Markov Random Fields, Gibbs States and Entropy Minimality

by

Nishant Chandgotia

B.Math., Indian Statistical Institute, 2009
M.Sc. in Mathematics, The University of British Columbia, 2011

A THESIS SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF

Doctor of Philosophy

in

FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES
(Mathematics)

The University of British Columbia
(Vancouver)

April 2015

Abstract

The well-known Hammersley-Clifford Theorem states (under certain conditions) that any Markov random field is a Gibbs state for a nearest neighbour interaction. Following Petersen and Schmidt we utilise the formalism of cocycles for the homoclinic relation and introduce "Markov cocycles", reparametrisations of Markov specifications. We exploit this formalism to deduce the conclusion of the Hammersley-Clifford Theorem for a family of Markov random fields which are outside the theorem's purview (including Markov random fields whose support is the d-dimensional "3-colored chessboard"). On the other extreme, we construct a family of shift-invariant Markov random fields which are not given by any finite range shift-invariant interaction.

The techniques that we use for this problem are further expanded upon to obtain the following results: Given a "four-cycle free" finite undirected graph $\mathcal{H}$ without self-loops, consider the corresponding 'vertex' shift, $Hom(\mathbb{Z}^d, \mathcal{H})$ denoted by $X_{\mathcal{H}}$. We prove that $X_{\mathcal{H}}$ has the pivot property, meaning that for all distinct configurations $x, y \in X_{\mathcal{H}}$ which differ only at finitely many sites there is a sequence of configurations $(x = x^1), x^2, \ldots, (x^n = y) \in X_{\mathcal{H}}$ for which the successive configurations (x^i, x^{i+1}) differ exactly at a single site. Further if $\mathcal{H}$ is connected then we prove that $X_{\mathcal{H}}$ is entropy minimal, meaning that every shift space strictly contained in $X_{\mathcal{H}}$ has strictly smaller entropy. The proofs of these seemingly disparate statements are related by the use of the 'lifts' of the configurations in $X_{\mathcal{H}}$ to their universal cover and the introduction of 'height functions' in this context.

Further we generalise the Hammersley-Clifford theorem with an added condition that the underlying graph is bipartite. Taking inspiration from Brightwell and Winkler we introduce a notion of folding for configuration spaces called strong config-folding to prove that if all Markov random fields supported on X are Gibbs with some nearest neighbour interaction so are Markov random fields supported on the "strong config-folds" and "strong config-unfolds" of X.

Preface

This thesis is a combination of three manuscripts: [13], [10] and [11].

The broad direction of study was suggested by Prof. Brian Marcus. Chapter 2 is based on the manuscript [13] and is joint with Dr. Tom Meyerovitch. Given the nature of this work, it is impossible to separate the individual contributions for this chapter. Chapter 3 is based on the manuscript [11]. Chapter 4 is based on the manuscript [10].

Table of Contents

List of Figures

Acknowledgments

I savour the sweet honey taste,
Yet I do not feel the spring gale,
See the bright flower
Or hear the busy bee.

A piece of work is a product of its milieu but the contribution of the milieu often remains hidden even to the discerning eye. I am not the centre of this work but a mere representative; it is therefore imperative that I acknowledge as far as my vision extends and this restricted space allows, the many faces behind this thesis.

I would like to acknowledge the great ergodic theory group I have had here at UBC. The facilitator of this group Brian is my advisor; I cannot thank him enough for supporting and nurturing me; he brought me up from the ebb of my confidence to being someone aware of his abilities. His gift of simplicity and humility is something I have always cherished and hope to attain with time. I was introduced to symbolic dynamics by Ronnie. Through the course of my studies he has been very helpful giving me questions to think about, suggesting problems, suggesting papers to read and critiquing my presentations. Discussions and collaborations with Tom have been very fruitful; I have learnt and continue to learn immensely from him. Felipe and Raimundo have been great company through the course of my PhD; whether it be travel guidance in Chile and Mexico, being part of the Great Annex Orchestra, deep philosophical discussions or visiting some interesting gun shops, they have always been there with me. The warmth of this group has kept the Vancouverite cold far away; I can't picture how life would be otherwise. I will also like to acknowledge the steady stream of visitors which helped create an intellectually stimulating atmosphere: numerous conversations with Prof. Klaus Schmidt, Prof. Mike Boyle and Prof. Michael Schraudner have helped me broaden my view point significantly. The active probability group here, discussions with Prof. Omer Angel and Prof. David Brydges has also been very helpful. I would also like to thank Prof. Lior Silberman, Prof. Peter Winkler and Prof. Anthony Quas for many useful discussions. Additionally I would like to thank Prof. Brian Marcus, Prof. Andrew Rechnitzer, Prof. Lior Silberman, Prof. Ed Perkins, Prof. Will Evans and Prof. Klaus Schmidt for being part of my examination comittee and giving many useful suggestions. I was funded by the Four-year Fellowship and the International Tuition award at the University of British Columbia.

I thank my parents for the privilege of a happy childhood and giving me opportunities which were not always within their reach. Rinku bhaiya, Bhabhi, Tini and Mini di, Ashu bhaiya, Uncle and Aunty, Harshu have all in their little ways been a constant source of support. Through the course of my thesis I have missed them the most. Akash, Gudia and Sonu have always held my back and supported my crazy dreams; their humour has kept my spirit up all my life.

I am also lucky to have been endowed with a great group of friends here in Vancouver. Vasu and Gourab have been the nearest and dearest. They willingly bore my nonsense all these years and continue to support me in my endeavours. I cannot forget my first few days at UBC; I was low and life seemed very hard. It was Vasu who helped me stay afloat and took me away from the sadness and misery. Terry has been like an elder brother to me; taking me for dinners, drafting my letters and extricating me from mires whenever I was stuck. His loud laughter still rings through the annex. Lalitha has been a very dear friend of mine; among many other things she has brought me closer to my country and culture, closer to myself. She continues to be a patient listener to my qualms and poetry, despite the distance. I hope that our endless conversation remain endless. Ankur, my current roommate has been a blessing as well. A great friend, he puts his life on line for my crazy streaks for hiking and driving each and every time. Our mutual love for good food, music and philosophy has made my stay in Vancouver very pleasant. Nithya, Swetha, Sneha, Pawan and Vignesh formed for me a great group of friends; our poker games, the drinking sessions were a treat despite my sobriety. I can only wish for more of our craziness; it pains me to see that we must now move away looking for another life at another place. Ajith and Monika have been my dear neighbours, excellent hosts, great friends and fine mentors; our chai sessions and discussions of Indian politics have always been very informative, the cakes quite delicious and company impeccable. In a short period of time, I have found a great friend in Rajarshi. He allows me to rumble around without much grumble and only minor voices of protest. The depth of his thinking has always been very inspiring. My previous roommate and officemate Marc has always been a good friend; his zest for mathematics, food and life in general was contagious and has truly caught on to me. I am glad that despite my fallacies Anujit has always been a great friend and a good roommate; I hope someday to make up for my misgivings. Subhajit, my dear junior continues to torment me whenever he gets the opportunity, the boiling blood has always been very entertaining! I would also like to thank the meek-mannered Tatchai for numerous discussions in ergodic theory and being a great friend (and the many dinners and snacks). Lastly I would like to thank Navid, Myrto, Amir and Haider for being around and helping in a multitude of ways.

Lastly I would like to thank the beautiful city of Vancouver. Most of the ideas in this thesis come from its beautiful summer months which were spent climbing its hills and watching the sea roll by.

Dedication

I dedicate this thesis to my parents, to my country, India and to the beautiful world of mathematics.

Chapter 1

Introduction

This thesis can be divided broadly into two areas of study: identifying conditions on the support of Markov random fields such that they are Gibbs for some nearest neighbour interaction and proving that a certain class of shift spaces is entropy minimal. In the introduction we will briefly discuss the results obtained in these areas (highlighted as theorems) during the course of my PhD and their interrelations. Further details and formal definitions can be found in the subsequent chapters.

$\mathcal{G} = (\mathcal{V}, \mathcal{E})$ will always refer to a locally finite countable undirected graph without multiple edges and self-loops, $\mathcal{H}$ will always refer to an undirected graph without multiple edges. By $Hom(\mathcal{G}, \mathcal{H})$ we will denote the space of graph homomorphisms from $\mathcal{G}$ to $\mathcal{H}$. Let $\vec{0}$ denote the origin and $\vec{e}_i$ denote the ith coordinate vector of $\mathbb{Z}^d$.

1.1 Markov Random Fields and Gibbs States

The *boundary* (or the external vertex boundary) of a set of vertices $F \subset \mathcal{V}$, denoted by ∂F, is the set of vertices outside F which are adjacent to F:

$$\partial F := \{v \in \mathcal{V} \setminus F \mid \exists w \in F \text{ s.t. } (v, w) \in \mathcal{E}\}.$$

Given a finite set $\mathcal{A}$, the space $\mathcal{A}^{\mathcal{V}}$ is a compact topological space with respect to the product topology, where the topology on $\mathcal{A}$ is discrete. For $F \subset \mathcal{V}$ finite and $a \in \mathcal{A}^F$, we denote by $[a]_F$ the *cylinder set*

$$[a]_F := \left\{x \in \mathcal{A}^{\mathcal{V}} \mid x|_F = a\right\}.$$

For $x \in \mathcal{A}^{\mathcal{V}}$ we use the notation $[x]_F$ for $[x|_F]_F$. The collection of cylinder sets generates the Borel σ-algebra on $\mathcal{A}^{\mathcal{V}}$.

A *Markov random field* (MRF) is a Borel probability measure μ on $\mathcal{A}^{\mathcal{V}}$ with the property that for all finite $A, B \subset \mathcal{V}$ such that $\partial A \subset B \subset A^c$ (as illustrated in Figure 1.1) and $a \in \mathcal{A}^A, b \in \mathcal{A}^B$

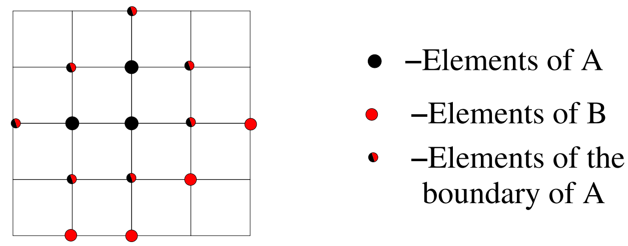

Figure 1.1: A, ∂A and B

satisfying $\mu([b]_B) > 0$

$$\mu\left([a]_A \,\middle|\, [b]_B\right) = \mu\left([a]_A \,\middle|\, [b|_{\partial A}]_{\partial A}\right).$$

Given an MRF μ the space of conditional probabilities or the "specification" is the system of conditional probabilities of the type $\mu([\cdot]_A \mid [x]_{\partial A})$ for all finite sets $A \subset \mathcal{V}$ and $x \in supp(\mu)$. In general infinitely many parameters may be required to describe a specification even if the graph $\mathcal{G}$ is $\mathbb{Z}^d$ and the measure is shift-invariant.

A closed configuration space is a closed subset $X \subset \mathcal{A}^{\mathcal{V}}$. In most cases X will be the (topological) support of some MRF μ denoted by $supp(\mu)$. Let us consider some examples:

1. *r-colourings of $\mathcal{G}$:* Let K_r denote the complete graph (without self-loops) with vertices $1, 2, \ldots r$. Then $X = Hom(\mathcal{G}, K_r)$ denotes the set of all r-colourings of $\mathcal{G}$.

2. *Homomorphisms to the n-cycle:* Let $d \geq 2$ and C_n denote the n-cycle with vertices $0, 1, 2, \ldots, n-1$. Then $X = Hom(\mathbb{Z}^2, C_n)$ will form an important class of closed configuration spaces for this thesis.

If $\mathcal{G} = \mathbb{Z}^2$, $r = 3$ and $n = 2$ then $Hom(\mathcal{G}, K_r)$ and $Hom(\mathcal{G}, C_n)$ is also known as the "3-coloured chessboard". Given a graph $\mathcal{H}$ and $d \in \mathbb{N}$, let $X^d_{\mathcal{H}} = Hom(\mathbb{Z}^d, \mathcal{H})$.

For all $W \subset \mathcal{V}$ let

$$\mathcal{L}_W(X) := \{w \in \mathcal{A}^W \mid \text{there exists } x \in X \text{ such that } x|_W = w\}.$$

The *language* of $X \subset \mathcal{A}^{\mathcal{V}}$ denoted by $\mathcal{L}(X)$ is defined as all finite patterns which occur in the elements of X:

$$\mathcal{L}(X) := \bigcup_{W \subset \mathcal{V} \text{ finite}} \mathcal{L}_W(X).$$

A *finite range interaction* is a function $\phi : \mathcal{L}(X) \longrightarrow \mathbb{R}$ such that for some $r \in \mathbb{N}$, $\phi(a) = 0$ for all $a \in \mathcal{L}_A(X)$ whenever $diam(A) > r$. A *nearest neighbour interaction* on X is a finite-range interaction where $r = 2$. When $\mathcal{G} = \mathbb{Z}^d$, an interaction ϕ is shift-invariant if for all $\vec{n} \in \mathbb{Z}^d$ and $a \in \mathcal{L}(X)$, $\phi(a) = \phi(\sigma^{\vec{n}}(a))$. Since the standard Cayley graph of $\mathbb{Z}^d$ has no triangles, a shift-invariant nearest neighbour interaction is uniquely determined by its values on patterns on $\{\vec{0}\}$

("single site potentials") and on patterns on pairs $\{\vec{0}, \vec{e}_i\}$ where $i = 1, \ldots, d$ ("edge interactions").

A *Gibbs state with a nearest neighbour interaction* ϕ is an MRF μ such that for all $x \in supp(\mu)$ and $A, B \subset \mathcal{V}$ finite satisfying $\partial A \subset B \subset A^c$,

$$\mu\left([x]_A \,\middle|\, [x]_B\right) = \frac{\prod\limits_{C \subset A \cup \partial A} e^{\phi(x|_C)}}{Z_{A, x|_{\partial A}}}$$

where $Z_{A,x|_{\partial A}}$ is the uniquely determined normalising factor so that $\mu(X \,\big|\, [x]_{\partial A}) = 1$ for all $x \in supp(\mu)$.

Some examples:

1. *(Ising Model)* Fix some $J, E \in \mathbb{R}$ and let $X = \{1, -1\}^{\mathbb{Z}^d}$ and ϕ be a shift-invariant nearest neighbour interaction ϕ given by

$$\begin{aligned} \phi([m,n]_{\vec{0},\vec{e}_i}) &= Jmn \\ \phi([m]_{\vec{0}}) &= Em \end{aligned}$$

   for all $m, n \in \{-1, 1\}$ and $1 \leq i \leq d$. The Ising model with "interaction" J and external field E is a Gibbs state supported on X with interaction ϕ.

2. *(Hard Square Model)* Fix some $\lambda \in \mathbb{R}$. The hard square model is the set $X \subset \{0,1\}^{\mathbb{Z}^d}$ which consists of configurations such that adjacent symbols cannot both be 1. Let ϕ be an interaction on X such that

$$\begin{aligned} \phi([0,0]_{\vec{0},\vec{e}_i}) = \phi([0,1]_{\vec{0},\vec{e}_i}) &= \phi([1,0]_{\vec{0},\vec{e}_i}) = 0 \\ \phi([0]_{\vec{0}}) = 0 \quad \text{and} & \quad \phi([1]_{\vec{0}}) = \lambda \end{aligned}$$

   for all $1 \leq i \leq d$. If μ is a Gibbs state supported on X with the shift-invariant nearest neighbour interaction ϕ then the specification takes a particularly nice form: For all $x \in X$ we get

$$\begin{aligned} \mu\left([x]_A \,\middle|\, [x]_B\right) &= \frac{\prod\limits_{C \subset A \cup \partial A} e^{\phi(x|_C)}}{Z_{A, x|_{\partial A}}} \\ &= \frac{e^{\lambda(\text{number of ones in } x|_{A \cup \partial A})}}{Z_{A, x|_{\partial A}}} \end{aligned}$$

   where $Z_{A,x|_{\partial A}}$ is the normalising factor.

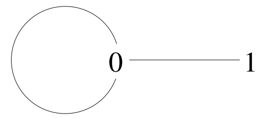

Figure 1.2: The Graph for the Hard Square Model

Note that the hard square model is $X^d_{\mathcal{H}}$ where $\mathcal{H}$ is the graph given by Figure 1.2.

If $\mathcal{G} = \mathbb{Z}^d$ we the specification of a Gibbs state with some shift-invariant nearest neighbour interaction is completely determined by the interaction and therefore by finitely many parameters. We want to address the question: Under what conditions on the support is every Markov random field a Gibbs measure for some nearest neighbour interaction?

The well-known Hammersley-Clifford theorem (Theorem 2.2.2) [3, 9, 23, 24, 54] gives one such condition, a positivity assumption on the MRF given by the presence of a safe symbol in the support which we explain next:

A closed configuration space $X \subset \mathcal{A}^{\mathcal{V}}$ is said to have a safe symbol $\star \in \mathcal{A}$ if for all $A \subset \mathcal{V}$ and $x \in X$ the configuration y given by

$$y_n = \begin{cases} x_n \text{ if } n \in A \\ \star \text{ if } n \notin A \end{cases}$$

is an element of X.

It is not hard to see that a space of configurations $X^d_{\mathcal{H}}$ has a safe symbol $\star$ if and only if $\star$ is adjacent to all the vertices of $\mathcal{H}$. For instance the symbol 0 is a safe symbol for the hard square model but $X^d_{K_r}$ and $X^d_{C_n}$ do not have a safe symbol for $r \geq 2$ and $n \geq 1$. The Hammersley-Clifford theorem does not apply to Markov random fields whose support is $X^d_{K_r}$ or $X^d_{C_n}$.

We will prove:

Theorem 1.1.1. *[13] Let $n \neq 1, 4$ and $d \geq 2$. If the support of a shift-invariant MRF is $X^d_{C_n}$ then it is a Gibbs state for some shift-invariant nearest neighbour interaction.*

On the other hand,

Theorem 1.1.2. *[13] For $\mathcal{G} = \mathbb{Z}^2$, there exists a shift-invariant MRF which is not Gibbs for any shift-invariant finite range interaction.*

It was proved in [9, 12] that if the underlying graph $\mathcal{G} = \mathbb{Z}$, then every shift-invariant MRF is a Gibbs state for some shift-invariant nearest neighbour interaction. Previously the same conclusion was obtained under certain mixing conditions with a more general alphabet (Theorems 10.25 and 10.35 in [22]); however there are MRFs which are not Gibbs states for any nearest neighbour interaction: when the alphabet is countable (Theorem 10.33 in [22]) and when the measure is not shift-invariant [16]. When $\mathcal{G}$ is finite, such MRFs were constructed in [34]. On the other hand,

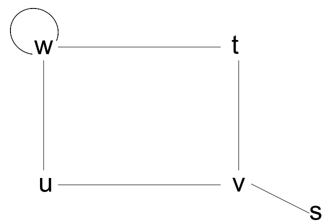

Figure 1.3: A Dismantlable Graph

there are algebraic conditions on the support [21] and conditions on the graph [28] which guarantee the conclusions of the Hammersley-Clifford Theorem.

MRFs and Gibbs States with nearest neighbour interactions will be defined in Subsection 2.1.1 and Subsection 2.1.4. The proof of Theorems 1.1.1 and 1.1.2 will be given in Sections 2.5 and 2.8 respectively.

1.2 A Generalisation of the Hammersley-Clifford Theorem for Bipartite Graphs

In the following $v \sim_{\mathcal{H}} w$ denotes that (v, w) is an edge in $\mathcal{H}$. Also $\mathcal{H}$ will denote both the graph and its set of vertices.

The following notions were introduced in [37] to classify cop-win graphs: Given a graph $\mathcal{H}$, we say that a vertex $v \in \mathcal{H}$ *folds* to $w \in \mathcal{H}$ if $u \sim_{\mathcal{H}} v$ implies that $u \sim_{\mathcal{H}} w$. A graph $\mathcal{H} \setminus \{v\}$ is called a *fold* of $\mathcal{H}$. A graph $\mathcal{H}$ is called *dismantlable* if there exists a sequence of folds $\mathcal{H}, \mathcal{H}_1, \ldots, \mathcal{H}_n$ such that $\mathcal{H}_n$ is a single vertex (with or without a loop).

For instance if $n \neq 4$ and $\mathcal{H} = C_n$ then $i \sim_{C_n} i+1, i-1 (\text{mod } n)$ for all $0 \leq i \leq n-1$ proving that C_n cannot be folded. We had remarked that for a graph $\mathcal{H}$, $X^d_{\mathcal{H}}$ has a safe symbol $\star$ if and only if $\star \sim_{\mathcal{H}} v$ for all $v \in \mathcal{H}$. Thus all vertices in such a graph $\mathcal{H}$ can be folded to $\star$ and so $\mathcal{H}$ is dismantlable. In particular, 0 is a safe symbol for the graph $\mathcal{H}$ in Figure 1.2. However there are many examples of dismantlable graphs $\mathcal{H}$ for which $X^d_{\mathcal{H}}$ does not have a safe symbol; for example, the graph $\mathcal{H}$ in Figure 1.3 can be folded to a single vertex by the sequence: s folds to t, v folds to w, u folds to w and t folds to w. Thus $\mathcal{H}$ is dismantlable but there is no vertex in $\mathcal{H}$ which is adjacent to all its other vertices; $X^d_{\mathcal{H}}$ does not have a safe symbol.

We will prove:

Theorem 1.2.1. *[11] Let $\mathcal{G}$ be a bipartite graph and let $\mathcal{H}$ be a dismantlable graph. Then any Markov random field on $Hom(\mathcal{G}, \mathcal{H})$ is a Gibbs state for some nearest neighbour interaction. Further if $\mathcal{G} = \mathbb{Z}^d$ and the Markov random field is shift-invariant then the corresponding interaction can be chosen to be shift-invariant as well.*

In fact we prove a more general result for configuration spaces which cannot be represented as the space of graph homomorphisms, generalising the Hammersley-Clifford theorem when the underlying graph $\mathcal{G}$ is bipartite. The details are part of Chapter 3 and for the statement of results in this direction look at Theorems 3.3.1, 3.3.2 and Corollary 3.3.8.

1.3 Pivot Property

For the study of Markov random fields we introduce the following combinatorial property on closed configuration spaces:

A closed configuration space X is said to have the *pivot property* if for all $(x, y) \in \Delta_X$ with $x \neq y$ there exists a finite sequence $x^{(1)} = x, x^{(2)}, \ldots, x^{(k)} = y \in X$ such that each $(x^{(i)}, x^{(i+1)})$ differ exactly at a single site.

The pivot property is useful to study Markov random fields for the following reason: Let $\mathcal{G} = \mathbb{Z}^d$. If μ is a shift-invariant Markov random field such that $supp(\mu)$ has the pivot property then for all $x, y \in supp(\mu)$ which differ at finitely many sites there exists a finite sequence $x^{(1)} = x, x^{(2)}, \ldots, x^{(k)} = y \in X$ such that each $(x^{(i)}, x^{(i+1)})$ differ exactly at a single site $\vec{n}_i \in \mathbb{Z}^d$. Let F be a set of sites such that $x|_{F^c} = y|_{F^c}$ and $\vec{n}_i \in F$ for all i. Then

$$\begin{aligned}\frac{\mu([x]_F \mid [x]_{\partial F})}{\mu([y]_F \mid [y]_{\partial F})} &= \prod_{i=1}^{k-1} \frac{\mu([x^{(i)}]_F \mid [x^{(i)}]_{\partial F})}{\mu([x^{(i+1)}]_F \mid [x^{(i+1)}]_{\partial F})} \\ &= \prod_{i=1}^{k-1} \frac{\mu([x^{(i)}]_{\vec{n}_i} \mid [x^{(i)}]_{\partial\{\vec{n}_i\}})}{\mu([x^{(i+1)}]_{\vec{n}_i} \mid [x^{(i+1)}]_{\partial\{\vec{n}_i\}})}.\end{aligned}$$

Since μ is shift-invariant these equations imply that the entire space of conditional probabilities is completely determined by finitely many parameters, viz., $\mu([x]_{\vec{0}} \mid [x]_{\partial\{\vec{0}\}})$ for $x \in supp(\mu)$.

Examples of Configuration Spaces with the Pivot Property:

1. Any closed configuration space with a safe symbol.
2. $X^d_{K_r}$ if $r \geq 2d + 2$ (Proposition 2.2.5) or $r = 3$ (Proposition 2.3.4).
3. $X^d_{\mathcal{H}}$ if $\mathcal{H}$ is a dismantlable graph [6].

On the other hand it is not true that all homomorphism spaces have the pivot property: It was observed by Brian Marcus that $X^2_{K_4}$ and $X^2_{K_5}$ do not have the pivot property. We will show that $X^2_{K_4}$ does not have the pivot property in Subsection 5.2.4.

We will prove:

Theorem 1.3.1. *[13] Let $n \neq 1, 4$ and $d \geq 2$. Then $X^d_{C_n}$ has the pivot property.*

Theorem 1.3.1 will be further generalised: A finite graph $\mathcal{H}$ is said to be four-cycle free if it has no self-loops and C_4 is not a subgraph of $\mathcal{H}$.

We will prove:

Theorem 1.3.2. *[10] Let $d \geq 2$. For all four-cycle free graphs $\mathcal{H}$, $X^d_{\mathcal{H}}$ has the pivot property.*

The pivot property will be defined and further examples will be given in Section 2.2.2. Theorems 1.3.1 and 1.3.2 will be restated as Proposition 2.3.4 and Theorem 4.1.4 respectively.

1.4 Entropy Minimality

While the proof of Theorem 1.2.1 is more combinatorial in nature the proof of Theorem 1.1.1 rests on some ergodic theory considerations which we will now touch upon:

Fix $d \geq 2$ and let $B_n = [-n, n]^d \cap \mathbb{Z}^d$. A *shift space* is a closed configuration space $X \subset \mathcal{A}^{\mathbb{Z}^d}$ which is shift-invariant. If $X = X^d_{\mathcal{H}}$ for some $\mathcal{H}$ then X is called a *hom-shift*. The *topological entropy* of a shift space X is

$$h_{top}(X) := \lim_{n \longrightarrow \infty} \frac{1}{|B_n|} \log |\mathcal{L}_{B_n}(X)|.$$

It measures the growth rate of the number of allowed patterns in X. It is easy to see that if $Y \subsetneq X$ then $h_{top}(Y) \leq h_{top}(X)$. As in [14] a shift space X is called entropy minimal if for any shift space $Y \subsetneq X$, $h_{top}(Y) < h_{top}(X)$; In other words, exclusion of any pattern in X causes a drop in its entropy.

There has been much work trying to establish some relationship between "mixing conditions" on the shift space and entropy minimality: For $d = 1$, all irreducible shifts of finite type are entropy minimal (Theorem 4.4.7 in [29]). Not much is known for higher dimensions; it is known that some strong mixing conditions like uniform filling imply entropy minimality [51] but weaker conditions like corner gluing do not [4]. There has been some recent work [52] which gives a condition equivalent to entropy minimality for shifts of finite type.

Our results in Chapter 2 imply that $X^d_{C_n}$ is entropy minimal for $n \neq 4$:

For proving this we will use the associated 'height functions': A *height function* is a function $h : \mathbb{Z}^d \longrightarrow \mathbb{Z}$ such that for adjacent vertices $\vec{i}, \vec{j} \in \mathbb{Z}^d$, $|h(\vec{i}) - h(\vec{j})| = 1$. Let Ht^d denote the set of height function on $\mathbb{Z}^d$. We will prove that the map $h \in Ht^d \longrightarrow h \mod n \in X^d_{C_n}$ is surjective when $n \neq 4$ (for the case when $n = 3$, also look at [48]) and that given any ergodic MRF μ (with some technical assumptions) on $X^d_{C_n}$ by the ergodic theorem there exists a notion of slope (average rate of increase of height) for every direction; if the slope is maximal in some direction then the support of μ is frozen, otherwise it is fully supported. From this, standard results in thermodynamic formalism (the Lanford-Ruelle Theorem [47]) imply that $X^d_{C_n}$ is entropy minimal. Have a look at Proposition 2.6.1 and Lemma 2.7.2.

These ideas will be further extended to prove:

Theorem 1.4.1. *[10] Let $\mathcal{H}$ be a four-cycle free connected graph. Then $X^d_{\mathcal{H}}$ is entropy minimal.*

In the following, by the *neighbourhood* of a vertex $v \in \mathcal{H}$ we will mean the set of all vertices adjacent to v. As in algebraic topology, there is a notion of covering spaces for graphs: C is a *covering space* of $\mathcal{H}$ if there is a graph homomorphism (called the *covering map*) $f : C \longrightarrow \mathcal{H}$ such that for every vertex $v \in \mathcal{H}$ the preimage of the neighbourhood N of v in $\mathcal{H}$ is a disjoint union of a constant number of neighbourhoods in C isomorphic to N via f. Given a graph $\mathcal{H}$, its universal cover (denoted by $E_{\mathcal{H}}$) is the unique covering space which is a tree (Look for instance in [1]). Denote the corresponding covering map by π.

For example, the universal cover of C_n is $\mathbb{Z}$ and the corresponding graph homomorphism $\pi : \mathbb{Z} \longrightarrow C_n$ is the map $\pi(r) = r \mod n$. On the other hand the universal cover of a tree is itself.

It is easy to see that for the induced map on $X^d_{E_{\mathcal{H}}}$ (also denoted by π) we have $\pi(X^d_{E_{\mathcal{H}}}) \subset X^d_{\mathcal{H}}$; we will prove in Proposition 4.4.2 that the map is surjective. Also the preimage of a configuration in $X^d_{\mathcal{H}}$ is unique if we fix the lift at a single vertex. Associate to every configuration $x \in X_{\mathcal{H}}$ a configuration $\tilde{x} \in X^d_{E_{\mathcal{H}}}$ such that $\pi(\tilde{x}) = x$. We can thus associate to the space $X^d_{\mathcal{H}}$ the *generalised height function* $h_{\mathcal{H}} : X^d_{\mathcal{H}} \times \mathbb{Z}^d \longrightarrow \mathbb{N} \cup \{0\}$ such that $h_{\mathcal{H}}(x, \vec{i})$ is the graph distance between $\tilde{x}_{\vec{0}}$ and $\tilde{x}_{\vec{i}}$. From here on, the steps for proving Theorem 1.4.1 are similar to those used in the proof of Proposition 2.6.1 but the proofs obtained for the individual steps in the latter are somewhat different. For instance now the existence of a slope for any ergodic measure on $X^d_{\mathcal{H}}$ in every direction will follow from the subadditive ergodic theorem instead of the ergodic theorem. The reason is that unlike height functions introduced earlier, $h_{\mathcal{H}}$ is not additive but is subadditive, meaning

$$h_{\mathcal{H}}(x, \vec{i} + \vec{j}) \leq h_{\mathcal{H}}(x, \vec{i}) + h_{\mathcal{H}}(\sigma^{\vec{i}}(x), \vec{j}).$$

Topological entropy will be defined in Section 2.7. A description of height functions will be given in Section 2.3 and that of the universal covers and generalised height functions will be given in Sections 4.4 and 4.5. A brief description of the tools from thermodynamic formalism will be given in Sections 2.7 and 4.2 respectively. Theorem 1.4.1 will be the same as Theorem 4.1.2.

Chapter 2

Markov Random Fields, Markov Cocycles and the 3-Coloured Chessboard

The main results of this chapter are Theorems 1.1.1 and 1.1.2. The results in this chapter are joint with Tom Meyerovitch [13].

2.1 Background and Notation

This section will recall the necessary concepts and introduce the basic notation.

2.1.1 Markov Random Fields and Topological Markov Fields

Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be a simple, undirected graph where the vertex set $\mathcal{V}$ is finite or countable. We always assume that $\mathcal{G}$ is *locally finite*, meaning all $v \in \mathcal{V}$ have a finite number of neighbours. The *boundary* of a set of vertices $F \subset \mathcal{V}$, denoted by ∂F, is the set of vertices outside F which are adjacent to F:

$$\partial F := \{v \in \mathcal{V} \setminus F \mid \exists w \in F \text{ s.t. } (v, w) \in \mathcal{E}\}.$$

Remark: Observe that in our notation $\partial F \subset F^c$. This is sometimes called the *outer boundary* of the set F. Consistent with our notation the *inner boundary* of F is $\partial(F^c)$.

Given a finite set $\mathcal{A}$, the space $\mathcal{A}^{\mathcal{V}}$ is a compact topological space with respect to the product topology, where the topology on $\mathcal{A}$ is discrete. For $F \subset \mathcal{V}$ finite and $a \in \mathcal{A}^F$, we denote by $[a]_F$ the *cylinder set*

$$[a]_F := \left\{x \in \mathcal{A}^{\mathcal{V}} \mid x|_F = a\right\}.$$

For $x \in \mathcal{A}^{\mathcal{V}}$ we use the notation $[x]_F$ for $[x|_F]_F$. The collection of cylinder sets generates the Borel σ-algebra on $\mathcal{A}^{\mathcal{V}}$.

A *Markov random field* is a Borel probability measure μ on $\mathcal{A}^{\mathcal{V}}$ with the property that for all finite $A, B \subset \mathcal{V}$ such that $\partial A \subset B \subset A^c$ and $a \in \mathcal{A}^A, b \in \mathcal{A}^B$ satisfying $\mu([b]_B) > 0$

$$\mu\left([a]_A \;\middle|\; [b]_B\right) = \mu\left([a]_A \;\middle|\; [b|_{\partial A}]_{\partial A}\right).$$

We say that the sets of vertices $A, B \subset \mathcal{V}$ are *separated* (in the graph $\mathcal{G}$) if they are disjoint and $(v, w) \notin \mathcal{E}$ whenever $v \in A$ and $w \in B$.

Here is an equivalent definition of a Markov random field: If x is a point chosen randomly according to the measure μ, and $A, B \subset \mathcal{V}$ are finite and separated, then conditioned on $x|_{\mathcal{V}\setminus(A\cup B)}$, $x|_A$ and $x|_B$ are independent random variables.

A Markov random field is called *global* if the conditional independence above holds for all separated sets $A, B \subset \mathcal{V}$ (finite or not) .

As in [9, 12], a *topological Markov field* is a compact set $X \subset \mathcal{A}^{\mathcal{V}}$ such that for all finite $F \subset \mathcal{V}$ and $x, y \in X$ satisfying $x|_{\partial F} = y|_{\partial F}$, there exist $z \in X$ satisfying

$$z_v := \begin{cases} x_v \text{ for } v \in F \\ y_v \text{ for } v \in \mathcal{V} \setminus F. \end{cases}$$

A topological Markov field is called *global* if we do not demand that F be finite.

The *support* of a Borel probability measure μ on $\mathcal{A}^{\mathcal{V}}$ denoted by *supp*(μ) is the intersection of all closed sets $Y \subset \mathcal{A}^{\mathcal{V}}$ for which $\mu(Y) = 1$. Equivalently,

$$supp(\mu) = \mathcal{A}^{\mathcal{V}} \setminus \bigcup_{[a]_A \in \mathcal{N}(\mu)} [a]_A,$$

where $\mathcal{N}(\mu)$ is the collection of all cylinder sets $[a]_A$ with $\mu([a]_A) = 0$. The support of a Markov random field is a topological Markov field (see Lemma 2.0.1 in [9]).

2.1.2 The Homoclinic Equivalence Relation of a Topological Markov Field and Adapted MRFs.

Following [42, 49], we denote by Δ_X the *asymptotic relation* of a topological Markov field $X \subset \mathcal{A}^{\mathcal{V}}$, which is given by

$$\Delta_X := \{(x, y) \in X \times X \mid x_n = y_n \text{ for all but finitely many } n \in \mathcal{V}\}. \tag{2.1.1}$$

Given a shift space X the *homoclinic relation* is defined to be the same as the asymptotic relation on X. We say that an MRF μ is *adapted* with respect to a topological Markov field X if *supp*$(\mu) \subset X$ and

$$x \in supp(\mu) \Longrightarrow \{y \in X \mid (x, y) \in \Delta_X\} \subset supp(\mu).$$

It follows that an MRF μ is adapted with respect to a TMF X if and only if all continuous functions $f : X \longrightarrow X$ satisfying $(x, f(x)) \in \Delta_X$ for all $x \in X$ are absolutely continuous with respect to μ. In this case μ is said to be *non-singular* with respect to Δ_X and it is possible to patch up all the Radon-Nikodym derivatives $\frac{df(\mu)}{d\mu}$ for all f as described above: The *Radon-Nikodym cocycle* of μ is a map $c_\mu : \Delta_X \longrightarrow \mathbb{R}^+$ such that for all functions f described above

$$c_\mu(x, f(x)) = \frac{df(\mu)}{d\,\mu}(x) \ \mu\text{-almost everywhere.}$$

If μ is an MRF then the Radon-Nikodym derivative takes a particularly nice form: If $(x, y) \in \Delta_{supp(\mu)}$ and a finite set $F \subset \mathcal{V}$ such that $x|_{F^c} = y|_{F^c}$ then

$$c_\mu(x, y) = \frac{\mu([y]_{F\cup\partial F})}{\mu([x]_{F\cup\partial F})}.$$

Since μ is a Markov random field the right hand side is independent of the choice of F. See [33] and references within for further details.

To illustrate adaptedness, if $supp(\mu) = X$ then μ is adapted with respect to X, and if $X = X_1 \cup X_2$ is the union of two topological Markov fields over disjoint alphabets and μ is a Markov random field with $supp(\mu) = X_1$ then μ is adapted with respect to X. On the other hand, the Bernoulli measure $(\frac{1}{2}\delta_0 + \frac{1}{2}\delta_1)^{\mathcal{V}}$ is *not* adapted with respect to $\{0, 1, 2\}^{\mathcal{V}}$. In fact, if $X = \mathcal{A}^{\mathcal{V}}$ for some finite alphabet $\mathcal{A}$ then any Markov random field which is adapted to X has $supp(\mu) = X$.

2.1.3 $\mathbb{Z}^d$-Shift Spaces and Shifts of Finite Type

For the Markov random fields we discuss in this chapter, the set of vertices of the underlying graph is the d-dimensional integer lattice. We identify $\mathbb{Z}^d$ with the set of vertices of the Cayley graph with respect to the standard generators. Rephrasing, $\vec{n}, \vec{m} \in \mathbb{Z}^d$ are adjacent iff $\|\vec{n} - \vec{m}\|_1 = 1$, where for all $\vec{n} = (n_1, \ldots, n_d) \in \mathbb{Z}^d$, $||\vec{n}||_1 := \sum_{r=1}^d |n_r|$ denotes the l^1 norm of $\vec{n}$. The *boundary* of a given finite set $F \subset \mathbb{Z}^d$ is thus given by:

$$\partial F = \{\vec{m} \in F^c \mid ||\vec{n} - \vec{m}||_1 = 1 \text{ for some } \vec{n} \in F\}.$$

On the compact space $\mathcal{A}^{\mathbb{Z}^d}$ (with the product topology over the discrete set $\mathcal{A}$) the maps $\sigma^{\vec{n}} : \mathcal{A}^{\mathbb{Z}^d} \to \mathcal{A}^{\mathbb{Z}^d}$ given by

$$(\sigma^{\vec{n}}(x))_{\vec{m}} := x_{\vec{m}+\vec{n}} \text{ for all } \vec{m}, \vec{n} \in \mathbb{Z}^d$$

define a $\mathbb{Z}^d$-action by homeomorphisms, called the *shift-action*. The pair $(\mathcal{A}^{\mathbb{Z}^d}, \sigma)$ is a topological dynamical system called the *d-dimensional full shift* on the alphabet $\mathcal{A}$. Note that σ acts on the Cayley graph of $\mathbb{Z}^d$ by graph isomorphisms.

A $\mathbb{Z}^d$-*shift space* or *subshift* is a dynamical system (X, σ) where $X \subset \mathcal{A}^{\mathbb{Z}^d}$ is closed and invariant under the map $\sigma^{\vec{n}}$ for each $\vec{n} \in \mathbb{Z}^d$.

A Borel probability measure μ on $\mathcal{A}^{\mathbb{Z}^d}$ is *shift-invariant* if $\mu \circ \sigma^{\vec{n}} = \mu$ for all $\vec{n} \in \mathbb{Z}^d$. It follows that the support of all shift-invariant measures μ is a subshift.

For $X \subset \mathcal{A}^{\mathcal{V}}$ and $W \subset \mathcal{V}$ let

$$\mathcal{L}_W(X) := \{w \in \mathcal{A}^W \mid \text{there exists } x \in X \text{ such that } x|_W = w\}.$$

The *language* of $X \subset \mathcal{A}^{\mathcal{V}}$ denoted by $\mathcal{L}(X)$ is defined as all finite patterns which occur in the elements of X:

$$\mathcal{L}(X) := \bigcup_{W \subset \mathcal{V} \text{ finite}} \mathcal{L}_W(X).$$

If $A, B \subset \mathcal{V}$ and $x \in \mathcal{A}^A, y \in \mathcal{A}^B$ such that $x|_{A\cap B} = y|_{A\cap B}$ then $x \vee y \in \mathcal{A}^{A\cup B}$ is given by

$$(x \vee y)_n := \begin{cases} x_n & n \in A \\ y_n & n \in B. \end{cases}$$

An alternative equivalent definition for a subshift is given by *forbidden patterns* as follows: Let

$$\mathcal{A}^\star := \bigcup_{W \subset \mathbb{Z}^d \text{ finite}} \mathcal{A}^W.$$

For all $\mathcal{F} \subset \mathcal{A}^\star$ let

$$X_{\mathcal{F}} = \{x \in \mathcal{A}^{\mathbb{Z}^d} | \text{ no translate of a subconfiguration of } x \text{ belongs to } \mathcal{F}\}.$$

It is well-known that a subset $X \subset \mathcal{A}^{\mathbb{Z}^d}$ is a subshift if and only if there exists $\mathcal{F} \subset \mathcal{A}^\star$ such that $X = X_{\mathcal{F}}$ (for $d = 1$ this is stated in [29] as Theorem 6.1.21; the proof for higher dimensions is similar). The set $\mathcal{F}$ is called the set of forbidden patterns for X. A subshift X is called a *shift of finite type* if $X = X_{\mathcal{F}}$ for some finite set $\mathcal{F}$. A shift of finite type is called a *nearest neighbour shift of finite type* if $X = X_{\mathcal{F}}$ where $\mathcal{F}$ consists of nearest neighbour constraints, i.e. $\mathcal{F}$ consists of patterns on single edges. When $d = 1$ nearest neighbour shifts of finite type are also called *topological Markov chains*. In fact the study of nearest neighbour shifts of finite type has been partly motivated by the fact that the support of stationary Markov chains are one-dimensional nearest neighbour shifts of finite type.

Every nearest neighbour $\mathbb{Z}^d$-shift of finite type is a shift-invariant topological Markov field. When $d = 1$ the converse is also true under the assumption that the subshift is non-wandering [12]. Without the non-wandering assumption, one-dimensional shift-invariant topological Markov fields are still so-called sofic shifts, but not necessarily of finite type [12]. This does not hold in higher dimensions ([9] and Section 2.8).

2.1.4 Gibbs States with Nearest Neighbour Interactions

For a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ and $A \subset \mathcal{V}$, let $diam(A)$ denote the diameter of A with respect to the graph distance (denoted by $d_{\mathcal{G}}$) in $\mathcal{G}$, that is, $diam(A) = \max_{i,j \in A} d(i, j)$.

Following [47], an *interaction* on X is a function ϕ from $\mathcal{L}(X)$ to $\mathbb{R}$, satisfying certain summability conditions. Here we will only consider finite range interactions, for which the summability conditions are automatically satisfied.

An interaction is of *range at most* k if $\phi(a) = 0$ for $a \in \mathcal{L}_A(X)$ whenever $diam(A) > k$. We will call an interaction of range 1 a *nearest neighbour interaction.* When $\mathcal{G} = \mathbb{Z}^d$, an interaction ϕ is shift-invariant if for all $\vec{n} \in \mathbb{Z}^d$ and $a \in \mathcal{L}(X)$, $\phi(a) = \phi(\sigma^{\vec{n}}(a))$. Since the standard Cayley graph of $\mathbb{Z}^d$ has no triangles, a shift-invariant nearest neighbour interaction is uniquely determined by its values on patterns on $\{\vec{0}\}$ ("single site potentials") and on patterns on pairs $\{\vec{0}, \vec{e_i}\}$, $i = 1, \ldots, d$ ("edge interactions"). We denote these by $\phi([a]_0)$ and $\phi([a, b]_i)$ respectively where $a, b \in \mathcal{A}$.

A *Gibbs state with a nearest neighbour interaction* ϕ is a Markov random field μ such that for all $x \in supp(\mu)$ and $A, B \subset \mathcal{V}$ finite satisfying $\partial A \subset B \subset A^c$,

$$\mu\left([x]_A \;\middle|\; [x]_B\right) \;=\; \frac{\prod\limits_{C \subset A \cup \partial A} e^{\phi(x|_C)}}{Z_{A, x|_{\partial A}}}$$

where $Z_{A,x|_{\partial A}}$ is the uniquely determined normalising factor so that $\mu(X \;\Big|\; [x]_{\partial A}) = 1$ for all $x \in supp(\mu)$.

2.1.5 Invariant Spaces, Measures and Interactions

An *automorphism* of the graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ is a bijection on the vertex set $g : \mathcal{V} \longrightarrow \mathcal{V}$ which preserves the adjacencies, that is, $(u, v) \in \mathcal{E}$ if and only if $(gu, gv) \in \mathcal{E}$. Let the group of all automorphisms of the graph $\mathcal{G}$ be denoted by $Aut(\mathcal{G})$.

There is a natural action of $Aut(\mathcal{G})$ on patterns and configurations: given $a \in \mathcal{A}^F$, $x \in \mathcal{A}^{\mathcal{V}}$ and $g \in Aut(\mathcal{G})$ we have $ga \in \mathcal{A}^{gF}$ and $gx \in \mathcal{A}^{\mathcal{V}}$ given by

$$(ga)_{gv} := a_v \text{ and}$$

$$(gx)_v := x_{g^{-1}v}.$$

This induces an action on measures on the space $\mathcal{A}^{\mathcal{V}}$ given by

$$(g\mu)(L) := \mu(g^{-1}L)$$

for all measurable sets $L \subset \mathcal{A}^{\mathcal{V}}$.

For a given subgroup $G \subset Aut(\mathcal{G})$, a set of configurations $X \subset \mathcal{A}^{\mathcal{V}}$ is said to be *G-invariant* if

$gX = X$ for all automorphisms $g \in G$. Similarly a measure μ on $\mathcal{A}^{\mathcal{V}}$ is said to be G-invariant if $g\mu = \mu$ for all $g \in G$. Note, for any subgroup $G \subset Aut(\mathcal{G})$, if μ is a G-invariant probability measure then $supp(\mu)$ is also a G-invariant configuration space. For $\mathcal{G} = \mathbb{Z}^d$ we abuse notation to denote the group of translations by $\mathbb{Z}^d$ as well. In this notation $\sigma^{\vec{n}}(x) = (-\vec{n})(x)$.

Let $X \subset \mathcal{A}^{\mathcal{V}}$ be a closed configuration space invariant under a subgroup $G \subset Aut(\mathcal{G})$. Then G acts on the interactions on X: Given an interaction ϕ on X for all $a \in \mathcal{A}^F$ and $g \in G$

$$g\phi([a]_F) := \phi([g^{-1}a]_{g^{-1}F}).$$

2.2 Markov Specifications and Markov Cocycles

Any Markov random field μ yields conditional probabilities of the form $\mu(x_F = \cdot \mid x_{\partial F} = \delta)$ for all finite $F \subset \mathcal{V}$ and *admissible* $\delta \in \mathcal{A}^{\partial F}$ (by admissible we mean $\mu([\delta]_{\partial F}) > 0$). We refer to such a collection of conditional probabilities as the *Markov specification* associated with μ. It may happen that two distinct Markov random fields have the same specification, as in the case of the 2-dimensional Ising model in low temperature [38]. In general it is a subtle and challenging problem to determine if a given Markov specification admits more than one Markov random field (the problem of uniqueness for the measure of maximal entropy of a $\mathbb{Z}^d$-shift of finite type is an instance of this problem [7]). For the purpose of our study and statement of our results, it would be convenient to have an intrinsic definition for a Markov specification, not involving a particular underlying Markov random field.

Let $X \subset \mathcal{A}^{\mathcal{V}}$ be a topological Markov field. A *Markov specification* on X is an assignment for each finite and non-empty $F \subset \mathcal{V}$ and $x \in \mathcal{L}_{\partial F}(X)$ of a probability measure $\Theta_{F,x}$ on $\mathcal{L}_F(X)$ satisfying the following conditions:

1. **Support condition**: For all finite and non-empty $F \subset \mathcal{V}$, $x \in \mathcal{L}_{\partial F}(X)$ and $y \in \mathcal{L}_F(X)$, $x \vee y \in \mathcal{L}_{F \cup \partial F}(X)$ if and only if $\Theta_{F,x}(y) > 0$. This condition can be written as follows:

$$supp(\Theta_{F,x}) = \{y \in \mathcal{L}_F(X) \mid x \vee y \in \mathcal{L}_{F \cup \partial F}(X)\}.$$

2. **Markovian condition**: For all finite and non-empty $F \subset \mathcal{V}$ and $x \in \mathcal{L}_{\partial F}(X)$, $\Theta_{F,x}$ is a Markov random field on the finite graph induced from $\mathcal{V}$ on F.

3. **Consistency condition**: If $F \subset H \subset \mathcal{V}$ are finite and non-empty, $x \in \mathcal{L}_{\partial F}(X)$, $y \in \mathcal{L}_{\partial H}(X)$ and $x_n = y_n$ for $n \in \partial F \cap \partial H$, then for all $z \in \mathcal{L}_F(X)$

$$\Theta_{F,x}(z) = \frac{\Theta_{H,y}([z \vee x]_{(F \cup \partial F) \cap H})}{\Theta_{H,y}([x]_{\partial F \cap H})}.$$

The definition above has been set up so that for any Markov random field μ with $X = supp(\mu)$, the assignment

$$(F, x) \mapsto \Theta_{F,x}(a) := \mu([a]_F \mid [x]_{\partial F})$$

is a Markov specification. Conversely, given any Markov specification Θ on X there exists a Markov random field μ on X compatible with Θ in the sense that $\mu([a]_F \mid [y]_{\partial F}) = \Theta_{F,y}(a)$ for all $a \in \mathcal{L}_F(X)$ whenever $\mu([y]_{\partial F}) > 0$ (Chapter 4 in [22]). Furthermore, when $X \subset \mathcal{A}^{\mathbb{Z}^d}$ is a subshift and the specification Θ is shift-invariant, it follows from amenability of $\mathbb{Z}^d$ that there exists a shift-invariant Markov random field μ compatible with Θ. However, in general it is possible that for a given specification Θ the support of any μ satisfying the above is a strict subset of X, in which case there exist certain finite $F \subset \mathcal{V}$ and $y \in \mathcal{L}_{\partial F}(X)$ for which the conditional probabilities $\mu([x]_F \mid [y]_{\partial F})$ are meaningless for all $x \in X$. We will provide such examples in Section 2.5. In such a case, according to our definition, the Markov specification associated with μ is the restriction of Θ to the support of μ, meaning the collection of conditional probabilities $\Theta_{F,x}$ for finite sets $F \subset \mathcal{V}$ and $x \in \mathcal{L}_{\partial F}(supp(\mu))$.

It will be convenient for our purposes to re-parameterize the set of Markov specifications on a given topological Markov field X. For this purpose we use the formalism of Δ_X-cocycles. To this well-known formalism we introduce an ad-hoc definition of a *Markov cocycle*, which we will explain now. As in [49], a (real-valued) *Δ_X-cocycle* is a function $M : \Delta_X \longrightarrow \mathbb{R}$ satisfying

$$M(x, z) = M(x, y) + M(y, z) \text{ whenever } (x, y), (y, z) \in \Delta_X. \tag{2.2.1}$$

Given $G \subset Aut(\mathcal{G})$, M is a G-invariant Δ_X-cocycle if $M(x, y) = M(g(x), g(y))$ for all $g \in G$. If $\mathcal{G} = \mathbb{Z}^d$ then $\mathbb{Z}^d$-invariant Δ_X-cocycles are said to be shift-invariant. We call M a *Markov cocycle* if it is a Δ_X-cocycle and satisfies: For all $(x, y) \in \Delta_X$ such that $x|_{F^c} = y|_{F^c}$ the value $M(x, y)$ is determined by $x|_{F\cup\partial F}$ and $y|_{F\cup\partial F}$.

There is a clear bijection between Markov cocycles and Markov specifications on X: If Θ is a Markov specification on X, the corresponding Markov cocycle is given by

$$M(x, y) := \log\left(\Theta_{F,y|_{\partial F}}(y|_F)\right) - \log\left(\Theta_{F,x|_{\partial F}}(x|_F)\right),$$

where $(x, y) \in \Delta_X$ and $F \subset \mathcal{V}$ finite such that $x|_{F^c} = y|_{F^c}$. Clearly M does not depend on the choice of F : Let $\tilde{F}$ is the set of sites on which x and y differ then

$$\begin{aligned}
M(x,y) &:= \log\left(\Theta_{F,y|_{\partial F}}(y|_F)\right) - \log\left(\Theta_{F,x|_{\partial F}}(x|_F)\right) \\
&= \log\left(\Theta_{F,y|_{\partial F}}([y]_{\tilde{F}} \,\Big|\, [y]_{F\setminus\tilde{F}})\right) - \log\left(\Theta_{F,x|_{\partial F}}([x]_F \,\Big|\, [x]_{F\setminus\tilde{F}})\right) \\
\text{(Markovian condition)} \quad &= \log\left(\Theta_{F,y|_{\partial F}}([y]_{\tilde{F}} \,\Big|\, [y]_{\partial\tilde{F}\cap F})\right) - \log\left(\Theta_{F,x|_{\partial F}}([x]_F \,\Big|\, [x]_{\partial\tilde{F}\cap F})\right) \\
\text{(Consistency condition)} \quad &= \log\left(\Theta_{F,y|_{\partial\tilde{F}}}(y|_{\tilde{F}})\right) - \log\left(\Theta_{\tilde{F},x|_{\partial\tilde{F}}}(x|_{\tilde{F}})\right).
\end{aligned}$$

Conversely, given a Markov cocycle M on X, the corresponding specification Θ is given by

$$\Theta_{F,a}(y) = \frac{1}{Z_{M,F,a,z}} e^{M(x \vee z, x \vee y)},$$

where $F \subset \mathcal{V}$ is a finite set, $a \in \mathcal{L}_{\partial F}(X)$, $y, z \in \mathcal{L}_F(X)$ are such that $a \vee y, a \vee z \in \mathcal{L}_{F \cup \partial F}(X)$ and $x \in \mathcal{L}_{F^c}(X)$ with $x|_{\partial F} = a$. The normalising constant $Z_{M,F,a,z}$ is given by:

$$Z_{M,F,a,z} = \sum_{y'} e^{M(x \vee z, x \vee y')},$$

where the sum is over all $y' \in \mathcal{L}_F(X)$ such that $y' \vee a \in \mathcal{L}_{F \cup \partial F}(X)$. Note that the expression for the specification is well defined: Since M is a Markov cocycle on the topological Markov field X the right-hand side does not depend on the particular choice of x. The choice of the auxiliary variable z on the right-hand side changes the normalising constant $Z_{M,F,a,z}$, but does not change $\Theta_{F,a}$.

When $X \subset \mathcal{A}^{\mathcal{V}}$ is G-invariant, this bijection maps G-invariant specifications to G-invariant Markov cocycles. Thus, a G-invariant Borel probability measure μ is a G-invariant Markov random field if and only if $X = supp(\mu)$ is a topological Markov field and the Radon-Nikodym cocycle of μ with respect to Δ_X is some G-invariant Markov cocycle M, that is, for all $(x, y) \in \Delta_X$

$$\frac{\mu([y]_\Lambda)}{\mu([x]_\Lambda)} = e^{M(x,y)}$$

for all $\Lambda \supset F \cup \partial F$ where F is the set of sites on which x, y differ.

Fix a topological Markov field X and a nearest neighbour interaction ϕ on X. The *Gibbs cocycle* corresponding to ϕ is given by:

$$M_\phi(x, y) = \sum_{W \subset \mathcal{V} \text{ finite}} \phi(y|_W) - \phi(x|_W).$$

Note that when ϕ is a nearest neighbour interaction, there are only finitely many non-zero terms in the sum whenever $(x, y) \in \Delta_X$ and so M_ϕ is well defined. Examination of the definitions verifies that under the above assumptions M_ϕ is a Markov cocycle. Our point of interest is the converse: When can a Markov cocycle be expressed in this form?

Proposition 2.2.1. *Let μ be an MRF and M be a Markov cocycle on $supp(\mu)$ given by*

$$M(x, y) = \log\left(\frac{\mu([y]_\Lambda)}{\mu([x]_\Lambda)}\right) \text{ for all } (x, y) \in \Delta_{supp(\mu)}$$

for any $\Lambda \supset F \cup \partial F$ where F is the set of vertices where x and y differ. Then μ is a Gibbs state with a nearest neighbour interaction ϕ if and only if $M = M_\phi$.

The proof of the proposition follows because μ is a Gibbs state with a nearest neighbour interaction ϕ if and only if

$$M(x,y) = \log\left(\frac{\mu\left([y]_F \mid [y]_{\partial F}\right)}{\mu\left([x]_F \mid [y]_{\partial F}\right)}\right) = \sum_{W\subset\mathcal{V}\text{ finite}} \phi(y|_W) - \phi(x|_W) = M_\phi(x,y)$$

for all $x, y \in \Delta_{supp(\mu)}$. Let X be a topological Markov field. Denote by $\mathbf{M}_X$ the set of all Markov cocycles and by $\mathbf{G}_X$ the set of all nearest neighbour Gibbs cocycles. In case X is G-invariant, we denote by $\mathbf{M}_X^G$ the set of G-invariant Markov cocycles and by $\mathbf{G}_X^G$ the set of Gibbs cocycles corresponding to a G-invariant nearest neighbour interaction.

The set $\mathbf{M}_X$ of Markov cocycles naturally carries a vector space structure: Given $M_1, M_2 \in \mathbf{M}_X$ and $c_1, c_2 \in \mathbb{R}$, $c_1M_1 + c_2M_2 \in \mathbf{M}_X$. The reader can easily verify that $\mathbf{G}_X$ is a linear subspace of $\mathbf{M}_X$. If X is a shift-space then the shift-invariant nearest neighbour interactions constitute a finite-dimensional vector space, and the map sending a nearest neighbour interaction ϕ to the cocycle M_ϕ is linear, it follows that $\mathbf{G}_X^{\mathbb{Z}^d} \subset \mathbf{M}_X$ has finite dimension.

For a topological Markov field X defined over a finite graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, $\mathcal{M}_X$ is finite dimensional; the problem of determining which Markov cocycles are Gibbs amounts to solving a finite (but possibly large) system of linear equations. The resulting equations are essentially the 'balanced conditions' mentioned in [34].

2.2.1 The "Safe Symbol Property" and the Hammersley-Clifford Theorem

A topological Markov field $X \subset \mathcal{A}^{\mathbb{Z}^d}$ is said to have a *safe symbol* if there exists an element $\star \in \mathcal{A}$ such that for all $x \in X$ and $A \subset \mathbb{Z}^d$, $y \in \mathcal{A}^{\mathbb{Z}^d}$ given by

$$y_n := \begin{cases} x_n & \text{for } n \in A \\ \star & \text{for } n \in A^c \end{cases}$$

is also an element of X.

A formulation of the Hammersley-Clifford Theorem states:

Theorem 2.2.2. ***(Hammersley-Clifford, weak version** [24]**)*** *Let X be a topological Markov field with a safe symbol. Then:*

1. *Any Markov random field with $supp(\mu) = X$ is a Gibbs state for a nearest neighbour interaction.*

2. *If X is G-invariant for some $G \subset Aut(\mathcal{G})$ then any G-invariant Markov random field with $supp(\mu) = X$ is a Gibbs state for some G-invariant nearest neighbour interaction.*

The second statement in the theorem above is not a part of the original formulation, but does follow since there is an explicit algorithm to produce a nearest neighbour interaction which is invariant under all graph automorphisms for which the original Markov random field was invariant

[9]. See also [2, 54, 55]. It is in general false that a G-invariant Markov random field whose (G-invariant) specification is compatible with some nearest neighbour interaction must also be compatible with some G-invariant nearest neighbour interaction (see Corollary 2.4.6 below). In particular, for a general topological Markov field X we have $\mathbf{G}_X^G \subset \mathbf{M}_X^G \cap \mathbf{G}_X$, but the inclusion may be strict.

An inspection of the original proof of the Hammersley-Clifford Theorem actually gives the following a priori stronger result:

Theorem 2.2.3. ***(Hammersley-Clifford Theorem, strong version)*** *Let X be a topological Markov field with a safe symbol. Then:*

1. *Any Markov cocycle on X is a Gibbs cocycle given by a nearest neighbour interaction. In our notation this is expressed by: $\mathbf{M}_X = \mathbf{G}_X$.*

2. *Given $G \subset Aut(\mathcal{G})$, if X is G-invariant then any G-invariant Markov cocycle on X is a Gibbs cocycle given by a G-invariant nearest neighbour interaction. In our notation this is expressed by: $\mathbf{M}_X^G = \mathbf{G}_X^G$.*

It is easily verified that any topological Markov field X which satisfies one of the conclusions of the "strong version" immediately satisfies the corresponding conclusion of the "weak version". We will demonstrate in the following section that the converse implication is false in general. The proof of the first part of this version follows from Theorem 2.2.2 with the additional knowledge that given a Markov cocycle on a topological Markov field X with a safe symbol there exists a corresponding MRF μ such that $supp(\mu) = X$. This in turn is implied by arguments very similar to those in the proof of the following proposition. The second part of the theorem can be proved using Theorem 2.0.6 in [9], noting that the conclusion holds even if the MRF is not invariant under G but the corresponding Markov cocycle is.

Proposition 2.2.4. *Let X be a topological Markov field with a safe symbol. Then any Markov random field μ adapted to X has $supp(\mu) = X$.*

Proof. Let μ be a Markov random field adapted to X. We need to show that for all finite $F \subset \mathcal{V}$ and $a \in \mathcal{L}_F(X)$, $\mu([a]_F) > 0$. Denote $F \cup \partial F$ by $\tilde{F}$. Let $b \in \mathcal{L}_{\tilde{F}}(X)$ and $c \in \mathcal{L}_{\partial\tilde{F}}(X)$ satisfy $\mu([b \vee c]_{\tilde{F}\cup\partial\tilde{F}}) > 0$. In particular, $\mu([c]_{\partial\tilde{F}}) > 0$. Let $\tilde{b} \in \mathcal{L}_{\tilde{F}}(X)$ be given by

$$\tilde{b}_n := \begin{cases} b_n & n \in F \\ \star & n \in \partial F. \end{cases}$$

Note that $\tilde{b} \vee c \in \mathcal{L}_{\tilde{F}\cup\partial\tilde{F}}(X)$ because $\star$ is a safe-symbol. Again, by the safe symbol property it follows that $\tilde{a} \in \mathcal{L}_{\tilde{F}}(X)$ where:

$$\tilde{a}_n := \begin{cases} a_n & n \in F \\ \star & n \in \partial F. \end{cases}$$

Since X is a topological Markov field, $\tilde{a} \vee c \in \mathcal{L}_{\tilde{F} \cup \partial \tilde{F}}(X)$. Since μ is an adapted Markov random field it follows that $\mu([\tilde{a}]_{\tilde{F}} \mid [c]_{\partial \tilde{F}}) > 0$, and since $\mu([c]_{\partial \tilde{F}}) > 0$ it follows that $\mu([\tilde{a} \vee c]_{\tilde{F} \cup \partial \tilde{F}}) > 0$; in particular we get that $\mu([a]_F) > 0$. □

Remark: Proposition 2.2.4 is a particular instance of the more general fact that all Δ_X-nonsingular measures μ satisfy $supp(\mu) = X$, whenever Δ_X is a topologically minimal. The latter condition means that for any $x \in X$, the Δ_X-orbit $\Delta_X(x) := \{y \in X \mid (x, y) \in \Delta_X\}$ is dense in X.

Remark: When the underlying graph is $\mathbb{Z}^d$, any shift-invariant topological Markov field X with a safe symbol is actually a nearest-neighbour shift of finite type; Proposition 3.1.2 proves a more general fact.

2.2.2 The Pivot Property

We shall now consider a weaker property than that of having a safe symbol. Let X be a topological Markov field. If $x, y \in X$ only differ at a single $v \in \mathcal{V}$, then the pair (x, y) will be called *a pivot move* in X. A topological Markov field X is said to have *the pivot property* if for all $(x, y) \in \Delta_X$ such that $x \neq y$ there exists a finite sequence of points $x^{(1)} = x, x^{(2)}, \ldots, x^{(k)} = y \in X$ such that each $(x^{(i)}, x^{(i+1)})$ is a pivot move. In this case we say $x^{(1)} = x, x^{(2)}, \ldots, x^{(k)} = y$ is a chain of pivots from x to y. Many properties similar to the pivot property have appeared in the literature (often by the name local-move connectedness), for instance look at [6, 36, 39, 53] Here are some examples of subshifts which have the pivot property:

1. Any topological Markov field with a trivial homoclinic relation.
2. Any topological Markov field with a safe symbol.
3. The r-colourings of $\mathbb{Z}^d$ (defined below) where $r = 3$ (generalised in Proposition 2.3.4 and Theorem 4.1.4) or $r \geq 2d + 2$ (Proposition 2.2.5).
4. The space of graph homomorphisms from $\mathbb{Z}^d$ to a "dismantlable graph", as in [6] (defined in the examples after the statement of Theorem 3.3.2).

Consider a countable graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ without multiple edges and self-loops. Let Col_n denote the proper vertex colourings of $\mathcal{G}$ with n colours:

$$Col_n := \left\{x \in \{1, \ldots, n\}^{\mathcal{V}} \mid x_v \neq x_w \text{ whenever } (v, w) \in \mathcal{E}\right\}.$$

The *r-colourings of* $\mathbb{Z}^d$ is the space Col_r when $\mathcal{G}$ is $\mathbb{Z}^d$. The fact that r-colourings of $\mathbb{Z}^d$ with $r \geq 2d + 2$ have the pivot property is a consequence of the following:

Proposition 2.2.5. *Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be a graph with all vertices of degree at most d and $n \geq d + 2$. Then Col_n is a topological Markov field that has the pivot property.*

Proof. Given $x, y \in \Delta_{Col_n}$ we will describe a sequence of pivot moves from x to y: Let $F := \{v \in \mathcal{V} \mid x_v \neq y_v\}$. Let $x^{(0)} := x$. For every $i \in \{1, \dots, n\}$, obtain a point $x^{(i)} \in Col_n$ by pivots moves starting from $x^{(i-1)}$ as follows:

- Step 1: For every $v \in F$ such that $x_v^{(i-1)} = i$, choose $j \in \{1, \dots, n\}$ such that

  $$j \notin \left\{x_v^{(i-1)} \mid v \in \partial\{v\} \cup \{v\}\right\}.$$

  Such j exists because by our assumption on n: $|\partial\{v\} \cup \{v\}| \leq d + 1 < n$. In particular $i \neq j$. Make a pivot move by changing $x_v^{(i-1)}$ from i to j. After at most $|F|$ pivot moves we obtain the point z such that $z_v \neq i$ for all $v \in F$ and $z_v = x_v^{(i-1)}$ unless $v \in F$ and $x_v^{(i-1)} = i$.

- Step 2: For every $v \in F$ such that $y_v = i$ apply a pivot move by changing z_v to i:

  $$x'_w := \begin{cases} z_w \text{ if } w \neq v \\ i \ \text{ if } w = v. \end{cases}$$

  To see that $x' \in Col_n$, we need to check that $z_w \neq i$ for any $w \in \partial\{v\}$. Indeed if $w \in \partial\{v\} \cap F$ then $z_w \neq i$ by step 1. If $w \in \partial\{v\} \cap F^c$ then $z_w = y_w$. Because $y_v = i$ it follows that $y_w \neq i$. Now iterate with z replaced by x'. After at most $|F|$ pivot moves we obtain the point $x^{(i)}$. This configuration has the property that $x_v^{(i)} = y_v$ unless $v \in F$ and $y_v > i$.

This describes a finite sequence of pivot moves from $x = x^{(0)}$ to $x^{(n)} = y$. □

Proposition 2.2.6. *Let $X \subset \mathcal{A}^{\mathbb{Z}^d}$ be a shift-invariant topological Markov field with the pivot property. Then the dimension of $\mathbf{M}_X^{\mathbb{Z}^d}$ is finite.*

Proof. Let $(x, y) \in \Delta_X$. Let $x = x^{(1)}, x^{(2)}, x^{(3)}, \dots, x^{(k)} = y$ be a chain of pivots from x to y. Then

$$M(x, y) = \sum_{i=1}^{k-1} M(x^{(i)}, x^{(i+1)}). \tag{2.2.2}$$

If $x^{(i)}, x^{(i+1)}$ differ only at $\vec{m}_i \in \mathbb{Z}^d$ then $M(x^{(i)}, x^{(i+1)}) = M(\sigma_{-\vec{m}_i}x^{(i)}, \sigma_{-\vec{m}_i}x^{(i+1)})$ depends only on $\sigma_{-\vec{m}_i}x^{(i)}|_{\{\vec{0}\}\cup\partial\{\vec{0}\}}$ and $\sigma_{-\vec{m}_i}x^{(i+1)}|_{\{\vec{0}\}\cup\partial\{\vec{0}\}}$. Therefore the dimension of the space of shift-invariant Markov cocycles is bounded by $|\mathcal{L}_{\{\vec{0}\}\cup\partial\{\vec{0}\}}(X)|^2$. □

2.3 $\mathbb{Z}_r$-Homomorphisms, 3-Coloured Chessboards and Height Functions

Recall that a *graph-homomorphism* from the graph $\mathcal{G}_1 = (\mathcal{V}_1, \mathcal{E}_1)$ to the graph $\mathcal{G}_2 = (\mathcal{V}_2, \mathcal{E}_2)$ is a function $f : \mathcal{V}_1 \to \mathcal{V}_2$ from the vertex set of $\mathcal{G}_1$ to the vertex set of $\mathcal{G}_2$ such that f sends edges in

$\mathcal{G}_1$ to edges in $\mathcal{G}_2$. Namely, if $(v, w) \in \mathcal{E}_1$ then $(f(v), f(w)) \in \mathcal{E}_2$. We consider $\mathbb{Z}^d$ as the vertex set of the standard Cayley graph, where an edge $(\vec{n}, \vec{m})$ is present if and only if $\|\vec{n} - \vec{m}\|_1 = 1$. Also a subset $A \subset \mathbb{Z}^d$ will also denote the induced subgraph of $\mathbb{Z}^d$ on A.

For the purposes of this chapter, a *height function* on $A \subset \mathbb{Z}^d$ is a graph homomorphism from A to the standard Cayley graph of $\mathbb{Z}$. We denote the set of height functions on $A \subset \mathbb{Z}^d$ by $Ht^{(d)}(A)$:

$$Ht^{(d)}(A) := \left\{ \hat{x} \in \mathbb{Z}^A \mid |\hat{x}_{\vec{n}} - \hat{x}_{\vec{m}}| = 1 \text{ whenever } \vec{n}, \vec{m} \in A \text{ and } \|\vec{n} - \vec{m}\|_1 = 1 \right\}. \tag{2.3.1}$$

In particular, we denote

$$Ht^{(d)} := Ht^{(d)}(\mathbb{Z}^d)$$

We now introduce a certain family of $\mathbb{Z}^d$-shifts of finite type $X_r^{(d)}$, where $r, d \in \mathbb{N}$, and $r > 1$: Denote by $\mathbb{Z}_r = \mathbb{Z}/r\mathbb{Z} \cong \{0, \dots, r-1\}$ the finite cyclic group of residues modulo r. Let $\phi_r : Ht^{(d)} \to (\mathbb{Z}_r)^{\mathbb{Z}^d}$ be defined by

$$(\phi_r(\hat{x}))_{\vec{n}} := \hat{x}_{\vec{n}} \mod r \text{ for all } \vec{n} \in \mathbb{Z}^d.$$

The $\mathbb{Z}^d$-subshift $X_r^{(d)}$ is defined by:

$$X_r^{(d)} := \phi_r(Ht^{(d)}).$$

For $r = 2$ it is easily verified that $X_2^{(d)}$ consists precisely of two points x^{even}, x^{odd}. These are "chessboard configurations", given by $x^{even}_{\vec{n}} = \|\vec{n}\|_1 \mod 2$ and $x^{odd}_{\vec{n}} = \|\vec{n}\|_1 + 1 \mod 2$.

In the following, to avoid cumbersome superscripts, we will fix some dimension $d \geq 2$, and denote $Ht := Ht^{(d)}$, $X_r := X_r^{(d)}$ and $Ht(A) := Ht^{(d)}(A)$ for all $A \subset \mathbb{Z}^d$.

For $r \neq 1, 4$, there is a direct and simple interpretation for the subshift X_r as the set of graph homomorphisms from the standard Cayley graph of $\mathbb{Z}^d$ to the standard Cayley graph of $\mathbb{Z}_r$ (Proposition 2.3.1 below). In the particular case $r = 3$ the standard Cayley graph of $\mathbb{Z}_r$ is the complete graph on 3 vertices, and so X_3 is the set of proper vertex-colourings of standard Cayley graph of $\mathbb{Z}^d$ with colours in $\{0, 1, 2\}$. This relation has certainly been noticed and recorded in the literature. For instance, it is stated without proof in [20]. The general method of using height functions is attributed to J.H. Conway [57]. For the sake of completeness, we bring a self-contained proof in Proposition 2.3.1 and Lemma 2.3.2 below. The proofs below essentially follow [48, 50], where the corresponding results are obtained for the case $d = 2$, $r = 3$. Also, many ideas going into the proofs of Propostion 2.3.1, Lemma 2.3.2 and Lemma 2.3.3 come from algebraic topology; we mention the connection briefly at the end of Section 4.4 (where we generalise some of these results). Within the proof we also define a function $grad : X_r \times \mathbb{Z}^d \to \mathbb{Z}$, which we use later on.

Proposition 2.3.1. *For any $d \geq 2$ and $r \in \mathbb{N} \setminus \{1, 4\}$, X_r is a nearest neighbour shift of finite type given by*

$$X_r = \{x \in (\mathbb{Z}_r)^{\mathbb{Z}^d} \mid x_{\vec{n}} - x_{\vec{m}} = \pm 1 \mod r, \text{ whenever } \|\vec{n} - \vec{m}\|_1 = 1\}.$$

Proof. When $r = 2$, by our previous remark $X_2 = \{x^{odd}, x^{even}\}$; the proposition is easily verified in this case. From now on assume $r \in \mathbb{N} \setminus \{1, 2, 4\}$. Temporarily, let us denote

$$Y_r := \{x \in (\mathbb{Z}_r)^{\mathbb{Z}^d} \mid x_{\vec{n}} - x_{\vec{m}} = \pm 1 \mod r, \text{ whenever } \|\vec{n} - \vec{m}\|_1 = 1\}.$$

We need to establish that $Y_r = X_r$.

For all $\hat{x} \in Ht$ and $\vec{m}, \vec{n} \in \mathbb{Z}^d$ with $\|\vec{n} - \vec{m}\|_1 = 1$, by definition of Ht, we have $|\hat{x}_{\vec{n}} - \hat{x}_{\vec{m}}| = 1$. Thus, $(\phi_r(\hat{x}))_{\vec{n}} - (\phi_r(\hat{x}))_{\vec{m}} = \pm 1 \mod r$, and so $\phi_r(\hat{x}) \in Y_r$. This establishes the inclusion $X_r \subset Y_r$.

To complete the proof, given $x \in Y_r$ we will exhibit $\hat{x} \in Ht$ so that $\phi_r(\hat{x}) = x$. Choose $\vec{n}_0 \in \mathbb{Z}^d$ and define $\hat{x}_{\vec{n}_0} := x_{\vec{n}_0}$. For all $\vec{n} \in \mathbb{Z}^d$, choose a path $\vec{n}_0, \vec{n}_1 \ldots, \vec{n}_k = n$ from $\vec{n}_0$ to $\vec{n}$ in the standard Cayley graph of $\mathbb{Z}^d$, meaning, $\vec{n}_{i+1} - \vec{n}_i \in \{\pm\vec{e}_1, \ldots, \pm\vec{e}_d\}$. Define

$$\hat{x}_{\vec{n}} := \hat{x}_{\vec{0}} + \sum_{i=1}^{k} [x_{\vec{n}_i} - x_{\vec{n}_{i-1}}],$$

where for $x \in Y_r$ and $\vec{m}, \vec{n} \in \mathbb{Z}^d$ with $\|\vec{m} - \vec{n}\|_1 = 1$,

$$[x_{\vec{n}} - x_{\vec{m}}] := \begin{cases} 1 & \text{if } x_{\vec{n}} - x_{\vec{m}} = 1 \mod r \\ -1 & \text{if } x_{\vec{n}} - x_{\vec{m}} = -1 \mod r \end{cases}. \tag{2.3.2}$$

We claim that for all $x \in Y_r$ the value of $\hat{x}_{\vec{n}}$ thus obtained is independent of the path chosen from $\vec{n}_0$ to $\vec{n}$. Another way to express this is as follows:

For $x \in Y_r$ and $\vec{n} \in \{\pm\vec{e}_1, \ldots, \pm\vec{e}_d\}$ let

$$grad(x, n) := [x_{\vec{n}} - x_{\vec{0}}]. \tag{2.3.3}$$

Extend *grad* to a map $grad : Y_r \times \mathbb{Z}^d \to \mathbb{Z}$ as follows: For an arbitrary $\vec{n} \in \mathbb{Z}^d$ write $\vec{n} = \sum_{j=1}^{M} \vec{s}_j$, with $\vec{s}_j \in \{\pm\vec{e}_1, \ldots, \pm\vec{e}_d\}$. Define:

$$grad(x, \vec{n}) := \sum_{k=1}^{M} grad(\sigma^{\vec{n}_{k-1}} x, \vec{s}_k), \tag{2.3.4}$$

where $\vec{n}_k := \sum_{j=1}^{k} \vec{s}_j$ and the expressions appearing in the sum on the right-hand side of (2.3.4) are defined by (2.3.3). We will now verify that $grad(x, \vec{n})$ is well defined, which means it does not depend on the representation $\vec{n} = \sum_{j=1}^{M} \vec{s}_j$. Specifically, we will check that for all $\vec{t}_1, \ldots, \vec{t}_N, \vec{s}_1, \ldots, \vec{s}_k \in \{\pm\vec{e}_1, \ldots, \pm\vec{e}_d\}$ satisfying $\sum_{i=1}^{N} \vec{t}_i = \sum_{j=1}^{k} \vec{s}_j$ and $x \in Y_r$

$$\sum_{j=1}^{k} [x_{\vec{n}_j} - x_{\vec{n}_{j-1}}] = \sum_{j=1}^{N} [x_{\vec{m}_j} - x_{\vec{m}_{j-1}}], \tag{2.3.5}$$

where $\vec{n}_j = \sum_{i=1}^{j} \vec{s}_i$ and $\vec{m}_j = \sum_{i=1}^{j} \vec{t}_i$.

From (2.3.2) it follows that $[x_{\vec{n}} - x_{\vec{m}}] = -[x_{\vec{m}} - x_{\vec{n}}]$ whenever $\|\vec{m} - \vec{n}\|_1 = 1$. Note that $\vec{0}, \vec{n}_1, \vec{n}_2, \ldots, \vec{n}_k, \vec{m}_{N-1}, \ldots, \vec{m}_1, \vec{0}$ forms a loop in $\mathbb{Z}^d$. Since loops in $\mathbb{Z}^d$ are generated by four-cycles of the type $\vec{0}, \vec{e}_i, \vec{e}_i + \vec{e}_j, \vec{e}_j, \vec{0}$ for all $1 \le i, j \le d$ to check (2.3.5), it thus suffices to verify for all $x \in Y_r$:

$$[x_{\vec{e}_j} - x_{\vec{0}}] + [x_{\vec{e}_j+\vec{e}_i} - x_{\vec{e}_j}] = [x_{\vec{e}_i} - x_{\vec{0}}] + [x_{\vec{e}_i+\vec{e}_j} - x_{\vec{e}_i}]. \tag{2.3.6}$$

From the definition (2.3.2), the equality in (2.3.6) holds modulo r. Also,

$$[x_{\vec{e}_j} - x_{\vec{0}}], [x_{\vec{e}_j+\vec{e}_i} - x_{\vec{e}_j}], [x_{\vec{e}_i} - x_{\vec{0}}], [x_{\vec{e}_j+\vec{e}_i} - x_{\vec{e}_i}] \in \{\pm 1\}.$$

Thus (2.3.6) is a consequence of the following simple exercise:

For all $A_1, A_2, A_3, A_4 \in \{\pm 1\}$ satisfying

$$A_1 + A_2 + A_3 + A_4 = 0 \mod r, \text{ where } r \in \mathbb{N} \setminus \{1, 2, 4\},$$

we have

$$A_1 + A_2 + A_3 + A_4 = 0.$$

It now follows that for all $x \in Y_r$ and $n \in \mathbb{Z}^d$

$$\mathit{grad}(x, \vec{n}) = \hat{x}_{\vec{n}} - \hat{x}_{\vec{0}}. \tag{2.3.7}$$

In particular it follows that for all $\vec{n}, \vec{m} \in \mathbb{Z}^d$ with $\|\vec{n} - \vec{m}\|_1 = 1$, $\hat{x}_{\vec{m}} - \hat{x}_{\vec{n}} \in \{\pm 1\}$. So indeed $\hat{x} \in Ht$. It is straightforward to check that $\phi_r(\hat{x}) = x$.

$\square$

Remark: From the proof above we see that map $\mathit{grad} : X_r \times \mathbb{Z}^d \to \mathbb{Z}$ satisfies the following relation:

$$\mathit{grad}(x, \vec{n} + \vec{m}) = \mathit{grad}(x, \vec{n}) + \mathit{grad}(\sigma^{\vec{n}}(x), \vec{m}). \tag{2.3.8}$$

This means that *grad* is a *cocycle* for the shift-action on X_r. See [48, 50] for more on cocycles for X_r and other subshifts.

In the particular case when $r = 3$, $X_3^{(d)}$ is a presentation of a shift of finite type known as the d-dimensional 3-*coloured chessboard*. The subshift $X_3^{(d)}$ is the set of proper vertex-colourings of standard Cayley graph of $\mathbb{Z}^d$ with colours in $\{0, 1, 2\}$. On the first reading of the following sections, we advise the reader to keep in mind the case $r = 3$ and $d = 2$, in which $X_r^{(d)}$ is the "2-dimensional 3-coloured chessboard".

Remark: For the "exceptional" case $r = 4$, X_4 is still a shift of finite type. This can be directly deduced from the following formula:

$$\begin{aligned} X_4^{(d)} \quad = \quad & \{x \in (\mathbb{Z}_r)^{\mathbb{Z}^d} \mid x_{\vec{n}} - x_{\vec{m}} = \pm 1 \mod 4, \text{ whenever } \|\vec{n} - \vec{m}\|_1 = 1 \\ & \text{and } [x_{\vec{p}} - x_{\vec{p}+\vec{e}_i}] + [x_{\vec{p}+\vec{e}_i} - x_{\vec{p}+\vec{e}_i+\vec{e}_j}] = [x_{\vec{p}} - x_{\vec{p}+\vec{e}_j}] + [x_{\vec{p}+\vec{e}_j} - x_{\vec{p}+\vec{e}_i+\vec{e}_i}] \\ & \text{for all } 1 \le i, j \le d \text{ and } \vec{p} \in \mathbb{Z}^d\}. \end{aligned}$$

However $X_4^{(d)}$ is not a topological Markov field, as we now explain. For simplicity assume $d = 2$. Let $x, y \in X_4^{(2)}$ be the periodic points satisfying

$$y_{\vec{n}} = \sum_{i=1}^{2} \vec{n}_i \mod 4 \text{ and } x_{\vec{n}} = 2 - \sum_{i=1}^{2} \vec{n}_i \mod 4.$$

That is:

$$y = \begin{matrix} \dots & \vdots & \vdots & \vdots & \dots \\ \dots & 1 & 2 & 3 & \dots \\ \dots & 0 & \underline{1} & 2 & \dots \\ \dots & 3 & \bar{0} & 1 & \dots \\ \dots & 2 & 3 & 0 & \dots \\ \dots & \vdots & \vdots & \vdots & \dots \end{matrix} \text{ and } x = \begin{matrix} \dots & \vdots & \vdots & \vdots & \dots \\ \dots & 1 & 0 & 3 & \dots \\ \dots & 2 & \underline{1} & 0 & \dots \\ \dots & 3 & \bar{2} & 1 & \dots \\ \dots & 0 & 3 & 2 & \dots \\ \dots & \vdots & \vdots & \vdots & \dots \end{matrix} \tag{2.3.9}$$

Observe that $x_{\vec{0}} = 2$ and $y_{\vec{0}} = 0$ and $x|_{\partial\{\vec{0}\}} = y|_{\partial\{\vec{0}\}}$. However the configuration z given by

$$z_{\vec{i}} = \begin{cases} x_{\vec{0}} \text{ if } \vec{i} = \vec{0} \\ y_{\vec{i}} \text{ otherwise} \end{cases}$$

is not an element of $X_4^{(2)}$ because

$$[z_{-\vec{e}_2} - z_{-\vec{e}_2+\vec{e}_1}] + [z_{-\vec{e}_2+\vec{e}_1} - z_{\vec{e}_1}] = -2$$

while

$$[z_{-\vec{e}_2} - z_{\vec{0}}] + [z_{\vec{0}} - z_{\vec{e}_1}] = 2.$$

A similar construction works for any $d > 2$.

Lemma 2.3.2. *Fix any $d \ge 2$ and $r \in \mathbb{N} \setminus \{1, 2, 4\}$. For $x \in X_r$ any two pre-images under ϕ_r differ by a constant integer multiple of r, that is, if $\hat{x}, \hat{y} \in Ht$ satisfy $\phi_r(\hat{x}) = \phi_r(\hat{y})$ then there exists $M \in \mathbb{Z}$ so that $\hat{x}_{\vec{n}} - \hat{y}_{\vec{n}} = rM$ for all $\vec{n} \in \mathbb{Z}^d$.*

Proof. Let $x \in X_r$, $\hat{x}, \hat{y} \in Ht$ satisfy $\phi_r(\hat{x}) = \phi_r(\hat{y}) = x$. We have

$$\hat{x}_{\vec{0}} \equiv \hat{y}_{\vec{0}} \equiv x_{\vec{0}} \mod r.$$

Thus there exists $M \in \mathbb{Z}$ so that $\hat{x}_{\vec{0}} - \hat{y}_{\vec{0}} = rM$. By (2.3.7) it follows that for all $n \in \mathbb{Z}^d$

$$\hat{x}_{\vec{n}} - \hat{x}_{\vec{0}} = \hat{y}_{\vec{n}} - \hat{y}_{\vec{0}} = grad(x, \vec{n}).$$

It follows that for all $\vec{n} \in \mathbb{Z}^d$

$$\hat{x}_{\vec{n}} - \hat{y}_{\vec{n}} = \hat{x}_{\vec{0}} - \hat{y}_{\vec{0}} = rM.$$

□

Lemma 2.3.3. *Let $d \geq 2$ and $r \in \mathbb{N} \setminus \{1, 2, 4\}$. Fix any $(x, y) \in \Delta_{X_r}$ and a finite $F \subset \mathbb{Z}^d$ so that $x_{\vec{n}} = y_{\vec{n}}$ for all $\vec{n} \in \mathbb{Z}^d \setminus F$.*

1. *There exists a finite set $\tilde{F}$ so that for all $\hat{x}, \hat{y} \in Ht$ such that $\phi_r(\hat{x}) = x$ and $\phi_r(\hat{y}) = y$, there exists $M \in \mathbb{Z}$ so that $\hat{x}_{\vec{n}} - \hat{y}_{\vec{n}} = rM$ for all $\vec{n} \in \mathbb{Z}^d \setminus \tilde{F}$.*

2. *We can choose $\hat{x}, \hat{y} \in Ht$ so that $M = 0$, that is, $(\hat{x}, \hat{y}) \in \Delta_{Ht}$.*

3. *If $(\hat{x}', \hat{y}'), (\hat{x}, \hat{y}) \in \Delta_{Ht}$ and $\phi_r(\hat{x}') = \phi_r(\hat{x}) = x$, $\phi_r(\hat{y}') = \phi_r(\hat{y}) = y$ then there exists $M \in \mathbb{Z}$ so that for all $\vec{n} \in \mathbb{Z}^d$, $\hat{y}'_{\vec{n}} = \hat{y}_{\vec{n}} + rM$ and $\hat{x}'_{\vec{n}} = \hat{x}_{\vec{n}} + rM$.*

4. *For any $(\hat{x}, \hat{y}) \in \Delta_{Ht}$ and all $\vec{n} \in \mathbb{Z}^d$, $(\hat{x}_{\vec{n}} - \hat{y}_{\vec{n}}) \in 2\mathbb{Z}$.*

5. *If $x, y \in X_r$ satisfy $x_{\vec{n}} = y_{\vec{n}}$ for all $\vec{n} \in \mathbb{Z}^d \setminus \{\vec{n}_0\}$ and $x_{\vec{n}_0} \neq y_{\vec{n}_0}$ then there exists $\hat{x}, \hat{y} \in Ht$ such that $\phi_r(\hat{x}) = x, \phi_r(\hat{y}) = y$ and*

$$|\hat{x}_{\vec{n}} - \hat{y}_{\vec{n}}| = \begin{cases} 2 \text{ if } \vec{n} = \vec{n}_0 \\ 0 \text{ otherwise }. \end{cases}$$

Proof. 1. Choose $\hat{x} \in \phi_r^{-1}(x)$ and $\hat{y} \in \phi_r^{-1}(y)$. Since the set F is finite, there is an infinite connected component $\tilde{A} \subset \mathbb{Z}^d \setminus F$ in the standard Cayley graph of $\mathbb{Z}^d$ so that $\tilde{F} := \mathbb{Z}^d \setminus \tilde{A}$ is finite. Fix some $\vec{n}_0 \in \tilde{A}$. For any $\vec{n} \in \tilde{A}$ choose a path $\vec{n}_0, \vec{n}_1 \ldots, \vec{n}_k = \vec{n} \in \tilde{A}$ from $\vec{n}_0$ to $\vec{n}$ in the standard Cayley graph of $\mathbb{Z}^d$, that is, $\vec{n}_{j+1} - \vec{n}_j \in \{\pm\vec{e}_1, \ldots, \vec{e}_d\}$ for all $1 \leq j \leq k-1$ and $x_j = y_j$ for all $1 \leq j \leq k$. Using (2.3.3) and (2.3.4) we conclude that

$$grad(\sigma^{\vec{n}_0}(x), \vec{n} - \vec{n}_0) = grad(\sigma^{\vec{n}_0}(y), \vec{n} - \vec{n}_0).$$

It now follows using (2.3.7) that

$$\hat{x}_{\vec{n}} - \hat{x}_{\vec{n}_0} = \hat{y}_{\vec{n}} - \hat{y}_{\vec{n}_0}.$$

Because $x_{\vec{n}_0} = y_{\vec{n}_0}$, $\hat{x}_{\vec{n}_0} - \hat{y}_{\vec{n}_0} \in r\mathbb{Z}$, and so there exists $M \in \mathbb{Z}$ so that $\hat{x}_{\vec{n}} - \hat{y}_{\vec{n}} = rM$ for all $\vec{n} \in \tilde{A} = \mathbb{Z}^d \setminus \tilde{F}$.

2. Define $\hat{z}$ by

$$\hat{z}_{\vec{n}} := \hat{y}_{\vec{n}} - (\hat{y}_{\vec{n}_0} - \hat{x}_{\vec{n}_0}) \text{ for all } \vec{n} \in \mathbb{Z}^d.$$

   Obviously $\hat{z} \in Ht$. Since $x_{\vec{n}_0} = y_{\vec{n}_0}$ it follows that $\hat{y}_{\vec{n}_0} - \hat{x}_{\vec{n}_0} \in r\mathbb{Z}$. Thus $\phi_r(\hat{z}) = \phi_r(\hat{y}) = y$. Also $\hat{z}_{\vec{n}} = \hat{y}_{\vec{n}}$ for all $\vec{n} \in \tilde{A}$. Thus, $(\hat{x}, \hat{z}) \in \Delta_{Ht}$, proving the second assertion.

3. By Lemma 2.3.2, for any choice of $\hat{x}' \in \phi_r^{-1}(x)$ there exists $M_1 \in \mathbb{Z}$ so that $\hat{x}'_{\vec{n}} = \hat{x}_{\vec{n}} + rM_1$ for all $\vec{n} \in \mathbb{Z}^d$. Similarly, for any choice of $\hat{y}' \in \phi_r^{-1}(y)$ there exists $M_2 \in \mathbb{Z}$ so that $\hat{y}'_{\vec{n}} = \hat{y}_{\vec{n}} + rM_2$ for all $\vec{n} \in \mathbb{Z}^d$. But if $(\hat{x}', \hat{y}') \in \Delta_{Ht}$ it follows that $M_1 = M_2$. This proves the third assertion.

4. From (2.3.3) it follows that $grad(x, \pm\vec{e}_j) = 1 \mod 2$ for all $x \in X_r$ and $j \in \{1, \ldots, d\}$. Thus by (2.3.4) the parity of $grad(x, \vec{n})$ is equal to the parity of $\|\vec{n}\|_1$ and does not depend on x. Therefore if $\hat{x}_{\vec{n}_0} = \hat{y}_{\vec{n}_0}$ for some $\vec{n}_0 \in \mathbb{Z}^d$, from (2.3.7) it follows that $\hat{x}_{\vec{n}} - \hat{y}_{\vec{n}} \in 2\mathbb{Z}$ for all $\vec{n} \in \mathbb{Z}^d$.

5. Suppose that $x, y \in X_r$ satisfy $x_{\vec{n}} = y_{\vec{n}}$ for all $\vec{n} \in \mathbb{Z}^d \setminus \{\vec{n}_0\}$ and $x_{\vec{n}_0} \neq y_{\vec{n}_0}$. Choose $(\hat{x}, \hat{y}) \in \Delta_{Ht}$ so that $\phi_r(\hat{x}) = x$ and $\phi_r(\hat{y}) = y$. By the argument above $\hat{x}_{\vec{n}} = \hat{y}_{\vec{n}}$ for all $\vec{n} \neq \vec{n}_0$. Let $\vec{m} := \vec{n}_0 + \vec{e}_1$. Since $x_{\vec{m}} = y_{\vec{m}}$ and $x_{\vec{n}_0} \neq y_{\vec{n}_0}$ it follows that $[x_{\vec{n}_0} - x_{\vec{m}}] \neq [y_{\vec{n}_0} - y_{\vec{m}}]$, thus

$$|[x_{\vec{n}_0} - x_{\vec{m}}] - [y_{\vec{n}_0} - y_{\vec{m}}]| = 2.$$

   Using (2.3.3) and (2.3.7), we conclude that $|\hat{x}_{\vec{n}_0} - \hat{y}_{\vec{n}_0}| = 2$.

□

It is a known and useful fact that the 3-coloured chessboard has the pivot property. This can be shown, for instance, using height functions. Essentially the same argument shows that X_r has the pivot property for all $r \in \mathbb{N} \setminus \{1, 2, 4\}$. We include a short proof below. Similar arguments appear in the proofs of certain claims in the subsequent sections.

Proposition 2.3.4. *For any $d \geq 2$ and $r \in \mathbb{N} \setminus \{1, 2, 4\}$ the subshift X_r has the pivot property. In other words, given any $(x, y) \in \Delta_{X_r}$ there exist $x = z^{(0)}, z^{(1)}, \ldots, z^{(N)} = y \in X_r$ such that for all $0 \leq k < N$, there is a unique $n_k \in \mathbb{Z}^d$ for which $z^{(k)}_{n_k} \neq z^{(k+1)}_{n_k}$.*

This will be further generalised as Theorem 4.1.4.

Proof. Fix $(x, y) \in \Delta_{X_r}$. By Lemma 2.3.3, we can choose $(\hat{x}, \hat{y}) \in \Delta_{Ht}$ with $\phi_r(\hat{x}) = x$ and $\phi_r(\hat{y}) = y$. We will proceed by induction on $D = \sum_{\vec{n} \in \mathbb{Z}^d} |\hat{x}_{\vec{n}} - \hat{y}_{\vec{n}}|$. Note that by Lemma 2.3.3 D is well defined, that is, the differences $(\hat{x}_{\vec{n}} - \hat{y}_{\vec{n}})$ in the sum do not depend on the choice of $(\hat{x}, \hat{y})$. When $D = 0$, then $x = y$ and the claim is trivial. Now, suppose $D > 0$. Let

$$\begin{aligned} F_+ &= \{\vec{n} \in \mathbb{Z}^d \mid (\hat{x}_{\vec{n}} - \hat{y}_{\vec{n}}) > 0\} \\ F_- &= \{\vec{n} \in \mathbb{Z}^d \mid (\hat{x}_{\vec{n}} - \hat{y}_{\vec{n}}) < 0\}, \end{aligned}$$

that is, $F_+ \subset \mathbb{Z}^d$ is the finite set of sites where $\hat{x}$ is strictly above $\hat{y}$ and $F_- \subset \mathbb{Z}^d$ is the finite set of sites where $\hat{y}$ is strictly above $\hat{x}$. Without loss of generality assume that F_+ is non-empty. Since $(\hat{x}_{\vec{n}} - \hat{y}_{\vec{n}}) \in 2\mathbb{Z}$, $\hat{x}_{\vec{n}} - \hat{y}_{\vec{n}} \geq 2$ for all $\vec{n} \in F_+$. Consider $\vec{n}_0 \in F_+$ such that $\hat{x}_{\vec{n}_0} = \max\{\hat{x}_{\vec{n}} \mid \vec{n} \in F_+\}$. It follows that $\hat{x}_{\vec{n}_0} - \hat{x}_{\vec{m}} = 1$ for all $\vec{m}$ neighbouring $\vec{n}_0$. We can thus define $\hat{z} \in Ht$ which is equal to $\hat{x}$ everywhere except at $\vec{n}_0$, where $\hat{z}_{\vec{n}_0} = \hat{x}_{\vec{n}_0} - 2$. Now set $z^{(1)} = \phi_r(\hat{z})$ and apply the induction hypothesis on $(z^{(1)}, y)$. □

2.4 Markov Cocycles on X_r

Our goal in the current section is to describe the space of shift-invariant Markov cocycles on X_r, when $r \in \mathbb{N} \setminus \{1, 2, 4\}$, and the subspace of Gibbs cocycles for shift-invariant nearest neighbour interactions.

In the following we assume $r \in \mathbb{N} \setminus \{1, 2, 4\}$ and $d > 1$, unless explicitly stated otherwise.

Lemma 2.4.1. *Fix $r \in \mathbb{N} \setminus \{1, 2, 4\}$ and $d > 1$. Let $F \subset \mathbb{Z}^d$ be a finite set and $x, y, z, w \in X_r$ such that $x|_{F^c} = y|_{F^c}$, $z|_{F^c} = w|_{F^c}$ and $x|_{F \cup \partial F} = z|_{F \cup \partial F}$, $y|_{F \cup \partial F} = w|_{F \cup \partial F}$. Consider $(\hat{x}, \hat{y}), (\hat{z}, \hat{w}) \in \Delta_{Ht}$ such that they are mapped by ϕ_r to the pairs $(x, y), (z, w) \in \Delta_{X_r}$ respectively. Then $\hat{x}_{\vec{n}} - \hat{y}_{\vec{n}} = \hat{z}_{\vec{n}} - \hat{w}_{\vec{n}}$ for all $\vec{n} \in \mathbb{Z}^d$.*

Proof. Let $F_0 \subset \mathbb{Z}^d$ denote the infinite connected component of F^c. For $\vec{n} \in F_0$, we clearly have $\hat{x}_{\vec{n}} - \hat{y}_{\vec{n}} = \hat{z}_{\vec{n}} - \hat{w}_{\vec{n}} = 0$. We can now prove by induction on the distance from $\vec{n} \in \mathbb{Z}^d$ to F_0 that $\hat{x}_{\vec{n}} - \hat{y}_{\vec{n}} = \hat{z}_{\vec{n}} - \hat{w}_{\vec{n}}$. Given $\vec{n} \in \mathbb{Z}^d \setminus F_0$, find a neighbour $\vec{m}$ of $\vec{n}$ which is closer to F_0. By the induction hypothesis, $\hat{x}_{\vec{m}} - \hat{y}_{\vec{m}} = \hat{z}_{\vec{m}} - \hat{w}_{\vec{m}}$.

If either $\vec{n} \in F$ or $\vec{m} \in F$, then both $\vec{m}$ and $\vec{n}$ are in $F \cup \partial F$ and so $x_{\vec{n}} - x_{\vec{m}} = z_{\vec{n}} - z_{\vec{m}}$ and $y_{\vec{n}} - y_{\vec{m}} = w_{\vec{n}} - w_{\vec{m}}$. (2.3.2) and (2.3.7) imply $\hat{x}_{\vec{n}} - \hat{x}_{\vec{m}} = \hat{z}_{\vec{n}} - \hat{z}_{\vec{m}}$ and $\hat{y}_{\vec{n}} - \hat{y}_{\vec{m}} = \hat{w}_{\vec{n}} - \hat{w}_{\vec{m}}$. Subtracting the equations and applying the induction hypothesis, we conclude in this case that $\hat{x}_{\vec{n}} - \hat{y}_{\vec{n}} = \hat{z}_{\vec{n}} - \hat{w}_{\vec{n}}$.

Otherwise, $\vec{n}, \vec{m} \in F^c$ and so $x_{\vec{n}} - x_{\vec{m}} = y_{\vec{n}} - y_{\vec{m}}$ and $z_{\vec{n}} - z_{\vec{m}} = w_{\vec{n}} - w_{\vec{m}}$, and again by (2.3.2) and (2.3.7) we conclude that $\hat{x}_{\vec{n}} - \hat{y}_{\vec{n}} = \hat{z}_{\vec{n}} - \hat{w}_{\vec{n}}$. □

For $i \in \mathbb{Z}_r$ and integers a, b with $a - b \in 2\mathbb{Z}$, let

$$N_i(a, b) := \begin{cases} |\{m \in (2\mathbb{Z} + a) \cap (r\mathbb{Z} + i) \mid m \in [a, b)\}| & \text{if } a \leq b \\ -N_i(b, a) & \text{otherwise} \end{cases}$$

Here N_i is the "net" number of crossings from $(i + r\mathbb{Z})$ to $(i + 2 + r\mathbb{Z})$ in a path going from a to b in steps of magnitude 2. Note that

$$N_i(a, b) = N_i(a + rn, b + rn) \tag{2.4.1}$$

for all $a, b, n \in \mathbb{Z}$ and

$$N_i(a, b) = N_i(a + c, b + c) \tag{2.4.2}$$

for all $a, b, c \in \mathbb{Z}$ such that $a - b \in r\mathbb{Z} \cap 2\mathbb{Z}$.

Proposition 2.4.2. *For any $r \in \mathbb{N} \setminus \{1, 2, 4\}$ and $d > 1$, the space $\mathcal{M}_{X_r}^{\mathbb{Z}^d}$ of shift-invariant Markov cocycles on X_r has a linear basis*

$$\{M_0, M_1, \ldots, M_{r-1}\},$$

where the cocycle M_i is given by

$$M_i(x, y) := \sum_{n \in \mathbb{Z}^d} N_i(\hat{x}_n, \hat{y}_n); \tag{2.4.3}$$

$(\hat{x}, \hat{y}) \in \Delta_{Ht}$ is any pair which is mapped to (x, y) by ϕ_r. In particular $\dim(\mathcal{M}_{X_r}^{\mathbb{Z}^d}) = r$.

Proof. By Lemma 2.3.3 different choices $(\hat{x}, \hat{y}) \in \Delta_{Ht}$ which map to (x, y) via ϕ_r differ by a fixed translation in $r\mathbb{Z}$. Thus by (2.4.1) the values $M_i(x, y)$ are independent of the choice of the corresponding height functions $(\hat{x}, \hat{y}) \in \Delta_{Ht}$ and hence are well defined.

We will show that for $i = 0, \ldots, r - 1$, M_i is indeed a Markov cocycle. Since $N_i(a, c) = N_i(a, b) + N_i(b, c)$ whenever $a \equiv b \equiv c \mod 2$, it follows that $M_i(x, z) = M_i(x, y) + M_i(y, z)$ whenever $x, y, z \in X_r$ are homoclinic. Thus M_i is a Δ_{X_r}-cocycle. Clearly, M_i is shift-invariant.

We now check that M_i satisfies the Markov property. This is equivalent to showing that $M_i(x, y) = M_i(z, w)$ whenever $x, y, z, w \in X_r$ satisfy the assumption in Lemma 2.4.1. In this case by Lemma 2.4.1, $\hat{x}_{\vec{n}} - \hat{y}_{\vec{n}} = \hat{z}_{\vec{n}} - \hat{w}_{\vec{n}}$ for all $\vec{n} \in \mathbb{Z}^d$. Also note that for any $\vec{n} \in \mathbb{Z}^d$, either $x_{\vec{n}} = z_{\vec{n}}$ and $y_{\vec{n}} = w_{\vec{n}}$ in which case $\hat{x}_{\vec{n}} - \hat{z}_{\vec{n}} = \hat{y}_{\vec{n}} - \hat{w}_{\vec{n}} \in r\mathbb{Z}$ or $x_{\vec{n}} = y_{\vec{n}}$ and $z_{\vec{n}} = w_{\vec{n}}$, in which case by Lemma 2.3.3 $\hat{x}_{\vec{n}} - \hat{y}_{\vec{n}} = \hat{z}_{\vec{n}} - \hat{w}_{\vec{n}} \in r\mathbb{Z} \cap 2\mathbb{Z}$. By (2.4.1) and (2.4.2) in either case $N_i(\hat{x}_{\vec{n}}, \hat{y}_{\vec{n}}) = N_i(\hat{z}_{\vec{n}}, \hat{w}_{\vec{n}})$ and summing over the $\vec{n}$'s, we get $M_i(x, y) = M_i(z, w)$ as required.

To complete the proof we need to show that all shift-invariant Markov cocycles on X_r can be uniquely written as a linear combination of $M_0, \ldots, M_{r-1}$. For $i \in \{0, \ldots, r-1\}$ let $(x^{(i)}, y^{(i)}) \in \Delta_{X_r}$ such that $x^{(i)}_{\vec{0}} = i$, $y^{(i)}_{\vec{0}} = i + 2 \mod r$ and $x^{(i)}_{\vec{n}} = y^{(i)}_{\vec{n}}$ for all $\vec{n} \in \mathbb{Z}^d \setminus \{\vec{0}\}$. Given a shift-invariant Markov cocycle M, let

$$\alpha_i := M(x^{(i)}, y^{(i)}) \tag{2.4.4}$$

We claim that for all $(x, y) \in \Delta_{X_r}$:

$$M(x, y) = \sum_{i=0}^{r-1} \alpha_i \cdot M_i(x, y). \tag{2.4.5}$$

Since X_r has the pivot property (Proposition 2.3.4), by (2.2.2) it is sufficient to show that (2.4.5) holds for all $(x, y) \in \Delta_{X_r}$ which differ only at a single site. By shift-invariance of M and the M_i's it is further enough to show this for (x, y) which differ only at the origin $\vec{0}$. In this case, we note that (x,y) coincide with either $(x^{(i)},y^{(i)})$ or $(y^{(i)},x^{(i)})$ on the sites $\{\vec{0}\}\cup\partial\{\vec{0}\}$ for some i. Without loss of generality assume that (x,y) coincide with $(x^{(i_0)},y^{(i_0)})$ on the sites $\{\vec{0}\}\cup\partial\{\vec{0}\}$. Since M and the M_i's are Markov cocycles we have $M(x,y)=M(x^{(i_0)},y^{(i_0)})=\alpha_{i_0}$ and

$$\sum_{j=0}^{r-1}\alpha_j\cdot M_j(x,y)=\sum_{j=0}^{r-1}\alpha_j M_j(x^{(i_0)},y^{(i_0)})=\sum_{j=0}^{r-1}\alpha_j\delta_{i_0,j}=\alpha_{i_0}.$$

$\square$

Remark: Without the assumption of shift-invariance, a similar argument shows that any Markov cocycle on X_r is of the following form:

$$M(x,y)=\sum_{i=0}^{r-1}\sum_{\vec{n}\in\mathbb{Z}^d}a_{i,\vec{n}}N_i(\hat{x}_{\vec{n}},\hat{y}_{\vec{n}})\text{ with }a_{i,\vec{n}}\in\mathbb{R}\text{ for all }\vec{n}\in\mathbb{Z}^d,\ 0\le i\le r-1.$$

We now describe the space $\mathbf{G}_{X_r}^{\mathbb{Z}^d}$ of Gibbs cocycles corresponding to shift-invariant nearest neighbour interactions for X_r.

Proposition 2.4.3. *Let $r\in\mathbb{N}\setminus\{1,2,4\}$ and $d>1$. A shift-invariant Markov cocycle on X_r is a Gibbs cocycle corresponding to a shift-invariant nearest neighbour interaction if and only if it is of the form $M=\sum_{i=0}^{r-1}\alpha_i M_i$, with $\sum_{i=0}^{r-1}\alpha_i=0$ and M_i's as in Proposition 2.4.2 and $\alpha_0,\dots,\alpha_{r-1}$ given by (2.4.4). In other words,*

$$\mathbf{G}_{X_r}^{\mathbb{Z}^d}=\left\{\sum_{i=0}^{r-1}\alpha_i M_i\mid\sum_{i=0}^{r-1}\alpha_i=0\right\}.$$

In particular, $\dim(\mathbf{G}_{X_r}^{\mathbb{Z}^d})=r-1$.

Proof. Let M be a Gibbs cocycle given by a shift-invariant nearest neighbour interaction ϕ. Choose $(x^{(i)},y^{(i)})\in\Delta_{X_r}$ as in the proof of Proposition 2.4.2 so that $\alpha_i=M(x^{(i)},y^{(i)})$. Expanding the Gibbs cocycle we have:

$$\begin{aligned}M(x^{(i)},y^{(i)}) &= \phi([i+2]_0)-\phi([i]_0)\\ &+\sum_{j=1}^{d}\left(\phi([i+2,i+1]_j)-(\phi([i,i+1]_j)+\phi([i+1,i+2]_j)-\phi([i+1,i]_j)\right).\end{aligned}$$

Summing these equations over i we get:

$$\sum_{i=0}^{r-1}\alpha_i=\sum_{i=0}^{r-1}M(x^{(i)},y^{(i)})=0$$

Conversely, for any values $\alpha_i = M(x^{(i)}, y^{(i)})$ such that $\sum_{i=0}^{r-1} \alpha_i = 0$ it is easy to see that there is a corresponding nearest neighbour shift-invariant interaction ϕ: For instance, set $\phi([i, i+1]_1) = -\sum_{k=i}^{r-1} \alpha_k$, $\phi([i, i+1]_j) = 0$ for $j = 2, \dots, d$ and $\phi([i+1, i]_j) = \phi([i]_0) = 0$ for $j = 1, \dots, d$ and $i = 0, \dots, r-1$. $\square$

Let $\hat{M} : \Delta_{X_r} \to \mathbb{R}$ be the Markov cocycle given by

$$\hat{M}(x, y) := \sum_{n \in \mathbb{Z}^d} \hat{y}_n - \hat{x}_n, \tag{2.4.6}$$

where $(\hat{x}, \hat{y}) \in \Delta_{Ht}$ satisfy $\phi_r(\hat{x}) = x$ and $\phi_r(\hat{y}) = y$. By the following we observe that $\hat{M}(x, y) = 2\sum_{i=0}^{r-1} M_i(x, y)$ for all $(x, y) \in \Delta_{X_r}$:

As in the proof of Proposition 2.4.2 it sufficient to verify this for $(x, y) = (x^{(i_0)}, y^{(i_0)})$ where $0 \le i_0 \le r-1$. In that case

$$\hat{M}(x^{(i_0)}, y^{(i_0)}) = \hat{y}_0^{(i_0)} - \hat{x}_0^{(i_0)} = 2$$

and

$$2\sum_{i=0}^{r-1} M_i(x^{(i_0)}, y^{(i_0)}) = 2\sum_{i=0}^{r-1} \delta_{i_0, i} = 2.$$

Corollary 2.4.4. *Let $r \in \mathbb{N} \setminus \{1, 2, 4\}$ and $d > 1$. Any shift-invariant Markov cocycle M on X_r can be uniquely written as*

$$M = M_0 + \alpha \hat{M}$$

where M_0 is some Gibbs cocycle, $\alpha \in \mathbb{R}$ and $\hat{M}$ is given by (2.4.6).

Thus, the conclusion of the second part of the strong version of the Hammersley-Clifford Theorem regarding shift-invariant Markov cocycles fails for X_r. Our next proposition asserts that the conclusion of the first part of the strong version of the Hammersley-Clifford Theorem still holds for X_r. This immediately implies the conclusion of the first part of the weak version of the Hammersley-Clifford Theorem of X_r.

Proposition 2.4.5. *($\mathbf{M}_{X_r} = \mathbf{G}_{X_r}$) Let $r \in \mathbb{N} \setminus \{1, 2, 4\}$ and $d > 1$. Let $M : \Delta_{X_r} \to \mathbb{R}$ be a Markov cocycle. There exists a nearest neighbour interaction ϕ, which is not necessarily shift-invariant, so that $M = M_\phi$.*

Proof. Given $M \in \mathbf{M}_{X_r}$, we will define a compatible nearest neighbour interaction ϕ as follows:

The interaction ϕ will assign 0 to any single site pattern. For $\vec{n} = (n_1 \dots, n_d) \in \mathbb{Z}^d$, and $1 \le j \le d$, let $\phi_{\vec{n},j}(a, b)$ denote the weight the interaction ϕ assigns to the pattern (a, b) on the edge $(\vec{n}, \vec{n} + e_j)$. Set $\phi_{\vec{n},j}(a, b) = 0$ whenever $j \in \{2, \dots, d\}$. For $\vec{n} = (n_1, \dots, n_d) \in \mathbb{Z}^d$ and $i \in \mathbb{Z}_r$ define recursively

$$\phi_{\vec{n},1}(i,i+1) := \begin{cases} 0 & n_1 \leq 0 \\ M(\sigma^{\vec{n}} y^{(i)}, \sigma^{\vec{n}} x^{(i)}) + \phi_{\vec{n}-\vec{e}_1,1}(i+1,i+2) & n_1 > 0 \end{cases}$$

$$\phi_{\vec{n},1}(i+1,i) := \begin{cases} 0 & n_1 \geq 0 \\ M(\sigma^{\vec{n}+\vec{e}_1} y^{(i)}, \sigma^{\vec{n}+\vec{e}_1} x^{(i)}) + \phi_{\vec{n}+\vec{e}_1,1}(i+2,i+1) & n_1 < 0 \end{cases}$$

whereas in the proof of Proposition 2.4.2, $(x^{(i)}, y^{(i)}) \in \Delta_{X_r}$ are such that $x^{(i)}_{\vec{0}} = i \mod r$, $y^{(i)}_{\vec{0}} = i+2 \mod r$ and $x^{(i)}_{\vec{n}} = y^{(i)}_{\vec{n}}$ for all $\vec{n} \in \mathbb{Z}^d \setminus \{\vec{0}\}$. To see that ϕ defines a nearest neighbour interaction for M, since X_r has the pivot property (Proposition 2.3.4) it is sufficient to verify by (2.2.2)

$$M(y,x) = \sum_{\vec{n} \in \mathbb{Z}^d} \sum_{j=1}^{d} \phi_{\vec{n},j}(x_{\vec{n}}, x_{\vec{n}+\vec{e}_j}) - \phi_{\vec{n},j}(y_{\vec{n}}, y_{\vec{n}+\vec{e}_j}). \tag{2.4.7}$$

for $(y,x) \in \Delta_{X_r}$ which differ at a single site $n' \in \mathbb{Z}^d$.

Then (y,x) coincide with either $(\sigma^{\vec{n}'} y^{(i)}, \sigma^{\vec{n}'} x^{(i)})$ or $(\sigma^{\vec{n}'} x^{(i)}, \sigma^{\vec{n}'} y^{(i)})$ on the sites $\{\vec{n}'\} \cup \partial\{\vec{n}'\}$ for some $0 \leq i \leq r-1$. Without loss of generality assume that (y,x) coincide with $(\sigma^{\vec{n}'} y^{(i_0)}, \sigma^{\vec{n}'} x^{(i_0)})$ on the sites $\{n'\} \cup \partial\{n'\}$ for some $0 \leq i_0 \leq r-1$. Since M is a Markov cocycle

$$\begin{aligned} M(y,x) &= M(\sigma^{\vec{n}'} y^{(i_0)}, \sigma^{\vec{n}'} x^{(i_0)}) \\ &= \begin{cases} \phi_{\vec{n}',1}(i_0, i_0+1) - \phi_{\vec{n}'-\vec{e}_1,1}(i_0+1, i_0+2) \text{ if } n'_1 > 0 \\ \phi_{\vec{n}'-\vec{e}_1,1}(i_0+1, i_0) - \phi_{\vec{n}',1}(i_0+2, i_0+1) \text{ if } n'_1 \leq 0 \end{cases} \\ &= \sum_{\vec{n} \in \mathbb{Z}^d} \sum_{j=1}^{d} \phi_{\vec{n},j}(x_{\vec{n}}, x_{\vec{n}+e_j}) - \phi_{\vec{n},j}(y_{\vec{n}}, y_{\vec{n}+e_j}). \end{aligned}$$

□

Combining the above results we obtain:

Corollary 2.4.6. *Let $r \in \mathbb{N} \setminus \{1,2,4\}$ and $d > 1$. There exists a shift-invariant Markov cocycle on X_r which is given by a nearest neighbour interaction but not by a shift-invariant nearest neighbour interaction, that is,*

$$\mathbf{G}^{\mathbb{Z}^d}_{X_r} \neq \mathbf{G}_{X_r} \cap \mathbf{M}^{\mathbb{Z}^d}_{X_r}.$$

2.5 Markov Random Fields on X_r Are Gibbs

Our main goal is to prove the following result:

Theorem 2.5.1. *Let $r \in \mathbb{N} \setminus \{1, 2, 4\}$ and $d > 1$. Any shift-invariant Markov random field adapted to X_r is a Gibbs state for some shift-invariant nearest neighbour interaction. In particular any shift-invariant Markov random field μ with supp$(\mu) = X_r$ is a Gibbs state for some shift-invariant nearest neighbour interaction.*

Theorem 2.5.1 implies that the conclusion of the second part of the weak version of the Hammersley-Clifford Theorem holds for X_r, that is, Theorem 1.1.1 although the argument is very different from the safe-symbol case.

For a subshift X, a point $x \in X$ will be called *frozen* if its homoclinic class is a singleton set. This notion coincides with the notion of frozen colouring in [6]. By Proposition 2.3.4, X_r has the pivot property so $x \in X_r$ is frozen if and only if for every $\vec{n} \in \mathbb{Z}^d$, $x_{\vec{j}} \neq x_{\vec{k}}$ for some $\vec{j}, \vec{k} \in \partial\{\vec{n}\}$, that is, any site is adjacent to at least two sites with distinct symbols. A subshift X will be called *frozen* if it consists of frozen points, equivalently Δ_X is the diagonal. A measure on a subshift X will be called *frozen* if its support consists of frozen points. Note that the collection of frozen points of a given topological Markov field X is itself a topological Markov field.

We derive Theorem 2.5.1 as an immediate corollary of the following proposition:

Proposition 2.5.2. *Let $r \in \mathbb{N} \setminus \{1, 2, 4\}$ and $d > 1$. Let μ be a shift-invariant Markov random field adapted to X_r with Radon-Nikodym cocycle equal to the restriction of e^M to its support where $M \in \mathbf{M}^{\mathbb{Z}^d}_{X_r} \setminus \mathbf{G}^{\mathbb{Z}^d}_{X_r}$ is a Markov cocycle which is not given by a shift-invariant nearest neighbour interaction. Then μ is frozen.*

Note that any frozen probability measure is Gibbs with any nearest neighbour interaction because the homoclinic relation of the support of the measure is trivial. Hence this proposition proves that any shift-invariant Markov random field adapted to X_r is Gibbs for some nearest neighbour interaction and thus implies Theorem 2.5.1. The intuition behind the proof of this proposition is the following: For a Markov cocycle $M = \sum_{i=1}^{r} \alpha_i M_i$ the condition $\sum_{i=1}^{r} \alpha_i > 0$ indicates an inclination to raise the height function. However shift-invariance implies the existence of a well defined "slope" for the height function in all directions. Unless this slope is extremal, that is, maximal ($\pm\|\vec{n}\|_1$) in some direction $\vec{n} \in \mathbb{Z}^d \setminus \{\vec{0}\}$, this will lead to a contradiction.

In preparation for the proof, we set up some auxiliary results.

2.5.1 Real Valued Cocycles for Measure-Preserving $\mathbb{Z}^d$-Actions

We momentarily pause our discussion about Markov random fields on X_r to discuss cocycles for measure-preserving $\mathbb{Z}^d$ actions. Subcocycles and further generalisations will be discussed in Section 4.5. Let $(X, \mathcal{F}, \mu, T)$ be an ergodic measure-preserving $\mathbb{Z}^d$-action. A measurable function $c : X \times \mathbb{Z}^d \to \mathbb{R}$ is called a *T-cocycle* if it satisfies the following equation μ-almost everywhere with respect to $x \in X$:

$$c(x, \vec{n} + \vec{m}) = c(x, \vec{n}) + c(T^{\vec{n}}x, \vec{m}) \ \forall \vec{n}, \vec{m} \in \mathbb{Z}^d. \tag{2.5.1}$$

By (2.3.8) the function *grad* : $X_r \times \mathbb{Z}^d \to \mathbb{R}$ defined in (2.3.3) and (2.3.4) is indeed a shift-cocycle with respect to the shift action on X_r.

We will use the following lemma:

Lemma 2.5.3. *Let $(X, \mathcal{F}, \mu, T)$ be an ergodic measure-preserving $\mathbb{Z}^d$ action and $c : X \times \mathbb{Z}^d \to \mathbb{R}$ be a measurable cocycle such that for all $\vec{n} \in \mathbb{Z}^d$ the function $f_{c,\vec{n}}(x) := c(x, \vec{n})$ is in $L^1(\mu)$, then for all $\vec{n} = (n_1, n_2, \ldots, n_d) \in \mathbb{Z}^d$*

$$\begin{aligned} \lim_{k\to\infty} \frac{c(x, k\vec{n})}{k} &= \int c(x, \vec{n}) d\mu(x) \\ &= \sum_{i=1}^{d} n_i \int c(x, \vec{e}_i) d\mu(x). \end{aligned}$$

The convergence holds almost everywhere with respect to μ and also in $L^1(\mu)$.

Proof. By the cocycle equation (2.5.1) for μ-almost every $x \in X$, any $k \in \mathbb{N}$ and $n \in \mathbb{Z}^d$ we have:

$$c(x, k \cdot \vec{n}) = \sum_{i=0}^{k-1} c(T^{i\vec{n}} x, \vec{n}).$$

The existence of almost everywhere and L^1 limit $\bar{f}(x) := \lim_{k\to\infty} \frac{c(x,k\vec{n})}{k}$ follows from the pointwise and L^1 ergodic theorems. To complete the proof we need to show that the limit is constant almost everywhere. We do this by showing that the limit is T-invariant. By the cocycle equation (2.5.1), for almost every $x \in X$ and any $m, n \in \mathbb{Z}^d$ and $k \in \mathbb{N}$ we have:

$$\begin{aligned} c(x, k\vec{n}) = & \quad c(x, \vec{m} + k\vec{n} - \vec{m}) \\ = & \quad c(x, \vec{m}) + c(T^{\vec{m}} x, k\vec{n} - \vec{m}) \\ = & \quad c(x, \vec{m}) + c\left(T^{\vec{m}} x, k\vec{n}\right) + c\left(T^{\vec{m}+k\vec{n}} x, -\vec{m}\right). \end{aligned}$$

Thus,

$$|\bar{f}(x) - \bar{f}(T^{\vec{m}}(x))| \leq \limsup_{k\to\infty} \frac{1}{k}\left(|c(x, \vec{m})| + |c(T^{\vec{m}+k\cdot\vec{n}} x, -\vec{m})|\right) \leq \limsup_{k\to\infty} \frac{1}{k}|f_{c,\vec{m}}(x)| + \frac{1}{k}|f_{c,-\vec{m}}(T^{\vec{n}k} T^{\vec{m}} x)|.$$

Since $f_{c,\vec{m}}, f_{c,-\vec{m}} \in L^1(\mu)$, $\limsup_{k\to\infty} \frac{1}{k}|f_{c,\vec{m}}(x)|$ and $\limsup_{k\to\infty} \frac{1}{k}|f_{c,-\vec{m}}(T^{k\vec{n}} T^{\vec{m}} x)|$ are both equal to 0 almost everywhere (the second term vanishes because $\lim_{k\to\infty} \frac{1}{k} g(S^k x) = 0$ a.e for $g \in L^1$ and S measure-preserving). Therefore,

$$\begin{aligned} \lim_{k\to\infty} \frac{c(x, k \cdot \vec{n})}{k} &= \int c(x, \vec{n}) d\mu(x) \\ &= \sum_{i=1}^{d} n_i \int c(x, \vec{e}_i) d\mu(x). \end{aligned}$$

$\square$

Remark: In the specific case that T is *totally ergodic*, meaning that the individual action of each $T^{\vec{n}}$ is ergodic for all $\vec{n} \in \mathbb{Z}^d \setminus \{\vec{0}\}$, the lemma above is completely obvious since $\frac{c(x,k\cdot\vec{n})}{k} = \frac{1}{k}\sum_{j=0}^{k-1} c(T^{j\vec{n}}x, \vec{n})$, which is an ergodic average. The point of Lemma 2.5.3 is that ergodicity of the $\mathbb{Z}^d$-action is sufficient for the limit to be constant.

The cocycle $grad : X_r \times \mathbb{Z}^d \to \mathbb{Z}$ is not only measurable but also continuous. We can use this, along with compactness of X_r and the unit ball in $\mathbb{R}^d$ to obtain uniformity of the convergence with respect to the "direction" on a set of full measure.

For convenience we extend the definition of $grad : X \times \mathbb{Z}^d \to \mathbb{R}$ given by (2.3.3) and (2.3.4) to a function $grad : X \times \mathbb{R}^d \longrightarrow \mathbb{R}$ as follows:

$$grad(x, \vec{w}) := grad(x, \lfloor \vec{w} \rfloor) \tag{2.5.2}$$

where $\vec{w} = (w_1, w_2, \ldots, w_d) \in \mathbb{R}^d$ and $\lfloor \vec{w} \rfloor$ denotes $(\lfloor w_1 \rfloor, \lfloor w_2 \rfloor, \ldots, \lfloor w_d \rfloor)$.

Lemma 2.5.4. *Let $r \in \mathbb{N} \setminus \{1, 2, 4\}$ and $d > 1$. Let μ be an ergodic measure on X_r. Then μ-almost surely*

$$\lim_{k\to\infty} \sup_{\|\vec{w}\|_1 = k} \frac{1}{k} |grad(x, \vec{w}) - \langle \vec{w}, \vec{v} \rangle| = 0$$

where

$$v_j := \int grad(x, \vec{e}_j) d\mu(x) \ , \ \vec{v} := (v_1, \ldots, v_d),$$

the supremum is over $\{\vec{w} \in \mathbb{R}^d \mid \|\vec{w}\|_1 = k\}$, and $\langle \vec{n}, \vec{v} \rangle = \sum_{i=1}^d n_i v_i$ is the standard inner product.

Proof. Let

$$E_\epsilon := \left\{ x \in X_r \mid \limsup_{k\to\infty} \sup_{\|\vec{w}\|_1 = k} \frac{1}{k} |grad(x, \vec{w}) - \langle \vec{w}, \vec{v} \rangle| > \epsilon \right\}.$$

We will prove the lemma by showing that $\mu(E_\epsilon) = 0$ for all $\epsilon > 0$.

Fix $\epsilon > 0$. Since $\mathbb{Q}^d$ is dense in $\mathbb{R}^d$, using compactness of the unit ball in $(\mathbb{R}^d, \|\cdot\|_1)$, we can find a finite $F \subset \mathbb{Q}^d$ which $\frac{\epsilon}{8}$-covers the unit ball. By this we mean that for all $\vec{w} \in \mathbb{R}^d$ such that $\|\vec{w}\|_1 = 1$ there exists $\vec{u} \in F$ so that $\|\vec{w} - \vec{u}\|_1 \le \frac{\epsilon}{8}$. Because $F \subset \mathbb{Q}^d$ is finite, there exists $M \in \mathbb{N}$ so that $M\vec{u} \in \mathbb{Z}^d$ for all $\vec{u} \in F$. By Lemma 2.5.3, there exists a measurable set $X' \subset X_r$ with $\mu(X') = 1$ so that for all $\vec{u} \in F$, $x \in X'$

$$\lim_{k\to\infty} \frac{1}{Mk} grad(x, Mk\vec{u}) = \langle \vec{u}, \vec{v} \rangle.$$

To complete the proof we will prove that $X' \subset E_\epsilon^c$. Given $x \in X'$ we can find an integer

$J > 8M\epsilon^{-1}$ so that for all $k > \frac{J}{M}$ and all $\vec{u} \in F$

$$\left| \frac{1}{Mk} grad(x, Mk\vec{u}) - \langle \vec{u}, \vec{v} \rangle \right| < \frac{\epsilon}{8}.$$

Note that for all $j > J$

$$\left| j - M \left\lfloor \frac{j}{M} \right\rfloor \right| \leq \frac{\epsilon}{8} j$$

Consider some $\vec{w} \in \mathbb{R}^d$ such that $\|\vec{w}\|_1 = j > J$. We can find $\vec{u} \in F$ so that $\| \frac{\vec{w}}{\|\vec{w}\|_1} - \vec{u}\|_1 \leq \frac{\epsilon}{8}$. Let $\tilde{k} := \lfloor \frac{j}{M} \rfloor$. Then

$$\|\vec{w} - \tilde{k}M\vec{u}\|_1 \leq \|\vec{w} - j\vec{u}\|_1 + |j - \tilde{k}M| \|\vec{u}\|_1 \leq \frac{\epsilon}{4} j. \tag{2.5.3}$$

Now observe that $|grad(x, \vec{w})| \leq \|\vec{w}\|_1$ for all $x \in X_r$, $\vec{w} \in \mathbb{Z}^d$. From the cocycle property (2.3.8) it follows that for all $\vec{n}, \vec{m} \in \mathbb{Z}^d$, $x \in X_r$:

$$grad(x, \vec{n}) = grad(x, \vec{m}) + grad(\sigma^{\vec{m}} x, \vec{n} - \vec{m}).$$

Therefore for all $\vec{w}', \vec{u}' \in \mathbb{R}^d$

$$|grad(x, \vec{w}') - grad(x, \vec{u}')| \leq \|\vec{w}' - \vec{u}'\|_1 + 2d.$$

Applying the above inequality with $\vec{w}' = \vec{w}$ and $\vec{u}' = \tilde{k}M\vec{u}$ it follows using (2.5.3)

$$\left| grad(x, \vec{w}) - grad(x, \tilde{k}M\vec{u}) \right| \leq \frac{\epsilon}{4} j + 2d.$$

Also, since $\|\vec{v}\|_\infty \leq 1$, it follows using (2.5.3) that

$$\left| \langle \vec{w}, \vec{v} \rangle - \langle \tilde{k}M\vec{u}, \vec{v} \rangle \right| < \frac{\epsilon}{4} j$$

which yields that for sufficiently large j,

$$\frac{1}{j} |grad(x, \vec{w}) - \langle \vec{w}, \vec{v} \rangle| < \epsilon.$$

This proves that $X' \subset E_\epsilon^c$. $\square$

2.5.2 Maximal Height Functions

For $\hat{x} \in Ht$ and a finite $F \subset \mathbb{Z}^d$, let

$$Ht_{\hat{x},F} := \{\hat{y} \in Ht \mid \hat{y}_{\vec{n}} = \hat{x}_{\vec{n}} \text{ if } \vec{n} \notin F\}.$$

Consider the partial ordering on $Ht_{\hat{x},F}$ given by $\hat{y} \geq \hat{z}$ if $\hat{y}_{\vec{n}} \geq \hat{z}_{\vec{n}}$ for all $\vec{n} \in \mathbb{Z}^d$.

Lemma 2.5.5. *Let $\hat{x} \in Ht$ and $N \in \mathbb{N}$ be given. Then $(Ht_{\hat{x},F}, \geq)$ has a maximum. If the maximum is attained by the height function $\hat{y}$ then for all $\vec{n} \in F$:*

$$\hat{y}_{\vec{n}} = \min\{\hat{x}_{\vec{k}} + ||\vec{n} - \vec{k}||_1 \mid \vec{k} \in \partial F\}.$$

Proof. We will first prove that given height functions $\hat{y}, \hat{z} \in Ht_{\hat{x},F}$ the function $\hat{w}$ defined by

$$\hat{w}_{\vec{i}} := \max(\hat{y}_{\vec{i}}, \hat{z}_{\vec{i}}).$$

is an element of $Ht_{\hat{x},F}$.

To see that $\hat{w}$ is a valid height function, we will show that $|\hat{w}_{\vec{i}} - \hat{w}_{\vec{j}}| = 1$ for all two adjacent sites $\vec{i}, \vec{j} \in \mathbb{Z}^d$. If $(\hat{w}_{\vec{i}}, \hat{w}_{\vec{j}}) = (\hat{y}_{\vec{i}}, \hat{y}_{\vec{j}})$ or $(\hat{w}_{\vec{i}}, \hat{w}_{\vec{j}}) = (\hat{z}_{\vec{i}}, \hat{z}_{\vec{j}})$ then $|\hat{w}_{\vec{i}} - \hat{w}_{\vec{j}}| = 1$ because $\hat{y}, \hat{z} \in Ht$. Otherwise, we can assume without the loss of generality that $\hat{w}_{\vec{i}} = \hat{y}_{\vec{i}} > \hat{z}_{\vec{i}}$ and $\hat{w}_{\vec{j}} = \hat{z}_{\vec{j}} > \hat{y}_{\vec{j}}$. By Lemma 2.3.3, because $(\hat{y}, \hat{z}) \in \Delta_{Ht}$ we have

$$\hat{y}_{\vec{i}} - \hat{z}_{\vec{i}}, \hat{y}_{\vec{j}} - \hat{z}_{\vec{j}} \in 2\mathbb{Z}.$$

Thus $\hat{y}_{\vec{i}} \geq \hat{z}_{\vec{i}} + 2$ and $\hat{z}_{\vec{j}} \geq \hat{y}_{\vec{j}} + 2$. Since $\hat{y}, \hat{z} \in Ht$ we have:

$$\hat{y}_{\vec{j}} + 1 \geq \hat{y}_{\vec{i}} \geq \hat{z}_{\vec{i}} + 2 \geq \hat{z}_{\vec{j}} + 1 \geq \hat{y}_{\vec{j}} + 3,$$

a contradiction.

We conclude that $\hat{w} \in Ht$. Also $\hat{w}_{\vec{i}} = \max(\hat{y}_{\vec{i}}, \hat{z}_{\vec{i}}) = \hat{x}_{\vec{i}}$ for all $\vec{i} \in F^c$. Hence $\hat{w} \in Ht_{\hat{x},F}$.

Since $Ht_{\hat{x},F}$ is finite, it has a maximum.

Suppose the maximum is attained by a height function $\hat{y}$. Let $\vec{i} \in F$, $\vec{k} \in \partial F$ and $(\vec{i}_1 = i), \vec{i}_2, \vec{i}_3, \ldots, \vec{i}_p, (\vec{i}_{p+1} = \vec{k})$ be a shortest path between $\vec{i}$ and $\vec{k}$. Then

$$\hat{y}_{\vec{i}} = \sum_{t=1}^{p} \hat{y}_{\vec{i}_t} - \hat{y}_{\vec{i}_{t+1}} + \hat{y}_{\vec{k}} = \sum_{t=1}^{p} \hat{y}_{\vec{i}_t} - \hat{y}_{\vec{i}_{t+1}} + \hat{x}_{\vec{k}}.$$

Therefore $\hat{y}_{\vec{i}} \leq ||\vec{i} - \vec{k}||_1 + \hat{x}_{\vec{k}}$ which proves that

$$\hat{y}_{\vec{i}} \leq \min\{\hat{x}_{\vec{k}} + ||\vec{i} - \vec{k}||_1 \mid \vec{k} \in \partial F\}. \tag{2.5.4}$$

For proving the reverse inequality, note that if $\hat{y}$ has a local minimum at some $\vec{n} \in F$ then the height at $\vec{n}$ can be increased. Since $\hat{y}$ is the maximum height function, for each $\vec{n} \in F$ at least one of the adjacent sites $\vec{m}$ must satisfy $\hat{y}_{\vec{n}} - \hat{y}_{\vec{m}} = 1$. Iterating this argument, for all $\vec{i} \in F$, we can choose a path $\vec{j}_1, \vec{j}_2, \vec{j}_3, \ldots, \vec{j}_{p+1}$, with $\vec{j}_1 = \vec{i}$, $\vec{j}_2, \ldots, \vec{j}_p \in F$, $\vec{j}_{p+1} \in \partial F$ along which $\hat{y}$ is increasing:

$\hat{y}_{\vec{j}_t} - \hat{y}_{\vec{j}_{t+1}} = 1$ for all $t \in \{1, 2, \ldots, p\}$. Then

$$\hat{y}_{\vec{i}} = \sum_{t=1}^{p} \hat{y}_{\vec{j}_t} - \hat{y}_{\vec{j}_{t+1}} + \hat{y}_{\vec{j}_{p+1}} \geq ||\vec{i} - \vec{j}_{p+1}||_1 + \hat{y}_{\vec{j}_{p+1}}.$$

Combining with the inequality (2.5.4), we get

$$\hat{y}_{\vec{i}} = \min\{\hat{x}_{\vec{k}} + ||\vec{i} - \vec{k}||_1 \mid \vec{k} \in \partial F\}.$$

□

Consider a shift-invariant Markov cocycle $M \in \mathbf{M}^{\mathbb{Z}^d}_{X_r}$. Recall that by Corollary 2.4.4 there exists $\alpha \in \mathbb{R}$ and a Gibbs cocycle $M_0 \in \mathbf{G}^{\mathbb{Z}^d}_{X_r}$ compatible with a shift-invariant nearest neighbour interaction such that $M = M_0 + \alpha \hat{M}$. The following lemma is based on the idea that in the "non-Gibbsian" case $\alpha \neq 0$, whenever $\hat{y}$ is much bigger than $\hat{x}$, $M(x, y)$ is roughly α times the 'volume' of the $(d+1)$-dimensional 'shape' bounded between the graphs of $\hat{y}$ and $\hat{x}$ in $\mathbb{Z}^d \times \mathbb{Z}$.

For $N \in \mathbb{N}$, let

$$D_N := \left\{\vec{n} \in \mathbb{Z}^d \mid \|\vec{n}\|_1 \leq N\right\} \tag{2.5.5}$$

be the ball of radius N centered at the origin in the standard Cayley graph of $\mathbb{Z}^d$. Also, denote:

$$S_N := \left\{\vec{n} \in \mathbb{Z}^d \mid \|\vec{n}\|_1 = N\right\} \tag{2.5.6}$$

Note that $S_N = \partial(D_N^c) = \partial D_{N-1}$.

Lemma 2.5.6. *Fix $r \in \mathbb{N} \setminus \{1, 2, 4\}$ and $d > 1$. Let $M = M_0 + \alpha\hat{M}$ be a shift-invariant Markov cocycle on $X_r^{(d)}$ where $M_0 \in \mathbf{G}^{\mathbb{Z}^d}_{X_r^{(d)}}$ is a Gibbs cocycle compatible with a shift-invariant nearest neighbour interaction, $\hat{M}$ is the Markov cocycle given by (2.4.6) and $\alpha > 0$. Then there exist a positive constant $c_1 > 0$ (depending only on d) and another positive constant $c_2 > 0$ (depending only on d and M_0) such that for all $N \in \mathbb{N}$*

$$M(x, y) \geq c_1 \alpha (\hat{y}_{\vec{0}} - \hat{x}_{\vec{0}})^{d+1} - c_2 \cdot N^d$$

for all $(\hat{x}, \hat{y}) \in \Delta_{Ht}$ satisfying $\hat{x} \leq \hat{y}$, $x = \phi(\hat{x})$, $y = \phi(\hat{y})$ and $x|_{D_N^c} = y|_{D_N^c}$.

Proof. Let $M = M_0 + \alpha\hat{M}$, $(x, y) \in \Delta_{X_r}$ and $(\hat{x}, \hat{y}) \in \Delta_{Ht}$ be as given in the lemma. First we show that there exists a suitable constant $c_1 > 0$ (depending on d) so that

$$\hat{M}(x, y) \geq c_1 (\hat{y}_{\vec{0}} - \hat{x}_{\vec{0}})^{d+1}. \tag{2.5.7}$$

Assume that $\hat{y}_{\vec{0}} - \hat{x}_{\vec{0}} > 0$. Denote:

$$K := \frac{\hat{y}_{\vec{0}} - \hat{x}_{\vec{0}}}{2} \tag{2.5.8}$$

Recall that $(\hat{x},\hat{y}) \in \Delta_{Ht}$, so by Lemma 2.3.3, $\hat{y}_{\vec{n}} - \hat{x}_{\vec{n}} \in 2\mathbb{Z}$ for all $\vec{n} \in \mathbb{Z}^d$. In particular, K is an integer. Since $\hat{y}_{\vec{n}} - \hat{x}_{\vec{n}} \geq 0$ we have:

$$\hat{M}(x,y) \geq \sum_{\vec{n}\in D_N} (\hat{y}_{\vec{n}} - \hat{x}_{\vec{n}}) \geq \sum_{j=0}^{K} \sum_{\vec{n}\in S_j} (\hat{y}_{\vec{n}} - \hat{x}_{\vec{n}}).$$

Since $\hat{y}_{\vec{n}} - \hat{x}_{\vec{n}} \geq \hat{y}_{\vec{0}} - \hat{x}_{\vec{0}} - 2\|\vec{n}\|_1$:

$$\sum_{j=0}^{K} \sum_{\vec{n}\in S_j} (\hat{y}_{\vec{n}} - \hat{x}_{\vec{n}}) \geq \sum_{j=0}^{K} |S_j|(\hat{y}_{\vec{0}} - \hat{x}_{\vec{0}} - 2j).$$

Finally the estimates

$$|S_j| \geq |\{\vec{n} \in S_j \mid \vec{n} > 0\}| = \binom{j+d-1}{d-1} \geq \frac{1}{d!} j^{d-1}$$

give

$$\begin{aligned} \hat{M}(x,y) \geq \quad & \tfrac{1}{d!}\textstyle\sum_{j=0}^{K} j^{d-1}(2K-2j) & \geq \frac{1}{d!}\sum_{j=\lceil K/3\rceil}^{\lfloor 2K/3\rfloor} j^{d-1}(2K-2j) \\ \geq \quad & \tfrac{1}{d!}\tfrac{K}{3}\left(\tfrac{K}{3}\right)^{d-1}\tfrac{2K}{3} & \geq (6d)^{-(d+1)}(\hat{y}_0 - \hat{x}_0)^{d+1} \end{aligned}$$

proving (2.5.7) with $c_1 = (6d)^{-(d+1)}$.

Let ϕ be the shift-invariant nearest neighbour interaction corresponding to M_0. We will show that there exists a suitable constant $c_2 > 0$ (depending on M_0 and d) so that $|M_0(x,y)| \leq c_2 N^d$. By expressing M_0 in terms of its interaction we see that

$$|M_0(x,y)| \leq \sum_{C\cap D_N \neq \emptyset} |\phi(y|_C) - \phi(x|_C)| \leq \sum_{\vec{n}\in D_N} \sum_{C\cap\{\vec{n}\}\neq\emptyset} |\phi(x|_C) - \phi(y|_C)|,$$

where C ranges over all the cliques (edges and vertices) in $\mathbb{Z}^d$. It follows that $|M_0(x,y)| \leq c_2'|D_N|$ where

$$c_2' := (4d+2)\sup\left\{|\phi(x|_C)| \mid x \in X_r \text{ and } C \text{ is a clique in } \mathbb{Z}^d\right\}.$$

Since $|D_N| \leq (2N+1)^d$, it follows that $|M_0(x,y)| \leq c_2 N^d$ with $c_2 := 3^d c_2'$.

Putting everything together, we conclude that

$$M(x,y) \geq \alpha\hat{M}(x,y) - |M_0(x,y)| \geq c_1\alpha(\hat{y}_{\vec{0}} - \hat{x}_{\vec{0}})^{d+1} - c_2 \cdot N^d.$$

$\square$

Under the same hypothesis except with $\alpha < 0$, we get,

$$M(x,y) \leq c_1\alpha \cdot (\hat{y}_{\vec{0}} - \hat{x}_{\vec{0}})^{d+1} + c_2 N^d$$

for the same constants $c_1, c_2 > 0$.

Lemma 2.5.7. *Let $r \in \mathbb{N} \setminus \{1,2,4\}$ and $d > 1$. Let μ be a shift-invariant measure on X_r and $\vec{n} \in \mathbb{Z}^d$ such that $\|\vec{n}\|_1 = 1$. If*

$$\left| \int grad(x, \vec{n}) d\mu(x) \right| = 1,$$

then μ is frozen.

Proof. If $\left|\int grad(x,\vec{n})d\mu(x)\right| = 1$ then either $\mu(\{x \in X_r \mid x_{\vec{0}} - x_{\vec{n}} = 1 \mod r\}) = 1$ or $\mu(\{x \in X_r \mid x_{\vec{0}} - x_{\vec{n}} = -1 \mod r\}) = 1$. In the first case it follows that μ-almost surely $x_{\vec{m}-\vec{n}} = x_{\vec{m}} + 1 \mod r$ and $x_{\vec{m}+\vec{n}} = x_{\vec{m}} - 1 \mod r$ for all $\vec{m} \in \mathbb{Z}^d$, so μ-almost surely x is frozen. The second case is similar. □

In the course of our proof, it will be convenient to restrict to ergodic shift-invariant Markov random fields. The following claim justifies this:

Theorem 2.5.8. *All shift-invariant Markov random fields μ with specification Θ are in the closure of the convex hull of the ergodic shift-invariant Markov random fields with specification Θ.*

Proof. See Theorem 14.14 in [22]. □

We now proceed to complete the proof of Proposition 2.5.2.

Proof. Since a convex combination of frozen measures is frozen, by Theorem 2.5.8 it suffices to prove that any ergodic Markov random field μ adapted to X_r with its Radon-Nikodym cocycle equal to e^M on its support where $M = M_0 + \alpha\hat{M}$ (as in Corollary 2.4.4) and $\alpha \neq 0$ is frozen.

Choose any μ satisfying the above assumptions, assuming without loss of generality that $\alpha > 0$. Let

$$v_j := \int grad(x, \vec{e}_j) d\mu(x) \text{ for } j = 1, \ldots, d. \tag{2.5.9}$$

If $|v_j| = 1$ for some $1 \leq j \leq d$, it follows from Lemma 2.5.7 that μ is frozen. We can thus assume that $|v_j| < 1$ for all $1 \leq j \leq d$. Choose $\epsilon > 0$ satisfying $\epsilon < \frac{1}{4}\min\{1 - |v_j| \mid 1 \leq j \leq d\}$.

For $k \in \mathbb{N}$, let

$$A_k = \left\{ x \in X_r \mid \sup_{\|\vec{w}\|_1 = k} \frac{1}{k} |grad(x, \vec{w}) - \langle \vec{w}, \vec{v} \rangle| < \epsilon \right\}.$$

By Lemma 2.5.4 for sufficiently large k, $\mu(A_k) > 1-\epsilon$. Consider $x \in A_k \cap supp(\mu)$ such that $\mu(A_k \mid [x]_{\partial D_{k-1}}) > 1-2\epsilon$. Then for all $y \in A_k$ satisfying $y|_{D^c_{k-1}} = x|_{D^c_{k-1}}$ and $\vec{n} \in \partial D_{k-1}$

$$-\,grad(y,\vec{n}) = \hat{y}_{\vec{0}} - \hat{y}_{\vec{n}} \leq -\langle \vec{n}, \vec{v}\rangle + \epsilon k < (1-\epsilon)k. \tag{2.5.10}$$

Choose $\hat{z}$ which is maximal in $Ht_{\hat{x},D_{k-1}}$ and let $z = \phi_r(\hat{z})$. It follows from Lemma 2.5.5 that for some $\vec{n} \in \partial D_{k-1}$,

$$\hat{z}_{\vec{0}} - \hat{x}_{\vec{n}} = \|\vec{n}\|_1 = k. \tag{2.5.11}$$

Since $x|_{D^c_{k-1}} = y|_{D^c_{k-1}} = z|_{D^c_{k-1}}$ we can assume by Lemma 2.3.3 that

$$\hat{x}|_{D^c_{k-1}} = \hat{y}|_{D^c_{k-1}} = \hat{z}|_{D^c_{k-1}}. \tag{2.5.12}$$

(2.5.10) together with (2.5.11) and (2.5.12) imply that $\hat{z}_{\vec{0}} - \hat{y}_{\vec{0}} \geq k\epsilon$ for all $y \in A_k$ satisfying $y|_{D^c_{k-1}} = z|_{D^c_{k-1}}$. Thus, by Lemma 2.5.6

$$M(y,z) > c_1\alpha(k\cdot\epsilon)^{d+1} - c_2k^d > c_3k^{d+1},$$

the last inequality holding for some $c_3 > 0$, when k is sufficiently large.

It follows that

$$\mu([z]_{D_{k-1}} \mid [x]_{\partial D_{k-1}}) \geq \mu([y]_{D_{k-1}} \mid [x]_{\partial D_{k-1}})e^{c_3k^{d+1}}.$$

We can write $A_k \cap [x]_{\partial D_{k-1}} = \bigcup_y ([y]_{D_{k-1}} \cap [x]_{\partial D_{k-1}})$, where the union is over all $y \in \mathcal{L}_{D_{k-1}}(X_r)$ such that $[y]_{D_{k-1}} \cap A_k \cap [x]_{\partial D_{k-1}} \neq \emptyset$. There are at most $|\mathcal{L}_{D_{k-1}\cup\partial D_{k-1}}(X_r)| = e^{O(k^d)}$ terms in the union above so

$$\mu(A_k \mid [x]_{\partial D_{k-1}}) \leq e^{O(k^d)}e^{-c_3k^{d+1}}\mu([z]_{D_{k-1}} \mid [x]_{\partial D_{k-1}}).$$

It follows that $\mu(A_k \mid [x]_{\partial D_{k-1}}) \to 0$ as $k \to \infty$. For k sufficiently large this would contradict our choice of x, for which $\mu(A_k \mid [x]_{\partial D_{k-1}}) > 1-2\epsilon$. □

2.6 Non-Frozen Adapted Shift-Invariant Markov Random Fields on X_r Are Fully-Supported

We have concluded that any shift-invariant Markov random field which is adapted with respect to X_r is a Gibbs measure for some shift-invariant nearest neighbour interaction. Our next goal is to show that any such measure must be fully-supported on X_r.

Proposition 2.6.1. *Let $r \in \mathbb{N} \setminus \{1,2,4\}$, $d > 1$ and μ be a shift-invariant MRF adapted with respect to X_r. Then either $supp(\mu) = X_r$ or μ is frozen.*

Roughly speaking we shall show that for non-frozen shift-invariant Markov random fields the height function corresponding to a typical point is "not very steep". Given a height function that is "not very steep", there is enough flexibility to "deform" the height function while keeping the values fixed outside some finite set. For an adapted Markov random field, the "deformed height function" corresponds to a point in the support as well, which will be the key to proving the required result. Somewhat related methods can be found in Section 4.3 of [48]. This is further generalised in Theorem 4.2.4.

We first introduce some more notation. For $x \in X_r$ and a finite set $F \subset \mathbb{Z}^d$ denote:

$$Range_F(x) := \max_{\vec{n}\in F} grad(x,\vec{n}) - \min_{\vec{n}\in F} grad(x,\vec{n}). \tag{2.6.1}$$

Given $A \subset \mathbb{Z}^d$, $\hat{x} \in Ht(A)$ and a finite set $F \subset A \subset \mathbb{Z}^d$, we define:

$$Range_F(\hat{x}) := \max_{\vec{n}\in F} \hat{x}_{\vec{n}} - \min_{\vec{n}\in F} \hat{x}_{\vec{n}}. \tag{2.6.2}$$

It follows that if $\hat{x} \in Ht$ and $x \in X_r$ are such that $x = \phi_r(\hat{x})$ then for all finite $F \subset \mathbb{Z}^d$, $Range_F(x) = Range_F(\hat{x})$.

Lemma 2.6.2. ***("Extremal values of height obtained on the boundary")*** *Let $F \subset \mathbb{Z}^d$ be a finite set and $\hat{x} \in Ht$ such that $Range_{\partial F}(\hat{x}) > 2$. Then there exists $\hat{y} \in Ht$ such that $\hat{y}_{\vec{n}} = \hat{x}_{\vec{n}}$ for all $\vec{n} \in F^c$ and*

$$Range_F(\hat{y}) = Range_{\partial F}(\hat{y}) - 2 = Range_{\partial F}(\hat{x}) - 2.$$

Proof. Denote

$$T := \max_{\vec{n}\in\partial F} \hat{x}_{\vec{n}} \text{ and } B := \min_{\vec{n}\in\partial F} \hat{x}_{\vec{n}}.$$

Let

$$\kappa = \kappa(\hat{x}, F) := \sum_{\vec{n}\in F} \max(\hat{x}_{\vec{n}} - T + 1, B - \hat{x}_{\vec{n}} + 1, 0).$$

The number κ is the absolute value for the deviations of $\hat{x}|_F$ from the (open) interval (B, T). We prove the claim by induction on κ. If $\kappa = 0$ then $y = x$ already satisfies the conclusion of this lemma because $B + 1 \leq \hat{x}_{\vec{n}} \leq T - 1$ for all $\vec{n} \in F$ which implies that

$$Range_F(\hat{x}) = \max_{\vec{m}\in F} \hat{x}_{\vec{m}} - \min_{\vec{m}\in F} \hat{x}_{\vec{m}} \leq (\max_{\vec{m}\in\partial F} \hat{x}_{\vec{m}} - 1) - (\min_{\vec{m}\in\partial F} \hat{x}_{\vec{m}} + 1)$$

$$= Range_{\partial F}(\hat{x}) - 2.$$

Now suppose $\kappa > 0$, and let $n \in F$ be a coordinate where $\hat{x}$ obtains an extremal value for $F \cup \partial F$. Without loss of generality suppose,

$$\hat{x}_{\vec{n}} = \max_{\vec{m}\in F\cup\partial F} \hat{x}_{\vec{m}}.$$

Since all neighbours of $\vec{n}$ are in $F \cup \partial F$, it follows that $\hat{x}_{\vec{m}} = \hat{x}_{\vec{n}} - 1$ for all $\vec{m}$ adjacent to $\vec{n}$.

Therefore we have $\hat{y} \in Ht$ given by

$$\hat{y}_{\vec{m}} := \begin{cases} \hat{x}_{\vec{m}} - 2 & \text{for } \vec{m} = \vec{n} \\ \hat{x}_{\vec{m}} & \text{otherwise.} \end{cases}$$

Since $Range_{\partial F}(\hat{x}) > 2$, it follows that $\hat{y}_n$ is neither a minimum nor a maximum for $\hat{y}$ in $F \cup \partial F$. Thus $\kappa(\hat{y}, F) < \kappa(\hat{x}, F)$ and so we can apply the induction hypothesis on $\hat{y}$ and conclude the proof. $\square$

Lemma 2.6.3. ***("Flat extension of an admissible pattern")*** *Let $\hat{x} \in Ht$ and $N \in \mathbb{N}$. Then there exists $\hat{y} \in Ht$ such that $\hat{y}_{\vec{n}} = \hat{x}_{\vec{n}}$ for $\vec{n} \in D_{N+1}$ and*

$$Range_{\partial D_{N+k}}(\hat{y}) = Range_{\partial D_N}(\hat{x}) - 2k,$$

for all $1 \leq k \leq \frac{Range_{\partial D_N}(\hat{x})}{2}$.

Proof. We will prove the following statement by induction on $M \in \mathbb{N}$: For all $N \in \mathbb{N}$ and height functions $\hat{x} \in Ht(D_{N+1+M})$ with $Range_{\partial D_N}\hat{x} = 2M$ there exists a height function $\hat{y} \in Ht(D_{N+1+M})$ such that $\hat{y}_{\vec{n}} = \hat{x}_{\vec{n}}$ for all $\vec{n} \in D_{N+1}$ and $1 \leq k \leq M$, $Range_{\partial D_{N+k}}(\hat{y}) = 2M - 2k$. Observe that the height function $\hat{y}$ satisfies in particular $Range_{\partial D_{N+M}}(\hat{y}) = 0$. Thus, the outermost boundary of $\hat{y}$ is flat and it can be extended to a height function on $\mathbb{Z}^d$, so the lemma will follow immediately once we prove the statement above for all $M \in \mathbb{N}$.

For the base case of the induction, there is nothing to prove.

Assume the result for some $M \in \mathbb{N}$. Let $\hat{x} \in Ht$ be a height function such that $Range_{\partial D_N}\hat{x} = 2(M+1)$. Denote $\tilde{N} := N + 1 + (M+1) = N + M + 2$. Let $\vec{n} \in D_{\tilde{N}} \setminus D_{N+1}$ be a site where $\hat{x}$ obtains an extremal value for $D_{\tilde{N}} \setminus D_N$. If there is no such site then

$$\begin{aligned} Range_{\partial D_{N+1}}(\hat{x}) &= \max_{\vec{m} \in \partial D_{N+1}} \hat{x}_{\vec{m}} - \min_{\vec{m} \in \partial D_{N+1}} \hat{x}_{\vec{m}} \\ &= (\max_{\vec{m} \in \partial D_N} \hat{x}_{\vec{m}} - 1) - (\min_{\vec{m} \in \partial D_N} \hat{x}_{\vec{m}} + 1) \\ &= 2M \end{aligned}$$

proving the induction step for that case. Without loss of generality we assume that it is a maximum, that is,

$$\hat{x}_{\vec{n}} = \max_{\vec{m} \in D_{\tilde{N}} \setminus D_N} \hat{x}_{\vec{m}}.$$

Then the function $\hat{\hat{y}}$ given by

$$\hat{\hat{y}}_{\vec{m}} = \begin{cases} \hat{x}_{\vec{n}} - 2 & \text{if } \vec{m} = \vec{n} \\ \hat{x}_{\vec{m}} & \text{otherwise} \end{cases}$$

is a valid height function on $D_{\tilde{N}}$. Hence we have lowered the height function at the site $\vec{n}$ of $\hat{x}$. Repeating the steps for sites with extremal height (formally, this is another internal induction, see for example the proof of Lemma 2.6.2), a height function $\hat{z}$ can be obtained on $D_{\tilde{N}}$ such that $\hat{z} = \hat{x}$ on D_{N+1} and

$$Range_{\partial D_{N+1}}(\hat{z}) = 2M.$$

Thus we can apply the induction hypothesis to $\hat{z}$, substituting $N+1$ for N to obtain a height function $\hat{y}$ on $D_{\tilde{N}}$ such that $\hat{y} = \hat{x}$ on D_{N+1} and

$$Range_{\partial D_{N+k}}(\hat{y}) = Range_{\partial D_N}(\hat{x}) - 2k$$

for $1 \leq k \leq \frac{Range_{\partial D_N}(\hat{x})}{2}$.

This completes the proof of the statement. $\square$

Lemma 2.6.4. ***("Patching an arbitrary finite pattern inside a non-steep point")*** *Let $r \neq 1, 2, 4$ be a positive integer, $d > 1$ and $N, k \in \mathbb{N}$. Choose $y \in X_r$ which satisfies $Range_{\partial D_{2N+2r+k+1}}(y) \leq 2k$ and some $x \in X_r$. Then:*

1. *If either r is odd or $x_{\vec{n}} - y_{\vec{n}}$ is even for all $\vec{n} \in \mathbb{Z}^d$, then there exists $z \in X_r$ such that*

$$z_{\vec{n}} = \begin{cases} x_{\vec{n}} & \text{if } \vec{n} \in D_N \\ y_{\vec{n}} & \text{if } \vec{n} \in D^c_{2N+2r+k+1} \end{cases}$$

2. *If r is even and $x_{\vec{n}} - y_{\vec{n}}$ is odd for all $\vec{n} \in \mathbb{Z}^d$, then there exists $z \in X_r$ such that*

$$z_{\vec{n}} = \begin{cases} x_{\vec{n}+\vec{e}_1} & \text{if } \vec{n} \in D_N \\ y_{\vec{n}} & \text{if } \vec{n} \in D^c_{2N+2r+k+1} \end{cases}$$

The idea of this proof lies in the use of Lemmata 2.6.2 and 2.6.3. Given any pattern on D_N we can extend it to a pattern on D_{2N} with flat boundary which can be then extended to y a little further away provided the range of y is not too large. There is a slight technical issue we deal with in the proof below in case r is even. This is essentially due to the fact that for r even we have

$$x_{\vec{n}} - x_{\vec{m}} \equiv \|\vec{n} - \vec{m}\|_1 \mod 2 \text{ for all } x \in X_r \text{ and } \vec{n}, \vec{m} \in \mathbb{Z}^d.$$

Proof. By applying Lemma 2.6.2 $k-1$ times we conclude that there exists $y' \in X_r$ such that $y' = y$ on $D^c_{2N+2r+k+1}$ and

$$Range_{\partial D_{2N+2r+2}}(y') = 2.$$

For all $\vec{n}, \vec{m} \in \partial D_{2N+2r+2}$, $||\vec{n}-\vec{m}||_1$ is even, therefore $\hat{y}'_{\vec{n}}$ and $\hat{y}'_{\vec{m}}$ have the same parity. Thus, there exist two values $c, d \in \mathbb{Z}_r$ which y' takes on $\partial D_{2N+2r+2}$. Consider $a \in \mathbb{Z}_r$ which is adjacent (for the Cayley graph of $\mathbb{Z}_r$) to both c and d. Then the configuration $y^{(1)}$ given by

$$y^{(1)}_{\vec{n}} := \begin{cases} y'_{\vec{n}} & \text{if } \|\vec{n}\|_1 \geq 2N+2r+3 \\ a & \text{if } \|\vec{n}\|_1 \leq 2N+2r+2 \text{ is even} \\ c & \text{if } \|\vec{n}\|_1 \leq 2N+2r+2 \text{ is odd} \end{cases}$$

is an element of X_r. By Lemma 2.6.3 choose $x^{(1)} \in X_r$ such that $x^{(1)}|_{D_N} = x|_{D_N}$ and

$$Range_{\partial D_{2N-1}}(x^{(1)}) = 0.$$

Equivalently, there exists $b \in \mathbb{Z}_r$ so that $x^{(1)}_n = b$ for all $n \in \partial D_{2N-1}$.

If we are in case (1), either r is even and $b \equiv a \mod 2$ or r is odd. Either way, there is some integer $k \in [0, \dots r-1]$ such that $a + k \equiv b - k \mod r$. Thus, we can find $y^{(2)}, x^{(2)} \in X_r$, so that $y^{(2)}$ agrees with $y^{(1)}$ in D^c_{2N+2r}, $x^{(2)}$ agrees with $x^{(1)}$ in D_{2N}, and so that both $x^{(2)}$ and $y^{(2)}$ have a common constant value $a + k = b + (r-k) \mod r$ on $\partial D_{2N+r-k-1}$. Thus, we get the required $z \in X_r$ by setting

$$z_{\vec{n}} := \begin{cases} x^{(2)}_{\vec{n}} & \text{for } \vec{n} \in D_{2N+r-k-1} \\ y^{(2)}_{\vec{n}} & \text{for } \vec{n} \in D^c_{2N+r-k-1} \end{cases}$$

To prove case (2) we follow the same procedure, substituting x by $\sigma^{\vec{e}_1}(x)$. □

We can now conclude the proof of Proposition 2.6.1:

Proof. Let μ be a shift-invariant Markov random field and $v_1, \dots, v_d$ be given by (2.5.9).

Assume that $supp(\mu)$ is not frozen. Then by Lemma 2.5.7, $|v_j| < 1$ for all $1 \leq j \leq d$. Again, choose $\epsilon > 0$ satisfying $\epsilon < \frac{1}{4}\min\{1 - |v_j| \mid 1 \leq j \leq d\}$.

We need to show that for all $N \in \mathbb{N}$ and patterns $c \in \mathcal{L}_{D_N}(X_r)$, $\mu([c]_{D_N}) > 0$.

From Lemma 2.5.4 it follows that for sufficiently large k,

$$\mu(\{y \in X_r \mid Range_{\partial D_k}(y) \leq 2(1-\epsilon)k\}) > 1 - \epsilon.$$

Now choose $k > (2N + 2r + 1)$ large enough so that there exists $y \in supp(\mu)$ with

$$Range_{\partial D_k}(y) \leq 2(1-\epsilon)k \leq 2(k - (2N + 2r + 1)).$$

By Lemma 2.6.4, it follows that there exists $z \in X_r$ with $z_{\vec{n}} = y_{\vec{n}}$ for $\vec{n} \in \mathbb{Z}^d \setminus D_k$ and $z_{\vec{n}} = c_{\vec{n}}$ for $\vec{n} \in D_N$. Since μ is an adapted Markov random field it follows that $z \in supp(\mu)$, in particular

$\mu([c]_{D_N}) > 0$.

□

2.7 Fully Supported Shift-Invariant Gibbs Measures on X_r

Next we demonstrate the existence of a fully-supported shift-invariant Gibbs measure for shift-invariant nearest neighbour interactions on X_r. We will obtain such measures by showing that equilibrium measures for certain interactions are non-frozen and thus are fully supported by Proposition 2.6.1. To state and prove this result, we need to introduce (measure-theoretic) pressure and equilibrium measures and apply a theorem of Lanford and Ruelle relating equilibrium measures and Gibbs measures. Our presentation is far from comprehensive, and is aimed to bring only definitions necessary for our current results. We refer readers seeking background on pressure and equilibrium measures to the many existing textbooks on the subject, for instance [47, 58].

Let μ be a shift-invariant probability measure on a shift of finite type X. The *measure theoretic entropy* can be defined by

$$h_\mu := \lim_{N\to\infty} \frac{1}{|D_N|} H_\mu^{D_N}, \tag{2.7.1}$$

where D_N was defined in (2.5.5) and

$$H_\mu^{D_N} := \sum_{a\in\mathcal{L}_{D_N}(X)} -\mu([a]_{D_N}) \log \mu([a]_{D_N}), \tag{2.7.2}$$

with the understanding that $0 \log 0 = 0$.

Given a continuous function $f : X \to \mathbb{R}$, the *measure-theoretic pressure* of f with respect to μ is given by

$$P_\mu(f) := \int f d\mu + h_\mu.$$

A shift-invariant probability measure μ is an *equilibrium state* for f if the maximum of $\nu \mapsto P_\nu(f)$ over all shift-invariant probability measures is attained at μ. The existence of an equilibrium state for any continuous f follows from upper-semi-continuity of the function $\nu \mapsto P_\nu(f)$ with respect to the weak-$*$ topology.

Let ϕ be a nearest neighbour interaction on X. As in [47] define a function $f_\phi : X \longrightarrow \mathbb{R}$ by

$$f_\phi(x) := \sum_{A \text{ finite } \mid \vec{0}\in A\subset\mathbb{Z}^d} \frac{1}{|A|}\phi(x|_A). \tag{2.7.3}$$

The following is a restricted case of a classical theorem by Lanford and Ruelle:

Theorem 2.7.1. ***(Lanford-Ruelle Theorem** [27, 47]**)*** *Let X be a $\mathbb{Z}^d$-shift of finite type and ϕ a shift-invariant nearest neighbour interaction. Then any equilibrium state μ for f_ϕ is a Gibbs state for the given interaction ϕ adapted to X.*

The statement of the Lanford-Ruelle theorem (in [27, 47]) does not explicitly mention the adaptedness assumption because for them Gibbs states are always adapted. The *topological entropy* of a $\mathbb{Z}^d$-subshift X is given by

$$h(X) := \lim_{k\to\infty} \frac{1}{|D_k|} \log |\mathcal{L}_{D_k}(X)|.$$

We recall the well known *variational principle* for topological entropy of $\mathbb{Z}^d$-actions, which (in particular) asserts that $h(X) = \sup_\nu h_\nu$ whenever X is a $\mathbb{Z}^d$-shift space and the supremum is over all probability measures on X (Theorem 3.12 in [47]).

Lemma 2.7.2. *Let μ be shift-invariant, frozen Markov random field on $\mathcal{A}^{\mathbb{Z}^d}$, then $h_\mu = 0$.*

Proof. Consider $X_\mu := supp(\mu)$. This is a shift-invariant topological Markov field, consisting of frozen points. Thus for all finite $F \subset \mathbb{Z}^d$, $|\mathcal{L}_F(X_\mu)| \le |\mathcal{L}_{\partial F}(X_\mu)|$. In particular,

$$\log |\mathcal{L}_{D_k}(X_\mu)| \le \log |\mathcal{L}_{\partial D_k}(X_\mu)| \le Ck^{d-1}.$$

It follows that $h(X_\mu) = 0$, so by the variational principle $h_\mu = 0$. $\square$

Lemma 2.7.3. *Let M be a Gibbs cocycle on X_r with a shift-invariant nearest neighbour interaction. Then there exists a shift-invariant nearest neighbour interaction ϕ such that $M = M_\phi$ and any equilibrium measure for f_ϕ is non-frozen.*

Proof. Let $(x^{(i)}, y^{(i)}) \in \Delta_{X_r}$ be as in the proof of Proposition 2.4.2. If $M \in \mathbf{G}_{X_r}$ then there exists a shift-invariant nearest neighbour interaction ϕ so that

$$\begin{aligned} M(x^{(i)}, y^{(i)}) &= \phi([i+2]_0) - \phi([i]_0) \\ &+ \sum_{j=1}^{d} \left(\phi([i+2, i+1]_j) - \phi([i+1, i]_j) + \phi([i+1, i+2]_j) - \phi([i, i+1]_j)\right). \end{aligned}$$

Consider the nearest neighbour interaction $\tilde{\phi}$ given by

$$\tilde{\phi}\left([i+1, i]_j\right) := \tilde{\phi}\left([i, i+1]_j\right) := \frac{1}{2d}\left[\sum_{j=1}^{d} \left(\phi([i+1, i]_j) + \phi([i, i+1]_j)\right) + \phi([i]_0)\right]$$

for all $i \in \mathbb{Z}_r$ and $j \in \{1, \ldots, d\}$, and $\tilde{\phi}([i]_0) = 0$ for all $i \in \mathbb{Z}_r$.

It follows that $M(x^{(i)}, y^{(i)}) = M_{\tilde{\phi}}(x^{(i)}, y^{(i)})$ for all $i \in \mathbb{Z}_r$ and so $M = M_{\tilde{\phi}}$.

Thus we can assume without loss of generality that $\phi = \tilde{\phi}$ satisfies

$$\phi([i, i+1]_j) = \phi([i+1, i]_j) = a_i \text{ for all } i \in \mathbb{Z}_r \text{ and } j \in \{1, 2, \ldots, d\}$$

and $\phi([i]_0) = 0$ for all $i \in \mathbb{Z}_r$.

By (2.7.3):

$$\int f_\phi(x)d\mu(x) = \int \frac{1}{2}\sum_{j=1}^{d}\left(\phi([x_{-e_j}, x_0]_j) + \phi([x_0, x_{e_j}]_j)\right)\mu(x)$$

Thus:

$$\begin{aligned}\int f_\phi(x)d\mu(x) &= \sum_{j=1}^{d}\sum_{i=0}^{r-1}\phi([i,i+1]_j)\mu([i,i+1]_j) + \phi([i+1,i]_j)\mu([i+1,i]_j)\\ &= \sum_{j=1}^{d}\sum_{i=0}^{r-1} a_i(\mu([i,i+1]_j) + \mu([i+1,i]_j)).\end{aligned}$$

Let $a = \max_{1\le i\le r} a_i$ attained by a_{i_0}. It follows that for any shift-invariant probability measure

$$\int f_\phi(x)d\mu(x) \le d\cdot a$$

with equality holding iff $\mu([i,i+1]_j) = \mu([i+1,i]_j) = 0$ for all $a_i < a$ and $j = 1,\ldots,d$.

For a frozen measure μ it follows that for some $j \in \{1,2,\ldots,d\}$, $\mu([i,i+1]_j) > 0$ or $\mu([i+1,i]_j) > 0$ for all $i \in \{0,1,\ldots,r-1\}$. Thus if $a_i < a$ for *some* $0 \le i \le r-1$, it follows that for any frozen measure μ,

$$\int f_\phi(x)d\mu(x) < \sup_\nu \int f_\phi(x)d\nu(x). \tag{2.7.4}$$

where the supremum is attained by the measure supported on the orbit of the point periodic point $x \in X_r$ given by

$$x_n := \begin{cases} i_0 & \|n\|_1 \text{ odd}\\ i_0+1 & \|n\|_1 \text{ even}\end{cases}$$

By Lemma 2.7.2, if μ is frozen then $h_\mu = 0$.

Thus in this case by (2.7.4) for any frozen probability measure μ

$$P_\mu(f_\phi) = \int f_\phi(x)d\mu(x) < \sup_\nu \int f_\phi(x)d\nu(x) \le \sup_\nu P_\nu(f_\phi)$$

and in particular any frozen measure μ can not be an equilibrium measure for f_ϕ.

The remaining case is when $a_i = a$ for all i, in which case $f_\phi(x) = d\cdot a$ is constant. Thus, by the variational principle $\sup_\nu P_\nu(f_\phi) = d\cdot a + \sup_\nu h_\nu = d\cdot a + h(X_r)$. Since $h(X_r) > 0$, it follows that the strict inequality $P_\mu(f_\phi) < \sup_\nu P_\nu(f_\phi)$ holds also in this case for any frozen measure μ. Thus we have the result that for a given Gibbs cocycle with a shift-invariant nearest neighbour interaction there exists an interaction for that cocycle such that the corresponding equilibrium state is not frozen. $\square$

Corollary 2.7.4. *Let $r \in \mathbb{N} \setminus \{1,2,4\}$ and $d > 1$. For all shift-invariant Gibbs cocycles M on X_r there exists a shift-invariant nearest neighbour interaction ϕ on X_r with $M = M_\phi$ and a corresponding shift-invariant Gibbs state ν with supp$(\nu) = X_r$.*

Proof. By Lemma 2.7.3, there exists a shift-invariant nearest neighbour interaction ϕ on X_r with $M = M_\phi$ and an equilibrium measure μ for f_ϕ which is non-frozen. By the Lanford-Ruelle Theorem such μ is a Gibbs state for ϕ and by Proposition 2.6.1 it is fully supported. $\square$

2.8 "Strongly" Non-Gibbsian Shift-Invariant Markov Random Fields

In this section we will prove Theorem 1.1.2, that is, demonstrate the existence of shift-invariant Markov random fields whose specification is not given by any shift-invariant finite range interaction. In contrast to Gibbs measures with shift-invariant finite range interaction, our example proves that generally the specification of a shift-invariant Markov random field cannot be "given by a finite number of parameters". Our construction is somewhat similar to the checkerboard island system as introduced in [44].

Let $\mathcal{A}$ be the alphabet consisting of the 18 'tiles' illustrated in Figure 2.1: A blank tile, a "seed" tile (marked with a "0"), 8 "interior arrow tiles"(4 of them have arrows in the coordinate directions and the rest are "corner tiles") and finally 8 additional "border arrow tiles"(marked by an extra symbol "B").

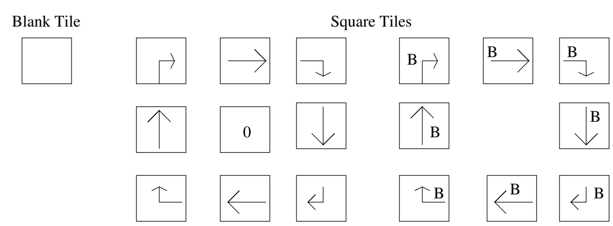

Figure 2.1: The Alphabet $\mathcal{A}$

All tiles other than the blank tile will be called *square tiles* and the tiles with arrows will be called *arrow tiles*. The arrow tiles with 'B' will be called *border tiles* and those without 'B' will be called *interior tiles*. Configurations of an $(2n+1) \times (2n+1)$ square shape whose inner boundary consists of border tiles as illustrated by the example in Figure 2.2 will be called an *n-square-island*. A square-island refers to an n-square island for some n.

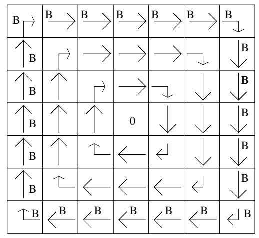

Figure 2.2: A 3-Square-Island

Informally, the idea is to have the square tiles form square-islands with the seed tile in the center and border tiles on their boundary "floating in the sea of the blank tiles". A 'generic' configuration can be seen in Figure 2.3.

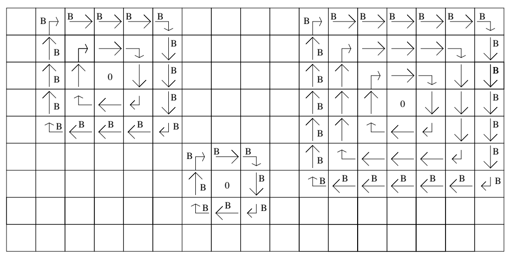

Figure 2.3: A 'Generic' Configuration

Let X be the nearest neighbour shift of finite type on the alphabet $\mathcal{A}$ with constraints:

1. Any "arrow head" must meet an "arrow tail" of matching type (interior or border) and vice versa.

2. Adjacent arrow tiles should not point in opposite directions.

3. Two corner direction tiles cannot be adjacent to one another.

4. The seed tile is only allowed to sit adjacent to straight arrow tiles.

5. An interior tile is always surrounded by other square tiles while a border tile has an interior tile on its right and the blank tile on its left (left and right here are taken from the point of view of the arrow).

Notice that the arrow tiles can turn only in the clock-wise direction. By Constraint (1), every arrow must either be a part of a bi-infinite path or a closed path. Any such path must either trace a straight line (vertical or horizontal) or an "L shape" or a "U shape", or a closed path which traces a rectangle. By Constraint (5) we find that tiles forming a rectangular path must have square tiles to their right (in the interior of the rectangle). These are confined to the interior of the rectangle and thus must themselves trace a smaller rectangle (smaller in terms of the area confined by the rectangle). Note that in constraint (5) we mean in particular that corner direction border tiles have blank tiles on the two sites on their left (where left is taken from the point of view of both the initial direction and the final direction of the arrow). By Constraints (3) and (2) we can conclude by induction on the length that closed paths must actually trace squares with a seed in the center. It follows that the finite connected components of the square tiles are square-islands. The length of these square-islands is $2N + 1$ for some $N \in \mathbb{N}$ because of Constraints (3) and (4). Two such square-islands cannot be adjacent because of Constraint (5).

We can also exclude the possibility of a "U shaped" path using Constraints (3) and (2), again by induction on the length of the "base of the U".

Because a border of blank tiles can be filled either with a square-island or with blank tiles, one can easily deduce that X has positive entropy.

For $r \in \mathbb{N}$ denote by $B_r \subset \mathbb{Z}^2$, the l^∞-ball of radius r in $\mathbb{Z}^2$, that is,

$$B_r := \{(i, j) \in \mathbb{Z}^2 \mid |i|, |j| \leq r\}.$$

Proposition 2.8.1. *Let μ be any measure of maximal entropy for X. Then μ is fully-supported.*

Proof. We will begin by proving that μ-almost surely any square tile is part of a square-island. Since a seed must be surrounded by arrow tiles, it suffices to prove that any arrow tile is part of a square-island. By the discussion above, any arrow tile is part of a path which is bi-infinite or traces a square. An infinite path is either "L shaped" or a straight vertical or horizontal line. It is easily verified that the appearance of an "L shape" in the origin is a transient event with respect to the horizontal or vertical shifts, so by Poincaré recurrence the probability of having a "L shaped path" is zero. An infinite horizontal line forces a half space of horizontal straight line. This forces either a transient event or periodicity. Because X has positive entropy, for a measure of maximal entropy for X, the measure of periodic points is 0 as well. Thus μ-almost surely, any arrow tile is part of a path which traces a square. By Constraint (5) either the square tile is contained in a square-island or there is an infinite sequence of nested square paths. The latter is again a transient event.

Let $x \in X$ be such that it does not have any infinite connected component composed of square tiles. We will now show for all $r \in \mathbb{N}$ that there exists a finite set A_r such that $B_r \subset A_r \subset \mathbb{Z}^2$ and x_i is the blank tile for all $i \in \partial A_r$. Let $Sq_1 \in \mathcal{A}^{C_1}, Sq_2 \in \mathcal{A}^{C_2}, \dots, Sq_k \in \mathcal{A}^{C_k}$ be an enumeration of the square-islands in x such that $C_i \cap B_{r+1} \neq \emptyset$. Let

$$A_r := \bigcup_{i=1}^{k} C_i \cup B_r.$$

Since every square-island is surrounded by the blank tile, A_r has the required properties.

Consider some $y \in X$ and $n \in \mathbb{N}$. We will prove that $\mu([y]_{B_n}) > 0$. Any incomplete square-island in $y|_{B_r}$ can be completed (possibly in multiple ways) in B_{4r}. By completing these square-islands we can obtain $z \in X$ such that it satisfies

$$z_i := \begin{cases} y_i & \text{for } i \in B_r \\ \text{blank tile} & \text{for } i \in B_{4r}^c. \end{cases}$$

Now choose any $x \in supp(\mu)$ which does not have any infinite connected component composed of square tiles. As previously discussed we can find $A_{4r} \subset \mathbb{Z}^2$ such that $B_{4r} \subset A_{4r}$ and x_i is the blank tile for all $i \in \partial A_{4r}$. Then $z|_{\partial A_{4r}} = x|_{\partial A_{4r}}$. By the Lanford-Ruelle Theorem μ is a Markov random field with the uniform specification adapted to X. Therefore

$$\frac{\mu([z]_{A_{4r} \cup \partial A_{4r}})}{\mu([x]_{A_{4r} \cup \partial A_{4r}})} = 1$$

proving $\mu([z]_{B_r}) = \mu([y]_{B_r}]) > 0$.

□

We now describe a subshift of finite type, Y, for which dimension of the space of Markov cocycles is infinite. Y is a nearest neighbour shift of finite type with the alphabet as in Figure 2.1 but now with two types of square tiles, type 1 and 2. The adjacency rules are as in the subshift X but also force adjacent square tiles to be of the same type, that is, any square-island in an element of Y will consist entirely of tiles of type 1 or of type 2. Let $\mathbf{p} = (p_i)_{i \in \mathbb{N}} \in (0,1)^{\mathbb{N}}$ and $\phi : Y \longrightarrow X$ be the map which forgets the type of square tiles. We will now construct a shift-invariant Markov random field $\mu_{\mathbf{p}}$ obtained by picking $x \in X$ according to a fixed measure of maximal entropy μ and then choosing the type of square-islands in x with the distribution: an i-square-island is of type 1 with probability p_i and 2 with probability $1 - p_i$. Precisely, let $\mathcal{F} := \phi^{-1}(Borel(X))$ be the pull-back of the Borel sigma-algebra on X. For all $y \in Y$, $i \in \mathbb{N}$ and $\Lambda \subset \mathbb{Z}^2$ finite consider the functions

$$\begin{aligned} m_\Lambda^i(y) &:= \text{ the number of } i\text{-square-islands of type 1 in } y \text{ intersecting } \Lambda \\ &\textit{and} \\ n_\Lambda^i(y) &:= \text{ the number of } i\text{-square-islands of type 2 in } y \text{ intersecting } \Lambda. \end{aligned}$$

$\mu_{\mathbf{p}}$ is the unique probability measure on Y satisfying $\mu_{\mathbf{p}}|_{\mathcal{F}} = \phi^{-1}\mu$, the pull-back of the measure μ and

$$\mu_{\mathbf{p}}([y]_\Lambda \mid \mathcal{F})(y) = \prod_{i=1}^{\infty} p_i^{m_\Lambda^i(y)} (1 - p_i)^{n_\Lambda^i(y)}$$

for all $\Lambda \subset \mathbb{Z}^2$ finite, $y \in Y$. In particular, if there are no square-islands in y intersecting both Λ and $\mathbb{Z}^2 \setminus \Lambda$ then $\mu_{\mathbf{p}}([y]_\Lambda)$ can be written as a product of the distribution $\phi^{-1}\mu$ and the distribution of colours, that is,

$$\mu_{\mathbf{p}}([y]_\Lambda) = \mu([\phi(y)]_\Lambda) \prod_{i=1}^{\infty} p_i^{m_\Lambda^i(y)} (1 - p_i)^{n_\Lambda^i(y)}$$

and further, for finite sets $\Gamma \supset \Lambda \cup \partial\Lambda$

$$\mu_{\mathbf{p}}\left([y]_\Lambda \middle| [y]_{\Gamma\setminus\Lambda}\right) \quad = \quad \mu\left([\phi(y)]_\Lambda \middle| [\phi(y)]_{\partial\Lambda}\right) \prod_{i=1}^{\infty} p_i^{m^i_\Lambda(y)}(1-p_i)^{n^i_\Lambda(y)}. \tag{2.8.1}$$

Let $(y, y') \in \Delta_Y$ and F be the set of sites on which (y, y') differ. We note that if $y|_C$ is an infinite connected component consisting of square tiles then $y|_C = y'|_C$, that is, $C \cap F = \emptyset$. Thus we can choose a finite set $\Lambda \subset \mathbb{Z}^2$ such that $F \subset \Lambda$ and $y_i = y'_i$ is a blank tile for all $i \in \partial\Lambda$. Since μ is a Markov random field with uniform specification, (2.8.1) implies for all finite sets $\Gamma \supset \Lambda \cup \partial\Lambda$ that

$$\begin{aligned} \frac{\mu_{\mathbf{p}}\left([y']_\Gamma\right)}{\mu_{\mathbf{p}}\left([y]_\Gamma\right)} \quad &= \quad \prod_{i=1}^{\infty} p_i^{m^i_\Lambda(y') - m^i_\Lambda(y)}(1-p_i)^{n^i_\Lambda(y') - n^i_\Lambda(y)} \\ &= \quad \prod_{i=1}^{\infty} p_i^{m^i_F(y') - m^i_F(y)}(1-p_i)^{n^i_F(y') - n^i_F(y)}. \end{aligned} \tag{2.8.2}$$

Let $M_{\mathbf{p}}$ be the shift-invariant cocycle on Y given by the following:

$$M_{\mathbf{p}}(y, y') = \sum_{i=1}^{\infty} (m^i_F(y') - m^i_F(y)) \log(p_i) + (n^i_F(y') - n^i_F(y)) \log(1-p_i)$$

for all $(y, y') \in \Delta_Y$ and the finite set $F \subset \mathbb{Z}^2$ on which they differ.

It follows from (2.8.2) that

$$\frac{\mu_{\mathbf{p}}([y']_\Gamma)}{\mu_{\mathbf{p}}([y]_\Gamma)} = e^{M_{\mathbf{p}}(y,y')} \tag{2.8.3}$$

whenever $(y, y') \in \Delta_Y$ and Γ satisfies the assumptions above. It follows that $M_{\mathbf{p}}$ is the logarithm of the Radon-Nikodym cocycle for μ.

We first state a small technical lemma:

Lemma 2.8.2. *Let $(y, y'), (z, z') \in \Delta_Y$ and F be the set of sites on which y and y' differ. Suppose*

$$\begin{aligned} y|_F = z|_F \quad &\textit{and} \quad y'|_F = z'|_F, \\ y|_{F^c} = y'|_{F^c} \quad &\textit{and} \quad z|_{F^c} = z'|_{F^c}. \end{aligned}$$

Let $C \subset \mathbb{Z}^2$ be a $(2i+1) \times (2i+1)$ square shape such that $y|_C$ is an i-square-island of type 1 and $C \cap F \neq \emptyset$. Then $z|_C = y|_C$.

For all $A \subset \mathbb{Z}^2$ let

$$\partial_2(A) := \{i \in \mathbb{Z}^2 \setminus A \mid \text{ there exists } j \in A \text{ such that } \|i-j\|_1 \leq 2\}.$$

Proof. Let $C_1, C_2, \ldots, C_n \subset \mathbb{Z}^2$ be the square shapes such that for all $1 \leq j \leq n$,

1. $y'|_{C_j}$ is a square-island.
2. There is a vertex $l \in C_j \cap C$ such that y_l is a border tile.

If there is no such square shape and E is the union of all the edges of C, then $E \subset F$ implying that $z|_C = y|_C$. Otherwise we note that by Constraint (5) $y'|_{\partial C_j}$ consists only of blank tiles for all $1 \leq j \leq n$ and divide the proof into the following cases:

1. *$n = 1$ and $C \subset C_1$:* Since $C \neq C_1$ are square shapes and $y|_{\partial(C^c)}$ consists of border tiles, F completely contains at least two edges of C. Thus $y|_F$ completely determines the square-island $y|_C$ and $z|_C = y|_C$.

2. *For some $1 \leq j \leq n$, C_j completely contains one of the edges of C and $C \not\subset C_j$:* Suppose C_j contains the top edge T of C. Since $C \not\subset C_j$, T does not intersect the top edge of C_j and thus $T \subset F$. Then $y|_F$ completely determines the square-island $y|_C$ and $z|_C = y|_C$.

3. *For all $1 \leq j \leq n$ the square shape C_j does not contain any of the edges of C completely:* Let E denote the union of all the edges of C. By Constraint (5), for all $1 \leq j \leq n$ and $l \in \partial_2 C_j$, y'_l is either a blank tile or a border tile implying that $C \cap \partial_2 C_j \subset F$. Also $y'|_{E \setminus \bigcup_{j=1}^n C_j}$ consists only of blank tiles therefore $E \setminus \cup_{j=1}^n C_j \subset F$. Thus we find that

$$\left(E \setminus \bigcup_{j=1}^{n} C_j\right) \cup \left(\bigcup_{j=1}^{n} C \cap \partial_2 C_j\right) \subset F$$

is a connected set in $\mathbb{Z}^2$ and touches all the edges of C. Thus $y|_F$ completely determines the square-island $y|_C$ proving that $y|_C = z|_C$.

$\square$

Proposition 2.8.3. *For any $\mathbf{p} \in (0,1)^{\mathbb{N}}$, the measure $\mu_{\mathbf{p}}$ defined above is a shift-invariant Markov random field with Radon-Nikodym cocycle $M_{\mathbf{p}}$.*

Proof. By (2.8.3) we are left to prove that the shift-invariant cocycle $M_{\mathbf{p}}$ is Markov. Let $(y, y'), (z, z') \in \Delta_Y$ and $F \subset \mathbb{Z}^2$ be as in Lemma 2.8.2. We will show that $m^i_F(y') - m^i_F(y) = m^i_F(z') - m^i_F(z)$: The inequality $m^i_F(z) \geq m^i_F(y)$ follows by Lemma 2.8.2. By interchanging z by y the reverse inequality holds, proving $m^i_F(z) = m^i_F(y)$, and similarly we obtain that $m^i_F(z') = m^i_F(y')$. It follows that $M_{\mathbf{p}}$ is Markov.

□

Proposition 2.8.3 above proves the existence of uncountably many linearly independent shift-invariant Markov cocycles which have corresponding fully-supported shift-invariant Markov random fields. Since the space of Markov cocycles which come from some shift-invariant finite range interaction is a union of finite dimensional vector spaces, this further implies that there exists a shift-invariant Markov random field which is not Gibbs for any shift-invariant finite range interaction proving Theorem 1.1.2. Alternatively, note that for any Gibbs cocycle with some shift-invariant finite range interaction the magnitude of the cocycle at a particular homoclinic pair is at most linear in the size of set of sites at which the two configurations differ. We can choose a $\mathbf{p} \in (0,1)^{\mathbb{N}}$ such that this does not happen.

A simple variation on the above construction yields topological Markov fields which are not sofic: Choose $\mathbf{p} \in [0,1]^{\mathbb{N}}$. If $p_i = 0$ or $p_i = 1$, this would disallow square-islands of certain types for specific sizes. Each such $\mathbf{p}$ would determine a shift-invariant Markov random field supported on a shift space contained in Y. Since there are uncountably many such subshifts a majority of such spaces will be not sofic. However it is easy to see from the proofs above that these are global topological Markov fields.

Chapter 3

Generalisation of the Hammersley-Clifford Theorem on Bipartite Graphs

In this chapter we will prove Theorem 1.2.1. Most of this chapter is a part of the submitted manuscript [11]. Taking inspiration from symbolic dynamics we will define n.n.constraint spaces in Section 3.1. Section 3.2 builds up the necessary background for this work. In Subsection 3.2.1 we introduce Hammersley-Clifford spaces and in Subsection 3.2.2 we introduce Markov-similarity and V-good pairs. In Subsection 3.2.3 we introduce folding and strong config-folding. Section 3.3 states and proves the main results of this chapter. Since the proofs are technical we work out a concrete example of our results in Subsection 3.3.1.

In this chapter $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ will always denote an undirected locally finite, countable, bipartite graph without self-loops and multiple edges and $\mathcal{A}$ will always denote a finite set and be referred to as the *alphabet*. $\mathcal{H} = (\mathcal{V}_{\mathcal{H}}, \mathcal{E}_{\mathcal{H}})$ will always denote a finite undirected graph with or without self-loops. Adjacency in a graph $\mathcal{S}$ will be denoted by $\sim_{S}$. We will often drop the subscript when the denotation is clear from the context. We remark that in this chapter given $A \subset B \subset \mathcal{V}$, given $b \in \mathcal{A}^B$ the cylinder set (as defined in Section 2.1.1) $[b]_A$ also represents the corresponding pattern $b|_A$.

3.1 N.N.Constraint Spaces

The following definitions take inspiration from symbolic dynamics ([29]): A *closed configuration space* is a closed subset of configurations contained in $\mathcal{A}^{\mathcal{V}}$. Let $\mathcal{F}$ be a given set of patterns on finite sets. Then *the configuration space with constraints* $\mathcal{F}$ is defined to be

$$X_{\mathcal{F}} := \{x \in \mathcal{A}^{\mathcal{V}} \mid \text{patterns from } \mathcal{F} \text{ do not appear in } x\}.$$

A set of constraints $\mathcal{F}$ is called *nearest neighbour* if $\mathcal{F}$ consists of patterns on cliques, that is, for all $[a]_F \in \mathcal{F}$, $diam(F) \leq 1$.

A *n.n.constraint space* is a configuration space with nearest neighbour constraints. Note that if $\mathcal{G}$ is bipartite then $\mathcal{F}$ consists of patterns on edges and vertices. These spaces correspond to nearest neighbour shifts of finite type as defined in Subsection 2.1.3.

Let $Hom(\mathcal{G},\mathcal{H})$ denote *the space of all graph homomorphisms* (defined in Section 2.3) from $\mathcal{G}$ to $\mathcal{H}$. For example it was mentioned in Section 2.3 that if C_3 is the 3-cycle with vertices 0, 1 and 2 then the space of 3-colourings is $Hom(\mathcal{G},C_3)$.

Given a graphs $\mathcal{G}$ and $\mathcal{H}$, $Hom(\mathcal{G},\mathcal{H})$ is an n.n.constraint space where the constraint is given by

$$\mathcal{F} := \{[a,b]_{v,w} \mid a \nsim b \in \mathcal{V}_{\mathcal{H}} \text{ and } v \sim w \in \mathcal{V}\}.$$

Then for all $x \in X_{\mathcal{F}}$ and vertices $v \sim w \in \mathcal{V}$, $x_v \sim x_w$ which implies $x \in Hom(\mathcal{G},\mathcal{H})$. Conversely for all homomorphisms $x \in Hom(\mathcal{G},\mathcal{H})$ and vertices $v \sim w \in \mathcal{V}$ we have $[x]_{\{v,w\}} \notin \mathcal{F}$ and hence $x \in X_{\mathcal{F}}$.

N.N.Constraint spaces arise naturally in the study of MRFs as is shown in the following propositions.

Proposition 3.1.1. *Let $\mathcal{A}$ be some finite set, $\mathcal{G} = (\mathcal{V},\mathcal{E})$ be a graph and $X \subset \mathcal{A}^{\mathcal{V}}$ be an n.n.constraint space. Then X is a topological Markov field.*

Proof. Consider $A \subset \mathcal{V}$ finite and $x, y \in X$ such that $x|_{\partial A} = y|_{\partial A}$. We want to prove that $z \in \mathcal{A}^{\mathcal{V}}$ defined by

$$z_v := \begin{cases} x_v \ if \ v \in A \cup \partial A \\ y_v \ if \ v \in A^c \end{cases}$$

is an element of X. Let $B \subset \mathcal{V}$ be a clique. If $B \cap A \neq \phi$ then $B \subset A \cup \partial A$ and $z|_B = x|_B \in \mathcal{L}_B(X)$ else $B \cap A = \phi$ implying $z|_B = y|_B \in \mathcal{L}_B(X)$. Since X is an n.n.constraint space $z \in X$. ☐

The following proposition completes the bridge between MRFs and n.n.constraint spaces.

Proposition 3.1.2. *Let $\mathcal{G} = (\mathcal{V},\mathcal{E})$ be a given graph and X be a topological Markov field on the graph $\mathcal{G}$ with a safe symbol. Then X is an n.n.constraint space.*

Remark: If μ is an MRF then $supp(\mu)$ is a topological Markov field. Thus this proposition implies that if a measure μ satisfies the hypothesis of the weak Hammersley-Clifford theorem (Theorem 2.2.2), that is, if μ is an MRF such that $supp(\mu)$ has a safe symbol then $supp(\mu)$ is an n.n.constraint space. The conclusion of this proposition does not hold without assuming presence of a safe symbol. (look at the comments following proof of Proposition 3.5 in [12])

Proof. Let $\star$ be a safe symbol for X. Consider the set

$$\mathcal{F} := \{a \in \mathcal{A}^A \mid A \subset \mathcal{V} \text{ forms a clique and there does not exist } x \in X \text{ such that } x|_A = a\}.$$

Note that $X \subset X_{\mathcal{F}}$ and if $A \subset \mathcal{V}$ is a clique then $\mathcal{L}_A(X_{\mathcal{F}}) = \mathcal{L}_A(X)$. We want to prove that $X_{\mathcal{F}} \subset X$. We will proceed by induction on $n \in \mathbb{N}$, the hypothesis being: Given $A \subset \mathcal{V}$ such that $|A| = n$, $\mathcal{L}_A(X_{\mathcal{F}}) \subset \mathcal{L}_A(X)$.

The base case follows immediately. Suppose for some $n \in \mathbb{N}$, given $A \subset \mathcal{V}$ satisfying $|A| \leq n$, $\mathcal{L}_A(X_{\mathcal{F}}) \subset \mathcal{L}_A(X)$.

For the induction step consider $A \subset \mathcal{V}$ such that $|A| = n+1$. There are two cases to consider: If A is a clique then $\mathcal{L}_A(X_{\mathcal{F}}) = \mathcal{L}_A(X)$. If A is not a clique then there exists $v \in A$ such that $|\partial\{v\} \cap A| < n$. Let $a \in \mathcal{L}_A(X_{\mathcal{F}})$. We will prove that $a \in \mathcal{L}_A(X)$. Now $|(\{v\} \cup \partial\{v\}) \cap A|, |A \setminus \{v\}| \leq n$, thus the induction hypothesis implies

$$a|_{(\{v\}\cup\partial\{v\})\cap A} \in \mathcal{L}_{(\{v\}\cup\partial\{v\})\cap A}(X)$$

and

$$a|_{A\setminus\{v\}} \in \mathcal{L}_{A\setminus\{v\}}(X).$$

Consider $x, y \in X$ such that

$$x|_{(\{v\}\cup\partial\{v\})\cap A} := a|_{(\{v\}\cup\partial\{v\})\cap A}$$

and

$$y|_{A\setminus\{v\}} := a|_{A\setminus\{v\}}.$$

Since $\star$ is a safe symbol for X therefore $x^\star, y^\star \in \mathcal{A}^{\mathcal{V}}$ given by

$$x^\star_w := \begin{cases} x_w \ if \ w \in (\{v\} \cup \partial\{v\}) \cap A \\ \star \quad otherwise \end{cases}$$

and

$$y^\star_w := \begin{cases} y_w \ if \ w \in A \setminus \{v\} \\ \star \quad otherwise \end{cases}$$

are elements of X. Note that $x^\star_w = x_w = a_w$, $y^\star_w = y_w = a_w$ if $w \in \partial\{v\} \cap A$ and $x^\star_w = y^\star_w = \star$ if $w \in A^c$. Therefore $x^\star|_{\partial\{v\}} = y^\star|_{\partial\{v\}}$. Since X is a topological Markov field, $z \in \mathcal{A}^{\mathcal{V}}$ defined by

$$z_w := \begin{cases} x^\star_w \ if \ w \in \{v\} \cup \partial\{v\} \\ y^\star_w \ otherwise \end{cases}$$

is an element of X. But $z_v = x^\star_v = x_v = a_v$ and $z_w = y^\star_w = y_w = a_w$ if $w \in A \setminus \{v\}$. Hence $z|_A = a \in \mathcal{L}_A(X)$. This completes the induction. Hence $X = X_{\mathcal{F}}$. $\square$

N.N.Constraint spaces allow us to change configurations one site at a time provided the edge-constraints are satisfied. To state this rigorously we define the following: given $x \in \mathcal{A}^{\mathcal{V}}$, and distinct vertices $w_1, w_2, \ldots, w_r \in \mathcal{V}$ and $c_1, c_2, \ldots, c_r \in \mathcal{A}$ we denote by $\theta^{w_1,w_2,\ldots,w_r}_{c_1,c_2,\ldots,c_r}(x)$ an element of $\mathcal{A}^{\mathcal{V}}$ given by

$$(\theta^{w_1,w_2,\ldots,w_r}_{c_1,c_2,\ldots,c_r}(x))_u := \begin{cases} x_u & \text{if } u \neq w_1, w_2 \ldots, w_r \\ c_i & \text{if } u = w_i \text{ for some } 1 \leq i \leq r. \end{cases}$$

Proposition 3.1.3. *Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be a bipartite graph. Suppose $X \subset \mathcal{A}^{\mathcal{V}}$ is an n.n.constraint space and $x \in X$. Let $w_1, w_2, \ldots, w_r \in \mathcal{V}$ be distinct vertices such that $w_i \nsim w_j$ for $1 \leq i, j \leq r$ and $c_1, c_2, \ldots, c_r \in \mathcal{A}$ such that $[c_i, x_{w'}]_{\{w_i,w'\}} \in \mathcal{L}_{\{w_i,w'\}}(X)$ for all $w' \sim w_i$ and $1 \leq i \leq r$. Then $\theta^{w_1,w_2,\ldots,w_r}_{c_1,c_2\ldots,c_r}(x) \in X$.*

Specialising to $r = 1$, if $X \subset \mathcal{A}^{\mathcal{V}}$ is an n.n.constraint space and $x \in X$ then for $v \in \mathcal{V}$ and $c \in \mathcal{A}$, $\theta^v_c(x) \in X$ if and only if $[x_w, c]_{\{w,v\}} \in \mathcal{L}_{\{w,v\}}(X)$ for all $w \sim v$.

Proof. The constraint set for X consists only of patterns on edges and vertices. Thus it is sufficient to check for all $v \sim w$ that

$$[\theta^{w_1,w_2,\ldots,w_r}_{c_1,c_2\ldots,c_r}(x)]_{\{v,w\}} \in \mathcal{L}_{\{v,w\}}(X).$$

Since $w_i \nsim w_j$ for all $1 \leq i, j \leq r$ at most one among v and w is w_i for some $1 \leq i \leq r$. If both of them are not equal to w_i then

$$[\theta^{w_1,w_2,\ldots,w_r}_{c_1,c_2\ldots,c_r}(x)]_{\{v,w\}} = [x]_{\{v,w\}} \in \mathcal{L}_{\{v,w\}}(X).$$

Otherwise we may assume $v = w_i$ for some $1 \leq i \leq r$ giving us

$$[\theta^{w_1,w_2,\ldots,w_r}_{c_1,c_2\ldots,c_r}(x)]_{\{v,w\}} = [c_i, x_w]_{\{w_i,w\}} \in \mathcal{L}_{\{w_i,w\}}(X).$$

□

3.2 Hammersley-Clifford Spaces and Strong Config-Folds

3.2.1 Hammersley-Clifford Spaces

A topological Markov field $X \subset \mathcal{A}^{\mathcal{V}}$ is called *Hammersley-Clifford* if the space of Markov cocycles on X is equal to the space of Gibbs cocycles on X, that is, $\mathbf{M}_X = \mathbf{G}_X$. If X is invariant under the some subgroup $G \subset Aut(\mathcal{G})$ then X is said to be *G-Hammersley-Clifford* if $\mathbf{M}^G_X = \mathbf{G}^G_X$.
Examples: (Further examples and explaination follow the statement of Theorem 3.3.2)

1. A frozen space of configurations.

2. A topological Markov field with a safe symbol.

3. $Hom(\mathcal{G}, Edge)$ where $Edge$ consists of two vertices 0 and 1 connected by a single edge.

4. $Hom(\mathbb{Z}^d, C_n)$ where C_n is an n-cycle, $d > 1$ and $n \neq 4$. (Proposition 2.4.5)

   This gives examples of Hammersley-Clifford spaces which are not G-Hammersley-Clifford spaces for some subgroup $G \subset Aut(\mathbb{Z}^d)$. It will follow from Theorem 3.3.2 below and example 3 above that $Hom(\mathcal{G}, C_4)$ is both Hammersley-Clifford and G-Hammersley-Clifford for all bipartite graphs $\mathcal{G}$ and subgroups $G \subset Aut(\mathcal{G})$.

3.2.2 Markov-Similar and V-Good Pairs

Suppose we are given a closed configuration space X, a Markov cocycle $M \in \mathbf{M}_X$ and an interaction V on X. If M is not Gibbs with the interaction V we might be still interested in the extent to which it is not. An asymptotic pair $(x, y) \in \Delta_X$ is called *(M,V)-good* if

$$M(x,y) = \sum_{S \subset \mathcal{V} \text{ finite}} \left(V([y]_S) - V([x]_S)\right).$$

In most cases the Markov cocycle M will be fixed, so we will drop M and call a pair V-good instead of (M, V)-good. An asymptotic pair $(x, y) \in \Delta_X$ is said to be *Markov-similar* to (z, w) if there is a finite set $A \subset \mathcal{V}$ such that

$$\begin{aligned} x_u &= y_u, \\ z_u &= w_u \text{ for } u \in A^c \end{aligned}$$

and

$$\begin{aligned} x_u &= z_u, \\ y_u &= w_u \text{ for } u \in A \cup \partial A. \end{aligned}$$

Being V-good is infectious.

Proposition 3.2.1. *Let X be an n.n.constraint space, M a Markov cocycle and V a nearest neighbour interaction on X. The set of V-good pairs is an equivalence relation on X. Additionally if $(x, y), (z, w) \in \Delta_X$ are Markov similar then (x, y) is V-good if and only if (z, w) is V-good.*

Proof. The reflexivity and symmetry of the relation V-good are immediate; the cocycle condition implies that the relation is transitive. Thus the relation is an equivalence relation.

Let $(x, y), (z, w) \in \Delta_X$ be Markov-similar pairs. Since M is a Markov cocycle

$$M(x,y) = M(z,w). \tag{3.2.1}$$

Let $A \subset \mathcal{V}$ be a finite set such that

$$x_u = z_u \text{ and } y_u = w_u$$

for $u \in A \cup \partial A$ and

$$x_u = y_u \text{ and } z_u = w_u$$

for $u \in A^c$. If $S \subset \mathcal{V}$ is a clique then either $S \subset A \cup \partial A$ or $S \subset A^c$. If $S \subset A \cup \partial A$ then

$$x|_S = z|_S \text{ and } y|_S = w|_S$$

implying

$$V([y]_S) - V([x]_S) = V([w]_S) - V([z]_S).$$

If $S \subset A^c$ then

$$x|_S = y|_S \text{ and } z|_S = w|_S$$

implying

$$V([y]_S) - V([x]_S) = V([w]_S) - V([z]_S) = 0.$$

Since V is a nearest neighbour interaction

$$\sum_{S \subset \mathcal{V} \text{ finite}} V([y]_S) - V([x]_S) = \sum_{S \subset \mathcal{V} \text{ finite}} V([w]_S) - V([z]_S).$$

Since (x, y) is a V-good pair by (3.2.1)

$$M(z, w) = M(x, y) = \sum_{S \subset \mathcal{V} \text{ finite}} (V([y]_S) - V([x]_S)) = \sum_{S \subset \mathcal{V} \text{ finite}} V([w]_S) - V([z]_S)$$

completing the proof. □

Corollary 3.2.2. *Let X be an n.n.constraint space, M a Markov cocycle and V a nearest neighbour interaction on X. Suppose for some $(x, y) \in \Delta_X$ there exists a chain $x = x_1, x_2, x_3, \ldots, x_n = y$ such that each $(x_i, x_{i+1}) \in \Delta_X$ and is Markov similar to a V-good pair. Then (x, y) is V-good.*

This follows from Lemma 3.2.1.

3.2.3 Strong Config-Folding

We shall now introduce graph folding and extract some of its properties so as to define a strong notion of folding for closed configuration spaces. Graph folding was introduced in [37] and used in [6] so as to prove a slew of properties which are satisfied by a given graph if and only if it is satisfied by its folds. Fix some finite undirected graph $\mathcal{H} = (\mathcal{V}_\mathcal{H}, \mathcal{E}_\mathcal{H})$ without multiple edges. For

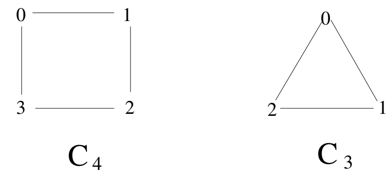

Figure 3.1: C_4 and C_3

any vertex $a \in \mathcal{H}$ we say that $\mathcal{H} \setminus \{a\}$ is a *fold* of the graph $\mathcal{H}$ if there exists $b \in \mathcal{H} \setminus \{a\}$ such that

$$\{c \in \mathcal{V}_{\mathcal{H}} \mid c \sim a\} \subset \{c \in \mathcal{V}_{\mathcal{H}} \mid c \sim b\}.$$

In such a case we say that a is folded into b.

For example in the 4-cycle C_4 the vertex 3 can be folded into the vertex 1. However no vertex can be folded in the 3-cycle C_3.

For any vertex $v \in \mathcal{V}$ the *n-ball around v* is given by

$$D_n(v) := \{w \in \mathcal{V} \mid d_{\mathcal{G}}(v, w) \leq n\}$$

where $d_{\mathcal{G}}$ is the graph distance on $\mathcal{G}$.

We wish to generalise the following property:

Proposition 3.2.3. *Consider a bipartite graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, a graph $\mathcal{H} = (\mathcal{V}_{\mathcal{H}}, \mathcal{E}_{\mathcal{H}})$ and vertices $a, b \in \mathcal{V}_{\mathcal{H}}$ where the vertex a can be folded into the vertex b. Let $X = Hom(\mathcal{G}, \mathcal{H})$. Then for all edges $(v_1, v_2), (v_2, v_3) \in \mathcal{E}$ and $c \in \mathcal{V}_{\mathcal{H}}$, $[a, c]_{\{v_1,v_2\}} \in \mathcal{L}_{\{v_1,v_2\}}(X)$ implies*

$$\begin{aligned} [b, c]_{\{v_1,v_2\}} &\in \mathcal{L}_{\{v_1,v_2\}}(X) \\ [c, b]_{\{v_2,v_3\}} &\in \mathcal{L}_{\{v_2,v_3\}}(X) \text{ and} \\ [b]_{\partial D_1(v_1)} &\in \mathcal{L}_{\partial D_1(v_1)}(X). \end{aligned}$$

Proof. Since $a \sim c$ and a can be folded into the vertex b we have $b \sim c$. Consider partite classes $P_1, P_2 \subset \mathcal{V}$ of $\mathcal{G}$ such that $v_1 \in P_1$. Then the configuration $x \in \mathcal{V}^{\mathcal{V}_{\mathcal{H}}}$ given by

$$x_v := \begin{cases} b \text{ if } v \in P_1 \\ c \text{ if } v \in P_2 \end{cases}$$

is an element of $Hom(\mathcal{G},\mathcal{H})$. Thus

$$[b,c]_{\{v_1,v_2\}} = [x]_{\{v_1,v_2\}} \in \mathcal{L}_{\{v_1,v_2\}}(X)$$
$$[c,b]_{\{v_2,v_3\}} = [x]_{\{v_2,v_3\}} \in \mathcal{L}_{\{v_2,v_3\}}(X) \text{ and}$$
$$[b]_{\partial D_1(v_1)} = [x]_{\partial D_1(v_1)} \in \mathcal{L}_{\partial D_1(v_1)}(X).$$

□

For the rest of the chapter fix a bipartite graph $\mathcal{G} = (\mathcal{V},\mathcal{E})$. Let $X \subset \mathcal{A}^{\mathcal{V}}$ be an n.n.constraint space. Given distinct symbols $a, b \in \mathcal{A}$, we say that a can be *strongly config-folded* into b if for all edges $(v_1, v_2), (v_2, v_3) \in \mathcal{E}$ and $c \in \mathcal{A}$, $[a,c]_{\{v_1,v_2\}} \in \mathcal{L}_{\{v_1,v_2\}}(X)$ implies

$$[b,c]_{\{v_1,v_2\}} \in \mathcal{L}_{\{v_1,v_2\}}(X), \tag{3.2.2}$$
$$[c,b]_{\{v_2,v_3\}} \in \mathcal{L}_{\{v_2,v_3\}}(X) \text{ and} \tag{3.2.3}$$
$$[b]_{\partial D_1(v_1)} \in \mathcal{L}_{\partial D_1(v_1)}(X). \tag{3.2.4}$$

In such a case, $X \cap (\mathcal{A} \setminus \{a\})^{\mathcal{V}}$ is called a *strong config-fold* of X and X is called a *strong config-unfold* of $X \cap (\mathcal{A} \setminus \{a\})^{\mathcal{V}}$. Note that $X \cap (\mathcal{A} \setminus \{a\})^{\mathcal{V}}$ is still an n.n.constraint space and is obtained by forbidding the symbol a in X. Further if X is invariant under a subgroup $G \subset Aut(\mathcal{G})$ then $X \cap (\mathcal{A} \setminus \{a\})^{\mathcal{V}}$ is also invariant under G. Let X_a denote the strong config-fold $X \cap (\mathcal{A} \setminus \{a\})^{\mathcal{V}}$. The idea of folding is captured by (3.2.2) while (3.2.3), (3.2.4) are reminiscent of homomorphism spaces. Indeed if an n.n.constraint space X satisfies (3.2.2) then for all $x \in X$ and $v \in \mathcal{V}$ such that $x_v = a$, the configuration $\theta^v_b(x) \in X$. Thus if a strongly config-folds into b then any appearance of a in any configuration in X can be replaced by b. Recall that a safe symbol can replace any other symbol. Thus the notion of strong config-folding generalises the notion of a safe symbol.

Proposition 3.2.4. *Let $\mathcal{G} = (\mathcal{V},\mathcal{E})$ be a bipartite graph. Let $X \subset \mathcal{A}^{\mathcal{V}}$ be an n.n.constraint space with a safe symbol $\star$. Then any symbol $a \in \mathcal{A} \setminus \{\star\}$ can be strongly config-folded into $\star$. The resulting strong config-fold X_a is also an n.n.constraint space with the same safe symbol $\star$.*

Indeed X_a is obtained just by forbidding the symbol a from X and $\star$ is still a safe symbol. In general it is not necessary that the symbol being strongly config-folded into has to be a safe symbol. For instance given any bipartite graph $\mathcal{G}$ the space $Hom(\mathcal{G}, C_4)$, can be strongly config-folded in two steps to $Hom(\mathcal{G}, Edge)$, yet C_4 does not have any safe symbol. Note that the strong config-unfold of an n.n.constraint space with a safe symbol need not have a safe symbol. For example if $\mathcal{H}$ is the graph given by Figure 3.2 then for any bipartite graph $\mathcal{G}$ the top vertex is a safe symbol in the space $Hom(\mathcal{G}, H)$.

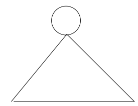

Figure 3.2: $\mathcal{H}$

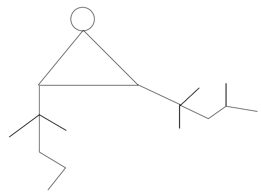

Figure 3.3: $\mathcal{H}'$

However if we attach trees to $\mathcal{H}$ to obtain $\mathcal{H}'$ given by Figure 3.3 then $Hom(\mathcal{G}, \mathcal{H}')$ does not have any safe symbol but can be strongly config-folded into $Hom(\mathcal{G}, \mathcal{H})$ by folding in the trees attached to $\mathcal{H}$.

Strong config-folding induces a natural map between the spaces of configurations and their cocycles as demonstrated by the following proposition.

Proposition 3.2.5. *Let $\mathcal{G}$ be a bipartite graph and $G \subset Aut(\mathcal{G})$ be a subgroup. Suppose $X \subset \mathcal{A}^{\mathcal{V}}$ is a G-invariant n.n.constraint space and let X_a be its strong config-fold. Then the linear map $F : \mathbf{M}^G_X \longrightarrow \mathbf{M}^G_{X_a}$ given by $F(M) := M|_{\Delta_{X_a}}$ is surjective and $F(\mathbf{G}^G_X) = \mathbf{G}^G_{X_a}$.*

Proof. If $M \in \mathbf{G}^G_X$ then the restriction of the G-invariant nearest neighbour interaction for M to X_a gives us a G-invariant nearest neighbour interaction for $F(M)$ proving that $F(M) \in \mathbf{G}^G_{X_a}$. Thus $F(\mathbf{G}^G_X) \subset \mathbf{G}^G_{X_a}$. We will construct a map $\phi^\star : \mathbf{M}^G_{X_a} \longrightarrow \mathbf{M}^G_X$ such that $\phi^\star(\mathbf{G}^G_{X_a}) \subset \mathbf{G}^G_X$ and $F \circ \phi^\star$ is the identity map on $\mathbf{M}^G_{X_a}$. Note that this is sufficient to conclude that F is surjective and $F(\mathbf{G}^G_X) = \mathbf{G}^G_{X_a}$ thereby completing the proof.

The strong config-folding induces a mapping $\phi : X \longrightarrow X_a$ given by

$$\phi(x)_v := \begin{cases} x_v \text{ if } x_v \neq a \\ b \quad \text{if } x_v = a \end{cases}$$

for all $x \in X$ and $v \in \mathcal{V}$. Let $g \in G$ and $x \in X$. Then

$$(\phi(gx))_v = \begin{cases} (gx)_v = x_{g^{-1}v} \text{ if } x_{g^{-1}v} \neq a \\ b \qquad\qquad\qquad \text{if } (gx)_v = x_{g^{-1}v} = a \end{cases}$$

and

$$(g(\phi(x))_v = (\phi(x))_{g^{-1}v} = \begin{cases} x_{g^{-1}v} \text{ if } x_{g^{-1}v} \neq a \\ b \qquad \text{if } x_{g^{-1}v} = a. \end{cases}$$

Therefore ϕ commutes with the action of G. Note that $\phi|_{X_a}$ is the identity.

The map ϕ in turn induces a map between the cocycles which we shall now describe. Let $M \in \mathbf{M}^G_{X_a}$ be a Markov cocycle. Consider $M' : \Delta_X \longrightarrow \mathbb{R}$ given by

$$M'(x,y) := M(\phi(x), \phi(y)).$$

We will prove that $M' \in \mathbf{M}^G_X$.

Cocycle condition: If $(x,y),(y,z) \in \Delta_X$ then

$$M'(x,y) + M'(y,z) = M(\phi(x),\phi(y)) + M(\phi(y),\phi(z)) = M(\phi(x),\phi(z)) = M'(x,z).$$

Markov condition: If $(x,y),(z,w) \in \Delta_X$ are Markov-similar then $(\phi(x),\phi(y)),(\phi(z),\phi(w)) \in \Delta_{X_a}$ are Markov-similar as well implying $M(\phi(x),\phi(y)) = M(\phi(z),\phi(w))$ and thus

$$\begin{aligned} M'(x,y) &= M(\phi(x),\phi(y)) \\ &= M(\phi(z),\phi(w)) \\ &= M'(z,w) \end{aligned}$$

which verifies the Markov condition for M'.

G-invariance condition: Since ϕ commutes with the action of G, for all $g \in G$

$$M'(gx,gy) = M(\phi(gx),\phi(gy)) = M(g(\phi(x)),g(\phi(y))) = M(\phi(x),\phi(y)) = M'(x,y).$$

Hence $M' \in \mathbf{M}^G_X$. Moreover if $M \in \mathbf{G}^G_{X_a}$ with a G-invariant nearest neighbour interaction V, then for all $(x,y) \in \Delta_X$

$$M'(x,y) = M(\phi(x),\phi(y)) = \sum_{A \subset \mathcal{V} \text{ finite}} V([\phi(y)]_A) - V([\phi(x)]_A)$$

proving that $V \circ \phi$ is a G-invariant nearest neighbour interaction for M'.

Thus the map $\phi^\star : \mathbf{M}^G_{X_a} \longrightarrow \mathbf{M}^G_X$ given by

$$\phi^\star(M)(x,y) := M(\phi(x),\phi(y))$$

satisfies $\phi^\star(\mathbf{G}^G_{X_a}) \subset \mathbf{G}^G_X$. Moreover since $\phi|_{X_a}$ is the identity map on X_a therefore $\phi^\star(M)|_{\Delta_{X_a}} = M$ for all $M \in \mathbf{M}^G_{X_a}$ proving $F \circ \phi^\star$ is the identity map on $\mathbf{M}_{X_a}$. $\square$

Given a G-invariant topological Markov field $Y \subset X$ there is always a linear map $F : \mathbf{M}^G_X \longrightarrow \mathbf{M}^G_Y$ given by $F(M) := M|_{\Delta_Y}$ and $F(\mathbf{G}^G_X) \subset \mathbf{G}^G_Y$. However if Y cannot be obtained by a sequence of strong config-folds starting with X, then this map need not be surjective. Indeed, consider the following example:

Let $\mathcal{H}$ be the graph given by Figure 3.2. Fix some $d \in \mathbb{N}$. Let $X := Hom(\mathbb{Z}^d, \mathcal{H})$ and $Y := Hom(\mathbb{Z}^d, C_3)$. Since there is a graph embedding from the 3-cycle C_3 to $\mathcal{H}$ it follows that $Hom(\mathbb{Z}^d, C_3) \subset Hom(\mathbb{Z}^d, \mathcal{H})$. Let $\mathbb{Z}^d$ denote the group of translations of the $\mathbb{Z}^d$ lattice. Since the top vertex of $\mathcal{H}$ is a safe symbol for $Hom(\mathbb{Z}^d, \mathcal{H})$ it follows from the strong Hammersley-Clifford theorem (Theorem 2.2.3) that $\mathbf{M}_X^{\mathbb{Z}^d} = \mathbf{G}_X^{\mathbb{Z}^d}$. Therefore $F(\mathbf{M}_X^{\mathbb{Z}^d}) \subset \mathbf{G}_Y^{\mathbb{Z}^d}$. However by Proposition 2.4.3, $\mathbf{G}_Y^{\mathbb{Z}^d} \subsetneq \mathbf{M}_Y^{\mathbb{Z}^d}$. It follows that $F(\mathbf{M}_X^{\mathbb{Z}^d}) \subsetneq \mathbf{M}_Y^{\mathbb{Z}^d}$.

3.3 The Main Results

Theorem 3.3.1. *Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be a bipartite graph, $\mathcal{A}$ a finite alphabet and $X \subset \mathcal{A}^{\mathcal{V}}$ a Hammersley-Clifford n.n.constraint space . Then the strong config-folds and strong config-unfolds of X are also Hammersley-Clifford.*

The G-invariant version of Theorem 3.3.1 holds as well.

Theorem 3.3.2. *Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be a bipartite graph, $\mathcal{A}$ a finite alphabet, $G \subset Aut(\mathcal{G})$ a subgroup and $X \subset \mathcal{A}^{\mathcal{V}}$ a G-Hammersley-Clifford n.n.constraint space. Then the strong config-folds and strong config-unfolds of X are also G-Hammersley-Clifford.*

We know that all frozen spaces of configurations are G-Hammersley-Clifford for all subgroups $G \subset Aut(\mathcal{G})$. We can construct many more examples of Hammersley-Clifford spaces by using these theorems.

1. N.N.Constraint space with a safe symbol.

   By Proposition 3.2.4 starting with an n.n.constraint space with a safe symbol $\star$ we can strong config-fold all the symbols one by one into the symbol $\star$ resulting in $\{\star\}^{\mathcal{V}}$ which is frozen. Thus these theorems generalise Theorem 2.2.3 in the case when $\mathcal{G}$ is a bipartite graph. Furthermore any closed configuration space which can be strongly config-folded into a space with a safe symbol is still Hammersley-Clifford. For instance given the graph $\mathcal{H}'$ in Figure 3.3, even though $Hom(\mathcal{G}, \mathcal{H}')$ does not have any safe symbol, it is G-Hammersley-Clifford for any subgroup $G \subset Aut(\mathcal{G})$.

2. $Hom(\mathcal{G}, Edge)$ where $Edge$ consists of two vertices 0 and 1 connected by a single edge.

   By these theorems a closed configuration space which can be strongly config-folded into $Hom(\mathcal{G}, Edge)$ is still Hammersley-Clifford. For example if $\mathcal{H}$ is the graph given by Figure 3.4 then it can be folded to the graph $Edge$ and hence $Hom(\mathcal{G}, \mathcal{H})$ is G-Hammersley-Clifford for any subgroup $G \subset Aut(\mathcal{G})$.

3. Consider the space $Hom(\mathcal{G}, \mathcal{H}_{n,m})$ where $\mathcal{H}_{n,m}$ is a graph with vertices $\mathcal{V}_{\mathcal{H}_{n,m}} := \{1, 2, \ldots, n\}$ and edges given by $(i, j) \in \mathcal{E}_{\mathcal{H}_{n,m}}$ if and only if $|i - j| \leq m$.

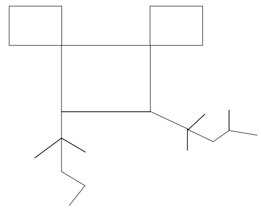

Figure 3.4: A Graph which Folds to the Edge Graph

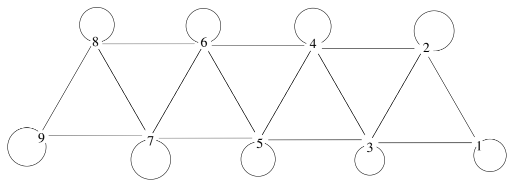

Figure 3.5: $\mathcal{H}_{9,2}$

> The sequence of folds 1 to 2, 2 to 3, 3 to 4, $\ldots$, $n-1$ to n yields the space $\{n\}^{\mathcal{G}}$ from $Hom(\mathcal{G}, \mathcal{H}_{n,m})$ proving that it is G-Hammersley-Clifford for any subgroup $G \subset Aut(\mathcal{G})$. A graph $\mathcal{H}$ is called *dismantlable* if there exists a sequence of folds on the graph leading to a single vertex. By these theorems, if $\mathcal{H}$ is dismantlable then $Hom(\mathcal{G}, \mathcal{H}_{n,m})$ is G-Hammersley-Clifford for any subgroup $G \subset Aut(\mathcal{G})$.

Note that although these are homomorphism spaces, the theorems are true in the general setting of closed configuration spaces. These specific examples have been chosen for convenience.

3.3.1 A Concrete Example

We will first work out the following example to illustrate the key ideas of the proof.

Suppose $\mathcal{H}$ and $\mathcal{H}'$ are graphs given by Figure 3.6. Let $X := Hom(\mathbb{Z}^2, \mathcal{H})$. Then by strong config-folding the vertex a into the vertex b we obtain the space $X_a := Hom(\mathbb{Z}^2, \mathcal{H}')$.

Note that X does not have any safe symbol but b is a safe symbol for X_a. Let $\mathbb{Z}^2$ denote the subgroup of all translations of $\mathbb{Z}^2$. By the strong Hammersley-Clifford theorem (Theorem 2.2.3) X_a is $\mathbb{Z}^2$-Hammersley-Clifford. We will prove that X is $\mathbb{Z}^2$-Hammersley-Clifford.

Let $M \in \mathbf{M}_X^{\mathbb{Z}^2}$ be a shift-invariant Gibbs cocycle. Then $M|_{X_a}$ is a shift-invariant Markov cocycle on X_a and hence a Gibbs cocycle with some shift-invariant nearest neighbour interaction, which

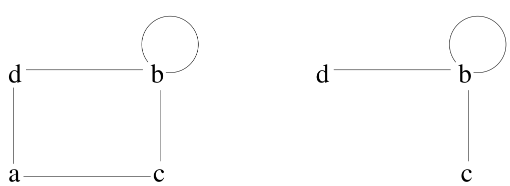

Figure 3.6: Graphs $\mathcal{H}$ and $\mathcal{H}'$

we will call V.

For $e,f,g,h,i \in \mathcal{V}_{\mathcal{H}}$ and $\vec{v} \in \mathbb{Z}^2$ consider the configuration $x = \left[\begin{smallmatrix} & e & \\ f & g & h \\ & i & \end{smallmatrix}\right]^{\vec{v}}$ given by

$$x_u := \begin{cases} g \text{ if } u = \vec{v} \\ e \text{ if } u = \vec{v} + (0,1) \\ f \text{ if } u = \vec{v} - (1,0) \\ h \text{ if } u = \vec{v} + (1,0) \\ i \text{ if } u = \vec{v} - (0,1) \\ b \text{ if } \vec{u} \in D_1(\vec{v})^c. \end{cases}$$

For all $\vec{v} \in \mathbb{Z}^2$ let $x^{\vec{v}} = \left[\begin{smallmatrix} & d & \\ d & a & d \\ & d & \end{smallmatrix}\right]^{\vec{v}}$. Consider a shift-invariant nearest neighbour interaction V' as follows:

1. If $\vec{v} \sim \vec{w} \in \mathbb{Z}^2$, $[e,f]_{\{\vec{v},\vec{w}\}} \in \mathcal{L}_{\{\vec{v},\vec{w}\}}(X_a)$ then

$$\begin{aligned} V'([e,f]_{\{\vec{v},\vec{w}\}}) &:= V([e,f]_{\{\vec{v},\vec{w}\}}) \text{ and} \\ V'([e]_{\vec{v}}) &:= V([e]_{\vec{v}}). \end{aligned} \tag{3.3.1}$$

2. The interaction between a and d is 0, that is, for all $\vec{v} \sim \vec{w} \in \mathbb{Z}^2$

$$V'([a,d]_{\{\vec{v},\vec{w}\}}) := 0. \tag{3.3.2}$$

3. The single site interaction for $[a]_{\vec{v}}$ for all $\vec{v} \in \mathbb{Z}^2$ is given by

$$\begin{aligned} V'([a]_{\vec{v}}) &:= M\left(\left[\begin{smallmatrix} & d & \\ d & b & d \\ & d & \end{smallmatrix}\right]^{\vec{v}}, \left[\begin{smallmatrix} & d & \\ d & a & d \\ & d & \end{smallmatrix}\right]^{\vec{v}}\right) + V([b]_{\vec{v}}) + V([b,d]_{\{\vec{v},\vec{v}+(1,0)\}}) + V([b,d]_{\{\vec{v},\vec{v}-(1,0)\}}) \\ &\quad + V([b,d]_{\{\vec{v},\vec{v}+(0,1)\}}) + V([b,d]_{\{\vec{v},\vec{v}-(0,1)\}}). \end{aligned}$$

   By (3.3.1) and (3.3.2) this implies that the pair $\left(\left[\begin{smallmatrix} & d & \\ d & b & d \\ & d & \end{smallmatrix}\right]^{\vec{v}}, \left[\begin{smallmatrix} & d & \\ d & a & d \\ & d & \end{smallmatrix}\right]^{\vec{v}}\right)$ is V'-good.

4. Let

$$\begin{aligned} V'([a,c]_{\{\vec{v},\vec{v}+(1,0)\}}) &:= M\left(\left[\begin{smallmatrix} & d & \\ d & a & d \\ & d & \end{smallmatrix}\right]^{\vec{v}}, \left[\begin{smallmatrix} & d & \\ d & a & c \\ & d & \end{smallmatrix}\right]^{\vec{v}}\right) + V([d]_{\vec{v}+(1,0)}) - V([c]_{\vec{v}+(1,0)}) \\ &\quad + V([d,b]_{\{\vec{v}+(1,0),\vec{v}+(1,1)\}}) + V([d,b]_{\{\vec{v}+(1,0),\vec{v}+(2,0)\}}) \\ &\quad + V([d,b]_{\{\vec{v}+(1,0),\vec{v}+(1,-1)\}}) - V([c,b]_{\{\vec{v}+(1,0),\vec{v}+(1,1)\}}) \\ &\quad - V([c,b]_{\{\vec{v}+(1,0),\vec{v}+(2,0)\}}) - V([c,b]_{\{\vec{v}+(1,0),\vec{v}+(1,-1)\}}). \end{aligned}$$

By (3.3.1) and (3.3.2) the previous equation implies that the pair $\left(\begin{bmatrix} & d & \\ d & a & d \\ & d & \end{bmatrix}^{\vec{v}}, \begin{bmatrix} & d & \\ d & a & c \\ & d & \end{bmatrix}^{\vec{v}}\right)$ is V'-good. Similarly we can define $V'([a,c]_{\{\vec{v},\vec{v}-(1,0)\}})$, $V'([a,c]_{\{\vec{v},\vec{v}+(0,1)\}})$ and $V'([a,c]_{\{\vec{v},\vec{v}-(0,1)\}})$, the corresponding expressions of which will imply that the pairs $\left(\begin{bmatrix} & d & \\ d & a & d \\ & d & \end{bmatrix}^{\vec{v}}, \begin{bmatrix} & d & \\ c & a & d \\ & d & \end{bmatrix}^{\vec{v}}\right)$, $\left(\begin{bmatrix} & d & \\ d & a & d \\ & d & \end{bmatrix}^{\vec{v}}, \begin{bmatrix} & c & \\ d & a & d \\ & d & \end{bmatrix}^{\vec{v}}\right)$, $\left(\begin{bmatrix} & d & \\ d & a & d \\ & d & \end{bmatrix}^{\vec{v}}, \begin{bmatrix} & d & \\ d & a & d \\ & c & \end{bmatrix}^{\vec{v}}\right)$ are V'-good.

Since V and M are shift-invariant it follows that V' is also shift-invariant. We want to prove that V' is an interaction for M. Equivalently we want to prove that all asymptotic pairs are V'-good. Let $(x,y) \in \Delta_X$. Since any appearance of a in the elements of X can be replaced by b, by replacing all the a's outside the set of sites where x and y differ and its boundary we can obtain a pair $(x^1, y^1) \in \Delta_X$ which is Markov-similar to (x,y) and has finitely many a's. Thus by Proposition 3.2.1 it is sufficient to prove that pairs $(x,y) \in \Delta_X$ with finitely many a's are V'-good. Since the a's can be replaced by b's one by one and any pair in Δ_{X_a} is V'-good by Lemma 3.2.2 it is sufficient to prove that pairs in X in which a single a is replaced by b are V'-good. Since a can be folded into b and $\partial\{a\} = \{c,d\}$ any such pair is Markov-similar to a pair of the type $\left(\begin{bmatrix} & e & \\ f & a & h \\ & i & \end{bmatrix}^{\vec{v}}, \begin{bmatrix} & e & \\ f & b & h \\ & i & \end{bmatrix}^{\vec{v}}\right)$ for some $\vec{v} \in \mathbb{Z}^2$ and $e,f,g,h,i \in \{c,d\}$.

The pairs

$$\left(\begin{bmatrix} & e & \\ f & a & h \\ & i & \end{bmatrix}^{\vec{v}}, \begin{bmatrix} & d & \\ f & a & h \\ & i & \end{bmatrix}^{\vec{v}}\right), \left(\begin{bmatrix} & d & \\ f & a & h \\ & i & \end{bmatrix}^{\vec{v}}, \begin{bmatrix} & d & \\ d & a & h \\ & i & \end{bmatrix}^{\vec{v}}\right), \left(\begin{bmatrix} & d & \\ d & a & h \\ & i & \end{bmatrix}^{\vec{v}}, \begin{bmatrix} & d & \\ d & a & d \\ & i & \end{bmatrix}^{\vec{v}}\right), \left(\begin{bmatrix} & d & \\ d & a & d \\ & i & \end{bmatrix}^{\vec{v}}, \begin{bmatrix} & d & \\ d & a & d \\ & d & \end{bmatrix}^{\vec{v}}\right)$$

are Markov-similar to

$$\left(\begin{bmatrix} & e & \\ d & a & d \\ & d & \end{bmatrix}^{\vec{v}}, \begin{bmatrix} & d & \\ d & a & d \\ & d & \end{bmatrix}^{\vec{v}}\right), \left(\begin{bmatrix} & d & \\ f & a & d \\ & d & \end{bmatrix}^{\vec{v}}, \begin{bmatrix} & d & \\ d & a & d \\ & d & \end{bmatrix}^{\vec{v}}\right), \left(\begin{bmatrix} & d & \\ d & a & h \\ & d & \end{bmatrix}^{\vec{v}}, \begin{bmatrix} & d & \\ d & a & d \\ & d & \end{bmatrix}^{\vec{v}}\right), \left(\begin{bmatrix} & d & \\ d & a & d \\ & i & \end{bmatrix}^{\vec{v}}, \begin{bmatrix} & d & \\ d & a & d \\ & d & \end{bmatrix}^{\vec{v}}\right)$$

respectively. Since $e,f,g,h,i \in \{c,d\}$, these pairs are V'-good. Thus each adjacent pair in the chain

$$\begin{bmatrix} & e & \\ f & a & h \\ & i & \end{bmatrix}^{\vec{v}}, \begin{bmatrix} & d & \\ f & a & h \\ & i & \end{bmatrix}^{\vec{v}}, \begin{bmatrix} & d & \\ d & a & h \\ & i & \end{bmatrix}^{\vec{v}}, \begin{bmatrix} & d & \\ d & a & d \\ & i & \end{bmatrix}^{\vec{v}}, \begin{bmatrix} & d & \\ d & a & d \\ & d & \end{bmatrix}^{\vec{v}}, \begin{bmatrix} & d & \\ d & b & d \\ & d & \end{bmatrix}^{\vec{v}}, \begin{bmatrix} & e & \\ f & b & h \\ & i & \end{bmatrix}^{\vec{v}}$$

is V'-good. By Corollary 3.2.2 the pair $\left(\begin{bmatrix} & e & \\ f & a & h \\ & i & \end{bmatrix}^{\vec{v}}, \begin{bmatrix} & e & \\ f & b & h \\ & i & \end{bmatrix}^{\vec{v}}\right)$ is V'-good. This completes the proof.

3.3.2 Proof of Theorems 3.3.1 and 3.3.2

We will now prove Theorems 3.3.1 and 3.3.2. The proof will give an explicit way of computing the interaction as well. It should also be noted that Theorem 3.3.1 is a special case of Theorem 3.3.2. Yet we separate the proofs so as to separate the various complications.

Proof of Theorem 3.3.1. The bulk of the proof lies in showing that the strong config-unfolds of

Hammersley-Clifford spaces are Hammersley-Clifford. We will first prove that the strong config-folds of a Hammersley-Clifford space are Hammersley-Clifford. Let $X \subset \mathcal{A}^{\mathcal{V}}$ be Hammersley-Clifford and X_a be its strong config-fold. Using Proposition 3.2.5 in the case where $G = \{id|_{\mathcal{G}}\}$ we obtain a surjective map $F : \mathbf{M}_X \longrightarrow \mathbf{M}_{X_a}$ such that $F(\mathbf{G}_X) = \mathbf{G}_{X_a}$. Since X is Hammersley-Clifford, $\mathbf{M}_X = \mathbf{G}_X$. Hence

$$\mathbf{M}_{X_a} = F(\mathbf{M}_X) = F(\mathbf{G}_X) = \mathbf{G}_{X_a}$$

proving that X_a is Hammersley-Clifford.

Now we will prove that strong config-unfolds of Hammersley-Clifford spaces are Hammersley-Clifford spaces as well. Let $X \subset \mathcal{A}^{\mathcal{V}}$ be an n.n.constraint space and X_a be a strong config-fold of X where a is strongly config-folded into b. Let the set of nearest neighbour constraints of X be given by the set $\mathcal{F}_X$. Suppose X_a is Hammersley-Clifford.

Let $M \in \mathbf{M}_X$ be a Markov cocycle. Since X_a is Hammersley-Clifford $M|_{\Delta_{X_a}} \in \mathbf{G}_{X_a}$. Let V be a corresponding nearest neighbour interaction. We shall now construct a nearest neighbour interaction V' for M. The idea is the following:

Since we have a nearest neighbour interaction for $M|_{\Delta_{X_a}}$ we will change asymptotic pairs in X to asymptotic pairs in X_a using the fewest possible distinct single site changes. These distinct single site changes will correspond to patterns on edges and vertices helping us build V'. If we use the single site changes which involve blindly changing the a's into b's we will incur a large number of such changes; instead we will use a smaller number as described by the following lemma.

Lemma 3.3.3 (Construction of special configurations)**.** *Let X be an n.n.constraint space and X_a be a strong config-fold of X where the symbol a is strongly config-folded into the symbol b. Let*

$$\mathcal{V}_1 := \left\{v \in \mathcal{V} \mid \textit{there exists } w \sim v \textit{ such that } [a,a]_{\{v,w\}} \in \mathcal{L}_{\{v,w\}}(X)\right\}$$

and

$$\mathcal{V}_2 := \left\{v \in \mathcal{V} \setminus \mathcal{V}_1 \mid [a]_{\{v\}} \in \mathcal{L}_{\{v\}}(X)\right\}.$$

For all $v \in \mathcal{V}_1 \cup \mathcal{V}_2$ there exists $x^v \in X$ such that

1. *If $v \in \mathcal{V}_1$ then $x^v_v := a$ and $x^v|_{D_2(v)\setminus\{v\}} := b$.*

2. *If $v \in \mathcal{V}_2$ then $x^v_v := a$ and $x^v|_{\partial D_1(v)} := b$.*

Moreover $\theta^v_b(x^v) \in X_a$ and if $w_1, w_2, w_3, \ldots w_r \sim v$ and $c_1, c_2, \ldots, c_r \in \mathcal{A}$ such that $[a, c_i]_{\{v,w_i\}} \in \mathcal{L}_{\{v,w_i\}}(X)$ then $\theta^{w_1,w_2,\ldots,w_r}_{c_1,c_2,\ldots,c_r}(x^v) \in X$.

Proof. Let $v \in \mathcal{V}_1$. By (3.2.3) $[a,b]_{\{v,w\}} \in \mathcal{L}_{\{v,w\}}(X)$ for all $w \sim v$. Again by (3.2.3) it follows that $[b,b]_{\{w,w_1\}} \in \mathcal{L}_{\{w,w_1\}}(X)$ for all $w, w_1 \in \mathcal{V}$ such that $w \sim v$ and $w_1 \sim w$. Then none of the patterns from $\mathcal{F}_X$, the nearest neighbour constraint set for X appear in $\alpha^v \in \mathcal{A}^{D_2(v)}$ given by

$$\alpha^v_u := \begin{cases} a & \text{if } u = v \\ b & \text{if } u \in D_2(v) \setminus \{v\}. \end{cases}$$

For $v \in \mathcal{V}_2$ there exists $x^1 \in X$ such that $x^1_v = a$. For all $w, w_1 \in \mathcal{V}$ such that $w \sim v$ and $w_1 \sim w$ (3.2.3) implies that $[x^1_w, b]_{\{w,w_1\}} \in \mathcal{L}_{\{w,w_1\}}(X)$. Then none of the patterns from $\mathcal{F}_X$ appear in $\alpha^v \in \mathcal{A}^{D_2(v)}$ given by

$$\alpha^v_u := \begin{cases} x^1_u & \text{if } u \in D_1(v) \\ b & \text{if } u \in D_2(v) \setminus D_1(v). \end{cases}$$

Fix $v \in \mathcal{V}_1 \cup \mathcal{V}_2$. By (3.2.4) there exists $x \in X$ such that $x|_{\partial D_1(v)} = b$. Moreover since a strongly config-folds into b we can assume that $x \in X_a$. Consider $x^v \in \mathcal{A}^{\mathcal{V}}$ given by

$$x^v_u := \begin{cases} \alpha^v_u \text{ if } u \in D_2(v) \\ x_u \text{ if } u \in D_1(v)^c. \end{cases}$$

The configurations x^v satisfy Conclusions (1) and (2) of this lemma. Since each edge in $\mathcal{G}$ either lies completely in $D_2(v)$ or in $D_1(v)^c$, no subpattern of x^v belongs to $\mathcal{F}_X$. Therefore $x^v \in X$.

Let $v \in \mathcal{V}_1 \cup \mathcal{V}_2$. Since $x \in X_a$, a appears in x^v only at v . Moreover since a strongly config-folds into b by (3.2.2), $\theta^v_b(x^v) \in X_a$. Let $w_1, w_2, w_3, \ldots w_r \sim v$ and $c_1, c_2, \ldots, c_r \in \mathcal{A}$ such that $[a, c_i]_{\{v,w_i\}} \in \mathcal{L}_{\{v,w_i\}}(X)$ for all $1 \leq i \leq r$. Because the graph is bipartite $w_i \nsim w_j$ for all $1 \leq i, j \leq r$. By (3.2.3) for all $w' \sim w_i$ and $1 \leq i \leq r$, $[c_i, b]_{\{w_i,w'\}} \in \mathcal{L}_{\{w_i,w'\}}(X)$. By Proposition 3.1.3 $\theta^{w_1,w_2,\ldots,w_r}_{c_1,c_2,\ldots,c_r}(x^v) \in X$. $\square$

We will now construct an interaction via the following technical lemma.

Lemma 3.3.4 (Construction of V')**.** *Let X be an n.n.constraint space and X_a be a strong config-fold of X where the symbol a is strongly config-folded into the symbol b. Consider sets $\mathcal{V}_1, \mathcal{V}_2 \subset \mathcal{V}$ and for all $v \in \mathcal{V}_1 \cup \mathcal{V}_2$, configurations $x^v \in X$ satisfying the conclusions of Lemma 3.3.3. Let $M \in \mathbf{M}_X$ be a Markov cocycle on X such that $M|_{\Delta_{X_a}}$ is a Gibbs cocycle with interaction V. Then there exists a unique nearest neighbour interaction V' on X which satisfies:*
If $v \sim w \in \mathcal{V}$ and $[c, d]_{\{v,w\}} \in \mathcal{L}_{\{v,w\}}(X_a)$ then

$$V'([c,d]_{\{v,w\}}) = V([c,d]_{\{v,w\}}) \text{ and} \tag{3.3.3}$$

$$V'([c])_{\{v\}} = V([c]_{\{v\}}). \tag{3.3.4}$$

For $v \in \mathcal{V}_1 \cup \mathcal{V}_2$ and $w \sim v$

$$V'([x^v]_{\{v,w\}}) = 0. \tag{3.3.5}$$

such that the following pairs are V'-good:

1. $(\tilde{x}, \tilde{y}) \in \Delta_{X_a}$.
2. $(\theta_b^v(x^v), x^v)$ *for* $v \in \mathcal{V}_1 \cup \mathcal{V}_2$.
3. $(\theta_c^w(x^v), x^v)$ *for* $v \in \mathcal{V}_1 \cup \mathcal{V}_2$, $w \sim v$ *and* $c \in \mathcal{A} \setminus \{a\}$ *satisfying* $[a, c]_{\{v,w\}} \in \mathcal{L}_{\{v,w\}}(X)$.
4. $(\theta_a^w(x^v), x^v)$ *for all* $v \in \mathcal{V}_1 \cap P_1$, $w \sim v$ *satisfying* $[a, a]_{\{v,w\}} \in \mathcal{L}_{\{v,w\}}(X)$.

In the following proof the reader is encouraged to refer to the statement of Lemma 3.3.3 for information about configurations x^v.

Proof of Lemma 3.3.4. We will begin by proving uniqueness of the interaction assuming its existence. Consider a nearest neighbour interaction V' on X which satisfies the conclusion of this lemma. We will prove the uniqueness by expressing V' in terms of the cocycle M and V.

Since V' satisfies (3.3.3), (3.3.4) and (3.3.5) we have to prove that the following can be expressed in terms of M and V:

(a) For all $v \in \mathcal{V}_1 \cup \mathcal{V}_2$, the value $V'([a]_v)$,

(b) For all $v \in \mathcal{V}_1 \cup \mathcal{V}_2$, $w \sim v$ and $c \in \mathcal{A} \setminus \{x_w^v, a\}$ such that $[a, c]_{\{v,w\}} \in \mathcal{L}_{\{v,w\}}(X)$, the value $V'([a, c]_{\{v,w\}})$ and

(c) For all $v \in \mathcal{V}_1 \cap P_1$, $w \sim v$ such that $[a, a]_{\{v,w\}} \in \mathcal{L}_{\{v,w\}}(X)$, the value $V'([a, a]_{\{v,w\}})$.

Proof for part (a): Let $v \in \mathcal{V}_1 \cup \mathcal{V}_2$. Since the pair $(\theta_b^v(x^v), x^v)$ ((2) in the statement of the lemma) is V'-good by rearranging the expression for $M(\theta_b^v(x^v), x^v)$ we get that

$$\begin{aligned} V'([a]_v) &= V'([x^v]_v) \\ &= M(\theta_b^v(x^v), x^v) + V'([\theta_b^v(x^v)]_v) + \sum_{w: w \sim v} V'\left([\theta_b^v(x^v)]_{\{v,w\}}\right) \\ &\quad -\Big(\sum_{w: w \sim v} V'\left([x^v]_{\{v,w\}}\right)\Big). \end{aligned} \tag{3.3.6}$$

Now we will express the right hand side of this expression in terms of M and V. Since $\theta_b^v(x^v) \in X_a$ $V'([\theta_b^v(x^v)]_{\{v,w\}}) = V([\theta_b^v(x^v)]_{\{v,w\}})$ and $V'([\theta_b^v(x^v)]_v) = V([\theta_b^v(x^v)]_v)$. By (3.3.5), $V'([x^v]_{\{v,w\}}) = 0$.

Putting all this together we get

$$V'([a]_v) = M(\theta_b^v(x^v), x^v) + V([\theta_b^v(x^v)]_v) + \sum_{w: w \sim v} V([\theta_b^v(x^v)]_{\{v,w\}}). \tag{3.3.7}$$

Proof for part (b): Consider $v \in \mathcal{V}_1 \cup \mathcal{V}_2$, $w \sim v$ and $c \in \mathcal{A} \setminus \{a, x_w^v\}$ such that $[a, c]_{\{v,w\}} \in \mathcal{L}_{\{v,w\}}(X)$. Since the pair $(\theta_c^w(x^v), x^v)$ ((3) in the statement of the lemma) is V'-good by rearranging the expression for $M(x^v, \theta_c^w(x^v))$ we get

$$\begin{aligned} V'([a,c]_{\{v,w\}}) &= V'([\theta_c^w(x^v)]_{\{v,w\}}) \\ &= M(x^v, \theta_c^w(x^v)) + \sum_{w':w'\sim w} V'([x^v]_{\{w',w\}}) + V'([x^v]_w) \\ &\quad - \left(\sum_{w':w'\sim w, w'\neq v} V'([\theta_c^w(x^v)]_{\{w',w\}}) \right) - V'([\theta_c^w(x^v)]_w). \end{aligned} \tag{3.3.8}$$

We will now express the right hand side of this expression in terms of M and V.

By (3.3.5), $V'([x^v]_{\{v,w\}}) = 0$. We know that $(\theta_c^w(x^v))_w$, $x_w^v \neq a$ and if $w' \sim w$, $w' \neq v$ then $w' \in \partial D_1(v)$ and so $(\theta_c^w(x^v))_{w'} = x_{w'}^v = b$. Therefore by (3.3.3) and (3.3.4)

$$V'([x^v]_{\{w',w\}}) = V([x^v]_{\{w',w\}}),\ V'([\theta_c^w(x^v)]_{\{w',w\}}) = V([\theta_c^w(x^v)]_{\{w',w\}})$$

and

$$V'([x^v]_w) = V([x^v]_w),\ V'([\theta_c^w(x^v)]_w) = V([\theta_c^w(x^v)]_w).$$

Putting all this together we get

$$\begin{aligned} V'([a,c]_{\{v,w\}}) &= M(x^v, \theta_c^w(x^v)) + \sum_{w':w'\sim w, w'\neq v} \big(V([x^v]_{\{w',w\}}) - V([\theta_c^w(x^v)]_{\{w',w\}})\big) \\ &\quad + V([x^v]_w) - V([\theta_c^w(x^v)]_w). \end{aligned} \tag{3.3.9}$$

Proof for part (c): Consider $v \in \mathcal{V}_1 \cap P_1$ and $w \sim v$ such that $[a,a]_{\{v,w\}} \in \mathcal{L}_{\{v,w\}}(X)$. Since the pair $(\theta_a^w(x^v), x^v)$ ((4) in the statement of the lemma) is V'-good by rearranging the expression for $M(x^v, \theta_a^w(x^v))$ we get that

$$\begin{aligned} V'([a,a]_{\{v,w\}}) &= V'([\theta_a^w(x^v)]_{\{v,w\}}) \\ &= M(x^v, \theta_a^w(x^v)) + \sum_{w':w'\sim w} V'([x^v]_{\{w',w\}}) + V'([x^v]_w) \\ &\quad - \left(\sum_{w':w'\sim w, w'\neq v} V'([\theta_a^w(x^v)]_{\{w',w\}}) \right) - V'([\theta_a^w(x^v)]_w). \end{aligned} \tag{3.3.10}$$

We will now express the right hand side of this expression in terms of M and V. By (3.3.5), $V'([x^v]_{\{v,w\}}) = 0$. Since $v \in \mathcal{V}_1$, for $w' \sim w$ such that $w' \neq v$ we know that $x_w^v = x_{w'}^v = b \neq a$. Therefore by (3.3.3) and (3.3.4)

$$V'([x^v]_{\{w',w\}}) = V([x^v]_{\{w',w\}})$$

and

$$V'([x^v]_w) = V([x^v]_w).$$

Since $[a,a] \in \mathcal{L}_{\{v,w\}}(X)$ therefore $v, w \in \mathcal{V}_1$ and $x^v_{w'} = x^w_{w'} = b$ for all $w' \sim w$, $w' \neq v$. Then by (3.3.5)

$$V'([\theta^w_a(x^v)]_{\{w',w\}}) = V'([b,a]_{\{w',w\}}) = V'([x^w]_{\{w',w\}}) = 0.$$

By (3.3.7) we get that

$$V'([\theta^w_a(x^v)]_w) = V'([a]_w) = M(\theta^w_b(x^w), x^w) + V([\theta^w_b(x^w)]_w) + \sum_{w':w'\sim w} V([\theta^w_b(x^w)]_{\{w,w'\}}).$$

Putting all this together, we get

$$\begin{aligned} V'([a,a]_{\{v,w\}}) &= M(x^v, \theta^w_a(x^v)) + \sum_{w':w'\sim w, w'\neq v} V([x^v]_{\{w',w\}}) + V([x^v]_w) - M(\theta^w_b(x^w), x^w) \\ &\quad -V([\theta^w_b(x^w)]_w) - (\sum_{w':w'\sim w} V([\theta^w_b(x^w)]_{\{w,w'\}})). \end{aligned} \tag{3.3.11}$$

This completes proof for uniqueness. It follows from the proofs that given an interaction V' which satisfies (3.3.3), (3.3.4) and (3.3.5), Equations(3.3.7), (3.3.9) and (3.3.11) are satisfied if and only if the pairs listed in (1), (2), (3) and (4) are V'-good.

Consider a nearest neighbour interaction V' on X given by the following:

(i) If $v \sim w \in \mathcal{V}$ and $[c,d]_{\{v,w\}} \in \mathcal{L}_{\{v,w\}}(X_a)$ then $V'([c,d]_{v,w})$ is given by (3.3.3)

(ii) and $V'([c]_v)$ is given by (3.3.4).

(iii) If $v \in \mathcal{V}_1 \cup \mathcal{V}_2$ and $w \sim v$, then $V'([x^v]_{\{v,w\}})$ is given by (3.3.5).

(iv) If $v \in \mathcal{V}_1 \cup \mathcal{V}_2$, the value $V'([a]_v)$ is given by (3.3.7).

(v) If $v \in \mathcal{V}_1 \cup \mathcal{V}_2$, $w \sim v$ and $c \in \mathcal{A} \setminus \{x^v_w, a\}$ such that $[a,c]_{\{v,w\}} \in \mathcal{L}_{\{v,w\}}(X)$, the value $V'([a,c]_{\{v,w\}})$ is given by (3.3.9).

(vi) If $v \in \mathcal{V}_1 \cap P_1$, $w \sim v$ such that $[a,a]_{\{v,w\}} \in \mathcal{L}_{\{v,w\}}(X)$, the value $V'([a,a]_{\{v,w\}})$ is given by (3.3.11).

By the preceding paragraph the proof is complete. $\square$

Now we will reap the benefits of the previous lemma. The following lemma explains why the weak conclusions of Lemma 3.3.4 are sufficient and completes the proof of Theorem 3.3.1.

Lemma 3.3.5. *Let X be an n.n.constraint space and X_a be a strong config-fold of X where the symbol a is strongly config-folded into the symbol b. Let P_1, P_2 be the partite classes of $\mathcal{V}$ and consider $\mathcal{V}_1, \mathcal{V}_2 \subset \mathcal{V}$ and $x^v \in X$ for all $v \in \mathcal{V}_1 \cup \mathcal{V}_2$ satisfying the conclusion of Lemma 3.3.3. Let $M \in \mathbf{M}_X$ be a Markov cocycle on X such that $M|_{\Delta_{X_a}}$ is a Gibbs cocycle with some nearest neighbour interaction V and V' be an interaction on X as obtained in Lemma 3.3.4. Then $M \in \mathbf{G}_X$ is Gibbs with nearest neighbour interaction V'.*

Proof. We will use the V'-good pairs guaranteed by Lemma 3.3.4 as steps in proving the following pairs are V'-good:

(a) Let $x \in X$ and $v \in \mathcal{V}_1 \cup \mathcal{V}_2$ such that $x_v = a$ and $x_w \neq a$ for all $w \sim v$. Then $(x, \theta_b^v(x))$ is V'-good.

(b) Let $x \in X$ and $v \in \mathcal{V}_1 \cap P_1$ and $w \sim v$ such that $x_v = x_w = a$. Then $(x, \theta_b^v(x))$ is V'-good.

(c) All asymptotic pairs $(x, y) \in \Delta_X$ are V'-good.

Given an asymptotic pair, Statements (a) and (b) allow replacement of the a's by b's giving us a pair in Δ_{X_a}. From Conclusion (1) in Lemma 3.3.4 we know that all pairs in Δ_{X_a} are V'-good. Since the relation V'-good is an equivalence relation this proves Statement (c) thereby completing the proof.

Consider any $v \in \mathcal{V}_1 \cup \mathcal{V}_2$ and $x \in X$ such that $x_v = a$. Let

$$\partial\{v\} = \{w_1, w_2, \ldots w_n\}.$$

Since for all $1 \leq r \leq n$, $[a, x_{w_r}]_{\{v, w_r\}} \in \mathcal{L}_{\{v, w_r\}}(X)$, Lemma 3.3.3 implies that

$$\theta^{w_r, w_{r+1} \ldots, w_n}_{x_{w_r}, x_{w_{r+1}}, \ldots, x_{w_n}}(x^v) \in X.$$

Let $x^1 = \theta^{w_1, w_2, \ldots, w_n}_{x_{w_1}, x_{w_2}, \ldots, x_{w_n}}(x^v)$. The pair $(x, \theta_b^v(x))$ is Markov-similar to $(x^1, \theta_b^v(x^1))$ with $A := \{v\}$. Note

$$\theta^{w_1, w_2, \ldots, w_r}_{x^v_{w_1}, x^v_{w_2}, \ldots, x^v_{w_r}}(x^1) = \theta^{w_{r+1} \ldots, w_n}_{x_{w_{r+1}}, \ldots, x_{w_n}}(x^v) \in X$$

and that x^1 and x^v differ only on $\partial\{v\}$.

In the following we will remove a's in configurations from those vertices which are isolated from other a's.

Proof of Statement (a): Consider the sequence

$$x^1, \theta^{w_1}_{x^v_{w_1}}(x^1), \theta^{w_1, w_2}_{x^v_{w_1}, x^v_{w_2}}(x^1), \ldots, \theta^{w_1, w_2, \ldots, w_n}_{x^v_{w_1}, x^v_{w_2}, \ldots, x^v_{w_n}}(x^1) = x^v, \theta_b^v(x^v), \theta_b^v(x^1).$$

Here single site changes have been made on $\partial\{v\}$ taking us from x^1 to x^v. Then the symbol at v

has been changed to obtain $\theta^v_b(x^v)$. In the last step $\theta^v_b(x^v)$ has been changed on $\partial\{v\}$ to obtain $\theta^v_b(x^1)$.

Note that each

$$(\theta^{w_1,w_2,\ldots,w_r}_{x^v_{w_1},x^v_{w_2},\ldots,x^v_{w_r}}(x^1), \theta^{w_1,w_2,\ldots,w_{r+1}}_{x^v_{w_1},x^v_{w_2},\ldots,x^v_{w_{r+1}}}(x^1))$$

is Markov-similar to $(\theta^{w_{r+1}}_{x^1_{w_{r+1}}}(x^v), x^v)$ for all $0 \leq r \leq n-1$ with $A := \{w_{r+1}\}$. By Conclusion (3) in Lemma 3.3.4, $(\theta^{w_{r+1}}_{x^1_{w_{r+1}}}(x^v), x^v)$ is V'-good for all $0 \leq r \leq n-1$. Thus by Corollary 3.2.2, we get that (x^1, x^v) is V'-good. By Conclusion (2) in Lemma 3.3.4 and symmetry of the relation V'-good we get that $(x^v, \theta^v_b(x^v))$ is V'-good. Since $\theta^v_b(x^v), \theta^v_b(x^1) \in X_a$, Conclusion (1) in Lemma 3.3.4 implies that $(\theta^v_b(x^v), \theta^v_b(x^1))$ is V'-good. Stringing these together by Corollary 3.2.2 we arrive at $(x^1, \theta^v_b(x^1))$ being V'-good. But $(x^1, \theta^v_b(x^1))$ is Markov-similar to $(x, \theta^v_b(x))$. Therefore by Proposition 3.2.1 we get that $(x, \theta^v_b(x))$ is V'-good.

In the next step we remove the a's which are not isolated.

Proof of Statement (b): We construct a sequence from x to $\theta^v_b(x)$ in three parts. In the first part single site changes will be made on $\partial\{v\}$ taking us from x^1 to x^v. In the second part the symbol at v will be changed to obtain $\theta^v_b(x^v)$. In the last part single site changes will be made on $\partial\{v\}$ to obtain $\theta^v_b(x^1)$ from $\theta^v_b(x^v)$.

Consider the sequence

$$(x^1, \theta^{w_1}_{x^v_{w_1}}(x^1), \theta^{w_1,w_2}_{x^v_{w_1},x^v_{w_2}}(x^1), \ldots, \theta^{w_1,w_2,\ldots,w_n}_{x^v_{w_1},x^v_{w_2},\ldots,x^v_{w_n}}(x^1) = x^v),$$
$$(x^v, \theta^v_b(x^v)),$$
$$(\theta^v_b(x^v), \theta^{w_1}_{x^1_{w_1}}(\theta^v_b(x^v)), \theta^{w_1,w_2}_{x^1_{w_1},x^1_{w_2}}(\theta^v_b(x^v)), \ldots, \theta^{w_1,w_2,\ldots,w_n}_{x^1_{w_1},x^1_{w_2},\ldots,x^1_{w_n}}(\theta^v_b(x^v)) = \theta^v_b(x^1)).$$

In the first part of the sequence notice that

$$(\theta^{w_1,w_2,\ldots,w_r}_{x^v_{w_1},x^v_{w_2},\ldots,x^v_{w_r}}(x^1), \theta^{w_1,w_2,\ldots,w_{r+1}}_{x^v_{w_1},x^v_{w_2},\ldots,x^v_{w_{r+1}}}(x^1))$$

is Markov-similar to $(\theta^{w_{r+1}}_{x^1_{w_{r+1}}}(x^v), x^v)$ for all $0 \leq r \leq n-1$ with $A := \{w_{r+1}\}$. If for some $0 \leq r \leq n-1$, $x^1_{w_{r+1}} \neq a$ then by Conclusion (3) in Lemma 3.3.4 we get that $(\theta^{w_{r+1}}_{x^1_{w_{r+1}}}(x^v), x^v)$ is V'-good. If for some $0 \leq r \leq n-1$, $x^1_{w_{r+1}} = a$ then by Conclusion (4) in Lemma 3.3.4 we get that $(\theta^{w_{r+1}}_{x^1_{w_{r+1}}}(x^v), x^v)$ is V'-good. Proposition 3.2.1 implies that

$$(\theta^{w_1,w_2,\ldots,w_r}_{x^v_{w_1},x^v_{w_2},\ldots,x^v_{w_r}}(x^1), \theta^{w_1,w_2,\ldots,w_{r+1}}_{x^v_{w_1},x^v_{w_2},\ldots,x^v_{w_{r+1}}}(x^1))$$

is V'-good for all $0 \leq r \leq n-1$. By Corollary 3.2.2, we get that (x^1, x^v) is V'-good and we are done with the first part of the sequence.

For the second part of the sequence by Conclusion (2) in Lemma 3.3.4 and symmetry of the relation V'-good we get that $(x^v, \theta_b^v(x^v))$ is V'-good.

For the third part of the sequence the asymptotic pair

$$(\theta^{w_1,w_2,\ldots,w_r}_{x^1_{w_1},x^1_{w_2},\ldots,x^1_{w_r}}(\theta_b^v(x^v)), \theta^{w_1,w_2,\ldots,w_{r+1}}_{x^1_{w_1},x^1_{w_2},\ldots,x^1_{w_{r+1}}}(\theta_b^v(x^v)))$$

is Markov-similar to $(\theta_b^v(x^v), \theta^{w_{r+1}}_{x^1_{w_{r+1}}}(\theta_b^v(x^v)))$ for all $0 \le r \le n-1$ with $A = \{w_{r+1}\}$.

If for some $0 \le r \le n-1$, $x^1_{w_{r+1}} \neq a$ then $(\theta_b^v(x^v), \theta^{w_{r+1}}_{x^1_{w_{r+1}}}(\theta_b^v(x^v)) \in X_a$ and by Conclusion (1) in Lemma 3.3.4 we get that $(\theta_b^v(x^v), \theta^{w_{r+1}}_{x^1_{w_{r+1}}}(\theta_b^v(x^v)))$ is V'-good. Since $v \in \mathcal{V}_1$, $x^v|_{D_2(v)\setminus\{v\}} = b$. Thus if for some $0 \le r \le n-1$, $x^1_{w_{r+1}} = a$ then

$$(\theta^{w_{r+1}}_{x^1_{w_{r+1}}}(\theta_b^v(x^v)), \theta_b^v(x^v)) = (\theta_a^{w_{r+1}}(\theta_b^v(x^v)), \theta_b^{w_{r+1}}(\theta_a^{w_{r+1}}(\theta_b^v(x^v)))$$

and $(\theta_a^{w_{r+1}}(\theta_b^v(x^v)))_{w'} = b \neq a$ for all $w' \sim w_{r+1}$. By Statement (a) in the proof of this lemma we get that $(\theta_a^{w_{r+1}}(\theta_b^v(x^v)), \theta_b^{w_{r+1}}(\theta_a^{w_{r+1}}(\theta_b^v(x^v)))$ is V'-good. By symmetry of the relation V'-good we get that $(\theta_b^v(x^v), \theta^{w_{r+1}}_{x^1_{w_{r+1}}}(\theta_b^v(x^v)))$ is V'-good in this case as well.

Thus for all $0 \le r \le n-1$ we find that $(\theta_b^v(x^v), \theta^{w_{r+1}}_{x^1_{w_{r+1}}}(\theta_b^v(x^v)))$ is V'-good. Using Corollary 3.2.2 we find that $(\theta_b^v(x^v), \theta_b^v(x^1))$ is V'-good.

So we have proven that $(x^1, x^v), (x^v, \theta_b^v(x^v)), (\theta_b^v(x^v), \theta_b^v(x^1))$ are V'-good. Stringing them by Corollary 3.2.2 we get that $(x^1, \theta_b^v(x^1))$ is V'-good. But $(x^1, \theta_b^v(x^1))$ is Markov-similar to $(x, \theta_b^v(x))$. Therefore by Proposition 3.2.1 we get that $(x, \theta_b^v(x))$ is V'-good.

The previous two statements give us the freedom to change the a's into b's. Now we will use them to prove the last statement.

Proof of Statement (c): Consider an asymptotic pair $(x, y) \in \Delta_X$. Let

$$F := \{v \in \mathcal{V} \mid x_v \neq y_v\}$$

and $x^1, y^1 \in \mathcal{A}^{\mathcal{V}}$ be obtained by replacing the a's outside $F \cup \partial F$ by b's, that is

$$x^1_u := \begin{cases} x_u \text{ if } u \in F \cup \partial F \text{ or } x_u \neq a \\ b \quad \text{otherwise} \end{cases}$$

and

$$y^1_u := \begin{cases} y_u \text{ if } u \in F \cup \partial F \text{ or } y_u \neq a \\ b \quad \text{otherwise.} \end{cases}$$

By (3.2.2) $x^1, y^1 \in X$. Since $x = y$ on F^c, it follows that (x, y) and (x^1, y^1) are Markov-similar but there are only finitely many vertices where x^1 and y^1 equal a. Let

$$\begin{aligned}
\{v_1, v_2, \ldots, v_r\} &:= \{v \in P_1 \mid x^1_v = a\}\\
\{w_1, w_2, \ldots, w_{r'}\} &:= \{w \in P_1 \mid y^1_w = a\}\\
\{v_{r+1}, v_{r+2} \ldots, v_{r+k}\} &:= \{v \in P_2 \mid x^1_v = a\}\\
\{w_{r'+1}, w_{r'+2} \ldots, w_{r'+k'}\} &:= \{w \in P_2 \mid y^1_w = a\}
\end{aligned}$$

index the vertices with a in x^1 and y^1. By Lemma 3.3.3 the configurations $\theta^{v_1,v_2\ldots,v_i}_{b,b,\ldots,b}(x^1)$ and $\theta^{w_1,w_2\ldots,w_{i'}}_{b,b,\ldots,b}(y^1)$ are elements of X for all $1 \leq i \leq r+k$ and $1 \leq i' \leq r'+k'$. Therefore we can consider the sequence (3.3.12 to 3.3.16) in X:
We begin by replacing the a's in x^1 from the partite class P_1 by b's.

$$(x^1, \theta^{v_1}_b(x^1), \theta^{v_1,v_2}_{b,b}(x^1), \ldots, \theta^{v_1,v_2,\ldots,v_r}_{b,b,\ldots,b}(x^1)). \tag{3.3.12}$$

In the resulting configuration $\theta^{v_1,v_2,\ldots,v_r}_{b,b,\ldots,b}(x^1)$ adjacent vertices cannot both have the symbol a; the a's left in the configuration x^1 are changed to b's.

$$(\theta^{v_1,v_2,\ldots,v_r}_{b,b,\ldots,b}(x^1), \theta^{v_1,v_2,\ldots,v_{r+1}}_{b,b,\ldots,b}(x^1), \ldots, \theta^{v_1,v_2,\ldots,v_{r+k}}_{b,b,\ldots,b}(x^1)). \tag{3.3.13}$$

After removing the a's from x^1 and y^1 the configurations obtained are elements of X_a.

$$(\theta^{v_1,v_2,\ldots,v_{r+k}}_{b,b,\ldots,b}(x^1), \theta^{w_1,w_2\ldots,w_{r'+k'}}_{b,b,\ldots,b}(y^1)). \tag{3.3.14}$$

Tactics from Sequences 3.3.12 and 3.3.13 are employed in reverse to obtain y^1 starting with $\theta^{w_1,w_2\ldots,w_{r'+k'}}_{b,b,\ldots,b}(y^1)$.

$$(\theta^{w_1,w_2\ldots,w_{r'+k'}}_{b,b,\ldots,b}(y^1), \theta^{w_1,w_2\ldots,w_{r'+k'-1}}_{b,b,\ldots,b}(y^1), \ldots, \theta^{w_1,w_2\ldots,w_{r'}}_{b,b,\ldots,b}(y^1)), \tag{3.3.15}$$
$$(\theta^{w_1,w_2\ldots,w_{r'}}_{b,b,\ldots,b}(y^1), \theta^{w_1,w_2\ldots,w_{r'-1}}_{b,b,\ldots,b}(y^1), \ldots, \theta^{w_1}_b(y^1), y^1). \tag{3.3.16}$$

For all $1 \leq i \leq r$, the vertex $v_i \in P_1$ and the symbol $x^1_{v_i} = a$. Thus by Statements (a) and (b) in this proof we get that

$$(\theta^{v_1,v_2,\ldots,v_i}_{b,b,\ldots,b}(x^1), (\theta^{v_1,v_2,\ldots,v_{i+1}}_{b,b,\ldots,b}(x^1))$$

is V'-good. Thus all adjacent pairs in the Sequence 3.3.12 are V'-good

Notice that $(\theta^{v_1,v_2,\ldots,v_r}_{b,b,\ldots,b}(x^1))_v \neq a$ for all $v \in P_1$ and hence $(\theta^{v_1,v_2,\ldots,v_{r+i}}_{b,b,\ldots,b}(x^1))_v \neq a$ for all $1 \leq i \leq k$ and $v \in P_1$. Now consider an adjacent pair in the Theorems 3.3.13,

$$(\theta^{v_1,v_2,\ldots,v_{r+i}}_{b,b,\ldots,b}(x^1), \theta^{v_1,v_2,\ldots,v_{r+i+1}}_{b,b,\ldots,b}(x^1))$$

for some $0 \leq i \leq k-1$. Since $v_{r+i+1} \in P_2$, $(\theta^{v_1,v_2,\ldots,v_{r+i}}_{b,b,\ldots,b}(x^1))_w \neq a$ for all $w \sim v_{r+i+1}$. But

$(\theta^{v_1,v_2,\ldots,v_{r+i}}_{b,b,\ldots,b}(x^1))_{v_{r+i+1}} = a$, therefore by Statement (a) we get that

$$(\theta^{v_1,v_2,\ldots,v_{r+i}}_{b,b,\ldots,b}(x^1), \theta^{v_1,v_2,\ldots,v_{r+i+1}}_{b,b,\ldots,b}(x^1))$$

is V'-good.

Notice that $(\theta^{v_1,v_2,\ldots,v_{r+k}}_{b,b,\ldots,b}(x^1)), (\theta^{w_1,w_2,\ldots,w_{r'+k'}}_{b,b,\ldots,b}(y^1)) \in X_a$. Thus by Conclusion (1) in Lemma 3.3.4, we get that the Pair 3.3.14 is V'-good.

The proof that the adjacent pairs listed in Sequences 3.3.15 and 3.3.16 are V'-good is identical to the proof for the Sequences 3.3.13 and 3.3.12 with an additional use of the symmetry of the relation V'-good.

Thus all adjacent pairs in Sequences 3.3.12-3.3.16 are V'-good. By Corollary 3.2.2 we get that (x^1, y^1) is V'-good. But (x^1, y^1) is Markov-similar to (x, y). By Proposition 3.2.1 we have that (x, y) is V'-good. This completes the proof. $\square$

If $\mathcal{G}$ is finite then Theorem 3.3.2 follows immediately from Theorem 3.3.1: if V' is a nearest neighbour interaction for a G-invariant Markov cocycle M then

$$\frac{\sum_{g\in G} gV'}{|G|}$$

is a G-invariant nearest neighbour interaction for M. We will prove the following result which along with Proposition 3.2.5 immediately implies Theorem 3.3.2.

Theorem 3.3.6. *Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be a bipartite graph and $\mathcal{A}$ a finite alphabet. Let $G \subset Aut(\mathcal{G})$ be a subgroup. Let X be a G-invariant n.n.constraint space and X_a be a strong config-fold of X. Suppose $M \in \mathbf{M}^G_X$ is a G-invariant Markov cocycle. Then $M \in \mathbf{G}^G_X$ if and only if $M|_{\Delta_{X_a}} \in \mathbf{G}^G_{X_a}$.*

Proof. By Proposition 3.2.5, $M \in \mathbf{G}^G_X$ implies $M|_{\Delta_{X_a}} \in \mathbf{G}^G_{X_a}$. We will prove the converse. Let $M \in \mathbf{M}^G_X$ such that $M|_{\Delta_{X_a}} \in \mathbf{G}^G_{X_a}$. Let V be a G-invariant nearest neighbour interaction for $M|_{\Delta_{X_a}}$.

Mimicking the proof of Lemma 3.3.3 we will now obtain special configurations x^v in a G-invariant way.

Lemma 3.3.7. *Let $G \subset Aut(\mathcal{G})$ be a subgroup, X be a G-invariant n.n.constraint space and X_a be a strong config-fold of X where the symbol a is strongly config-folded into the symbol b. Let*

$$\mathcal{V}_1 := \{v \in \mathcal{V} \mid \text{ there exists } w \sim v \text{ such that } [a,a]_{\{v,w\}} \in \mathcal{L}_{\{v,w\}}(X)\}$$

and

$$\mathcal{V}_2 := \{v \in \mathcal{V} \setminus \mathcal{V}_1 \mid [a]_v \in \mathcal{L}_v(X)\}.$$

Then $\mathcal{V}_1$ and $\mathcal{V}_2$ are invariant under the action of G. Moreover for all $v \in \mathcal{V}_1 \cup \mathcal{V}_2$ there exists $x^v \in X$ satisfying the conclusions of Lemma 3.3.3 such that $(gx^v)|_{gD_2(v)} = x^{gv}|_{gD_2(v)}$ for all $g \in G$.

Proof. Since X is G-invariant it follows that the sets $\mathcal{V}_1$ and $\mathcal{V}_2$ are G-invariant.

Consider some $v \in \mathcal{V}_1$ and $g \in G$. Then by Lemma 3.3.3 there exists $x^v, x^{gv} \in X$ such that $x^v_v = x^{gv}_{gv} = a$ and $x^v|_{D_2(v)\setminus\{v\}} = x^{gv}|_{D_2(gv)\setminus\{gv\}} = b$. Thus we find that $(gx^v)|_{gD_2(v)} = x^{gv}|_{gD_2(v)}$.

Let $v \in \mathcal{V}_2$. Then for all $w \sim v$, $g \in G$ and $c \in \mathcal{A} \setminus \{a\}$ the pattern $[a, c]_{\{v,w\}} \in \mathcal{L}_{\{v,w\}}(X)$ if and only if $[a, c]_{\{gv,gw\}} \in \mathcal{L}_{\{gv,gw\}}(X)$. Thus for all $w \sim v$ we can choose $c_{v,w} \in \mathcal{A} \setminus \{a\}$ such that $[a, c_{v,w}]_{\{v,w\}} \in \mathcal{L}_{\{v,w\}}(X)$ and $c_{v,w} = c_{gv,gw}$ for all $g \in G$. Note that since $v \in \mathcal{V}_2$ we know that $c_{v,w} \neq a$.

By (3.2.4) there exists $x^{1,v} \in X$ such that $x^{1,v}|_{\partial D_1(v)} = b$. Since a can be strongly config-folded into the symbol b we can assume that $x^{1,v} \in X_a$. Consider $x^v \in \mathcal{A}^{\mathcal{V}}$ defined by

$$x^v_u := \begin{cases} a & \text{if } u = v \\ c_{v,u} & \text{if } u \sim v \\ x^{1,v}_u & \text{if } u \in D_1(v)^c. \end{cases}$$

Note that a appears in x^v only at the vertex v. Any edge (u_1, u_2) in $\mathcal{G}$ lies either completely in $D_2(v)$ or in $D_1(v)^c$. If the edge lies in $D_1(v)^c$ then $[x^v]_{\{u_1,u_2\}} = x^{1,v}_{\{u_1,u_2\}} \in \mathcal{L}_{\{u_1,u_2\}}(X)$. If the edge is of the form (v, w) then $[x^v]_{\{v,w\}} = [a, c_{v,w}]_{\{v,w\}} \in \mathcal{L}_{\{v,w\}}(X)$. If the edge is of the form (w, w') where $w \in \partial\{v\}$ and $w' \in \partial D_1(v)$ then $[x^v]_{\{w,w'\}} = [c_{v,w}, b]_{\{w,w'\}}$. Since (v, w) and (w, w') are edges in the graph $\mathcal{G}$ and $[a, c_{v,w}] \in \mathcal{L}_{\{v,w\}}(X)$ by (3.2.3) we know that $[c_{v,w}, b]_{\{w,w'\}} \in \mathcal{L}_{\{w,w'\}}(X)$.

Thus we have proved for every edge (u_1, u_2) in $\mathcal{G}$ that $[x^v]_{\{u_1,u_2\}} \in \mathcal{L}_{\{u_1,u_2\}}(X)$. Since X is an n.n.constraint space we get that $x^v \in X$.

Moreover for all $v \in \mathcal{V}_2$ and $g \in G$

$$(gx^v)_u = \begin{cases} a & \text{if } u = gv \\ c_{v,g^{-1}u} & \text{if } u \sim gv \\ b & \text{if } u \in \partial D_1(gv) \end{cases}$$

and

$$(x^{gv})_u = \begin{cases} a & \text{if } u = gv \\ c_{gv,u} = c_{v,g^{-1}u} & \text{if } u \sim gv \\ b & \text{if } u \in \partial D_1(gv), \end{cases}$$

that is, $(gx^v)|_{gD_2(v)} = x^{gv}|_{gD_2(v)}$.

Thus the configurations x^v satisfy Conclusions (1) and (2) of Lemma 3.3.3 and $(gx^v)|_{gD_2(v)} = x^{gv}|_{gD_2(v)}$ for all $g \in G$ and $v \in \mathcal{V}_1 \cup \mathcal{V}_2$. The rest follows exactly as in the proof of Lemma 3.3.3. □

Consider sets $\mathcal{V}_1, \mathcal{V}_2 \subset \mathcal{V}$ and for all $v \in \mathcal{V}$ configurations $x^v \in X$ as obtained by Lemma 3.3.7. Then by Lemma 3.3.4 there exists a unique nearest neighbour interaction V' on X such that the pairs listed in (1), (2), (3) and (4) listed in Lemma 3.3.4 are V'-good. By Lemma 3.3.5 we get that

V' is a nearest neighbour interaction for M. We will prove that the interaction V' is G-invariant. For this we will invoke the uniqueness of the interaction satisfying the conclusions of Lemma 3.3.4.

Let $g \in G$. gV' is a nearest neighbour interaction corresponding to $gM = M$. Thus the pairs listed in (1), (2), (3) and (4) in Lemma 3.3.4 are gV'-good. Since V is G-invariant, $gV'|_{\mathcal{L}_{X_a}} = gV = V$. Hence gV' satisfies (3.3.3), (3.3.4).

If $v \in \mathcal{V}_1 \cup \mathcal{V}_2$ then we know from Lemma 3.3.7 that $(gx^v)|_{gD_1(v)} = x^{gv}|_{gD_1(v)}$. Thus if $w \sim v$ since V' satisfies (3.3.5) we get that

$$gV'([x^v]_{\{v,w\}}) = V'([x^v_v, x^v_w]_{\{g^{-1}v, g^{-1}w\}}) = V'([x^{g^{-1}v}_{g^{-1}v}, x^{g^{-1}v}_{g^{-1}w}]_{\{g^{-1}v, g^{-1}w\}}) = 0.$$

Thus the interaction gV' satisfies (3.3.5). We have seen that the interaction gV' is a nearest neighbour interaction which satisfies (3.3.3), (3.3.4) and (3.3.5) such that the pairs listed in (1), (2), (3) and (4) in Lemma 3.3.4 are gV'-good. By Lemma 3.3.4 we know that such an interaction is unique. Thus $gV' = V'$ and $M \in \mathbf{G}^G_X$.

$\square$

This leads us to the following corollary:

Corollary 3.3.8. *Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be a bipartite graph and $\mathcal{A}$ a finite alphabet. Let $G \subset Aut(\mathcal{G})$ be a subgroup. Let X be a G-invariant n.n.constraint space and X_a be a strong config-fold of X. Then $\mathbf{M}^G_X / \mathbf{G}^G_X$ is isomorphic to $\mathbf{M}^G_{X_a} / \mathbf{G}^G_{X_a}$.*

Clearly this corollary subsumes Theorem 3.3.2 and implies Theorem 1.2.1. Thereby to understand the difference between Markov and Gibbs cocycles it is sufficient to study the cocycles over closed configuration spaces which cannot be strongly config-folded any further.

Also this corollary is most relevant when the dimension of the quotient space $\mathbf{M}^G_X / \mathbf{G}^G_X$ is finite. For example this holds in the following two situations:

1. The underlying graph $\mathcal{G}$ is finite.

2. The underlying graph $\mathcal{G}$ is $\mathbb{Z}^d$ for some dimension d, G is the group of translations on $\mathbb{Z}^d$ and the space X has the pivot property (Proposition 2.2.6).

Proof. By Proposition 3.2.5 the map $F : \mathbf{M}^G_X \longrightarrow \mathbf{M}^G_{X_a}$ given by

$$F(M) := M|_{\Delta_{X_a}} \text{ for all } M \in \mathbf{M}^G_{X_a}$$

is surjective. By Theorem 3.3.6 we know that for a Markov cocycle $M \in \mathbf{M}^G_X$, $M \in \mathbf{G}^G_X$ if and only if $M|_{\Delta_{X_a}} \in \mathbf{G}^G_{X_a}$. Thus $F^{-1}(\mathbf{G}^G_{X_a}) = \mathbf{G}^G_X$.

Via the second isomorphism theorem for vector spaces the map

$$\tilde{F} : \mathbf{M}_X^G / F^{-1}(\mathbf{G}_{X_a}^G) \longrightarrow \mathbf{M}_{X_a}^G / \mathbf{G}_{X_a}^G$$

given by

$$\tilde{F}(M \mod F^{-1}(\mathbf{G}_{X_a}^G)) := F(M) \mod \mathbf{G}_{X_a}^G$$

is an isomorphism. Since $F^{-1}(\mathbf{G}_{X_a}^G) = \mathbf{G}_X^G$ the proof is complete. $\square$

Chapter 4

Four-Cycle Free Graphs, the Pivot Property and Entropy Minimality

By $\mathcal{H}$ we will always denote an undirected graph without multiple edges and single isolated vertices. The main aim of this chapter is to prove Theorem 1.3.2 (Theorem 4.1.4) and Theorem 1.4.1 (Theorem 4.1.2). Most of this chapter is part of the submitted manuscript [10].

Hom-shifts will be introduced in Section 4.1. Some aspects of thermodynamic formalism will be stated in Section 4.2. Some technical details regarding folding will be discussed in Section 4.3. Universal covers will be defined in Section 4.4 and the generalised height functions, subcocycles will be described in Section 4.5. The proof of the main results can be found in Section 4.6.

4.1 Hom-Shifts

For a graph $\mathcal{H}$ we will denote the adjacency relation by $\sim_{\mathcal{H}}$ and the set of vertices of $\mathcal{H}$ by $\mathcal{H}$ (abusing notation). We identify $\mathbb{Z}^d$ with the set of vertices of the Cayley graph with respect to the standard generators $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_d$, that is, $\vec{i} \sim_{\mathbb{Z}^d} \vec{j}$ if and only if $\|\vec{i} - \vec{j}\|_1 = 1$ where $\|\cdot\|_1$ is the l^1 norm. We drop the subscript in $\sim_{\mathcal{H}}$ when $\mathcal{H} = \mathbb{Z}^d$. Let D_n and B_n denote the $\mathbb{Z}^d$-balls of radius n around $\vec{0}$ in the l^1 and the l^∞ norm respectively. The graph C_n will denote the n-cycle where the set of vertices is $\{0, 1, 2, \ldots, n-1\}$ and $i \sim_{C_n} j$ if and only if $i \equiv j \pm 1 \mod n$. The graph K_n will denote the complete graph with n vertices where the set of vertices is $\{1, 2, \ldots, n\}$ and $i \sim_{K_n} j$ if and only if $i \neq j$.

A *sliding block code* from a shift space X to a shift space Y is a continuous map $f : X \longrightarrow Y$ which commutes with the shifts, that is, $\sigma^{\vec{i}} \circ f = f \circ \sigma^{\vec{i}}$ for all $\vec{i} \in \mathbb{Z}^d$. A surjective sliding block code is called a *factor map* and a bijective sliding block code is called a *conjugacy*. We note that a conjugacy defines an equivalence relation; in fact, it has a continuous inverse since it is a continuous bijection between compact sets.

In this chapter, we will focus on a special class of nearest neighbour shifts of finite type where the forbidden patterns are the same in every 'direction':

Given a graph $\mathcal{H}$ let

$$X^d_{\mathcal{H}} := Hom(\mathbb{Z}^d, \mathcal{H}) = \{x \in \mathcal{H}^{\mathbb{Z}^d} \mid x_{\vec{i}} \sim_{\mathcal{H}} x_{\vec{j}} \text{ for all } \vec{i} \sim \vec{j}\}.$$

Such spaces will be called *hom-shifts*. We fix some dimension $d \geq 2$ and thereafter drop the superscript in $X^d_{\mathcal{H}}$. If $\mathcal{H}$ is finite and

$$\mathcal{F}_{\mathcal{H}} := \{[v, w]_{\vec{0}, \vec{e}_j} \mid v \nsim_{\mathcal{H}} w, 1 \leq j \leq d\}$$

then we noted in Section 3.1 that $X_{\mathcal{H}} = X_{\mathcal{F}_{\mathcal{H}}}$ and hence is nearest neighbour shift of finite type. These are exactly the nearest neighbour shifts of finite type with symmetric and isotropic constraints. For example if the graph $\mathcal{H}$ is given by Figure 4.1 then $X_{\mathcal{H}}$ is the hard square shift, that is, configurations with alphabet $\{0, 1\}$ such that adjacent symbols cannot both be 1. X_{K_n} is the space of n-colourings of the graph, that is, configurations with alphabet $\{1, 2, \ldots, n\}$ where all adjacent colours are distinct. We note that the properties, symmetry and isotropy, are not invariant under conjugacy. In this new notation the spaces X_n introduced in Chapter 2 are the hom-shifts X_{C_n} for $n \neq 1, 4$.

$\mathcal{F}$ will always denote a set of patterns and $\mathcal{H}$ will always denote a graph, there will not be any ambiguity in the notations $X_{\mathcal{F}}, X_{\mathcal{H}}$.

A finite graph $\mathcal{H}$ is called *four-cycle free* if it is finite, it has no self-loops and C_4 is not a subgraph of $\mathcal{H}$. For instance K_4 is not a four-cycle free graph.

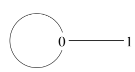

Figure 4.1: Graph for the Hard Square Shift

Hom-shifts form a special class of shifts of finite type. In general the set of globally allowed pattern is different from the set of locally allowed patterns: Let X be a shift space with a forbidden list $\mathcal{F}$. Given a finite set A, a pattern $a \in \mathcal{A}^A$ is said to be *locally allowed* if no pattern from $\mathcal{F}$ appears in a. In general it is undecidable for shifts of finite type whether a locally allowed pattern belongs to $\mathcal{L}(X)$ [46]; however it is decidable when X is a hom-shift where it is sufficient to check whether the pattern extends to a locally allowed pattern on B_n for some n.

It is well known that the topological entropy (introduced in Section 2.7) of a shift space X can be calculated via the following equation:

$$h_{top}(X) = \lim_{n \longrightarrow \infty} \frac{\log |\mathcal{L}_{B_n}(X)|}{|B_n|}.$$

The existence of the limit follows from subadditivity arguments via the well-known multivariate version of Fekete's Lemma [8]. Moreover the topological entropy is an invariant under conjugacy (for $d = 1$ look at Proposition 4.1.9 in [29]; the proof extends to higher dimensions). We remark that the computation of this invariant for shifts of finite type in $d > 1$ is a hard problem and very little is known [40], however there are algorithms to compute approximating upper and lower bounds of the topological entropy of the hom-shifts [19, 30]. When $\mathcal{H}$ is a finite connected graph with at least two edges, then $h_{top}(X) > 0$:

Proposition 4.1.1. *Let $\mathcal{H}$ be a finite graph with vertices a, b and c such that $a \sim_{\mathcal{H}} b$ and $b \sim_{\mathcal{H}} c$. Then $h_{top}(X_{\mathcal{H}}) \geq \frac{\log 2}{2}$.*

Proof. It sufficient to see this for a graph $\mathcal{H}$ with exactly three vertices a, b and c such that $a \sim_{\mathcal{H}} b$ and $b \sim_{\mathcal{H}} c$. For such a graph any configuration in $X_{\mathcal{H}}$ is composed of b on one partite class of $\mathbb{Z}^d$ and a free choice between a and c for vertices on the other partite class. Then

$$|\mathcal{L}_{B_n}(X)| = 2^{\lfloor \frac{(2n+1)^d}{2} \rfloor} + 2^{\lceil \frac{(2n+1)^d}{2} \rceil}$$

proving that $h_{top}(X_{\mathcal{H}}) = \frac{\log 2}{2}$. □

A shift space X is called *entropy minimal* if for all shift spaces $Y \subsetneq X$, $h_{top}(X) > h_{top}(Y)$. In other words, a shift space X is entropy minimal if forbidding any word causes a drop in entropy. From [45] we know that every shift space contains an entropy minimal shift space with the same entropy and also a characterisation of same entropy factor maps on entropy minimal shifts of finite type.

One of the main results of this chapter is the following:

Theorem 4.1.2. *Let $\mathcal{H}$ be a connected four-cycle free graph. Then $X_{\mathcal{H}}$ is entropy minimal.*

For $d = 1$ all irreducible shifts of finite type are entropy minimal [29]. A necessary condition for the entropy minimality of $X_{\mathcal{H}}$ is that $\mathcal{H}$ has to be connected.

Proposition 4.1.3. *Suppose $\mathcal{H}$ is a finite graph with connected components $\mathcal{H}_1, \mathcal{H}_2, \ldots \mathcal{H}_r$. Then $h_{top}(X_{\mathcal{H}}) = \max_{1 \leq i \leq r} h_{top}(X_{\mathcal{H}_i})$.*

This follows from the observation that

$$\max_{1 \leq i \leq r} |\mathcal{L}_{B_n}(X_{\mathcal{H}_i})| \leq |\mathcal{L}_{B_n}(X_{\mathcal{H}})| = \sum_{i=1}^{r} |\mathcal{L}_{B_n}(X_{\mathcal{H}_i})| \leq r \max_{1 \leq i \leq r} |\mathcal{L}_{B_n}(X_{\mathcal{H}_i})|.$$

The following theorem is another main result in this chapter. We refer the reader to Section 2.2.2 so as to review the pivot property.

Theorem 4.1.4. *For all four-cycle free graphs $\mathcal{H}$, $X_{\mathcal{H}}$ has the pivot property.*

It is sufficient to prove this theorem for four-cycle free graphs $\mathcal{H}$ which are connected because of the following proposition:

Proposition 4.1.5. *Let $X_1, X_2, \dots, X_n$ be shift spaces on disjoint alphabets such that each of them has the pivot property. Then $\cup_{i=1}^n X_i$ also has the pivot property.*

This is true since $(x, y) \in \Delta_{\cup_{i=1}^n X_i}$ implies $(x, y) \in \Delta_{X_i}$ for some $1 \leq i \leq n$.

4.2 Thermodynamic Formalism

In this chapter we will introduce several special cases of theorems introduced in Section 2.7 which will be useful in this chapter. We refer to it in case the reader would like to review the definitions of measure theoretic entropy and equilibrium states. Adapted measures can be reviewed from Subsection 2.1.2.

Given a shift-invariant probability measure ν let h_ν denote the measure theoretic entropy of ν. A shift-invariant probability measure μ is a *measure of maximal entropy* of X if the maximum of $\nu \mapsto h_\nu$ over all shift-invariant probability measures on X is obtained at μ. In other words, measures of maximal entropy are equilibrium states for the function $f \equiv 0$. The well-known *variational principle* for topological entropy of $\mathbb{Z}^d$-actions asserts that if μ is a measure of maximal entropy $h_{top}(X) = h_\mu$ whenever X is a shift space.

The following is a well-known characterisation of entropy minimality (it is used for instance in the proof of Theorem 4.1 in [32]):

Proposition 4.2.1. *A shift space X is entropy minimal if and only if every measure of maximal entropy for X is fully supported.*

We understand this by the following: Suppose X is entropy minimal and μ is a measure of maximal entropy for X. Then by the variational principle for X and $supp(\mu)$ we get

$$h_{top}(X) = h_\mu \leq h_{top}(supp(\mu)) \leq h_{top}(X)$$

proving that $supp(\mu) = X$. To prove the converse, suppose for contradiction that X is not entropy minimal and consider $Y \subsetneq X$ such that $h_{top}(X) = h_{top}(Y)$. Then by the variational principle there exists a measure μ on Y such that $h_\mu = h_{top}(X)$. Thus μ is a measure of maximal entropy for X which is not fully supported.

Further is known if X is a nearest neighbour shift of finite type: Given a set $A \subset \mathbb{Z}^d$ we denote the *r-boundary* of A by $\partial_r A$, that is,

$$\partial_r A = \{\vec{w} \in \mathbb{Z}^d \setminus A \;\Big|\; \|\vec{w} - \vec{v}\|_1 \leq r \text{ for some } \vec{v} \in A\}.$$

Note that $\partial_1 A = \partial A$. A *uniform Markov random field* is a Markov random field μ such that

$$\mu([a]_A \;\Big|\; [b]_{\partial A}) = \frac{1}{n_{A, b|_{\partial A}}}$$

where $n_{A,b|_{\partial A}} = |\{a \in \mathcal{A}^A \mid \mu([a]_A \cap [b]_{\partial A}) > 0\}|$.

The following is a special case of Theorem 2.7.1:

Theorem 4.2.2. *All measures of maximal entropy on a nearest neighbour shift of finite type X are shift-invariant uniform Markov random fields μ adapted to X.*

The converse is also true under further mixing assumptions on the shift space X (called the D-condition).

We will often restrict our proofs to the ergodic case. We can do so via the following standard facts implied by Theorem 2.5.8 and Theorem 4.3.7 in [26]:

Theorem 4.2.3. *Let μ be a shift-invariant uniform Markov random field adapted to a shift space X. Let its ergodic decomposition be given by a measurable map $x \longrightarrow \mu_x$ on X, that is, $\mu = \int_X \mu_x d\mu$. Then μ-almost everywhere the measures μ_x are shift-invariant uniform Markov random fields adapted to X such that $supp(\mu_x) \subset supp(\mu)$. Moreover $\int h_{\mu_x} d\mu(x) = h_\mu$.*

We will prove the following:

Theorem 4.2.4. *Let $\mathcal{H}$ be a connected four-cycle free graph. Then every ergodic probability measure adapted to $X_\mathcal{H}$ with positive entropy is fully supported.*

This implies Theorem 4.1.2 by the following: The Lanford-Ruelle theorem implies that every measure of maximal entropy on $X_\mathcal{H}$ is a uniform shift-invariant Markov random field adapted to $X_\mathcal{H}$. By Proposition 4.1.1 and the variational principle we know that these measures have positive entropy. By Theorems 4.2.3 and 4.2.4 they are fully supported. Finally by Proposition 4.2.1, $X_\mathcal{H}$ is entropy minimal.

Alternatively, the conclusion of Theorem 4.2.4 can be obtained via some strong mixing conditions on the shift space; we will describe one such assumption. A shift space X is called *strongly irreducible* if there exists $g > 0$ such that for all $x, y \in X$ and $A, B \subset \mathbb{Z}^d$ satisfying $\min_{\vec{i} \in A, \vec{j} \in B} \|\vec{i} - \vec{j}\|_1 \geq g$, there exists $z \in X$ such that $z|_A = x|_A$ and $z|_B = y|_B$. For such a space, the homoclinic relation is minimal implying the conclusion of Theorem 4.2.4 and further, that every probability measure adapted to X is fully supported. Note that this does not prove that X is entropy minimal unless we assume that X is a nearest neighbour shift of finite type. Such an argument is used in the proof of Lemma 4.1 in [32] which implies that every strongly irreducible shift of finite type is entropy minimal. A more combinatorial approach was used in [51] to show that shift spaces (and not just shifts of finite type) with a more general mixing property called uniform filling are entropy minimal.

4.3 Folding, Entropy Minimality and the Pivot Property

Recall, as in Subsection 3.2.3 given a graph $\mathcal{H}$ we say that v *folds* into w if and only if $u \sim_\mathcal{H} v$ implies $u \sim_\mathcal{H} w$. In this case the graph $\mathcal{H} \setminus \{v\}$ is called a *fold* of $\mathcal{H}$. This map gives rise to a 'retract', that is a graph homomorphism from $\mathcal{H}$ to $\mathcal{H} \setminus \{v\}$ which is the identity on $\mathcal{H} \setminus \{v\}$ and sends v to w. This was introduced in [37] to help characterise cop-win graphs and used in [6] to establish many properties which are preserved under 'folding' and 'unfolding'. Given a finite tree $\mathcal{H}$ with more than two vertices note that a leaf vertex (vertex of degree 1) can always be folded to some other vertex of the tree. Thus starting with $\mathcal{H}$, there exists a sequence of folds resulting in a single edge. In fact using a similar argument we can prove the following proposition.

Proposition 4.3.1. *Let $\mathcal{H} \subset \mathcal{H}'$ be trees. Then there is a graph homomorphism $f : \mathcal{H}' \longrightarrow \mathcal{H}$ such that $f|_{\mathcal{H}}$ is the identity map.*

To show this, first note that if $\mathcal{H} \subsetneq \mathcal{H}'$ then there is a leaf vertex in $\mathcal{H}'$ which is not in $\mathcal{H}$. This leaf vertex can be folded into some other vertex in $\mathcal{H}'$. Thus by induction on $|\mathcal{H}' \setminus \mathcal{H}|$ we can prove that there is a sequence of folds from $\mathcal{H}'$ to $\mathcal{H}$. Corresponding to this sequence of folds we obtain a graph homomorphism from $\mathcal{H}'$ to $\mathcal{H}$ which is the identity on $\mathcal{H}$.

Here we consider a related notion for shift spaces. Given a nearest neighbour shift of finite type $X \subset \mathcal{A}^{\mathbb{Z}^d}$, *the neighbourhood* of a symbol $v \in \mathcal{A}$ is given by

$$N_X(v) := \{a \in \mathcal{A}^{\partial\vec{0}} \mid [v]_{\vec{0}} \cap [a]_{\partial\vec{0}} \in \mathcal{L}_{D_1}(X)\},$$

that is the collection of all patterns which can 'surround' v in X. We will say that v *config-folds* into w in X if $N_X(v) \subset N_X(w)$. In such a case we say that X *config-folds* to $X \cap (\mathcal{A} \setminus \{v\})^{\mathbb{Z}^d}$. Note that $X \cap (\mathcal{A} \setminus \{v\})^{\mathbb{Z}^d}$ is obtained by forbidding v from X and hence it is also a nearest neighbour shift of finite type. Also if $X = X_{\mathcal{H}}$ for some graph $\mathcal{H}$ then v config-folds into w in $X_{\mathcal{H}}$ if and only if v folds into w in $\mathcal{H}$. Thus if $\mathcal{H}$ is a tree then there is a sequence of folds starting at $X_{\mathcal{H}}$ resulting in the two checkerboard configurations with two symbols (the vertices of the edge which $\mathcal{H}$ folds into). This property is weaker than the notion of folding introduced in Subsection 3.2.3.

The main thrust of this property in our context is: if v config-folds into w in X then given any $x \in X$, every appearance of v in x can be replaced by w to obtain another configuration in X. This replacement defines a factor (surjective, continuous and shift-invariant) map $f : X \longrightarrow X \cap (\mathcal{A} \setminus \{v\})^{\mathbb{Z}^d}$ given by

$$(f(x))_{\vec{i}} := \begin{cases} x_{\vec{i}} & \text{if } x_{\vec{i}} \neq v \\ w & \text{if } x_{\vec{i}} = v. \end{cases}$$

Note that the map f defines a 'retract' from X to $X \cap (\mathcal{A} \setminus \{v\})^{\mathbb{Z}^d}$. Frequently we will config-fold more than one symbol at once (especially in Section 4.6):

Distinct symbols $v_1, v_2, \ldots, v_n$ *config-fold disjointly* into $w_1, w_2, \ldots, w_n$ in X if v_i config-folds into w_i and $v_i \neq w_j$ for all $1 \leq i, j \leq n$. In this case the symbols $v_1, v_2, \ldots, v_n$ can be replaced by $w_1, w_2, \ldots, w_n$ simultaneously for all $x \in X$. Suppose $v_1, v_2, \ldots v_n$ is a maximal set of symbols which can be config-folded disjointly in X. Then $X \cap (\mathcal{A} \setminus \{v_1, v_2, \ldots, v_n\})^{\mathbb{Z}^d}$ is called a *full config-fold* of

X. Let v_i config-fold into w_i for all $1 \leq i \leq n$. Consider $f_X : \mathcal{A} \longrightarrow \mathcal{A} \setminus \{v_1, v_2, \ldots, v_n\}$ given by

$$f_X(v) := \begin{cases} v & \text{if } v \neq v_j \text{ for all } 1 \leq j \leq n \\ w_j & \text{if } v = v_j \text{ for some } 1 \leq j \leq n. \end{cases}$$

This defines a factor map $f_X : X \longrightarrow X \cap (\mathcal{A} \setminus \{v_1, v_2, \ldots, v_n\})^{\mathbb{Z}^d}$ given by $(f_X(x))_{\vec{i}} := f_X(x_{\vec{i}})$ for all $\vec{i} \in \mathbb{Z}^d$. f_X denotes both the factor map and the map on the alphabet, it should be clear from the context which function is being used.

For example consider a tree $\mathcal{H} := (\mathcal{V}, \mathcal{E})$ where $\mathcal{V} := \{v_1, v_2, v_3, \ldots, v_{n+1}\}$ and $\mathcal{E} := \{(v_i, v_{n+1}) | 1 \leq i \leq n\}$. Then $\{v_1, v_2, \ldots, v_{n-1}\}$ is a maximal set of symbols which config-folds disjointly into v_n in $X_{\mathcal{H}}$ resulting in the checkerboard patterns with the symbols v_n and v_{n+1}. Though the full config-fold is not necessarily unique, we choose a full config-fold for every shift space and use it to construct the corresponding function f_X.

In many cases we will fix a configuration on a set $A \subset \mathbb{Z}^d$ and apply a config-fold on the rest. Hence we define the map $f_{X,A} : X \longrightarrow X$ given by

$$(f_{X,A}(x))_{\vec{i}} := \begin{cases} x_{\vec{i}} & \text{if } \vec{i} \in A \\ f_X(x_{\vec{i}}) & \text{otherwise.} \end{cases}$$

The map $f_{X,A}$ can be extended beyond X:

Proposition 4.3.2. *Let $X \subset Y$ be nearest neighbour shifts of finite type, Z be a full config-fold of X and $y \in Y$ such that for some $A \subset \mathbb{Z}^d$, $y|_{A^c \cup \partial(A^c)} \in \mathcal{L}_{A^c \cup \partial(A^c)}(X)$. Then the configuration z given by*

$$z_{\vec{i}} := \begin{cases} y_{\vec{i}} & \textit{if } \vec{i} \in A \\ f_X(y_{\vec{i}}) & \textit{otherwise} \end{cases}$$

is an element of Y. Moreover $z|_{A^c} \in \mathcal{L}_{A^c}(Z)$.

Abusing the notation, in such cases we shall denote the configuration z by $f_{X,A}(y)$.

If A^c is finite, then $f_{X,A}$ changes only finitely many coordinates. These changes can be applied one by one, that is, there is a chain of pivots in Y from y to $f_{X,A}(y)$.

A nearest neighbour shift of finite type which cannot be config-folded is called a *stiff shift*. As in the case for graphs where all the stiff graphs obtained by a sequence of folds of a given graph are isomorphic [6], all the stiff shifts obtained by a sequence of config-folds of a given nearest neighbour shift of finite type are topologically conjugate via one-block maps; the proof is similar and we omit it. Starting with a nearest neighbour shift of finite type X the *radius* of X is the smallest number of full config-folds required to obtain a stiff shift. If $\mathcal{H}$ is a tree then the radius of $X_{\mathcal{H}}$ is equal to

$$\left\lfloor \frac{diameter(\mathcal{H})}{2} \right\rfloor.$$

Thus for every nearest neighbour shift of finite type X there is a sequence of full config-folds (not necessarily unique) which starts at X and ends at a stiff shift of finite type. Let the radius of X be r and $X = X_0, X_1, X_2, \ldots, X_r$ be a sequence of full config-folds where X_r is stiff. This generates a sequence of maps $f_{X_i} : X_i \longrightarrow X_{i+1}$ for all $0 \leq i \leq r-1$. In many cases we will fix a pattern on D_n or D_n^c and apply these maps on the rest of the configuration. Consider the maps $I_{X,n} : X \longrightarrow X$ and $O_{X,n} : X \longrightarrow X$ (for $n > r$) given by

$$I_{X,n}(x) := f_{X_{r-1}, D_{n+r-1}} \left(f_{X_{r-2}, D_{n+r-2}} \left(\ldots \left(f_{X_0, D_n}(x)\right) \ldots\right)\right) \text{ (Inward Fixing Map)}$$

and

$$O_{X,n}(x) := f_{X_{r-1}, D^c_{n-r+1}} \left(f_{X_{r-2}, D^c_{n-r+2}} \left(\ldots \left(f_{X_0, D^c_n}(x)\right) \ldots\right)\right) \text{ (Outward Fixing Map).}$$

Similarly we consider maps which do not fix anything, $F_X : X \longrightarrow X_r$ given by

$$F_X(x) := f_{X_{r-1}} \left(f_{X_{r-2}} \left(\ldots \left(f_{X_0}(x)\right) \ldots\right)\right).$$

Note that $D_k \cup \partial D_k = D_{k+1}$ and $D_k^c \cup \partial (D_k^c) = D_{k-1}^c$. This along with repeated application of Proposition 4.3.2 implies that the image of $I_{X,n}$ and $O_{X,n}$ lie in X. This also implies the following proposition:

Proposition 4.3.3 (The Onion Peeling Proposition)**.** *Let $X \subset Y$ be nearest neighbour shifts of finite type with radius r, Z be a stiff shift obtained by a sequence of config-folds starting with X and $y^1, y^2 \in Y$ such that $y^1|_{D^c_{n-1}} \in \mathcal{L}_{D^c_{n-1}}(X)$ and $y^2|_{D_{n+1}} \in \mathcal{L}_{D_{n+1}}(X)$. Let $z^1, z^2 \in Y$ be given by*

$$\begin{aligned} z^1 &:= f_{X_{r-1}, D_{n+r-1}} \left(f_{X_{r-2}, D_{n+r-2}} \left(\ldots \left(f_{X_0, D_n}(y^1)\right) \ldots\right)\right) \\ z^2 &:= f_{X_{r-1}, D^c_{n-r+1}} \left(f_{X_{r-2}, D^c_{n-r+2}} \left(\ldots \left(f_{X_0, D^c_n}(y^2)\right) \ldots\right)\right) \text{ for } n > r. \end{aligned}$$

The patterns $z^1|_{D^c_{n+r-1}} \in \mathcal{L}_{D^c_{n+r-1}}(Z)$ and $z^2|_{D_{n-r+1}} \in \mathcal{L}_{D_{n-r+1}}(Z)$. If $y^1, y^2 \in X$ then in addition

$$\begin{aligned} z^1|_{D^c_{n+r-1}} &= F_X(y^1)|_{D^c_{n+r-1}} \text{ and} \\ z^2|_{D_{n-r+1}} &= F_X(y^2)|_{D_{n-r+1}}. \end{aligned}$$

Abusing the notation, in such cases we shall denote the configurations z^1 and z^2 by $I_{X,n}(y^1)$ and $O_{X,n}(y^2)$ respectively. Note that $I_{X,n}(y^1)|_{D_n} = y^1|_{D_n}$ and $O_{X,n}(y^2)|_{D^c_n} = y^2|_{D^c_n}$. Also, $O_{X,n}$ is a composition of maps of the form $f_{X,A}$ where A^c is finite; there is a chain of pivots in Y from y to

$O_{X,n}(y)$.

There are two kind of stiff shifts which will be of interest to us: A configuration $x \in \mathcal{A}^{\mathbb{Z}^d}$ is called *periodic* if there exists $n \in \mathbb{N}$ such that $\sigma^{n\vec{e_i}}(x) = x$ for all $1 \leq i \leq d$. A configuration $x \in X$ is called *frozen* if its homoclinic class is singleton. This notion coincides with the notion of frozen coloring in [6]. A subshift X will be called *frozen* if it consists of frozen configurations, equivalently Δ_X is the diagonal. A measure on X will be called *frozen* if its support is frozen. Note that any shift space consisting just of periodic configurations is frozen. All frozen nearest neighbour shifts of finite type are stiff.

Proposition 4.3.4. *Let X be a nearest neighbour shift of finite type such that a sequence of config-folds starting from X results in the orbit of a periodic configuration. Then every shift-invariant probability measure adapted to X is fully supported.*

Proposition 4.3.5. *Let X be a nearest neighbour shift of finite type such that a sequence of config-folds starting from X results in a frozen shift. Then X has the pivot property.*

Examples:

1. $X := \{0\}^{\mathbb{Z}^d} \cup \{1\}^{\mathbb{Z}^d}$ is a frozen shift space but not the orbit of a periodic configuration. Clearly the delta measure $\delta_{\{0\}^{\mathbb{Z}^d}}$ is a shift-invariant probability measure adapted to X but not fully supported. A more non-trivial example of the nearest neighbour shift of finite type which is frozen but not the orbit of a periodic configuration is the set of the Robinson tilings [46]. It is well known that it is uniquely ergodic and the unique measure is an (adapted) uniform Markov random field which is not fully supported.

2. A shift space $X \subset \mathcal{A}^{\mathbb{Z}^d}$ is said to have a safe symbol $\star$ if for all $x \in X$ and $A \subset \mathbb{Z}^d$ the configuration $z \in \mathcal{A}^{\mathbb{Z}^d}$ given by

$$z_{\vec{i}} := \begin{cases} x_{\vec{i}} & \text{if } \vec{i} \in A \\ \star & \text{if } \vec{i} \in A^c \end{cases}$$

   is also an element of X. Then any symbol in X can be config-folded into the safe symbol. By config-folding the symbols one by one we obtain a fixed point $\{\star\}^{\mathbb{Z}^d}$. Thus any nearest neighbour shift of finite type with a safe symbol satisfies the hypothesis of both the propositions.

3. Suppose $\mathcal{H}$ is a graph which folds into a single edge (denoted by *Edge*) or a single vertex v with a loop. Then the shift space $X_{\mathcal{H}}$ can be config-folded to X_{Edge} (which consists of two periodic configurations) or the fixed point $\{v\}^{\mathbb{Z}^d}$ respectively. In the latter case, the graph $\mathcal{H}$ is called *dismantlable* [37]. Note that finite trees and the graph C_4 fold into an edge. Thus in this class of examples $\mathcal{H}$ may have C_4 as a subgraph or self-loops. For dismantlable graphs $\mathcal{H}$ Theorem 4.1 in [6] implies the conclusion of Propositions 4.3.4 and 4.3.5 for $X_{\mathcal{H}}$.

Proof of Proposition 4.3.4. Let μ be a shift-invariant probability measure adapted to X. To prove that $supp(\mu) = X$ it is sufficient to prove that for all $n \in \mathbb{N}$ and $x \in X$ that $\mu([x]_{D_n}) > 0$. Let $X_0 = X$, X_1, $X_2, \ldots, X_r$ be a sequence of full config-folds where $X_r := \{\sigma^{\vec{i}_1}(p), \sigma^{\vec{i}_2}(p), \ldots, \sigma^{\vec{i}_{k-1}}(p)\}$ is the orbit of a periodic point. For any two configurations $z, w \in X$ there exists $\vec{i} \in \mathbb{Z}^d$ such that $F_X(z) = F_X(\sigma^{\vec{i}}(w))$. Since μ is shift-invariant we can choose $y \in supp(\mu)$ such that $F_X(x) = F_X(y)$. Consider the configurations $I_{X,n}(x)$ and $O_{X,n+2r-1}(y)$. By Proposition 4.3.3 they satisfy the equations

$$\begin{aligned} I_{X,n}(x)|_{D^c_{n+r-1}} &= F_X(x)|_{D^c_{n+r-1}} \text{ and} \\ O_{X,n+2r-1}(y)|_{D_{n+r}} &= F_X(y)|_{D_{n+r}}. \end{aligned}$$

Then $I_{X,n}(x)|_{\partial D_{n+r-1}} = O_{X,n+2r-1}(y)|_{\partial D_{n+r-1}}$. Since X is a nearest neighbour shift of finite type, the configuration z given by

$$\begin{aligned} z|_{D_{n+r}} &:= I_{X,n}(x)|_{D_{n+r}} \\ z|_{D^c_{n+r-1}} &:= O_{X,n+2r-1}(y)|_{D^c_{n+r-1}} \end{aligned}$$

is an element of X. Moreover

$$\begin{aligned} z|_{D_n} &= I_{X,n}(x)|_{D_n} = x|_{D_n} \\ z|_{D^c_{n+2r-1}} &= O_{X,n+2r-1}(y)|_{D^c_{n+2r-1}} = y|_{D^c_{n+2r-1}}. \end{aligned}$$

Thus $(y, z) \in \Delta_X$. Since μ is adapted we get that $z \in supp(\mu)$. Finally

$$\mu([x]_{D_n}) = \mu([z]_{D_n}) > 0.$$

$\square$

Note that all the maps being discussed here, f_X, $f_{X,A}$, F_X, $I_{X,n}$ and $O_{X,n}$ are (not necessarily shift-invariant) single block maps, that is, maps f where $(f(x))_{\vec{i}}$ depends only on $x_{\vec{i}}$. Thus if f is one such map and $x|_A = y|_A$ for some set $A \subset \mathbb{Z}^d$ then $f(x)|_A = f(y)|_A$; they map homoclinic pairs to homoclinic pairs.

Proof of Proposition 4.3.5. Let $X_0 = X$, X_1, $X_2, \ldots, X_r$ be a sequence of full config-folds where X_r is frozen. Let $(x, y) \in \Delta_X$. Since X_r is frozen, $F_X(x) = F_X(y)$. Suppose $x|_{D^c_n} = y|_{D^c_n}$ for some $n \in \mathbb{N}$. Then $O_{X,n+r-1}(x)|_{D^c_n} = O_{X,n+r-1}(y)|_{D^c_n}$. Also by Proposition 4.3.3,

$$O_{X,n+r-1}(x)|_{D_n} = F_X(x)|_{D_n} = F_X(y)|_{D_n} = O_{X,n+r-1}(y)|_{D_n}.$$

This proves that $O_{X,n+r-1}(x) = O_{X,n+r-1}(y)$. In fact it completes the proof since for all $z \in X$ there exists a chain of pivots in X from z to $O_{X,n+r-1}(z)$. $\square$

4.4 Universal Covers

Most cases will not be as simple as in the proof of Propositions 4.3.4 and 4.3.5. We wish to prove the conclusions of these propositions for hom-shifts $X_{\mathcal{H}}$ when $\mathcal{H}$ is a connected four-cycle free graph. Many ideas carry over from the proofs of these results because of the relationship of such graphs with their universal covers; we describe this relationship next. The results in this section are not original; look for instance in [56]. We mention them for completeness.

Let $\mathcal{H}$ be a finite connected graph with no self-loops. We denote by $d_{\mathcal{H}}$ the ordinary graph distance on $\mathcal{H}$ and by $D_{\mathcal{H}}(u)$, the *ball of radius 1* around u. A graph homomorphism $\pi : \mathcal{C} \longrightarrow \mathcal{H}$ is called a *covering map* if for some $n \in \mathbb{N} \cup \{\infty\}$ and all $u \in \mathcal{H}$, there exist disjoint sets $\{C_i\}_{i=1}^n \subset \mathcal{C}$ such that $\pi^{-1}(D_{\mathcal{H}}(u)) = \cup_{i=1}^n C_i$ and $\pi|_{C_i} : C_i \longrightarrow D_{\mathcal{H}}(u)$ is an isomorphism of the induced subgraphs for $1 \leq i \leq n$. A *covering space* of a graph $\mathcal{H}$ is a graph $\mathcal{C}$ such that there exists a covering map $\pi : \mathcal{C} \longrightarrow \mathcal{H}$.

A *universal covering space* of $\mathcal{H}$ is a covering space of $\mathcal{H}$ which is a tree. Unique up to graph isomorphism [56], these covers can be described in multiple ways. Their standard construction uses non-backtracking walks [1]: A *walk* on $\mathcal{H}$ is a sequence of vertices $(v_1, v_2, \ldots, v_n)$ such that $v_i \sim_{\mathcal{H}} v_{i+1}$ for all $1 \leq i \leq n-1$. The *length* of a walk $p = (v_1, v_2, \ldots, v_n)$ is $|p| = n-1$, the number of edges traversed on that walk. It is called *non-backtracking* if $v_{i-1} \neq v_{i+1}$ for all $2 \leq i \leq n-1$, that is, successive steps do not traverse the same edge. Choose a vertex $u \in \mathcal{H}$. The vertex set of the universal cover is the set of all non-backtracking walks on $\mathcal{H}$ starting from u; there is an edge between two such walks if one extends the other by a single step. The choice of the starting vertex u is arbitrary; choosing a different vertex gives rise to an isomorphic graph. We denote the universal cover by $E_{\mathcal{H}}$. The covering map $\pi : E_{\mathcal{H}} \longrightarrow \mathcal{H}$ maps a walk to its terminal vertex. Usually, we will denote by $\tilde{u}, \tilde{v}$ and $\tilde{w}$ the vertices of $E_{\mathcal{H}}$ such that $\pi(\tilde{u}) = u$, $\pi(\tilde{v}) = v$ and $\pi(\tilde{w}) = w$.

This construction shows that the universal cover of a graph is finite if and only if it is a finite tree. To see this if the graph has a cycle then the finite segments of the walk looping around the cycle give us infinitely many vertices for the universal cover. If the graph is a finite tree, then all walks must terminate at the leaves and their length is bounded by the diameter of the tree. In fact, the universal cover of a tree is itself while the universal cover of a cycle (for instance C_4) is $\mathbb{Z}$ obtained by finite segments of the walks $(1, 2, 3, 4, 1, 2, 3, 4, \ldots)$ and $(1, 4, 3, 2, 1, 4, 3, 2, \ldots)$.

Following the ideas of homotopies in algebraic topology, there is a natural operation on the set of walks: two walks can be joined together if one begins where the other one ends. More formally, given two walks $p = (v_1, v_2, \ldots, v_n)$ and $q = (w_1, w_2, \ldots, w_m)$ where $v_n = w_1$, consider $p \star q = (v_1, v_2, \ldots, v_n, w_2, w_3, \ldots, w_m)$. However even when p and q are non-backtracking $p \star q$ need not be non-backtracking. So we consider the walk $[p \star q]$ instead which erases the backtracking segments of $p \star q$, that is, if for some $i \in \mathbb{N}$, $v_{n-i+1} \neq w_i$ and $v_{n-j+1} = w_j$ for all $1 \leq j \leq i-1$ then

$$[p \star q] := (v_1, v_2, \ldots, v_{n-i+1}, w_{i-1}, w_i, \ldots, w_m).$$

This operation of erasing the backtracking segments is called *reduction*, look for instance in [56]. The following proposition is well-known (Section 4 of [56]) and shall be useful in our context as well.

Proposition 4.4.1. *Let $\mathcal{H}$ be a finite connected graph without any self-loops. Then for all $\tilde{v}, \tilde{w} \in E_{\mathcal{H}}$ satisfying $\pi(\tilde{v}) = \pi(\tilde{w})$ there exists a graph isomorphism $\phi : E_{\mathcal{H}} \longrightarrow E_{\mathcal{H}}$ such that $\phi(\tilde{v}) = \tilde{w}$ and $\pi \circ \phi = \pi$.*

To see how to construct this isomorphism, consider as an example (u), the empty walk on $\mathcal{H}$ and $(v_1, v_2, \ldots, v_n)$, some non-backtracking walk such that $v_1 = v_n = u$. Then the map $\phi : E_{\mathcal{H}} \longrightarrow E_{\mathcal{H}}$ given by

$$\phi(\tilde{w}) := [(v_1, v_2, \ldots, v_n) \star \tilde{w}].$$

is a graph isomorphism which maps (u) to $(v_1, v_2, \ldots, v_n)$; its inverse is $\psi : E_{\mathcal{H}} \longrightarrow E_{\mathcal{H}}$ given by

$$\psi(\tilde{w}) := [(v_n, v_{n-1}, \ldots, v_1) \star \tilde{w}].$$

The maps ϕ, π described above give rise to natural maps, also denoted by ϕ and π where

$$\phi : X_{E_{\mathcal{H}}} \longrightarrow X_{E_{\mathcal{H}}}$$

is given by $\phi(\tilde{x})_{\vec{i}} := \phi(\tilde{x}_{\vec{i}})$ and

$$\pi : X_{E_{\mathcal{H}}} \longrightarrow X_{\mathcal{H}}$$

is given by $\pi(\tilde{x})_{\vec{i}} := \pi(\tilde{x}_{\vec{i}})$ for all $\vec{i} \in \mathbb{Z}^d$ respectively. A *lift* of a configuration $x \in X_{\mathcal{H}}$ is a configuration $\tilde{x} \in X_{E_{\mathcal{H}}}$ such that $\pi \circ \tilde{x} = x$.

Now we shall analyse some consequences of this formalism in our context. More general statements (where $\mathbb{Z}^d$ is replaced by a different graph) are true (under a different hypothesis on $\mathcal{H}$), but we restrict to the four-cycle free condition. We noticed in Section 4.3 that if $\mathcal{H}$ is a tree then $X_{\mathcal{H}}$ satisfies the conclusions of Theorems 4.2.4 and 4.1.4. Now we will draw a connection between the four-cycle free condition on $\mathcal{H}$ and the formalism in Section 4.3.

Proposition 4.4.2 (Existence of Lifts)**.** *Let $\mathcal{H}$ be a connected four-cycle free graph. For all $x \in X_{\mathcal{H}}$ there exists $\tilde{x} \in X_{E_{\mathcal{H}}}$ such that $\pi(\tilde{x}) = x$. Moreover the lift $\tilde{x}$ is unique up to a choice of $\tilde{x}_{\vec{0}}$.*

Proof. We will begin by constructing a sequence of graph homomorphisms $\tilde{x}^n : D_n \longrightarrow E_{\mathcal{H}}$ such that $\pi \circ \tilde{x}^n = x|_{D_n}$ and $\tilde{x}^m|_{D_n} = \tilde{x}^n$ for all $m > n$. Then by taking the limit of these graph homomorphisms we obtain a graph homomorphism $\tilde{x} \in X_{E_{\mathcal{H}}}$ such that $\pi \circ \tilde{x} = x$. It will follow that given $\tilde{x}^0$ the sequence $\tilde{x}^n$ is completely determined proving that that the lifting is unique up to a choice of $\tilde{x}_{\vec{0}}$.

The recursion is the following: Let $\tilde{x}^n : D_n \longrightarrow E_{\mathcal{H}}$ be a given graph homomorphism for some $n \in \mathbb{N} \cup \{0\}$ such that $\pi \circ \tilde{x}^n = x|_{D_n}$. For any $\vec{i} \in D_{n+1} \setminus D_n$, choose a vertex $\vec{j} \in D_n$ such that $\vec{j} \sim \vec{i}$. Then $\pi(\tilde{x}^n_{\vec{j}}) = x_{\vec{j}} \sim x_{\vec{i}}$. Since π defines a local isomorphism between $E_{\mathcal{H}}$ and $\mathcal{H}$, there exists a unique vertex $\tilde{v}_{\vec{i}} \sim \tilde{x}^n_{\vec{j}} \in E_{\mathcal{H}}$ such that $\pi(\tilde{v}_{\vec{i}}) = x_{\vec{i}}$. Define $\tilde{x}^{n+1} : D_{n+1} \longrightarrow E_{\mathcal{H}}$ by

$$\tilde{x}^{n+1}_{\vec{i}} := \begin{cases} \tilde{x}^n_{\vec{i}} & \text{if } \vec{i} \in D_n \\ \tilde{v}_{\vec{i}} & \text{if } \vec{i} \in D_{n+1} \setminus D_n. \end{cases}$$

Then clearly $\pi \circ \tilde{x}^{n+1} = x|_{D_{n+1}}$ and $\tilde{x}^{n+1}|_{D_n} = \tilde{x}^n$. Note that the extension $\tilde{x}^{n+1}$ is uniquely defined given $\tilde{x}^n$.

We need to prove that this defines a valid graph homomorphism from D_{n+1} to $E_{\mathcal{H}}$. Let $\vec{i} \in D_{n+1} \setminus D_n$ and $\vec{j} \in D_n$ be chosen as described above. Consider if possible any $\vec{j}' \neq \vec{j} \in D_n$ such that $\vec{j}' \sim \vec{i}$. To prove that $\tilde{x}^{n+1}$ is a graph homomorphism we need to verify that $\tilde{x}^{n+1}_{\vec{j}'} \sim \tilde{x}^{n+1}_{\vec{i}}$.

Consider $\vec{i}' \in D_n$ such that $\vec{i}' \sim \vec{j}$ and $\vec{j}'$. Then $\vec{i}', \vec{j}, \vec{i}$ and $\vec{j}'$ form a four-cycle. Since $\mathcal{H}$ is four-cycle free either $x_{\vec{i}'} = x_{\vec{i}}$ or $x_{\vec{j}'} = x_{\vec{j}}$.

Suppose $x_{\vec{i}'} = x_{\vec{i}}$; the other case is similar. Since π is a local isomorphism and $\tilde{x}^{n+1}_{\vec{i}}, \tilde{x}^{n+1}_{\vec{i}'} \sim \tilde{x}^{n+1}_{\vec{j}}$, we get that $\tilde{x}^{n+1}_{\vec{i}} = \tilde{x}^{n+1}_{\vec{i}'}$. But $\vec{i}', \vec{j}' \in D_n$ and $\tilde{x}^{n+1}|_{D_n} = \tilde{x}^n$ is a graph homomorphism; therefore $\tilde{x}^{n+1}_{\vec{i}} = \tilde{x}^{n+1}_{\vec{i}'} \sim \tilde{x}^{n+1}_{\vec{j}'}$. $\square$

Corollary 4.4.3. *Let $\mathcal{H}$ be a connected four-cycle free graph and $x, y \in X_{\mathcal{H}}$. Consider some lifts $\tilde{x}, \tilde{y} \in X_{E_{\mathcal{H}}}$ such that $\pi(\tilde{x}) = x$ and $\pi(\tilde{y}) = y$. If for some $\vec{i}_0 \in \mathbb{Z}^d$, $\tilde{x}_{\vec{i}_0} = \tilde{y}_{\vec{i}_0}$ then $\tilde{x} = \tilde{y}$ on the connected subset of*

$$\{\vec{j} \in \mathbb{Z}^d \mid x_{\vec{j}} = y_{\vec{j}}\}$$

which contains $\vec{i}_0$.

Proof. Let D be the connected component of $\{\vec{i} \in \mathbb{Z}^d \mid x_{\vec{i}} = y_{\vec{i}}\}$ and $\tilde{D}$ be the connected component of $\{\vec{i} \in \mathbb{Z}^d \mid \tilde{x}_{\vec{i}} = \tilde{y}_{\vec{i}}\}$ which contain $\vec{i}_0$.

Clearly $\tilde{D} \subset D$. Suppose $\tilde{D} \neq D$. Since both D and $\tilde{D}$ are non-empty, connected sets there exist $\vec{i} \in D \setminus \tilde{D}$ and $\vec{j} \in \tilde{D}$ such that $\vec{i} \sim \vec{j}$. Then $x_{\vec{i}} = y_{\vec{i}}$, $x_{\vec{j}} = y_{\vec{j}}$ and $\tilde{x}_{\vec{j}} = \tilde{y}_{\vec{j}}$. Since π is a local isomorphism, the lift must satisfy $\tilde{x}_{\vec{i}} = \tilde{y}_{\vec{i}}$ implying $\vec{i} \in \tilde{D}$. This proves that $D = \tilde{D}$. $\square$

The following corollary says that any two lifts of the same graph homomorphism are 'identical'.

Corollary 4.4.4. *Let $\mathcal{H}$ be a connected four-cycle free graph. Then for all $\tilde{x}^1, \tilde{x}^2 \in X_{E_{\mathcal{H}}}$ satisfying $\pi(\tilde{x}^1) = \pi(\tilde{x}^2) = x$ there exists an isomorphism $\phi : E_{\mathcal{H}} \longrightarrow E_{\mathcal{H}}$ such that $\phi \circ \tilde{x}^1 = \tilde{x}^2$.*

Proof. By Proposition 4.4.1 there exists an isomorphism $\phi : E_{\mathcal{H}} \longrightarrow E_{\mathcal{H}}$ such that $\phi(\tilde{x}^1_{\vec{0}}) = \tilde{x}^2_{\vec{0}}$ and $\pi \circ \phi = \pi$. Then $(\phi \circ \tilde{x}^1)_{\vec{0}} = \tilde{x}^2_{\vec{0}}$ and $\pi(\phi \circ \tilde{x}^1) = (\pi \circ \phi)(\tilde{x}^1) = \pi(\tilde{x}^1) = x$. By Proposition 4.4.2 $\phi \circ \tilde{x}^1 = \tilde{x}^2$. $\square$

It is worth noting at this point the relationship of the universal cover described here with the universal cover in algebraic topology. Undirected graphs can be identified with 1 dimensional CW-complexes where the set of vertices correspond to the 0-cells, the edges to the 1-cells of the complex and the attaching map sends the end-points of the edges to their respective vertices. With this correspondence in mind the (topological) universal covering space coincides with the (combinatorial) universal covering space described above; indeed a 1 dimensional CW-complex is simply connected if and only if it does not have any loops, that is, the corresponding graph does not have any cycles; it is a tree. The results in the section are well known in much greater generality. Look for instance in Chapter 13 in [35] or Chapters 5 and 6 in [31].

4.5 Generalised Height Functions and Sub-Cocycles

Existence of lifts as described in the previous section enables us to measure the 'rigidity' of configurations. In this section we define generalised height functions and subsequently the slope of configurations, where steepness corresponds to this 'rigidity'.

Fix a connected four-cycle free graph $\mathcal{H}$. Given $x \in X_{\mathcal{H}}$ we can define the corresponding *generalised height function* $h_x : \mathbb{Z}^d \times \mathbb{Z}^d \longrightarrow \mathbb{Z}$ given by $h_x(\vec{i}, \vec{j}) := d_{E_{\mathcal{H}}}(\tilde{x}_{\vec{i}}, \tilde{x}_{\vec{j}})$ where $\tilde{x}$ is a lift of x. It follows from Corollary 4.4.4 that h_x is independent of the lift $\tilde{x}$.

Given a finite subset $A \subset \mathbb{Z}^d$ and $x \in X_{\mathcal{H}}$ we define the *range of x on A* as

$$Range_A(x) := \max_{\vec{j}_1, \vec{j}_2 \in A} h_x(\vec{j}_1, \vec{j}_2).$$

For all $x \in X_{\mathcal{H}}$

$$Range_A(x) \leq Diameter(A)$$

and more specifically

$$Range_{D_n}(x) \leq 2n \tag{4.5.1}$$

for all $n \in \mathbb{N}$. Since $\tilde{x} \in X_{E_{\mathcal{H}}}$ is a map between bipartite graphs it preserves the parity of the distance function, that is, if $\vec{i}, \vec{j} \in \mathbb{Z}^d$ and $x \in X_{\mathcal{H}}$ then the parity of $\|\vec{i} - \vec{j}\|_1$ is the same as that of $h_x(\vec{i}, \vec{j})$. As a consequence it follows that $Range_{\partial D_n}(x)$ is even for all $x \in X_{\mathcal{H}}$ and $n \in \mathbb{N}$. We note that

$$Range_A(x) = diameter(Image(\tilde{x}|_A)).$$

The generalised height function h_x is subadditive, that is,

$$h_x(\vec{i}, \vec{j}) \leq h_x(\vec{i}, \vec{k}) + h_x(\vec{k}, \vec{j})$$

for all $x \in X_{\mathcal{H}}$ and $\vec{i}, \vec{j}$ and $\vec{k} \in \mathbb{Z}^d$. This is in contrast with the usual height function (as in Subsection 2.3 and [41]) where there is an equality instead of the inequality. This raises some technical difficulties which are partly handled by the subadditive ergodic theorem.

The following terminology is not completely standard: Given a shift space X a *sub-cocycle* is a measurable map $c : X \times \mathbb{Z}^d \longrightarrow \mathbb{N} \cup \{0\}$ such that for all $\vec{i}, \vec{j} \in \mathbb{Z}^d$

$$c(x, \vec{i} + \vec{j}) \leq c(x, \vec{i}) + c(\sigma^{\vec{i}}(x), \vec{j}).$$

Sub-cocycles arise in a variety of situations; look for instance in [25]. We are interested in the case $c(x, \vec{i}) = h_x(\vec{0}, \vec{i})$ for all $x \in X_{\mathcal{H}}$ and $\vec{i} \in \mathbb{Z}^d$. The measure of 'rigidity' lies in the asymptotics of this sub-cocycle, the existence of which is provided by the subadditive ergodic theorem. Given a set X if $f : X \longrightarrow \mathbb{R}$ is a function then let $f^+ := \max(0, f)$.

Theorem 4.5.1 (Subadditive Ergodic Theorem)**.** *[58] Let $(X, \mathcal{B}, \mu)$ be a probability space and let $T : X \longrightarrow X$ be measure preserving. Let $\{f_n\}_{n=1}^{\infty}$ be a sequence of measurable functions $f_n : X \longrightarrow \mathbb{R} \cup \{\infty\}$ satisfying the conditions:*

(a) $f_1^+ \in L^1(\mu)$

(b) for each m, $n \geq 1$, $f_{n+m} \leq f_n + f_m \circ T^n$ μ-almost everywhere.

Then there exists a measurable function $f \longrightarrow \mathbb{R} \cup \{-\infty\}$ such that $f^+ \in L^1(\mu)$, $f \circ T = f$, $\lim_{n \to \infty} \frac{1}{n} f_n = f$, μ-almost everywhere and

$$\lim_{n \longrightarrow \infty} \frac{1}{n} \int f_n d\mu = \inf_n \frac{1}{n} \int f_n d\mu = \int f d\mu.$$

Given a direction $\vec{i} = (i_1, i_2, \ldots, i_d) \in \mathbb{R}^d$ let $\lfloor \vec{i} \rfloor = (\lfloor i_1 \rfloor, \lfloor i_2 \rfloor, \ldots, \lfloor i_d \rfloor)$. We define for all $x \in X_{\mathcal{H}}$ the *slope of x in the direction $\vec{i}$* as

$$sl_{\vec{i}}(x) := \lim_{n \longrightarrow \infty} \frac{1}{n} h_x(\vec{0}, \lfloor n\vec{i} \rfloor)$$

whenever it exists.

If $\vec{i} \in \mathbb{Z}^d$ we note that the sequence of functions $f_n : X_{\mathcal{H}} \longrightarrow \mathbb{N} \cup \{\vec{0}\}$ given by

$$f_n(x) = h_x(\vec{0}, n\vec{i})$$

satisfies the hypothesis of this theorem for any shift-invariant probability measure on $X_{\mathcal{H}}$: $|f_1| \leq \|\vec{i}\|_1$ and the subadditivity condition in the theorem is just a restatement of the sub-cocycle condition described above, that is, if $T = \sigma^{\vec{i}}$ then

$$f_{n+m}(x) = h_x(\vec{0}, (n+m)\vec{i}) \leq h_x(\vec{0}, n\vec{i}) + h_{\sigma^{n\vec{i}}x}(\vec{0}, m\vec{i}) = f_n(x) + f_m(T^n(x)).$$

The asymptotics of the generalised height functions (or more generally the sub-cocycles) are a consequence of the subadditive ergodic theorem as we will describe next. In the following by an ergodic measure on $X_{\mathcal{H}}$, we mean a probability measure on $X_{\mathcal{H}}$ which is ergodic with respect to the $\mathbb{Z}^d$-shift action on $X_{\mathcal{H}}$.

Proposition 4.5.2 (Existence of Slopes). *Let $\mathcal{H}$ be a connected four-cycle free graph and μ be an ergodic measure on $X_{\mathcal{H}}$. Then for all $\vec{i} \in \mathbb{Z}^d$*

$$sl_{\vec{i}}(x) = \lim_{n \longrightarrow \infty} \frac{1}{n} h_x(\vec{0}, n\vec{i})$$

exists almost everywhere and is independent of x. Moreover if $\vec{i} = (i_1, i_2 \ldots, i_d)$ then

$$sl_{\vec{i}}(x) \leq \sum_{k=1}^{d} |i_k| sl_{\vec{e}_k}(x).$$

Proof. Fix a direction $\vec{i} \in \mathbb{Z}^d$. Consider the sequence of functions $\{f_n\}_{n=1}^{\infty}$ and the map $T : X_{\mathcal{H}} \longrightarrow X_{\mathcal{H}}$ as described above. By the subadditive ergodic theorem there exists a function $f : X_{\mathcal{H}} \longrightarrow \mathbb{R} \cup \{-\infty\}$ such that

$$\lim_{n \to \infty} \frac{1}{n} f_n = f \text{ almost everywhere.}$$

Note that $f = sl_{\vec{i}}$. Since for all $x \in X_{\mathcal{H}}$ and $n \in \mathbb{N}$, $0 \leq f_n \leq n\|\vec{i}\|_1$, $0 \leq f(x) \leq \|i\|_1$ whenever it exists. Fix any $\vec{j} \in \mathbb{Z}^d$. Then

$$f_n(\sigma^{\vec{j}}(x)) \quad = \quad h_{\sigma^{\vec{j}}(x)}(\vec{0}, n\vec{i}) = h_x(\vec{j}, n\vec{i} + \vec{j})$$

and hence

$$\begin{aligned} -h_x(\vec{j}, \vec{0}) + h_x(\vec{0}, n\vec{i}) - h_x(n\vec{i}, n\vec{i} + \vec{j}) \quad &\leq \quad f_n(\sigma^{\vec{j}}(x)) \\ &\leq \quad h_x(\vec{j}, \vec{0}) + h_x(\vec{0}, n\vec{i}) + h_x(n\vec{i}, n\vec{i} + \vec{j}) \end{aligned}$$

implying

$$-2\|\vec{j}\|_1 + f_n(x) \leq \quad f_n(\sigma^{\vec{j}}(x)) \quad \leq 2\|\vec{j}\|_1 + f_n(x)$$

implying

$$f(x) = \lim_{n \longrightarrow \infty} \frac{1}{n} f_n(x) = \lim_{n \longrightarrow \infty} \frac{1}{n} f_n(\sigma^{\vec{j}} x) = f(\sigma^{\vec{j}}(x))$$

almost everywhere. Since μ is ergodic $sl_{\vec{i}} = f$ is constant almost everywhere. Let $\vec{i}^k = (i_1, i_2, \ldots, i_k, 0, \ldots, 0) \in$

$\mathbb{Z}^d$. By the subadditive ergodic theorem

$$\begin{aligned} sl_{\vec{i}}(x) = \int sl_{\vec{i}}(x) d\mu &= \lim_{n \longrightarrow \infty} \frac{1}{n} \int h_x(\vec{0}, n\vec{i}) d\mu \\ &\leq \sum_{k=1}^{d} \lim_{n \longrightarrow \infty} \frac{1}{n} \int h_{\sigma^{n\vec{i}^{k-1}}(x)}(\vec{0}, n i_k \vec{e}_k) d\mu \\ &= \sum_{k=1}^{d} \lim_{n \longrightarrow \infty} \frac{1}{n} \int h_x(\vec{0}, n i_k \vec{e}_k) d\mu \\ &\leq \sum_{k=1}^{d} |i_k| \lim_{n \longrightarrow \infty} \frac{1}{n} \int h_x(\vec{0}, n\vec{e}_k) d\mu \\ &= \sum_{k=1}^{d} |i_k| sl_{\vec{e}_k}(x). \end{aligned}$$

almost everywhere. □

Corollary 4.5.3. *Let $\mathcal{H}$ be a connected four-cycle free graph. Suppose μ is an ergodic measure on $X_{\mathcal{H}}$. Then for all $\vec{i} \in \mathbb{R}^d$*

$$sl_{\vec{i}}(x) = \lim_{n \longrightarrow \infty} \frac{1}{n} h_x(\vec{0}, \lfloor n\vec{i} \rfloor)$$

exists almost everywhere and is independent of x. Moreover if $\vec{i} = (i_1, i_2, \ldots, i_d)$ then

$$sl_{\vec{i}}(x) \leq \sum_{k=1}^{d} |i_k| sl_{\vec{e}_k}(x).$$

Proof. Let $\vec{i} \in \mathbb{Q}^d$ and $N \in \mathbb{N}$ such that $N\vec{i} \in \mathbb{Z}^d$. For all $n \in \mathbb{N}$ there exists $k \in \mathbb{N} \cup \{0\}$ and $0 \leq m \leq N-1$ such that $n = kN + m$. Then for all $x \in X_{\mathcal{H}}$

$$h_x(\vec{0}, kN\vec{i}) - N\|\vec{i}\|_1 \leq h_x(\vec{0}, \lfloor n\vec{i} \rfloor) \leq h_x(\vec{0}, kN\vec{i}) + N\|\vec{i}\|_1$$

proving

$$sl_{\vec{i}}(x) = \lim_{n \longrightarrow \infty} \frac{1}{n} h_x(\vec{0}, \lfloor n\vec{i} \rfloor) = \frac{1}{N} \lim_{k \longrightarrow \infty} \frac{1}{k} h_x(\vec{0}, kN\vec{i}) = \frac{1}{N} sl_{N\vec{i}}(x)$$

almost everywhere. Since $sl_{N\vec{i}}$ is constant almost everywhere, we have that $sl_{\vec{i}}$ is constant almost everywhere as well; denote the constant by $c_{\vec{i}}$. Also

$$sl_{\vec{i}}(x) \leq \frac{1}{N} \sum_{l=1}^{d} |N i_l| sl_{\vec{e}_l}(x) = \sum_{l=1}^{d} |i_l| sl_{\vec{e}_l}(x).$$

Let $X \subset X_{\mathcal{H}}$ be a set of configurations x such that

$$\lim_{n \longrightarrow \infty} \frac{1}{n} h_x(\vec{0}, \lfloor n\vec{i} \rfloor) = c_{\vec{i}}$$

for all $\vec{i} \in \mathbb{Q}^d$. We have proved that $\mu(X) = 1$.

Fix $x \in X$. Let $\vec{i}, \vec{j} \in \mathbb{R}^d$ such that $\|\vec{i} - \vec{j}\|_1 < \epsilon$. Then

$$\left| \frac{1}{n} h_x(\vec{0}, \lfloor n\vec{i} \rfloor) - \frac{1}{n} h_x(\vec{0}, \lfloor n\vec{j} \rfloor) \right| \leq \frac{1}{n} \| \lfloor n\vec{i} \rfloor - \lfloor n\vec{j} \rfloor \|_1 \leq \epsilon + \frac{2d}{n}.$$

Thus we can approximate $\frac{1}{n} h_x(\vec{0}, \lfloor n\vec{i} \rfloor)$ for $\vec{i} \in \mathbb{R}^d$ by $\frac{1}{n} h_x(\vec{0}, \lfloor n\vec{j} \rfloor)$ for $\vec{j} \in \mathbb{Q}^d$ to prove that $\lim_{n \longrightarrow \infty} \frac{1}{n} h_x(\vec{0}, \lfloor n\vec{i} \rfloor)$ exists for all $\vec{i} \in \mathbb{R}^d$, is independent of $x \in X$ and satisfies

$$sl_{\vec{i}}(x) \leq \sum_{k=1}^{d} |i_k| sl_{\vec{e}_k}(x).$$

□

The existence of slopes can be generalised from generalised height functions to continuous sub-cocycles; the same proofs work:

Proposition 4.5.4. *Let $c : X \times \mathbb{Z}^d \longrightarrow \mathbb{R}$ be a continuous sub-cocycle and μ be an ergodic measure on X. Then for all $\vec{i} \in \mathbb{R}^d$*

$$sl^c_{\vec{i}}(x) := \lim_{n \longrightarrow \infty} \frac{1}{n} c(x, \lfloor n\vec{i} \rfloor)$$

exists almost everywhere and is independent of x. Moreover if $\vec{i} = (i_1, i_2 \ldots, i_d)$ then

$$sl^c_{\vec{i}}(x) \leq \sum_{k=1}^{d} |i_k| sl^c_{\vec{e}_k}(x).$$

Let C_X be the space of continuous sub-cocycles on a shift space X. C_X has a natural vector space structure: given $c_1, c_2 \in C_X$, $(c_1 + \alpha c_2)$ is also a continuous sub-cocycle on X for all $\alpha \in \mathbb{R}$ where addition and scalar multiplication is point-wise. The following is not hard to prove and follows directly from definition.

Proposition 4.5.5. *Let X, Y be conjugate shift spaces. Then every conjugacy $f : X \longrightarrow Y$ induces an isomorphism $f^\star : C_Y \longrightarrow C_X$ given by*

$$f^\star(c)(x, \vec{i}) := c(f(x), \vec{i})$$

for all $c \in C_Y$, $x \in X$ and $\vec{i} \in \mathbb{Z}^d$. Moreover $sl^c_{\vec{i}}(y) = sl^{f^\star(c)}_{\vec{i}}(f^{-1}(y))$ for all $y \in Y$ and $\vec{i} \in \mathbb{R}^d$ for which the slope $sl^c_{\vec{i}}(y)$ exists.

4.6 Proofs of the Main Theorems

Proof of Theorem 4.2.4. If $\mathcal{H}$ is a single edge, then $X_{\mathcal{H}}$ is the orbit of a periodic configuration; the result follows immediately. Suppose this is not the case. The proof follows loosely the proof of Theorem 4.3.4 and morally the ideas from [52]: We prove existence of two kind of configurations in $X_{\mathcal{H}}$, ones which are 'poor' (Lemma 4.6.1), in the sense that they are frozen and others which are 'universal' (Lemma 4.6.2), for which the homoclinic class is dense.

Ideas for the following proof were inspired by discussions with Anthony Quas. A similar result in a special case is contained in Lemma 2.7.2.

Lemma 4.6.1. *Let $\mathcal{H}$ be a connected four-cycle free graph and μ be an ergodic probability measure on $X_{\mathcal{H}}$ such that $sl_{\vec{e}_k}(x) = 1$ almost everywhere for some $1 \leq k \leq d$. Then μ is frozen and $h_\mu = 0$.*

Proof. Without loss of generality assume that $sl_{\vec{e}_1}(x) = 1$ almost everywhere. By the subadditivity of the generalised height function for all $k, n \in \mathbb{N}$ and $x \in X_{\mathcal{H}}$ we know that

$$\frac{1}{kn}h_x(\vec{0}, kn\vec{e}_1) \leq \frac{1}{kn}\sum_{m=0}^{n-1} h_x(km\vec{e}_1, k(m+1)\vec{e}_1) = \frac{1}{n}\sum_{m=0}^{n-1}\frac{1}{k}h_{\sigma^{km\vec{e}_1}(x)}(\vec{0}, k\vec{e}_1) \leq 1.$$

Since $sl_{\vec{e}_1}(x) = 1$ almost everywhere, we get that

$$\lim_{n\longrightarrow\infty}\frac{1}{n}\sum_{m=0}^{n-1}\frac{1}{k}h_{\sigma^{km\vec{e}_1}(x)}(\vec{0}, k\vec{e}_1) = 1$$

almost everywhere. By the ergodic theorem

$$\int \frac{1}{k}h_x(\vec{0}, k\vec{e}_1)d\mu = 1.$$

Therefore $h_x(\vec{0}, k\vec{e}_1) = k$ almost everywhere which implies that

$$h_x(\vec{i}, \vec{i} + k\vec{e}_1) = k \tag{4.6.1}$$

for all $\vec{i} \in \mathbb{Z}^d$ and $k \in \mathbb{N}$ almost everywhere. Let $X \subset supp(\mu)$ denote the set of such configurations.

For some $n \in \mathbb{N}$ consider two patterns $a, b \in \mathcal{L}_{B_n \cup \partial_2 B_n}(supp(\mu))$ such that $a|_{\partial_2 B_n} = b|_{\partial_2 B_n}$. We will prove that then $a|_{B_n} = b|_{B_n}$. This will prove that μ is frozen, and so $|\mathcal{L}_{B_n}(supp(\mu))| = |\mathcal{L}_{\partial_2 B_n}(supp(\mu))| \leq |\mathcal{A}|^{|\partial_2 B_n|}$ implying that $h_{top}(supp(\mu)) = 0$. By the variational principle this implies that $h_\mu = 0$.

Consider $x, y \in X$ such that $x|_{B_n \cup \partial_2 B_n} = a$ and $y|_{B_n \cup \partial_2 B_n} = b$. Noting that $\partial_2 B_n$ is connected, by Corollary 4.4.3 we can choose lifts $\tilde{x}, \tilde{y} \in X_{E_{\mathcal{H}}}$ such that $\tilde{x}|_{\partial_2 B_n} = \tilde{y}|_{\partial_2 B_n}$. Consider any $\vec{i} \in B_n$ and choose $k \in -\mathbb{N}$ such that $\vec{i} + k\vec{e}_1, \vec{i} + (2n + 2 + k)\vec{e}_1 \in \partial B_n$. Then by (4.6.1)

$d_{E_{\mathcal{H}}}(\tilde{x}_{\vec{i}+k\vec{e}_1}, \tilde{x}_{\vec{i}+(2n+2+k)\vec{e}_1}) = 2n+2$. But

$$(\tilde{x}_{\vec{i}+k\vec{e}_1}, \tilde{x}_{\vec{i}+(k+1)\vec{e}_1}, \ldots, \tilde{x}_{\vec{i}+(2n+2+k)\vec{e}_1}) \text{ and}$$

$$(\tilde{y}_{\vec{i}+k\vec{e}_1}, \tilde{y}_{\vec{i}+(k+1)\vec{e}_1}, \ldots, \tilde{y}_{\vec{i}+(2n+2+k)\vec{e}_1})$$

are walks of length $2n+2$ from $\tilde{x}_{\vec{i}+k\vec{e}_1}$ to $\tilde{x}_{\vec{i}+(2n+2+k)\vec{e}_1}$. Since $E_{\mathcal{H}}$ is a tree and the walks are of minimal length, they must be the same. Thus $\tilde{x}|_{B_n} = \tilde{y}|_{B_n}$. Taking the image under the map π we derive that

$$a|_{B_n} = x|_{B_n} = y|_{B_n} = b|_{B_n}.$$

□

This partially justifies the claim that steep slopes lead to greater 'rigidity'. We are left to analyse the case where the slope is submaximal in every direction. As in the proof of Proposition 2.6.1 we will now prove a certain mixing result for the shift space $X_{\mathcal{H}}$.

Lemma 4.6.2. *Let $\mathcal{H}$ be a connected four-cycle free graph and $|\mathcal{H}| = r$. Consider any $x \in X_{\mathcal{H}}$ and some $y \in X_{\mathcal{H}}$ satisfying $Range_{\partial D_{(d+1)n+3r+k}}(y) \leq 2k$ for some $n \in \mathbb{N}$. Then*

1. *If either $\mathcal{H}$ is not bipartite or $x_{\vec{0}}, y_{\vec{0}}$ are in the same partite class of $\mathcal{H}$ then there exists $z \in X_{\mathcal{H}}$ such that*
$$z_{\vec{i}} = \begin{cases} x_{\vec{i}} & if\ \vec{i} \in D_n \\ y_{\vec{i}} & if\ \vec{i} \in D^c_{(d+1)n+3r+k}. \end{cases}$$

2. *If $\mathcal{H}$ is bipartite and $x_{\vec{0}}, y_{\vec{0}}$ are in different partite classes of $\mathcal{H}$ then there exists $z \in X_{\mathcal{H}}$ such that*
$$z_{\vec{i}} = \begin{cases} x_{\vec{i}+\vec{e}_1} & if\ \vec{i} \in D_n \\ y_{\vec{i}} & if\ \vec{i} \in D^c_{(d+1)n+3r+k}. \end{cases}$$

The distance $dn + 3r + k$ is not optimal, but sufficient for our purposes.

Proof. We will construct the configuration z only in the case when $\mathcal{H}$ is not bipartite. The construction in the other cases is similar; the differences will be pointed out in the course of the proof.

1. **Boundary patterns with non-maximal range to monochromatic patterns inside.** Let $\tilde{y}$ be a lift of y and $\mathcal{T}'$ be the image of $\tilde{y}|_{D_{(d+1)n+3r+k+1}}$. Let $\mathcal{T}$ be a minimal subtree of $E_{\mathcal{H}}$ such that
$$Image(\tilde{y}|_{\partial D_{(d+1)n+3r+k}}) \subset \mathcal{T} \subset \mathcal{T}'.$$

Since $Range_{\partial D_{(d+1)n+3r+k}}(y) \leq 2k$, $diameter(\mathcal{T}) \leq 2k$. By Proposition 4.3.1 there exists a graph homomorphism $f : \mathcal{T}' \longrightarrow \mathcal{T}$ such that $f|_{\mathcal{T}}$ is the identity. Consider the configuration $\tilde{y}^1$ given by

$$\tilde{y}^1_{\vec{i}} = \begin{cases} f(\tilde{y}_{\vec{i}}) & \text{if } \vec{i} \in D_{(d+1)n+3r+k+1} \\ \tilde{y}_{\vec{i}} & \text{otherwise.} \end{cases}$$

The pattern

$$\tilde{y}^1|_{D_{(d+1)n+3r+k+1}} \in \mathcal{L}_{D_{(d+1)n+3r+k+1}}(X_{\mathcal{T}}) \subset \mathcal{L}_{D_{(d+1)n+3r+k+1}}(X_{E_{\mathcal{H}}}).$$

Moreover since $f|_{\mathcal{T}}$ is the identity map,

$$\tilde{y}^1|_{D^c_{(d+1)n+3r+k}} = \tilde{y}|_{D^c_{(d+1)n+3r+k}} \in \mathcal{L}_{D^c_{(d+1)n+3r+k}}(X_{E_{\mathcal{H}}}).$$

Since $X_{E_{\mathcal{H}}}$ is given by nearest neighbour constraints $\tilde{y}^1 \in X_{E_{\mathcal{H}}}$.

Recall that the radius of a nearest neighbour shift of finite type (in our case $X_{\mathcal{T}}$) is the total number of full config-folds required to obtain a stiff shift. Since $diameter(\mathcal{T}) \leq 2k$ the radius of $X_{\mathcal{T}} \leq k$. Let a stiff shift obtained by a sequence of config-folds starting at $X_{\mathcal{T}}$ be denoted by Z. Since $\mathcal{T}$ folds into a graph consisting of a single edge, Z consists of two checkerboard patterns in the vertices of an edge in $\mathcal{T}$, say $\tilde{v}_1$ and $\tilde{v}_2$. Corresponding to such a sequence of full config-folds, we had defined in Section 4.3 the outward fixing map $O_{X_{\mathcal{T}},(d+1)n+3r+k}$. By Proposition 4.3.3 the configuration $O_{X_{\mathcal{T}},(d+1)n+3r+k}(\tilde{y}^1) \in X_{E_{\mathcal{H}}}$ satisfies

$$\begin{aligned} O_{X_{\mathcal{T}},(d+1)n+3r+k}(\tilde{y}^1)|_{D_{(d+1)n+3r+1}} &\in \mathcal{L}_{D_{(d+1)n+3r+1}}(Z) \\ O_{X_{\mathcal{T}},(d+1)n+3r+k}(\tilde{y}^1)|_{D^c_{(d+1)n+3r+k}} &= \tilde{y}^1|_{D^c_{(d+1)n+3r+k}} = \tilde{y}|_{D^c_{(d+1)n+3r+k}}. \end{aligned}$$

Note that the pattern $O_{X_{\mathcal{T}},(d+1)n+3r+k}(\tilde{y}^1)|_{\partial D_{(d+1)n+3r}}$ uses a single symbol, say $\tilde{v}_1$. Let $\pi(\tilde{v}_1) = v_1$. Then the configuration $y' = \pi(O_{X_{\mathcal{T}},(d+1)n+3r+k}(\tilde{y}^1)) \in X_{\mathcal{H}}$ satisfies

$$\begin{aligned} y'|_{\partial D_{(d+1)n+3r}} &= v_1 \\ y'|_{D^c_{(d+1)n+3r+k}} &= y|_{D^c_{(d+1)n+3r+k}}. \end{aligned}$$

2. **Constant extension of an admissible pattern.** Consider some lift $\tilde{x}$ of x. We begin by extending $\tilde{x}|_{B_n}$ to a periodic configuration $\tilde{x}^1 \in X_{E_{\mathcal{H}}}$. Consider the map $f : [-n, 3n] \longrightarrow [-n, n]$ given by

$$f(k) = \begin{cases} k & \text{if } k \in [-n, n] \\ 2n - k & \text{if } k \in [n, 3n]. \end{cases}$$

Then we can construct the pattern $\tilde{a} \in \mathcal{L}_{[-n,3n]^d}(X_{E_\mathcal{H}})$ given by

$$\tilde{a}_{i_1,i_2,\ldots i_d} = \tilde{x}_{f(i_1),f(i_2),\ldots,f(i_d)}.$$

Given $k, l \in [-n, 3n]$ if $|k-l| = 1$ then $|f(k) - f(l)| = 1$. Thus $\tilde{a}$ is a locally allowed pattern in $X_{E_\mathcal{H}}$. Moreover since $f(-n) = f(3n)$ the pattern $\tilde{a}$ is 'periodic', meaning,

$$\tilde{a}_{i_1,i_2,\ldots,i_{k-1},-n,i_{k+1},\ldots,i_d} = \tilde{a}_{i_1,i_2,\ldots,i_{k-1},3n,i_{k+1},\ldots,i_d}$$

for all $i_1, i_2, \ldots, i_d \in [-n, 3n]$. Also $\tilde{a}|_{B_n} = \tilde{x}|_{B_n}$. Then the configuration $\tilde{x}^1$ obtained by tiling $\mathbb{Z}^d$ with $\tilde{a}|_{[-n,3n-1]^d}$, that is,

$$\tilde{x}^1_{\vec{i}} = \tilde{a}_{(i_1 \mod 4n,\ i_2 \mod 4n,\ \ldots,\ i_d \mod 4n)-(n,n,\ldots,n)} \text{ for all } \vec{i} \in \mathbb{Z}^d$$

is an element of $X_{E_\mathcal{H}}$. Moreover $\tilde{x}^1|_{B_n} = \tilde{a}|_{B_n} = \tilde{x}|_{B_n}$ and $Image(\tilde{x}^1) = Image(\tilde{x}|_{B_n})$. Since $diameter(B_n) = 2dn$, $diameter(Image(\tilde{x}^1)) \leq 2dn$. Let $\tilde{\mathcal{T}} = Image(\tilde{x}^1)$. Then $radius(X_{\tilde{\mathcal{T}}}) \leq dn$. Let a stiff shift obtained by a sequence of config-folds starting at $X_{\tilde{\mathcal{T}}}$ be denoted by Z'. Since $\tilde{\mathcal{T}}$ folds into a graph consisting of a single edge, Z' consists of two checkerboard patterns in the vertices of an edge in $\tilde{T}$, say $\tilde{w}_1$ and $\tilde{w}_2$. Then by Proposition 4.3.3

$$I_{X_{\tilde{\mathcal{T}}},n}(\tilde{x}^1)|_{D_n} = \tilde{x}^1|_{D_n} = \tilde{x}|_{D_n}$$
$$I_{X_{\tilde{\mathcal{T}}},n}(\tilde{x}^1)|_{D^c_{(d+1)n-1}} \in \mathcal{L}_{D^c_{(d+1)n-1}}(Z').$$

We note that $I_{X_{\tilde{\mathcal{T}}},n}(\tilde{x}^1)|_{\partial D_{(d+1)n-1}}$ consists of a single symbol, say $\tilde{w}_1$. Let $\pi(\tilde{w}_1) = w_1$. Then the configuration $x' = \pi(I_{X_{\tilde{\mathcal{T}}},n}(\tilde{x}^1)) \in X_\mathcal{H}$ satisfies

$$x'|_{D_n} = x|_{D_n} \text{ and}$$
$$x'|_{\partial D_{(d+1)n-1}} = w_1.$$

3. **Patching of an arbitrary pattern inside a configuration with non-maximal range.** We will first prove that there exists a walk on $\mathcal{H}$ from w_1 to v_1, $((w_1 = u_1), u_2, \ldots, (u_{3r+2} = v_1))$. Since the graph is not bipartite, it has a cycle p_1 such that $|p_1| \leq r-1$ and is odd. Let v' be a vertex in p_1. Then there exist walks p_2 and p_3 from w_1 to v' and from v' to v_1 respectively such that $|p_2|, |p_3| \leq r-1$. Consider any vertex $w' \sim_\mathcal{H} v_1$. If $3r+1-|p_2|-|p_3|$ is even then the walk

$$p_2 \star p_3 (\star(v_1, w', v_1))^{\frac{3r+1-|p_2|-|p_3|}{2}}$$

and if not then the walk

$$p_2 \star p_1 \star p_3(\star(v_1, w', v_1))^{\frac{3r+1-|p_1|-|p_2|-|p_3|}{2}}$$

is a walk of length $3r+1$ in $\mathcal{H}$ from w_1 to v_1. This is the only place where we use the fact that $\mathcal{H}$ is not bipartite. If it were bipartite, then we would require that $x'_{\vec{0}}$ and $y'_{\vec{0}}$ have to be in the same partite class to construct such a walk.

Given such a walk the configuration z given by

$$\begin{aligned} z|_{D_{(d+1)n}} &= x'|_{(d+1)n} \\ z|_{D^c_{(d+1)n+3r}} &= y'|_{D^c_{(d+1)n+3r}} \\ z|_{\partial D_{(d+1)n+i-2}} &= u_i \text{ for all } 1 \leq i \leq 3r+2 \end{aligned}$$

is an element of $X_{\mathcal{H}}$ for which $z|_{D_n} = x'|_{D_n} = x|_{D_n}$ and $z|_{D^c_{(d+1)n+3r+k}} = y'|_{D^c_{(d+1)n+3r+k}} = y|_{D^c_{(d+1)n+3r+k}}$.

□

We now return to the proof of Theorem 4.2.4. Let μ be an ergodic probability measure adapted to $X_{\mathcal{H}}$ with positive entropy.

Suppose $sl_{\vec{e}_i}(x) = \theta_i$ almost everywhere. By Lemma 4.6.1, $\theta_i < 1$ for all $1 \leq i \leq d$. Let $\theta = \max_i \theta_i$ and $0 < \epsilon < \frac{1}{4}(1-\theta)$. Denote by S^{d-1}, the sphere of radius 1 in $\mathbb{R}^d$ for the l^1 norm. Since S^{d-1} is compact in $\mathbb{R}^d$ we can choose a finite set $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_t\} \subset S^{d-1}$ such that for all $\vec{v} \in S^{d-1}$ there exists some $1 \leq i \leq t$ satisfying $\|\vec{v}_i - \vec{v}\|_1 < \epsilon$. By Corollary 4.5.3 for all $\vec{v} \in S^{d-1}$

$$\lim_{n \longrightarrow \infty} \frac{1}{n} h_x(\vec{0}, \lfloor n\vec{v} \rfloor) \leq \theta$$

almost everywhere. By Egoroff's theorem [18] there exists $N_0 \in \mathbb{N}$ such that for all $n \geq N_0$ and $1 \leq i \leq t$

$$\mu(\{x \in X_{\mathcal{H}} \mid h_x(\vec{0}, \lfloor n\vec{v}_i \rfloor) \leq n\theta + n\epsilon \ for\ all\ 1 \leq i \leq t\}) > 1 - \epsilon. \tag{4.6.2}$$

Let $\vec{v} \in \partial D_{n-1}$ and $1 \leq i_0 \leq t$ such that $|\frac{1}{n}\vec{v} - \vec{v}_{i_0}| < \epsilon$. If for some $x \in X_{\mathcal{H}}$ and $n \in \mathbb{N}$

$$h_x(\vec{0}, \lfloor n\vec{v}_{i_0} \rfloor) \leq n\theta + n\epsilon$$

then

$$h_x(\vec{0}, \lfloor \vec{v} \rfloor) \leq h_x(\vec{0}, \lfloor n\vec{v}_{i_0} \rfloor) + \lceil n\epsilon \rceil \leq n\theta + 2n\epsilon + 1.$$

By Inequality (4.6.2) we get

$$\mu\left(\{x \in X_{\mathcal{H}} \mid h_x\left(\vec{0}, \lfloor \vec{v} \rceil\right) \leq n\theta + 2n\epsilon + 1 \ for\ all\ \vec{v} \in \partial D_{n-1}\}\right) > 1 - \epsilon$$

for all $n \geq N_0$. Therefore for all $n \geq N_0$ there exists $x^{(n)} \in supp(\mu)$ such that

$$Range_{\partial D_{n-1}}\left(x^{(n)}\right) \leq 2n\theta + 4n\epsilon + 2 < 2n(1-\epsilon) + 2.$$

Let $x \in X_{\mathcal{H}}$ and $n_0 \in \mathbb{N}$. It is sufficient to prove that $\mu([x]_{D_{n_0-1}}) > 0$. Suppose $r := |\mathcal{H}|$. Choose $k \in \mathbb{N}$ such that

$$\begin{aligned} n_0(d+1) + 3r + k + 1 &\geq N_0 \\ 2\left(n_0(d+1) + 3r + k + 1\right)(1-\epsilon) + 2 &\leq 2k. \end{aligned}$$

Then by Lemma 4.6.2 there exists $z \in X_{\mathcal{H}}$ such that either

$$z_{\vec{j}} = \begin{cases} x_{\vec{j}} & if\ \vec{j} \in D_{n_0} \\ x_{\vec{j}}^{(n_0(d+1)+3r+k+1)} & if\ \vec{j} \in D^c_{n_0(d+1)+3r+k} \end{cases}$$

or

$$z_{\vec{j}} = \begin{cases} x_{\vec{j}+\vec{e}_1} & if\ \vec{j} \in D_{n_0} \\ x_{\vec{j}}^{(n_0(d+1)+3r+k+1)} & if\ \vec{j} \in D^c_{n_0(d+1)+3r+k}. \end{cases}$$

In either case $(z, x^{(n_0(d+1)+3r+k+1)}) \in \Delta_{X_{\mathcal{H}}}$. Since μ is adapted to $X_{\mathcal{H}}$, $z \in supp(\mu)$. In the first case we get that $\mu([x]_{D_{n_0-1}}) = \mu([z]_{D_{n_0-1}}) > 0$. In the second case we get that

$$\mu([x]_{D_{n_0-1}}) = \mu(\sigma^{\vec{e}_1}([x]_{D_{n_0-1}})) = \mu([z]_{D_{(n_0-1)-\vec{e}_1}}) > 0.$$

This completes the proof.

□

Every shift space conjugate to an entropy minimal shift space is entropy minimal. However a shift space X which is conjugate to $X_{\mathcal{H}}$ for $\mathcal{H}$ which is connected and four-cycle free need not even be a hom-shift. By following the proof carefully it is possible to extract a condition for entropy minimality which is conjugacy-invariant:

Theorem 4.6.3. *Let X be a shift of finite type and c a continuous sub-cocycle on X with the property that $c(\cdot, \vec{i}) \leq \|\vec{i}\|_1$ for all $\vec{i} \in \mathbb{Z}^d$ and for every ergodic probability measure μ adapted to X*

1. *If $sl^c_{\vec{e}_i}(x) = 1$ almost everywhere for some $1 \leq i \leq d$ then $h_\mu < h_{top}(X)$.*
2. *If $sl^c_{\vec{e}_i}(x) < 1$ almost everywhere for all $1 \leq i \leq d$ then $supp(\mu) = X$.*

Then X is entropy minimal.

Here is a sketch: By Proposition 4.2.1 and Theorems 4.2.2, 4.2.3 it is sufficient to prove that every ergodic measure of maximal entropy is fully supported. If X is a shift of finite type satisfying the hypothesis of Theorem 4.6.3 then it is entropy minimal because every ergodic measure of maximal entropy of X is an ergodic probability measure adapted to X; its entropy is either smaller than $h_{top}(X)$ or it is fully supported. To see why the condition is conjugacy invariant suppose that $f : X \longrightarrow Y$ is a conjugacy and $c \in C_Y$ satisfies the hypothesis of the theorem. Then by Proposition 4.5.5 it follows that $f^{\star}(c) \in C_X$ satisfies the hypothesis as well.

Proof of Theorem 4.1.4. By Proposition 4.1.5 we can assume that $\mathcal{H}$ is connected. Consider some $(x, y) \in \Delta_{X_{\mathcal{H}}}$. By Corollary 4.4.3 there exist $(\tilde{x}, \tilde{y}) \in \Delta_{X_{E_{\mathcal{H}}}}$ such that $\pi(\tilde{x}) = x$ and $\pi(\tilde{y}) = y$. It is sufficient to prove that there is a chain of pivots from $\tilde{x}$ to $\tilde{y}$. We will proceed by induction on $\sum_{\vec{i}\in\mathbb{Z}^d} d_{E_{\mathcal{H}}}(\tilde{x}_{\vec{i}}, \tilde{y}_{\vec{i}})$. The induction hypothesis (on M) is : If $\sum_{\vec{i}\in\mathbb{Z}^d} d_{E_{\mathcal{H}}}(\tilde{x}_{\vec{i}}, \tilde{y}_{\vec{i}}) = 2M$ then there exists a chain of pivots from $\tilde{x}$ to $\tilde{y}$.

We note that $d_{E_{\mathcal{H}}}(\tilde{x}_{\vec{i}}, \tilde{y}_{\vec{i}})$ is even for all $\vec{i} \in \mathbb{Z}^2$ since there exists $\vec{i'} \in \mathbb{Z}^d$ such that $\tilde{x}_{\vec{i'}} = \tilde{y}_{\vec{i'}}$ and hence $\tilde{x}_{\vec{i}}$ and $\tilde{y}_{\vec{i}}$ are in the same partite class of $E_{\mathcal{H}}$ for all $\vec{i} \in \mathbb{Z}^d$.

The base case ($M = 1$) occurs exactly when $\tilde{x}$ and $\tilde{y}$ differ at a single site; there is nothing to prove in this case. Assume the hypothesis for some $M \in \mathbb{N}$.

Consider $(\tilde{x}, \tilde{y}) \in \Delta_{X_{E_{\mathcal{H}}}}$ such that

$$\sum_{\vec{i}\in\mathbb{Z}^d} d_{E_{\mathcal{H}}}(\tilde{x}_{\vec{i}}, \tilde{y}_{\vec{i}}) = 2M + 2.$$

Let

$$B = \{\vec{j} \in \mathbb{Z}^d \mid \tilde{x}_{\vec{j}} \neq \tilde{y}_{\vec{j}}\}$$

and a vertex $\tilde{v} \in E_{\mathcal{H}}$. Without loss of generality we can assume that

$$\max_{\vec{i}\in B} d_{E_{\mathcal{H}}}(\tilde{v}, \tilde{x}_{\vec{i}}) \geq \max_{\vec{i}\in B} d_{E_{\mathcal{H}}}(\tilde{v}, \tilde{y}_{\vec{i}}). \tag{4.6.3}$$

Consider some $\vec{i}_0 \in B$ such that

$$d_{E_{\mathcal{H}}}(\tilde{v}, \tilde{x}_{\vec{i}_0}) = \max_{\vec{i}\in B} d_{E_{\mathcal{H}}}(\tilde{v}, \tilde{x}_{\vec{i}}).$$

Consider the shortest walks $(\tilde{v} = \tilde{v}_1, \tilde{v}_2, \ldots, \tilde{v}_n = \tilde{x}_{\vec{i}_0})$ from $\tilde{v}$ to $\tilde{x}_{\vec{i}_0}$ and $(\tilde{v} = \tilde{v}'_1, \tilde{v}'_2, \ldots, \tilde{v}'_{n'} = \tilde{y}_{\vec{i}_0})$ from $\tilde{v}$ to $\tilde{y}_{\vec{i}_0}$. By Assumption (4.6.3), $n' \leq n$. Since these are the shortest walks on a tree, if $\tilde{v}'_k = \tilde{v}_{k'}$ for some $1 \leq k \leq n'$ and $1 \leq k' \leq n$ then $k = k'$ and $\tilde{v}_l = \tilde{v}'_l$ for $1 \leq l \leq k$. Let

$$k_0 = \max\{1 \leq k \leq n' \mid \tilde{v}'_k = \tilde{v}_k\}.$$

Then the shortest walk from $\tilde{x}_{\vec{i}_0}$ to $\tilde{y}_{\vec{i}_0}$ is given by $\tilde{x}_{\vec{i}_0} = \tilde{v}_n, \tilde{v}_{n-1}, \tilde{v}_{n-2}, \ldots, \tilde{v}_{k_0}, \tilde{v}'_{k_0+1}, \ldots, \tilde{v}'_{n'} = \tilde{y}_{\vec{i}_0}$.

We will prove for all $\vec{i} \sim \vec{i}_0$, $\tilde{x}_{\vec{i}} = \tilde{v}_{n-1}$. This is sufficient to complete the proof since then the configuration

$$\tilde{x}^{(1)}_{\vec{j}} = \begin{cases} \tilde{x}_{\vec{j}} & if\ \vec{j} \neq \vec{i}_0 \\ \tilde{v}_{n-2} & if\ \vec{j} = \vec{i}_0, \end{cases}$$

is an element of $X_{E_{\mathcal{H}}}$, $(\tilde{x}, \tilde{x}^{(1)})$ is a pivot and

$$n + n' - 2k_0 - 2 = d_{E_{\mathcal{H}}}(\tilde{x}^{(1)}_{\vec{i}_0}, \tilde{y}_{\vec{i}_0}) < d_{E_{\mathcal{H}}}(\tilde{x}_{\vec{i}_0}, \tilde{y}_{\vec{i}_0}) = n + n' - 2k_0$$

giving us a pair $(\tilde{x}^{(1)}, \tilde{y})$ such that

$$\sum_{\vec{i} \in \mathbb{Z}^d} d_{E_{\mathcal{H}}}(\tilde{x}^{(1)}_{\vec{i}}, \tilde{y}_{\vec{i}}) = \sum_{\vec{i} \in \mathbb{Z}^d} d_{E_{\mathcal{H}}}(\tilde{x}_{\vec{i}}, \tilde{y}_{\vec{i}}) - 2 = 2M.$$

There are two possible cases:

1. $\vec{i} \in B$: Then $d_{E_{\mathcal{H}}}(\tilde{v}, \tilde{x}_{\vec{i}}) = d_{E_{\mathcal{H}}}(\tilde{v}, \tilde{x}_{\vec{i}_0}) - 1$ and $\tilde{x}_{\vec{i}} \sim_{E_{\mathcal{H}}} \tilde{x}_{\vec{i}_0}$. Since $E_{\mathcal{H}}$ is a tree, $\tilde{x}_{\vec{i}} = \tilde{v}_{n-1}$.

2. $\vec{i} \notin B$: Then $\tilde{x}_{\vec{i}} = \tilde{y}_{\vec{i}}$ and we get that $d_{E_{\mathcal{H}}}(\tilde{x}_{\vec{i}_0}, \tilde{y}_{\vec{i}_0}) = 2$. Since $\tilde{x}_{\vec{i}} \sim_{E_{\mathcal{H}}} \tilde{x}_{\vec{i}_0}$, the shortest walk joining $\tilde{v}$ and $\tilde{x}_{\vec{i}}$ must either be $\tilde{v} = \tilde{v}_1, \tilde{v}_2, \ldots, \tilde{v}_{n-1} = \tilde{x}_{\vec{i}}$ or $\tilde{v} = \tilde{v}_1, \tilde{v}_2, \ldots, \tilde{v}_n = \tilde{x}_{\vec{i}_0}, \tilde{v}_{n+1} = \tilde{x}_{\vec{i}}$. We want to prove that the former is true. Suppose not.

   Since $\tilde{y}_{\vec{i}_0} \sim_{E_{\mathcal{H}}} \tilde{x}_{\vec{i}}$ and $\vec{i}_0 \in B$, the shortest walk from $\tilde{v}$ and $\tilde{y}_{\vec{i}_0}$ is $\tilde{v} = \tilde{v}_1, \tilde{v}_2, \ldots, \tilde{v}_n = \tilde{x}_{\vec{i}_0}, \tilde{v}_{n+1} = \tilde{x}_{\vec{i}}, \tilde{v}_{n+2} = \tilde{y}_{\vec{i}_0}$. This contradicts Assumption (4.6.3) and completes the proof.

$\square$

Chapter 5

Further Directions

5.1 Markov Random Fields and Gibbs States with Nearest Neighbour Interactions

In Chapter 2 we introduced Markov cocycles and used them to study Markov random fields and Gibbs states with nearest neighbour interactions. When the underlying configuration space is a shift space with the pivot property the space of shift-invariant Markov cocycles is finite dimensional; it can act as a substitute for shift-invariant nearest neighbour interactions for giving a description of the specification "using finitely many parameters". In Chapter 3 we introduced a new notion of folding called strong config-folding and used it to generalise the Hammersley-Clifford theorem when the underlying graph is bipartite.

5.1.1 Supports of Markov Random Fields

Which shift spaces can be the support of a shift-invariant Markov random field on the Cayley graph of $\mathbb{Z}^d$? A necessary condition is that they must be topological Markov fields which are the support of some shift-invariant probability measure. For $d = 1$ this condition is sufficient as well and the support of shift-invariant Markov random fields are further characterised as the non-wandering nearest neighbour shifts of finite type [12]. However we do not know whether the condition is sufficient in higher dimensions. Also: Suppose a shift of finite type is the support of some shift-invariant Markov random field. Must it also be the support of a shift-invariant Gibbs state for some shift-invariant nearest neighbour interaction?

5.1.2 Algorithmic Aspects

Suppose we are given a nearest neighbour shift of finite type $X \subset \mathcal{A}^{\mathbb{Z}^d}$ with the pivot property along with its globally allowed patterns on $\{\vec{0}\} \cup \partial\{\vec{0}\}$. Is there an algorithm to determine the dimension of $\mathbf{M}_X^{\mathbb{Z}^d}$? If so, is there a way to decide which of the shift-invariant Markov cocycles have a corresponding fully supported shift-invariant probability measure on the subshift? In case the subshift has a safe symbol, such an algorithm can be derived from the proof of the Hammersley-Clifford Theorem [43] and also from Lemma 3.1 in [15]. More generally, in case the subshift strongly config-folds to $\{\star\}^{\mathbb{Z}^d}$ for some $\star \in \mathcal{A}$ such an algorithm can be derived from the proof of Theorem 3.3.2.

5.1.3 Markov Random Fields for other Models

There are other models for which it would be interesting to get a good description of the Markov random fields and Gibbs states. One such family is the r-colourings of $\mathbb{Z}^d$ with $r \geq 2d+2$. The space of domino tilings of $\mathbb{Z}^2$ is another interesting model. For these examples we know the generalised pivot property (defined in Subsection 5.2.4) holds, so the space of shift-invariant Markov cocycles is finite dimensional. We believe that some of the techniques developed Chapters 2 and 3 will be useful in studying other systems.

5.1.4 Changing the Underlying Graph

Our results in Chapter 2 were dependent on the structure of the standard Cayley graph of $\mathbb{Z}^d$. How does the underlying graph affect our results? What happens if we choose a different set of generators?

Further, by Theorems 3.3.1 and 3.3.2 we have generalised the Hammersley-Clifford Theorem, but only when the graph $\mathcal{G}$ is bipartite. Can this be generalised beyond the bipartite case? Note that $\mathcal{G}$ being bipartite is used in many critical parts of the proof e.g. the construction of the elements x^v in Lemma 3.3.3, construction of the interaction in Lemma 3.3.4 etc.

5.1.5 Mixing Properties of Subshifts and the Dimension of the Invariant Markov Cocycles

In Section 2.8, we constructed a subshift such that the dimension of the space of shift-invariant Markov cocycles is uncountable. However this subshift has poor mixing properties (See [4] for a discussion of some mixing properties of $\mathbb{Z}^d$-subshifts). On the other hand the safe symbol assumption implies that the space of shift-invariant Markov cocycles is the same as the space of Gibbs cocycles with shift-invariant nearest neighbour interactions (and hence finite dimensional). Are there some natural mixing conditions for shift spaces which imply that the space of shift-invariant Markov cocycles is finite-dimensional?

5.1.6 Identifying Hammersley-Clifford Spaces

Suppose a finite graph $\mathcal{H}$ can be folded into a single vertex (with or without a loop) or an edge. We have proved that for any bipartite graph $\mathcal{G}$ the space $Hom(\mathcal{G}, \mathcal{H})$ is Hammersley-Clifford. We have also shown that the strong config-folds and strong config-unfolds of Hammersley-Clifford spaces are Hammersley-Clifford spaces as well. Following [6] we will call a graph $\mathcal{H}$ *stiff* if it cannot be folded anymore. Fixing a particular domain graph say $\mathbb{Z}^2$, is it possible to classify all stiff graphs $\mathcal{H}$ for which $Hom(\mathbb{Z}^2, \mathcal{H})$ is Hammersley-Clifford? What if we want $Hom(\mathcal{G}, \mathcal{H})$ to be Hammersley-Clifford for all bipartite graphs $\mathcal{G}$?

5.2 Hom-shifts, Entropy Minimality and the Pivot Property

In Chapter 4 we proved entropy minimality and the pivot property for a special class of hom-shifts viz. $X_{\mathcal{H}}$ where $\mathcal{H}$ is a finite connected four-cycle free graph.

5.2.1 Identification of Hom-Shifts

Can we determine when a shift space is conjugate to a hom-shift?

Being conjugate to a hom-shift lays many restrictions on the shift space, for instance on its periodic configurations. Consider a conjugacy $f : X \longrightarrow X_{\mathcal{H}}$ where $\mathcal{H}$ is a finite undirected graph. Let $Z \subset X_{\mathcal{H}}$ be the set of configurations invariant under $\{\sigma^{2\vec{e}_i}\}_{i=1}^{d}$. Then there is a bijection between Z and $\mathcal{L}_A(X_{\mathcal{H}})$ where A is the rectangular shape

$$A := \{\sum_{i=1}^{d} \delta_i \vec{e}_i \mid \delta_i \in \{0, 1\}\}$$

because every pattern in $\mathcal{L}_A(X_{\mathcal{H}})$ extends to a unique configuration in Z. More generally given a graph $\mathcal{H}$ it is not hard to compute the number of periodic configurations for a specific finite-index subgroup of $\mathbb{Z}^d$. Moreover periodic points are dense in these shift spaces and there are algorithms to compute approximating upper and lower bounds of their entropy [19, 30]. Thus the same holds for the shift space X as well. We are not familiar with nice decidable conditions which imply that a shift space is conjugate to a hom-shift.

5.2.2 Hom-Shifts and Strong Irreducibility

Which hom-shifts are strongly irreducible?

We know two such conditions:

1. [6] If $\mathcal{H}$ is a finite graph which folds into $\mathcal{H}'$ then $X_{\mathcal{H}}$ is strongly irreducible if and only if $X_{\mathcal{H}'}$ is strongly irreducible. This reduces the problem to graphs $\mathcal{H}$ which are stiff. For instance if $\mathcal{H}$ is dismantlable, then $X_{\mathcal{H}}$ is strongly irreducible.

2. [5] $X_{\mathcal{H}}$ is single site fillable. A shift space $X_{\mathcal{F}} \subset \mathcal{A}^{\mathbb{Z}^d}$ is said to be *single site fillable* if for all patterns $a \in \mathcal{A}^{\partial\{\vec{0}\}}$ there exists a locally allowed pattern in $X_{\mathcal{F}}$, $b \in \mathcal{A}^{D_1}$ such that $b|_{\partial\{\vec{0}\}} = a$. In case $X_{\mathcal{F}} = X_{\mathcal{H}}$ for some graph $\mathcal{H}$ then it is single site fillable if and only if given vertices $v_1, v_2, \ldots, v_{2d} \in \mathcal{H}$ there exists a vertex $v \in \mathcal{H}$ adjacent to all of them.

It follows that X_{K_5} is single site fillable and hence strongly irreducible for $d = 2$. In fact strong irreducibility has been proved in [5] for shifts of finite type under a property weaker than single site fillability called TSSM. This does not cover all possible examples. For instance it was proved in [5] that X_{K_4} is strongly irreducible for $d = 2$ even though it is not TSSM and K_4 is stiff. We do not know if it is possible to verify whether a given hom-shift is TSSM.

We remark that results about TSSM are in contrast to the results obtained in Chapter 4. It can be concluded from the results in Chapter 4 that $X_{\mathcal{H}}$ is not even block-gluing (a weaker mixing property than strong irreducibility) when $\mathcal{H}$ is four-cycle free.

5.2.3 Hom-Shifts and Entropy Minimality

Given a finite connected graph $\mathcal{H}$ when is $X_{\mathcal{H}}$ entropy minimal?

We have provided some examples in Chapter 4:

1. $\mathcal{H}$ can be folded to a single vertex with a loop or a single edge. (Proposition 4.3.4)

2. $\mathcal{H}$ is connected and four-cycle free. (Theorem 4.1.2)

This does not provide the complete picture. For instance X_{K_4} is strongly irreducible when $d = 2$ and hence entropy minimal even though K_4 is stiff and not four-cycle free. A possible approach might be via identifying the right sub-cocycle and Theorem 4.6.3.

Conjecture: Let $d = 2$ and $\mathcal{H}$ be a finite connected graph. Then $X_{\mathcal{H}}$ is entropy minimal.

5.2.4 Hom-Shifts and the Pivot Property

We had provided some examples of graphs $\mathcal{H}$ for which the shift space $X_{\mathcal{H}}$ has the pivot property in Section 2.2.2. In Chapter 4 we gave two further sets of examples:

1. $\mathcal{H}$ can be folded to a single vertex with a loop or a single edge. (Proposition 4.3.5)

2. $\mathcal{H}$ is four-cycle free. (Theorem 4.1.4)

It is not true that all hom-shifts have the pivot property. The following was observed by Brian

1	2	3	4	5
3	$\boxed{4}$	$\boxed{5}$	1	2
5	1	2	3	4
2	3	4	5	1
4	5	1	2	3

Figure 5.1: Frozen Pattern

Marcus: Recall that K_n denotes the complete graph with n vertices. X_{K_4}, X_{K_5} do not possess the pivot property if the dimension is 2. For instance consider a configuration in X_{K_5} which is obtained by tiling the plane with the pattern given in Figure 5.1. It is clear that the symbols in the box can be interchanged but no individual symbol can be changed. Therefore X_{K_5} does not have the pivot property. However both X_{K_4} and X_{K_5} satisfy a more general property:

A shift space X is said to have the *generalised pivot property* if there is an $r \in \mathbb{N}$ such that for all $(x, y) \in \Delta_X$ there exists a chain $(x^1 = x), x^2, x^3, \ldots, (y = x^n) \in X$ such that x^i and x^{i+1} differ at most on some translate of D_r.

If a shift space X satisfies this generalised property then $\mathbf{M}_X^{\mathbb{Z}^d}$ is a finite-dimensional vector space. It can be shown that that any nearest neighbour shift of finite type $X \subset \mathcal{A}^{\mathbb{Z}}$ has the generalised pivot property. In higher dimensions this is not always true; consider the subshift Y constructed in Section 2.8; since $\mathbf{M}_Y^{\mathbb{Z}^d}$ is infinite dimensional it follows that Y does not have the generalised pivot property. It is not hard to prove that any single site fillable nearest neighbour shift of finite type has the generalised pivot property. This can be generalised further: in [5] it is proven that every shift space satisfying TSSM has the generalised pivot property. The space of domino tilings forms another interesting and well known example for a subshift with the generalised pivot property [17].

For which graphs $\mathcal{H}$ does $X_{\mathcal{H}}$ satisfy the pivot property? What about the generalised pivot property?

Bibliography

[1] D. Angluin. Local and global properties in networks of processors (extended abstract). In *Proceedings of the 12th Annual ACM Symposium on Theory of Computing, April 28-30, 1980, Los Angeles, California, USA*, pages 82–93, 1980.

[2] M. B. Averincev. Description of Markov random fields by means of Gibbs conditional probabilities. *Teor. Verojatnost. i Primenen.*, 17:21–35, 1972.

[3] J. Besag. Spatial interaction and the statistical analysis of lattice systems. *J. Roy. Statist. Soc. Ser. B*, 36:192–236, 1974. With discussion by D. R. Cox, A. G. Hawkes, P. Clifford, P. Whittle, K. Ord, R. Mead, J. M. Hammersley, and M. S. Bartlett and with a reply by the author.

[4] M. Boyle, R. Pavlov, and M. Schraudner. Multidimensional sofic shifts without separation and their factors. *Trans. Amer. Math. Soc.*, 362(9):4617–4653, 2010.

[5] R. Briceño. The topological strong spatial mixing property and new conditions for pressure approximation. http://arxiv.org/abs/1411.2289, 2014.

[6] G. R. Brightwell and P. Winkler. Gibbs measures and dismantlable graphs. *J. Combin. Theory Ser. B*, 78(1):141–166, 2000.

[7] R. Burton and J. E. Steif. Non-uniqueness of measures of maximal entropy for subshifts of finite type. *Ergodic Theory Dynam. Systems*, 14(2):213–235, 1994.

[8] S. Capobianco. Multidimensional cellular automata and generalization of Fekete's lemma. *Discrete Math. Theor. Comput. Sci.*, 10(3):95–104, 2008.

[9] N. Chandgotia. Markov random fields and measures with nearest neighbour potentials. *MSc Thesis, Mathematics Department, the University of British Columbia*, 2011.

[10] N. Chandgotia. Four-cycle free graphs, the pivot property and entropy minimality. http://arxiv.org/abs/1411.4029, 2014.

[11] N. Chandgotia. Generalisation of the Hammersley-Clifford theorem on bipartite graphs. http://arxiv.org/abs/1406.1849, 2014.

[12] N. Chandgotia, G. Han, B. Marcus, T. Meyerovitch, and R. Pavlov. One-dimensional Markov random fields, Markov chains and topological Markov fields. *Proc. Amer. Math. Soc.*, 142(1):227–242, 2014.

[13] N. Chandgotia and T. Meyerovitch. Markov random fields, Markov cocycles and the 3-colored chessboard. http://arxiv.org/abs/1305.0808, 2013.

[14] E. M. Coven and J. Smítal. Entropy-minimality. *Acta Math. Univ. Comenian. (N.S.)*, 62(1):117–121, 1993.

[15] S. Dachian and B. S. Nahapetian. Description of random fields by means of one-point conditional distributions and some applications. *Markov Process. Related Fields*, 7(2):193–214, 2001.

[16] R. L. Dobrušin. Description of a random field by means of conditional probabilities and conditions for its regularity. *Teor. Verojatnost. i Primenen*, 13:201–229, 1968.

[17] N. Elkies, G. Kuperberg, M. Larsen, and J. Propp. Alternating-sign matrices and domino tilings. I. *J. Algebraic Combin.*, 1(2):111–132, 1992.

[18] G. B. Folland. *Real analysis*. Pure and Applied Mathematics (New York). John Wiley & Sons, Inc., New York, second edition, 1999. Modern techniques and their applications, A Wiley-Interscience Publication.

[19] S. Friedland. On the entropy of $\mathbf{Z}^d$ subshifts of finite type. *Linear Algebra Appl.*, 252:199–220, 1997.

[20] D. Galvin. On homomorphisms from the Hamming cube to $\mathbb{Z}$. *Israel J. Math.*, 138:189–213, 2003.

[21] D. Geiger, C. Meek, and B. Sturmfels. On the toric algebra of graphical models. *Ann. Statist.*, 34(3):1463–1492, 2006.

[22] H.-O. Georgii. *Gibbs measures and phase transitions*, volume 9 of *de Gruyter Studies in Mathematics*. Walter de Gruyter & Co., Berlin, 1988.

[23] G. R. Grimmett. A theorem about random fields. *Bull. London Math. Soc.*, 5:81–84, 1973.

[24] J. Hammersley and P. Clifford. Markov fields on finite graphs and lattices. *Unpublished notes*, 1971.

[25] J. M. Hammersley and D. J. A. Welsh. First-passage percolation, subadditive processes, stochastic networks, and generalized renewal theory. In *Proc. Internat. Res. Semin., Statist. Lab., Univ. California, Berkeley, Calif*, pages 61–110. Springer-Verlag, New York, 1965.

[26] G. Keller. *Equilibrium states in ergodic theory*, volume 42 of *London Mathematical Society Student Texts*. Cambridge University Press, Cambridge, 1998.

[27] O. E. Lanford, III and D. Ruelle. Observables at infinity and states with short range correlations in statistical mechanics. *Comm. Math. Phys.*, 13:194–215, 1969.

[28] S. L. Lauritzen. *Graphical models*, volume 17 of *Oxford Statistical Science Series*. The Clarendon Press, Oxford University Press, New York, 1996. Oxford Science Publications.

[29] D. Lind and B. Marcus. *An introduction to symbolic dynamics and coding*. Cambridge University Press, 1995, reprinted 1999.

[30] E. Louidor and B. H. Marcus. Improved lower bounds on capacities of symmetric 2D constraints using Rayleigh quotients. *IEEE Trans. Inform. Theory*, 56(4):1624–1639, 2010.

[31] W. S. Massey. *Algebraic topology: an introduction.* Springer-Verlag, New York-Heidelberg, 1977. Reprint of the 1967 edition, Graduate Texts in Mathematics, Vol. 56.

[32] R. Meester and J. E. Steif. Higher-dimensional subshifts of finite type, factor maps and measures of maximal entropy. *Pacific J. Math.*, 200(2):497–510, 2001.

[33] T. Meyerovitch. Gibbs and equilibrium measures for some families of subshifts. *Ergodic Theory Dynam. Systems*, 33(3):934–953, 2013.

[34] J. Moussouris. Gibbs and Markov random systems with constraints. *J. Statist. Phys.*, 10:11–33, 1974.

[35] J. R. Munkres. *Topology: a first course.* Prentice-Hall, Inc., Englewood Cliffs, N.J., 1975.

[36] F. Nakano, H. Ono, and T. Sadahiro. Local move connectedness of domino tilings with diagonal impurities. *Discrete Math.*, 310(13-14):1918–1931, 2010.

[37] R. Nowakowski and P. Winkler. Vertex-to-vertex pursuit in a graph. *Discrete Math.*, 43(2-3):235–239, 1983.

[38] L. Onsager. Crystal statistics. I. A two-dimensional model with an order-disorder transition. *Phys. Rev. (2)*, 65:117–149, 1944.

[39] E. Ordentlich and R. M. Roth. When data must satisfy constraints upon writing. 2014.

[40] R. Pavlov. Approximating the hard square entropy constant with probabilistic methods. *Ann. Probab.*, 40(6):2362–2399, 2012.

[41] R. Peled. High-dimensional Lipschitz functions are typically flat. http://arxiv.org/abs/1005.4636, 2010.

[42] K. Petersen and K. Schmidt. Symmetric Gibbs measures. *Trans. Amer. Math. Soc.*, 349(7):2775–2811, 1997.

[43] C. J. Preston. *Gibbs states on countable sets.* Cambridge University Press, London-New York, 1974. Cambridge Tracts in Mathematics, No. 68.

[44] A. Quas and A. A. Şahin. Entropy gaps and locally maximal entropy in $\mathbb{Z}^d$ subshifts. *Ergodic Theory Dynam. Systems*, 23(4):1227–1245, 2003.

[45] A. N. Quas and P. B. Trow. Subshifts of multi-dimensional shifts of finite type. *Ergodic Theory Dynam. Systems*, 20(3):859–874, 2000.

[46] R. M. Robinson. Undecidability and nonperiodicity for tilings of the plane. *Invent. Math.*, 12:177–209, 1971.

[47] D. Ruelle. *Thermodynamic formalism.* Cambridge Mathematical Library. Cambridge University Press, Cambridge, second edition, 2004. The mathematical structures of equilibrium statistical mechanics.

[48] K. Schmidt. The cohomology of higher-dimensional shifts of finite type. *Pacific J. Math.*, 170(1):237–269, 1995.

[49] K. Schmidt. Invariant cocycles, random tilings and the super-K and strong Markov properties. *Trans. Amer. Math. Soc.*, 349(7):2813–2825, 1997.

[50] K. Schmidt. Tilings, fundamental cocycles and fundamental groups of symbolic $\mathbb{Z}^d$-actions. *Ergodic Theory Dynam. Systems*, 18(6):1473–1525, 1998.

[51] M. Schraudner. Projectional entropy and the electrical wire shift. *Discrete Contin. Dyn. Syst.*, 26(1):333–346, 2010.

[52] M. Schraudner and S. Lightwood. Entropy minimality of Zd shifts of finite type and the family of wire shifts. Work in progress.

[53] S. Sheffield. Ribbon tilings and multidimensional height functions. *Trans. Amer. Math. Soc.*, 354(12):4789–4813 (electronic), 2002.

[54] S. Sherman. Markov random fields and Gibbs random fields. *Israel J. Math.*, 14:92–103, 1973.

[55] F. Spitzer. Markov random fields and Gibbs ensembles. *Amer. Math. Monthly*, 78:142–154, 1971.

[56] J. R. Stallings. Topology of finite graphs. *Invent. Math.*, 71(3):551–565, 1983.

[57] W. P. Thurston. Conway's tiling groups. *Amer. Math. Monthly*, 97(8):757–773, 1990.

[58] P. Walters. *An introduction to ergodic theory*, volume 79 of *Graduate Texts in Mathematics*. Springer-Verlag, New York-Berlin, 1982.