper semi-continuous, the supremum is attained again. The measures for which the supremum is attained are the equilibrium states for $\beta\varphi$. As a convex function, p_φ has left and right derivatives $p'_{\varphi,-}(\beta)$ and $p'_{\varphi,+}(\beta)$ at each $\beta > 0$, and

$$(22) \quad p'_{\varphi,-}(\beta) \leq \int \varphi d\mu_{\beta\varphi} \leq p'_{\varphi,+}(\beta) \text{ for any } \mu_{\beta\varphi} \in \text{ES}(\beta\varphi).$$

(When $p'_{\varphi,-}(\beta) = p'_{\varphi,+}(\beta)$, that is, when the pressure function is differentiable at β , we thus have $p'_\varphi(\beta) = \int \varphi d\mu_{\beta\varphi}$.) These facts follow from a combination of [27, Chapter 16] and [46, Theorems 3.4 and 3.12].

Theorem 6.2. *Let $\varphi : \mathcal{A}^{\mathbb{Z}^d} \rightarrow \mathbb{R}$ be a continuous function. We have the following facts:*

- (i) $\text{ZTA}(\varphi) \subseteq \text{MM}(\varphi)$, and if $\mu \in \text{ZTA}(\varphi)$ then $h(\mu) = \sup\{h(\nu) : \nu \in \text{MM}(\varphi)\}$.
- (ii) *The pressure function has a slant asymptote. More precisely,*

$$\lim_{\beta \rightarrow +\infty} (p_\varphi(\beta) - (s_\varphi\beta + h_\varphi)) = 0.$$

Proof. Firstly, if $\eta \in \text{MM}(\varphi)$, the variational principle yields the bounds

$$(23) \quad h(\eta) + \beta s_\varphi \leq p(\beta\varphi) \leq \log |\mathcal{A}| + \beta s_\varphi.$$

Now let $\mu^* \in \text{ZTA}(\varphi)$ and let $\beta_n \rightarrow \infty$ and (μ_n) be such that μ_n is an equilibrium state for $\beta_n\varphi$ and $\mu_n \rightsquigarrow \mu^*$. By (23), $\int \beta_n\varphi d\mu_n + h(\mu_n) \geq \beta_n s_\varphi + h(\eta)$ for all n . Dividing by β_n and taking the limit, we obtain $\int \varphi d\mu^* \geq s_\varphi$, so that $\mu^* \in \text{MM}(\varphi)$. That is, we have shown $\text{ZTA}(\varphi) \subseteq \text{MM}(\varphi)$.

Now we prove that, if $\mu^* \in \text{ZTA}(\varphi)$ then it has the maximal entropy over the set of maximizing measures for φ . Along the way, we will also prove that $\lim_{\beta \rightarrow +\infty} (p_\varphi(\beta) - s_\varphi\beta) = h_\varphi$, which will complete the proof of statement (ii). Define the auxiliary function $g_\varphi(\beta) := p_\varphi(\beta) - s_\varphi\beta$ (for $\beta > 0$). From (23), we see

$$(24) \quad g_\varphi(\beta) \geq h(\eta), \quad \forall \beta > 0.$$

We can compute the left derivative of g_φ to show that it is a decreasing function. Indeed,

$$g'_{\varphi,-}(\beta) = p'_{\varphi,-} - s_\varphi \leq \int \varphi d\mu_{\beta\varphi} - s_\varphi \leq 0.$$

Thus, the limit of $g_\varphi(\beta)$ as $\beta \rightarrow \infty$ exists, so using (24), we see

$$(25) \quad \lim_{\beta \rightarrow \infty} g_\varphi(\beta) \geq h(\eta).$$

Taking a supremum of (25) over $\eta \in \text{MM}(\varphi)$, we see

$$(26) \quad \lim_{\beta \rightarrow \infty} g_\varphi(\beta) \geq \sup\{h(\eta) : \eta \in \text{MM}(\varphi)\}.$$

Let $\alpha > 0$ and let $\mu_{\alpha\varphi}$ be any equilibrium state for $\alpha\varphi$. We see from the variational principle and the definition of s_φ that

$$g_\varphi(\alpha) = h(\mu_{\alpha\varphi}) + \alpha \left(\int \varphi d\mu_{\alpha\varphi} - s_\varphi \right) \leq h(\mu_{\alpha\varphi}),$$

so that since g_φ is decreasing,

$$h(\mu_{\alpha\varphi}) \geq \lim_{\beta \rightarrow \infty} g_\varphi(\beta) \text{ for all } \alpha.$$

Let $\mu^* \in \text{ZTA}(\varphi)$ and let $\beta_n \rightarrow \infty$ and (μ_n) be a sequence of equilibrium states for $\beta_n\varphi$ such that $\mu_n \rightsquigarrow \mu^*$. By upper semi-continuity of the entropy function we deduce that

$$(27) \quad h(\mu^*) \geq \limsup_n h(\mu_{\beta_n\varphi}) \geq \lim_{\beta \rightarrow \infty} g_\varphi(\beta)$$

Since we have shown $\mu^* \in \text{MM}(\varphi)$, combining (26) with (27), we see

$$\lim_{\beta \rightarrow \infty} g_\varphi(\beta) \geq \sup\{h(\eta) : \eta \in \text{MM}(\varphi)\} \geq h(\mu^*) \geq \lim_{\beta \rightarrow \infty} g_\varphi(\beta).$$

This finishes the proof of statement (ii) and the second part of statement (i). \square

6.2. When the pressure function is affine on a segment. The following proposition provides insight into the affine parts of the pressure function graph.

Proposition 6.3. *Let $\varphi : \mathcal{A}^{\mathbb{Z}^d} \rightarrow \mathbb{R}$ be a continuous function. Then, for any $\beta_1 < \beta_2$, the following assertions are equivalent:*

- (1) $\text{ES}(\beta_1\varphi) \cap \text{ES}(\beta_2\varphi) \neq \emptyset$;
- (2) p_φ is affine on $[\beta_1, \beta_2]$;
- (3) $\text{ES}(\beta\varphi) = \text{ES}(\beta_1\varphi) \cap \text{ES}(\beta_2\varphi)$ for all $\beta \in (\beta_1, \beta_2)$.

Proof. Suppose (1) holds and let $\mu \in \text{ES}(\beta_1\varphi) \cap \text{ES}(\beta_2\varphi)$ and $s \in (0, 1)$ so that $(1-s)\beta_1 + s\beta_2 \in (\beta_1, \beta_2)$. Then since p_φ is convex, $p_\varphi((1-s)\beta_1 + s\beta_2) \leq (1-s)p_\varphi(\beta_1) + sp_\varphi(\beta_2)$. On the other hand, $p_\varphi((1-s)\beta_1 + s\beta_2) \geq h(\mu) + ((1-s)\beta_1 + s\beta_2) \int \varphi d\mu = (1-s)p_\varphi(\beta_1) + sp_\varphi(\beta_2)$, so that p_φ is affine on $[\beta_1, \beta_2]$ as claimed and $\mu \in \text{ES}(((1-s)\beta_1 + s\beta_2)\varphi)$.

Now suppose that p_φ is affine on $[\beta_1, \beta_2]$ and let $\beta = (1-s)\beta_1 + s\beta_2$ with $0 < s < 1$ and $\mu \in \text{ES}(\beta\varphi)$. Then $p'_\varphi(\beta) = \int \varphi d\mu$, so that $h(\mu) + t \int \varphi d\mu = h(\mu) + \beta \int \varphi d\mu + (t-\beta) \int \varphi d\mu = p_\varphi(\beta) + (t-\beta)p'_\varphi(\beta) = p_\varphi(t)$ for all $t \in [\beta_1, \beta_2]$. It follows that $\mu \in \text{ES}(\beta_1\varphi) \cap \text{ES}(\beta_2\varphi)$. Conversely, if $\mu \in \text{ES}(\beta_1\varphi) \cap \text{ES}(\beta_2\varphi)$ the first part implies that $\mu \in \text{ES}(\beta\varphi)$.

The implication (3) implies (1) is trivial. \square

6.3. Proof of Proposition 2.3. By Proposition 6.3, p_φ is affine on $[\beta_1, \beta_2]$ if and only if $\text{ES}(\beta\varphi)$ is equal for all $\beta \in (\beta_1, \beta_2)$. If φ freezes at β_c , by continuity of p_φ , we see p_φ is affine on $[\beta_c, \infty)$ but no larger interval.

For the converse, using Proposition 6.3, if p_φ is affine on $[\beta_c, \infty)$, then $\text{ES}(\beta\varphi) = \text{ES}(\beta'\varphi)$ for all $\beta, \beta' > \beta_c$. Meanwhile, if $\beta < \beta_c < \beta'$, since p_φ is not affine on $[\beta, \beta']$, we see $\text{ES}(\beta\varphi) \cap \text{ES}(\beta'\varphi) = \emptyset$. Hence φ has a freezing phase transition at β_c .

By Theorem 6.2, $p_\varphi(\beta)$ approaches the slant asymptote $s_\varphi\beta + h_\varphi$ as $\beta \rightarrow \infty$, so if p_φ is affine on $[\beta_c, \infty)$, in fact $p_\varphi(\beta) = s_\varphi\beta + h_\varphi$ on that range.

6.4. Gibbs states have full support. For completeness, and since we were unable to find a suitable reference, we provide a proof of the following well-known result invoked in Section 4.2.

We begin by quickly recalling the definition of a Gibbs state associated with a given interaction $\Phi \in \mathcal{S}$ (see (14)). For a comprehensive treatment, we refer to [27].

Let \mathfrak{B} denote the Borel σ -algebra on X , which coincides with the Borel σ -algebra generated by cylinder sets. Given a subset $\Lambda \subset \mathbb{Z}^d$, we define \mathfrak{B}_Λ as the σ -algebra generated by the coordinate maps $x \mapsto x_i$ for $i \in \Lambda$.

The Gibbsian specification for Φ is given by

$$\gamma_\Lambda^\Phi(x|y) := \frac{e^{U_\Lambda^\Phi(x|y)}}{Z_\Lambda^\Phi(y)}$$

where

$$U_\Lambda^\Phi(x|y) := \sum_{\substack{\Lambda' \in \mathbb{Z}^d \\ \Lambda' \cap \Lambda \neq \emptyset}} \Phi_{\Lambda'}(x|_\Lambda y|_{\Lambda^c}),$$

and $Z_\Lambda^\Phi(y)$ is the normalization factor, $y|_{\Lambda^c}$ represents the boundary condition, and $\Lambda \Subset \mathbb{Z}^d$. (For each y , $(\gamma_\Lambda^\Phi(\cdot|y))_{\Lambda \Subset \mathbb{Z}^d}$ is a family of probability kernels on (X, \mathfrak{B}) subject to natural conditions we don't list here.)

A probability measure μ on (X, \mathfrak{B}) is said to be a Gibbs state for Φ if, for every finite subset $\Lambda \Subset \mathbb{Z}^d$, the conditional probability $\gamma_\Lambda^\Phi(x|\cdot)$ provides a version of

$$\mu(x|_\Lambda | \mathfrak{B}_{\Lambda^c}).$$

Equivalently, this means that for all $B \in \mathfrak{B}$ and all finite $\Lambda \Subset \mathbb{Z}^d$, we have

$$(28) \quad \mu(B) = \int \sum_{x \in \mathcal{A}^\Lambda} \gamma_\Lambda^\Phi(x|y) \mathbf{1}_B(x|_\Lambda y|_{\Lambda^c}) d\mu(y).$$

(This is called the DLR equation – after Dobrushin, Lanford and Ruelle.) It is well known that at least one Gibbs state for Φ always exists.

Finally, for $x \in X$ and $n \geq 1$, we define the cylinder set centered at x with support on Λ_n as

$$C_n(x) := \{y \in X : y|_{\Lambda_n} = x|_{\Lambda_n}\}.$$

Proposition 6.4. *Let $\Phi \in \mathcal{S}$ and μ be a Gibbs state for Φ . Then there exists $\rho = \rho(\Phi, \mathcal{A}) \in (0, 1)$ such that, for all $x \in X, n \geq 1$,*

$$(29) \quad \mu(C_n(x)) \geq \rho^{|\Lambda_n|}.$$

Proof. Since $\Phi \in \mathcal{S}$, we have the bound

$$|U_\Lambda^\Phi(x|y)| \leq |\Lambda| \|\Phi\|_{\mathcal{S}},$$

which is true for all $x, y \in X$, hence

$$e^{-|\Lambda| \|\Phi\|_{\mathcal{S}}} \leq e^{U_\Lambda^\Phi(x|y)} \leq e^{|\Lambda| \|\Phi\|_{\mathcal{S}}} \quad \text{and} \quad Z_\Lambda^\Phi(y) \leq |\mathcal{A}|^{|\Lambda|} e^{|\Lambda| \|\Phi\|_{\mathcal{S}}}.$$

Therefore

$$\gamma_\Lambda^\Phi(x|y) \geq |\mathcal{A}|^{-|\Lambda|} e^{-2|\Lambda| \|\Phi\|_{\mathcal{S}}},$$

which by (28) yields

$$\mu(C_n(x)) = \int \gamma_{\Lambda_n}^\Phi(x|y) d\mu(y) \geq |\mathcal{A}|^{-|\Lambda_n|} e^{-2|\Lambda_n| \|\Phi\|_{\mathcal{S}}}$$

since $\mu(X) = 1$. Letting $\rho := |\mathcal{A}|^{-1} e^{-2\|\Phi\|_{\mathcal{S}}}$, this concludes the proof. \square

REFERENCES

- [1] M. Aizenman, E.H. Lieb, *The third law of thermodynamics and the degeneracy of the ground state for lattice systems*, J. Stat. Phys. 24 (1981), 279–297.
- [2] J. Aaronson and B. Weiss, *A \mathbb{Z}^d ergodic theorem with large normalising constants*, in Convergence in ergodic theory and probability (Columbus, OH, 1993), de Gruyter, 1–13.
- [3] S. Barbieri, R. Bissacot, G. Dalle Vedove and Ph. Thieullen, *Zero-temperature chaos in bidimensional models with finite-range potentials*, Adv. in Math. 457 (2024), 109906, 51 pages.
- [4] N. Bédaride, J. Cassaigne, P. Hubert and R. Leplaideur, *Thermodynamic formalism and the Thue Morse substitution: Phase transition*. Preprint (2025).
- [5] H. Bruin and R. Leplaideur, *Renormalization, thermodynamic formalism and quasicrystals in subshifts*, Commun. Math. Phys. 321 (2013), 209–247.
- [6] H. Bruin and R. Leplaideur, *Renormalization, freezing phase transitions and Fibonacci quasicrystals*. Ann. Sc. École Normale Supérieure 48 (2015), no. 3, 739–763.
- [7] J. Buzzi, B. Kloeckner and R. Leplaideur, *Nonlinear thermodynamical formalism*. Annales Henri Lebesgue 6 (2023), 1429–1477.
- [8] T. Bousch, *La condition de Walters*. Ann. Sci. École Norm. Sup. 34 (2001), 287–311.
- [9] R. Burton, J. Steif, *Non-uniqueness of measures of maximal entropy for subshifts of finite type*, Ergodic Theory and Dynamical Systems 14 (1994), 213–235.
- [10] J.-R. Chazottes, J.-M. Gambaudo and E. Ugalde, *Zero-temperature limit of one-dimensional Gibbs states via renormalization: the case of locally constant potentials*. Ergodic Theory and Dynamical Systems 31 (2011), 1109–1161.
- [11] J.-R. Chazottes, M. Hochman, *On the zero-temperature of Gibbs states*, Commun. Math. Phys. 297 (2010), no. 1, 265–281.
- [12] J.-R. Chazottes, G. Keller, *Pressure and Equilibrium States in Ergodic Theory*, pp 6939–6955. In: R. Meyers (eds), Encyclopedia of Complexity and Systems Science. Springer, 2009.
- [13] J.-R. Chazottes, M. Shinoda, *On the absence of zero-temperature limit of equilibrium states for finite-range interactions on the lattice \mathbb{Z}^2* , preprint (2022).

- [14] N. Chernov, *Invariant measures for hyperbolic dynamical systems*, pp 321-407. In: B. Hasselblatt, A. Katok (eds), *Handbook of Dynamical Systems*, Elsevier Science, Volume 1, Part A, 2002.
- [15] D. Coronel, J. Rivera-Letelier, *Sensitive dependence of Gibbs measures at low temperatures*, J. of Stat. Phys. 160 (2015), 1658-1683.
- [16] B. Durand, G. Gamard, A. Grandjean, *Aperiodic tilings and entropy*, Theoretical Computer Science 666 (2017), 36-47.
- [17] A. van Enter, *Aperiodicity in equilibrium systems: between order and disorder*. Proceedings of the 12th International Conference on Quasicrystals (ICQ12). Acta Physica Polonica A 126 (2), 621-624.
- [18] A. C. D. van Enter, R. Fernández, and A. D. Sokal, *Regularity properties and pathologies of position-space renormalization-group transformations: Scope and limitations of Gibbsian theory*, J. Stat. Phys. 72 (1994), 879-1167.
- [19] A. van Enter, H. Koivusalo and J. Miękisz, *Sturmian ground states in classical lattice-gas models*. J. Stat. Phys. 178 (2020), 832-844.
- [20] A. van Enter and J. Miękisz, *How should one define a (weak) crystal?*, J. of Stat. Phys. 66 (1992), 1147-1153.
- [21] A. C. D. van Enter, J. Miękisz and M. Zahradník, *Nonperiodic long-range order for fast-decaying interactions at positive temperatures*, J. Stat. Phys. 90 (1998), 1441-1447.
- [22] M. E. Fisher and B. U. Felderhof, *Phase transitions in one-dimensional cluster-interaction fluids*. IA, IB, II, III, Ann. Physics 58 (1970), 176-216, 217-267, 268-280, 281-300.
- [23] S. Friedli and Y. Velenik. *Statistical Mechanics of Lattice Systems: A Concrete Mathematical Introduction* Cambridge University Press, 2017.
- [24] C. Gardner, J. Miękisz, C. Radin and A. van Enter, *Fractal symmetry in an Ising model*, J. Phys. A: Math. Gen. 22 (1989) L1019-L1023.
- [25] E. Garibaldi and P. Thieullen, *An Ergodic Description of Ground States*. J. Stat. Phys. 158 (2015), 359-371.
- [26] L. Gayral, M. Sablik, S. Taati, *Characterisation of the set of ground states of uniformly chaotic finite-range lattice models*, preprint (2023).
- [27] H.-O. Georgii. *Gibbs Measures and Phase Transitions*, Berlin, New York: De Gruyter, 2011.
- [28] E. Glasner. *Ergodic Theory via joinings*, Math. Surveys and Monographs, AMS, 101, 2003.
- [29] C. E. Hedges, *On the equivalence of equilibrium and freezing states in dynamical systems*, preprint arXiv (2024).
- [30] F. Hofbauer, *Examples for the nonuniqueness of the equilibrium state*, Trans. Amer. Math. Soc. 228 (1977), 223-241.
- [31] R. B. Israel. *Convexity in the Theory of Lattice Gases*. Princeton Series in Physics. Princeton Legacy Library Princeton University Press, 2015.
- [32] O. Jenkinson, *Ergodic optimization in dynamical systems*, Ergod. Th. and Dynam. Sys. 39 (2019), 2593-2618.
- [33] G. Keller. *Equilibrium states in ergodic theory*, volume 42, London Mathematical Society Student Texts. Cambridge University Press, Cambridge, 1998.
- [34] T. Kucherenko, A. Quas and C. Wolf, *Multiple phase transitions on compact symbolic systems*, Advances in Mathematics 385 (2021), 107768.
- [35] T. Kucherenko and A. Quas, *Flexibility of the pressure function*. Commun. Math. Phys. 395 (2022), 1431-1461.
- [36] T. Kucherenko and D. Thompson, *Measures of maximal entropy for suspension flows over the full shift*, Math. Z. 294 (2020), 769-781.
- [37] T. Kucherenko and D. Thompson, *Measures of maximal entropy on subsystems of topological suspension semi-flows*, Studia Mathematica 260 (2) (2021), 229-240.

- [38] S. Labbé, *Metallic mean Wang tiles I: self-similarity, aperiodicity and minimality*, preprint arXiv (2023).
- [39] S. Labbé, *Metallic mean Wang tiles II: the dynamics of an aperiodic computer chip*, preprint arXiv (2024).
- [40] F. Ledrappier, *Un exemple de transition de phase*, Monatshefte für Mathematik 83 (1977), 147–153.
- [41] D. Lind, K. Schmidt, *Symbolic and algebraic dynamical systems*, Chapter 10 in *Handbook of Dynamical Systems*, B. Hasselblatt, A. Katok (Eds.), Elsevier Science, Vol. 1, Part A, 2002, 765–812.
- [42] A. Lopes, *The Zeta Function, non-differentiability of the pressure, and the critical exponent of transition*, Adv. in Math., **101** (1993), 133–165.
- [43] T. Meyerovitch, B. Weiss, *Kac's Lemma and countable generators for actions of countable groups*, preprint (2024), arXiv:2410.18488.
- [44] E. Olivier, *Multifractal analysis in symbolic dynamics and distribution of pointwise dimension for g -measures*, Nonlinearity 12(6), 1571–1584, 1999.
- [45] E. A. Robinson, *Symbolic dynamics and tilings of \mathbb{R}^d* , in Symbolic dynamics and its applications, Amer. Math. Soc.(2004), 81–119.
- [46] D. Ruelle. Thermodynamic Formalism: The Mathematical Structure of Equilibrium Statistical Mechanics. Cambridge University Press, 2004
- [47] P. Walters, *A necessary and sufficient condition for a two-sided continuous function to be cohomologous to a one-sided continuous function*, Dyn. Syst. **18** (2003), no. 2, 131–138.

CPHT, CNRS, ÉCOLE POLYTECHNIQUE, INSTITUT POLYTECHNIQUE DE PARIS, 91120 PALAISEAU, FRANCE

Email address: `jeanrene.chazottes@cnrs.fr`

DEPARTMENT OF MATHEMATICS, THE CITY COLLEGE OF NEW YORK, NEW YORK, NY, 10031, USA

Email address: `tkucherenko@ccny.cuny.edu`

DEPARTMENT OF MATHEMATICS & STATISTICS, UNIVERSITY OF VICTORIA, VICTORIA BC, CANADA V8W 2Y2

Email address: `aquas@uvic.ca`