



HAL
open science

Vitesse de convergence de l'échantillonneur de Gibbs appliqué à des modèles de la physique statistique

Amine Helali

► **To cite this version:**

Amine Helali. Vitesse de convergence de l'échantillonneur de Gibbs appliqué à des modèles de la physique statistique. Analyse numérique [math.NA]. Université de Bretagne occidentale - Brest; Université de Sfax (Tunisie), 2019. Français. NNT : 2019BRES0002 . tel-02292870

HAL Id: tel-02292870

<https://theses.hal.science/tel-02292870>

Submitted on 20 Sep 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



THESE DE DOCTORAT DE

L'UNIVERSITE DE BRETAGNE OCCIDENTALE

COMUE UNIVERSITE BRETAGNE LOIRE

EN COTUTELLE AVEC
L'UNIVERSITE DE SFAX - TUNISIE
ECOLE DOCTORALE N° 601

*Mathématiques et Sciences et Technologies
de l'Information et de la Communication*
Spécialité : *Mathématiques*

Par

Amine HELALI

Vitesse de convergence de l'échantillonneur de Gibbs appliqué à des modèles de physique statistique

Thèse présentée et soutenue à Brest, le 11/01/2019

Unités de recherche : Laboratoire de Mathématiques de Bretagne Occidentale, LMBA UMR 6205

Laboratoire d'algèbre, géométrie et théorie spectrale, LAGTS

Rapporteurs avant soutenance :

Nabila TORKI HAMZA Professeur des Universités, Université de Mahdia, Tunisie
Arnaud LE NY Professeur des Universités, Université de Paris Est Créteil, France

Composition du Jury :

Président du Jury :	Arnaud LE NY	Professeur des Universités, Université de Paris Est Créteil, France
Examineurs :	Matthias LÔWE Nabila TORKI HAMZA Catherine RAINER	Chaire de probabilité, Université de Münster, Allemagne Professeur des Universités, Université de Mahdia, Tunisie Maitre de conférences, HDR, Université de Bretagne Occidentale, France
Dir. de thèse :	Brice FRANKE	Professeur des Universités, Université de Bretagne Occidentale, France
Co-dir. de thèse :	Mondher DAMAK	Professeur des Universités, Université de Sfax, Tunisie

Invité :

Afif MASMOUDI Professeur des Universités, Université de Sfax, Tunisie

Acknowledgements

It is with great pleasure that I reserve this page as a sign of deep gratitude to all those who have kindly provided the necessary support for the smooth running of this thesis.

I present my thanks to Prof. Mathias Löwe for the honor he had accorded me for agreeing to be the committee chair of my thesis. I also thank Prof. Catherine Rainer for the valuable service to examine my thesis and to be a member of the committee. My distinguished thanks go also to Prof. Nabila Turki Hamza and Prof. Arnaud Le Ny for taking their time to review my dissertation and for their relevant comments.

I would like to express my deep gratitude to my supervisors Prof. Mondher Damak and Prof. Brice Franke for their outstanding commitment to this thesis. I am also grateful for the support they gave me. Their professionalism, friendliness and pedagogical and scientific qualities have been invaluable.

I am thankful to all of my colleagues at both LMBA Laboratory (Brest-France) and LAGTS Laboratory (Sfax-Tunisia). Not only were our technical discussions very interesting, but it was wonderful working with them.

I owe thanks to the people who mean a lot to me, my mother Monjia and my father Houcine. I would not have made it this far without you raising me with a love of science and knowledge. You have always encouraged me to explore new directions in life and seek my own destiny. Thank you also to my sister Asma and my brother Mohamed, for your love and encouragement even with the thousands miles that separate us.

Lastly, a special person deserves to be thanked. Thank you my fiancée Saoussen. Your true and unconditional support was essential in completing my Ph.D. Thank you for giving me power every time I was powerless. We have together a bright future worth attaining.

Contents

Abstract	1
Résumé	1
Motivation	3
Introduction	6
Introduction	14
1 Preliminaries	22
1.1 Markov chains and some properties	22
1.2 Diaconnis and Stroock's approach	23
1.3 Ingrassia's approach	25
1.3.1 Upper bound for β_1	26
1.3.2 Lower bound for $\beta_{ \mathcal{X} -1}$	27
1.4 Shiu and Chen's result	27
1.4.1 Description of the model	27
1.4.2 Main result	28
1.5 Alison Gibbs's approach	30
1.5.1 Explanation of the approach	30
1.5.2 Main results	31
2 The convergence rate of the Gibbs sampler for generalized 1-D Ising model	32
2.1 Introduction	33
2.2 Main result	34
2.2.1 Selection of paths	34
2.2.2 Geometric bound for the second largest eigenvalue	35
2.2.3 Bound for the absolute value of the second largest eigenvalue	37
2.3 Discussion	37
2.4 Proofs of the main results	38
2.4.1 Proof of theorem 2.2.1	38
2.4.2 Proof of theorem 2.2.2	47

3	The Convergence rate of the Gibbs sampler for the 2–D Ising model via a geometric bound	49
3.1	Introduction	50
3.2	Main result	52
3.2.1	Selection of paths	52
3.2.2	Geometric bound for the second largest eigenvalue	53
3.2.3	Bound for the second largest eigenvalue in absolute value	55
3.3	Comparison	55
3.4	conclusion	57
3.5	Proof of the main result	57
3.5.1	Proof of Lemma 3.2.1	57
3.5.2	Proof of proposition 3.2.2	58
3.5.3	Proof of Theorem 3.2.3	66
3.6	Asymptotic study	67
3.6.1	For $(i, j) = \{2, \dots, n - 1\} \times \{2, \dots, n - 1\}$	67
3.6.2	For $(i, j) = \{2, \dots, n - 1\} \times \{1\}$	70
3.6.3	For $(i, j) = (1, n)$	71
4	The convergence rate of some perturbed Gibbs sampler for the 1–D Ising model	73
4.1	Introduction	74
4.2	Main result	76
4.2.1	Selection of paths	76
4.2.2	Geometric bound for the second largest eigenvalue	77
4.2.3	Summary	84
4.2.4	Bound for the second largest eigenvalue in absolute value	84
4.3	Comparison	85
4.4	Overall conclusion	86
4.5	Discussion of the convergence rate at low temperature	86
4.6	Appendix	86
	Discussion	88
	Summary	88
	Open questions	88
	Bibliography	90

Abstract

Monte Carlo Markov chain methods MCMC are mathematical tools used to simulate probability measures π defined on state spaces of high dimensions. The idea is to construct a Markov chain P easy to simulate which converges to the measure π . The speed of convergence of this Markov chain X to its invariant state π is a natural question to study in this context.

To measure the convergence rate of a Markov chain some authors use the asymptotic variance, while other authors use the total variation distance. It is well known that the convergence rate of a reversible Markov chain depends on its second largest eigenvalue in absolute value denoted by β^* . An important part in the estimation of β^* is the estimation of the second largest eigenvalue which is denoted by β_1 . Diaconis and Stroock (1991) introduced a method based on Poincaré inequality to obtain a bound for β_1 for general finite state reversible Markov chains.

Recently Shiu and Chen (2015) are able to obtain explicit value for the bound of Diaconis and Stroock in case of the 1–D Ising model with two states (spin down, spin up). This result improves the one introduced by Ingrassia (1994) and it is considered to be the best existing bound in the literature for this model.

In the second chapter of this thesis we use the Chen and Shiu approach to study the case of the Gibbs sampler for the 1–D Ising model with three and more states which is also called Potts model. Then, we generalize in the third chapter the result of Shiu and Chen to the case of the 2–D Ising model with two states. Our method is based on some generalization of the techniques introduced by Shiu and Chen (2015). The results we obtain improve the ones obtained by Ingrassia (1994).

In the last chapter of this thesis, we introduce some method to disrupt the Gibbs sampler in order to improve its convergence rate to equilibrium. From this idea we are able to improve the result introduced by Shiu and Chen (2015) at high temperatures. However, at low temperature we did not succeed to improve the geometric bound, although an improvement should be expected.

Résumé

Les méthodes de Monte Carlo par chaînes de Markov MCMC sont des outils mathématiques utilisées pour simuler des mesures de probabilités π définies sur des espaces de grandes dimensions. Le principe consiste à construire une chaîne de Markov P facile à simuler qui converge vers la mesure π . Une des questions les plus importantes dans ce contexte est de savoir à quelle vitesse converge la chaîne de Markov P vers la mesure invariante π .

Pour mesurer la vitesse de convergence de la chaîne de Markov P vers sa mesure invariante π certains auteurs utilisent la variance asymptotique alors que d'autres auteurs utilisent la distance de la variation totale. Il est bien connu que la vitesse de convergence d'une chaîne de Markov réversible P dépend de la deuxième plus grande valeur propre en valeur absolue de la matrice P notée β^* . Une partie importante dans l'estimation de β^* consiste à estimer la deuxième plus grande valeur propre de la matrice P , qui est notée β_1 . Diaconis et Stroock (1991) ont introduit une méthode basée sur l'inégalité de Poincaré pour estimer β_1 pour le cas général des chaînes de Markov réversibles avec un nombre fini d'état. Récemment Shiu et Chen (2015) ont réussi à obtenir une expression explicite de la borne de Diaconis et Stroock dans le cas d'un modèle d'Ising unidimensionnel avec deux états. Ce résultat améliore celui introduit par Ingrassia (1994) et il est considéré comme la meilleure borne existante dans la littérature pour ce modèle.

Dans le chapitre deux de cette thèse, nous utilisons la méthode de Shiu et Chen pour étudier le cas de l'algorithme de l'échantillonneur de Gibbs pour le modèle d'Ising unidimensionnel avec trois états ou plus appelé aussi modèle de Potts. Puis, nous généralisons dans le troisième chapitre le résultat de Shiu et Chen au cas du modèle d'Ising deux-dimensionnel avec deux états. Notre méthode est basée sur une généralisation des techniques introduites par Shiu et Chen (2015). Les résultats obtenus minorent ceux introduits par Ingrassia (1994).

Dans le dernier chapitre de la thèse nous avons pensé à perturber l'échantillonneur de Gibbs afin d'améliorer sa vitesse de convergence vers l'équilibre. En partant de cette idée, nous avons réussi à améliorer le résultat introduit par Shiu et Chen (2015) à haute température. Cependant, à basse température, nous n'avons pas réussi à améliorer la borne géométrique, même si une amélioration est à prévoir.

Motivation

Image processing and dynamics of magnetic materials are studied by means of similar theoretical structures. This similarity follows from a common underlying situation. They are both constructed from many elements which interact between each others. On one hand, the classical spin system has a lot of spinning particles and each spin interacts with its neighboring spins. The ordered state is determined from interactions and external fields. On the other hand, images are constructed from a lot of pixels. Output of each pixel is determined from the values of the neighboring pixels. The resulting images are determined from a priori information and given data. As a result, the conventional image processing theory and the classical spin system have some common structure which in both cases can be described by the Ising model.

The Ising model is a model from statistical mechanics, which is used to model different phenomena in which collective effects are produced by local interactions between neighboring spins on the sites of a lattice. The measure π which represents the probability of occurrence of a given configuration x of spins in the space of all possible configurations χ is given by:

$$\pi(x) = \frac{1}{Z_T} \exp \left\{ \frac{1}{T} H(x) \right\},$$

where

- $T > 0$ is the temperature of the system,
- Z_T is the normalizing constant,
- $H : \chi \rightarrow \mathbb{R}$ is the Hamiltonian function.

In the case of configurations with spins $x \in \{-1, 1\}^N$, which describe spins in a solid, the Hamiltonian H has the form:

$$H(x) = \sum_{i \sim j} x_i x_j - \sum_i \Phi(x_i)$$

where

- $\sum_{i \sim j} x_i x_j$ introduces dependence between neighboring spins favoring equal spins,
- $\sum_i \Phi(x_i)$ introduces an external force field (magnetic field).

One major task in statistical mechanics is to estimate the normalizing constant Z_T . In image analysis, let $y \in \chi$ be a data sample representing the colors of the pixels of a given image. The task is to clean up the image from transmission errors. One wants to assign a likelihood to each configuration $x \in \chi$ according to its conformity with the data y and some further quality requirement.

The maximum likelihood estimator is then the configuration $x \in \chi$ which maximizes

$$\pi(x) = \frac{1}{Z_T} e^{\frac{1}{T} H(x)}$$

with a suitably chosen hamiltonian function. In a two color situation $\chi = \{-1, 1\}^N$ the Hamiltonian may take the form:

$$H(x) = \sum_{i \sim j} x_i x_j - c \sum_{i \sim j} x_i y_j - h \sum_{i \sim j} x_i,$$

where

- $\sum_{i \sim j} x_i x_j$ gives a penalty to configurations with too rough behavior,
- $c \sum_{i \sim j} x_i y_j$ ensures proximity of the configuration x to the data sample y ,
- $h \sum_{i \sim j} x_i$ drives towards one of the colors black or white.

To solve the inverse problem to find the maximum of the function H one uses the fact that this maximum coincides with the maximum of π for any choice of T .

If T is chosen sufficiently small the the probability law π concentrates around its maximum x_0 . Direct sampling is impossible in the case of image analysis because the set χ is too large and since the partition function is computationally intractable. In this context the idea of Monte Carlo Markov chain method, MCMC was introduced in statistical mechanics to numerically integrate the normalizing constant Z_T . This method consists in constructing a Markov chain which converges to the measure π we want to simulate.

Thus a Markov chain which has π as its invariant law should stay a long proportion of time in the neighborhood of x_0 after time τ . One can then use a large number of independent realizations $X^{(1)}, \dots, X^{(n)}$ of the Markov chain and the law of large numbers to compute:

$$\begin{aligned} \mathbb{E}_\pi [f] &= \lim_{\tau \rightarrow \infty} \mathbb{E}_\pi [f(X_\tau)] \\ &= \lim_{\tau \rightarrow \infty} \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n f(X_\tau^{(i)}). \end{aligned}$$

A natural question in this context is to study the distance between $\mathbb{E}_\pi [f(X_\tau)] = P^\tau f$ and $\mathbb{E}_\pi [f]$. The convergence rate of the Markov chain to its equilibrium depends on the number of sites n . There are many approaches to measure the convergence rate of the resulting Markov chain to its equilibrium π .

Frigessi and his collaborators proved in [9] that the convergence rate of the MCMC method for the 1-D Ising model (without phase transition) is $O(n \log(n))$. But for example in the

case of 2–D Ising model or the Curie Weiss model the convergence rate is $O(e^n)$ at low temperatures ($T < T_c$) and it is of order $O(n \log(n))$ for $T > T_c$ where T_c is the critical temperature and n is the size of the lattice. In these results the constant of proportionality isn't known which means that we can't compute a precise bound for the convergence time. In the following we will study existing results that include proportionality constants. Diaconis and Stroock (see [5]) considered the total variation distance and prove that the rate of convergence of some reversible Markov chain on a graph (Γ, E) which is associated to some transition matrix P to its equilibrium π can be controlled as follows:

$$4\|P^k(x, \cdot) - \pi\|_{var}^2 \leq \frac{1 - \pi(x)}{\pi(x)} (\beta^*)^{2k},$$

where $\beta^* = \max \left\{ |\beta_i|; \beta_i \in \text{Spec}(P) \setminus \{1\} \right\}$.

Later, they use Poincaré inequality to bound the second largest eigenvalue of the matrix P as follows:

$$\beta_1 \leq 1 - \frac{1}{\kappa}$$

where

$$\kappa = \max_{e \in E} Q(e)^{-1} \sum_{\gamma_{xy} \ni e} |\gamma_{xy}| \pi(x) \pi(y).$$

Here E is the set of all edges e linking two neighbors in the graph and $|\gamma_{xy}|$ designates the length of the path γ_{xy} .

Shiu and Chen (see [25]) study the case of the Gibbs sampler for the 1 – D Ising model where the Hamiltonian function H for a configuration $x \in \chi = \{\pm 1\}^n$ is:

$$H(x) = \sum_{i=1}^{n-1} x_i x_{i+1}.$$

They use some symmetry argument in order to compute the Diaconis and Stroock bound for β_1 explicitly in this case. They obtain the smallest bound in the literature for this model. A natural question is whether one can use this technique in the case of other statistical mechanics models. In this thesis we will study this question for the case of the 1–D Potts model, where for a configuration $x \in \chi = \{c_1, \dots, c_N\}^n$ the Hamiltonian is:

$$H(x) = \sum_{k=1}^{n-1} (\mathbb{1}_{\{x_k = x_{k+1}\}} - \mathbb{1}_{\{x_k \neq x_{k+1}\}})$$

and for the 2–D Ising model where for $x \in \chi = \{\pm 1\}^{n^2}$ the Hamiltonian is:

$$H(x) = \sum_{j=1}^n \sum_{i=1}^{n-1} x_i^j x_{i+1}^j + \sum_{i=1}^n \sum_{j=1}^{n-1} x_i^j x_i^{j+1}.$$

Introduction

The generalized Ising model which is also called Potts model is a spin model that can be defined on any graph. An atom is assumed to lie on each vertex of the graph and to have a spin of either $x_i = c_1, \dots, c_N$ where N is the number of orientations that each site can take; the case when $N = 2$ is called the Ising model with spins equal to $+1$ or -1 . Atoms are allowed to interact only with their nearest neighbors on the graph as well as with an external magnetic field. The Hamiltonian function H is defined on the state space $\chi = \{c_1, \dots, c_N\}^\Lambda$ as follows:

$$\begin{aligned} H : \chi &\longrightarrow \mathbb{R} \\ x &\longrightarrow \sum_{i \sim j} (\mathbb{1}_{\{x_i = x_j\}} - \mathbb{1}_{\{x_i \neq x_j\}}) \end{aligned}$$

where $i \sim j$ means that i and j are neighbors and Λ is the size of the lattice. This Hamiltonian describes a situation where no external field is assumed.

We define a probability measure on the space of all possible configurations depending on the configurations energy level. According to this the probability of occurrence of a given configuration $x \in \chi$ is:

$$\pi(x) = \frac{1}{Z_T} \exp \left\{ \frac{1}{T} H(x) \right\}, \quad (1)$$

where

- T is the temperature;
- $Z_T = \sum_{x \in \chi} \exp \left\{ \frac{1}{T} H(x) \right\}$ is the normalizing constant.

Many relevant physical quantities are integrals of certain functionals with respect to the distribution π . However this integration is difficult to do explicitly due to the high dimensionality of the state space χ .

One way to solve this problem is to simulate the probability measure π through a Markov chain which distributes according to the measure π and which is easy to simulate. This method is called Monte Carlo Markov Chain method (MCMC). MCMC method is a very useful technique to draw samples from the Ising model. The Gibbs sampler introduced by Geman and Geman (see [11]) and the Metropolis-Hastings algorithm introduced by Metropolis et al. (see [21]) and Hastings (see [13]) are the most popular Monte Carlo Markov chain methods used in this context. Those two algorithms use an aperiodic

Introduction

and irreducible Markov chain which is reversible with respect to the measure π . The reversibility follows from the following *detailed balance equation*:

$$Q(x, y) := \pi(x)P(x, y) = \pi(y)P(y, x) := Q(y, x) \quad \forall x, y \in \chi.$$

Under the previous conditions (aperiodicity, irreducibility and reversibility) the measure π is the unique invariant measure for the matrix P . The eigenvalues of P can be arranged as follows:

$$1 = \beta_0 > \beta_1 \geq \beta_2 \geq \dots \geq \beta_{|\chi|-1} > -1.$$

There are several criteria to measure the convergence rate of a Markov chain. For example the asymptotic variance as in Sinclair (see [26]), Frigessi et al. (see [10]), Ingrassia (see [20]), Chen et al. (see [3]) and Chen and Hwang (see [4]) and the total variation distance as in Diaconis and Stroock (see [5]). Diaconis and Stroock proof the following inequality:

$$4\|P^k(x, \cdot) - \pi\|_{var}^2 \leq \frac{1 - \pi(x)}{\pi(x)} (\beta^*)^{2k}$$

where $\beta^* = \max\{\beta_1, |\beta_{|\chi|-1}|\}$.

Many authors as Sinclair and Jerrum (see [27]), Diaconis and Stroock (see [5]) and Sinclair (see [26]) introduced bounds for the eigenvalues β_1 and $\beta_{|\chi|-1}$. Ingrassia (see [20]) shows that the result of Diaconis and Stroock (see [5]) leads to the tightest bound. This motivates us to have a closer look at the explicit bounds that one can obtain from the approach of Diaconis and Stroock.

In order to do so, let us first remind the upper bound for β_1 that was obtained in Diaconis and Stroock (see [5]) for finite state reversible Markov chains with transition matrix P . Let $\mathcal{G}(P) = (\chi, E)$ be the graph constructed from the transition matrix P , where the state space χ is the vertex set and $E = \{(x, y) \mid P(x, y) > 0\}$ is the set of edges. Then, for each pair of distinct configurations x and y we choose a path γ_{xy} in $\mathcal{G}(P)$ linking x to y . The irreducibility of the matrix P guarantees that such paths exist. Finally, we define the set $\Gamma = \{\gamma_{xy} : x, y \in \chi\}$. Note that only one path γ_{xy} for each pair of configurations x, y is chosen. According to Diaconis and Stroock the second largest eigenvalue β_1 is then bounded from above as follows:

$$\beta_1 \leq 1 - \frac{1}{\kappa} \tag{2}$$

where

$$\kappa = \max_{e \in E} Q(e)^{-1} \sum_{\gamma_{xy} \ni e} |\gamma_{xy}| \pi(x) \pi(y), \tag{3}$$

and where $|\gamma_{xy}|$ designates the length of the path γ_{xy} .

In equation (3) the maximum is over all directed edges in the graph $\mathcal{G}(P)$ and the sum is over all paths from the set Γ passing through the fixed edge e . It is clear that κ measures the bottlenecks (charged edges) in the graph $\mathcal{G}(P)$. We notice that a small κ gives a better result. So, in order to obtain the tightest bound for β_1 on one hand we should choose the shortest path to link some pair of configurations (x, y) and on the other hand we must avoid that many paths pass through the same edge.

In this thesis we consider the situation, where a probability measure π is fixed on the

configurations over a finite set. We study the Gibbs sampler which chooses a random coordinate to be updated according to the conditional probabilities given the other coordinates.

The associated matrix for the Ising model is defined as follows:

$$P(x, y) = \begin{cases} \frac{1}{\Lambda} \pi(y_i|x) & \text{if } x \sim y \\ 1 - \frac{1}{\Lambda} \sum_{i=1}^{\Lambda} \pi(y_i|x) & \text{if } x=y \\ 0 & \text{else} \end{cases}$$

where

$$\pi(y_i|x) = \frac{\pi(x_1, \dots, x_{i-1}, y_i, x_{i+1}, \dots, x_{\Lambda})}{\sum_{y_i \in \{c_1, \dots, c_N\}} \pi(x_1, \dots, x_{i-1}, y_i, x_{i+1}, \dots, x_{\Lambda})}$$

is the probability to move from a given configuration x to its nearest neighbor y . Note that x and y differ only in one coordinate.

To be able to use the result of Diaconis and Stroock [5] to calculate the bound of the second largest eigenvalue, we should fix a collection of paths. For a pair of configurations (x, y) let d_1, \dots, d_m be an increasing sequence such that:

- $x_i \neq y_i$ for $i \in \{d_1, \dots, d_m\}$;
- $x_i = y_i$ otherwise.

We use the technique introduced by Shiu and Chen in [25] to define a path from x to y as follows:

$$\begin{aligned} (x_1, \dots, x_n) &= (y_1, \dots, y_{d_1-1}, x_{d_1}, x_{d_1+1}, \dots, x_{d_2-1}, x_{d_2}, x_{d_2+1}, \dots, x_{\Lambda}) \\ &\rightarrow (y_1, \dots, y_{d_1-1}, y_{d_1}, x_{d_1+1}, \dots, x_{d_2-1}, x_{d_2}, x_{d_2+1}, \dots, x_{\Lambda}) \\ &= (y_1, \dots, y_{d_1-1}, y_{d_1}, y_{d_1+1}, \dots, y_{d_2-1}, x_{d_2}, x_{d_2+1}, \dots, x_{\Lambda}) \\ &\rightarrow (y_1, \dots, y_{d_1-1}, y_{d_1}, y_{d_1+1}, \dots, y_{d_2-1}, y_{d_2}, x_{d_2+1}, \dots, x_{\Lambda}) \\ &\vdots \\ &\rightarrow (y_1, \dots, y_{\Lambda}). \end{aligned}$$

After fixing the paths connecting each pair of configurations (x, y) in the state space χ we turn to the computation of κ defined in (3). We begin by introducing the result of Shiu and Chen (see [25]) who study the case of the Gibbs sampler of the 1-D Ising model with two states. The measure π given in (1) becomes:

$$\pi(x) = \frac{1}{Z_T} \exp \left\{ \frac{1}{T} \sum_{k=1}^{n-1} x_k x_{k+1} \right\} \quad \forall x = (x_1, \dots, x_n) \in \chi,$$

where

- $\chi = \{-1, 1\}^n$ is the state space,

Introduction

- T is the temperature,
- $Z_T = \sum_{x \in \chi} \exp \left\{ \frac{1}{T} \sum_{k=1}^{n-1} x_k x_{k+1} \right\}$ is the normalizing constant.

The probability to move from x to its nearest neighbor y is:

$$\pi(y_i | x) = \frac{\pi(x_1, \dots, x_{i-1}, y_i, x_{i+1}, \dots, x_n)}{\pi(x_1, \dots, x_{i-1}, y_i, x_{i+1}, \dots, x_n) + \pi(x_1, \dots, x_{i-1}, -y_i, x_{i+1}, \dots, x_n)}.$$

Shiu and Chen (see [25]) give the following result:

Theorem 0.0.1. [Shiu and Chen 2015] *The second largest eigenvalue of the Gibbs sampler on the 1-D Ising model satisfies:*

$$\beta_1 \leq 1 - e^{-\frac{2}{T}n^{-2}}.$$

We turn now to present the results obtained in this thesis. In the second chapter we are interested in the study of the Gibbs sampler for the 1-D Ising model with three states. The measure π given in (1) becomes:

$$\pi(x) = \frac{1}{Z_T} \exp \left\{ \frac{1}{T} \sum_{k=1}^{n-1} (\mathbb{1}_{\{x_k=x_{k+1}\}} - \mathbb{1}_{\{x_k \neq x_{k+1}\}}) \right\} \quad \forall x = (x_1, \dots, x_n) \in \chi,$$

where

- $\chi = \{c_1, c_2, c_3\}^n$ is the state space,
- T is the temperature,
- $Z_T = \sum_{x \in \chi} \exp \left\{ \frac{1}{T} \sum_{k=1}^{n-1} (\mathbb{1}_{\{x_k=x_{k+1}\}} - \mathbb{1}_{\{x_k \neq x_{k+1}\}}) \right\}$ is the normalizing constant.

The probability to move between two neighbors of configurations x and y which differ by only one site is:

$$\pi(y_i | x) = \frac{\pi(x_1, \dots, x_{i-1}, y_i, x_{i+1}, \dots, x_n)}{\sum_{l=1}^3 \pi(x_1, \dots, x_{i-1}, c_l, x_{i+1}, \dots, x_n)}.$$

Then the first result of this thesis is:

Theorem 0.0.2. [Helali 2018] *The second largest eigenvalue of the Gibbs sampler for the one-dimensional Ising model with three states satisfies :*

$$\beta_1 < 1 - 3 \times n^{-2} \frac{e^{-\frac{4}{T}}}{1 + 2e^{-\frac{4}{T}}}. \quad (4)$$

After that, we think to generalize the previous result to the case of the 1–D Ising model with multiple colors called also Potts model. The measure π given in (1) can be written as:

$$\pi(x) = \frac{1}{Z_T} \exp \left\{ \frac{1}{T} \sum_{k=1}^{n-1} (\mathbb{1}_{\{x_k=x_{k+1}\}} - \mathbb{1}_{\{x_k \neq x_{k+1}\}}) \right\} \quad \forall x = (x_1, \dots, x_n) \in \chi ,$$

where

- $\chi = \{c_1, \dots, c_N\}^n$ is the state space where N is the number of colors,
- T is the temperature,
- $Z_T = \sum_{x \in \chi} \exp \left\{ \frac{1}{T} \sum_{k=1}^{n-1} (\mathbb{1}_{\{x_k=x_{k+1}\}} - \mathbb{1}_{\{x_k \neq x_{k+1}\}}) \right\}$ is the normalizing constant.

Similarly, the conditional probability $\pi(y_i|x)$ is:

$$\pi(y_i|x) = \frac{\pi(x_1, \dots, x_{i-1}, y_i, x_{i+1}, \dots, x_n)}{\sum_{l=1}^N \pi(x_1, \dots, x_{i-1}, c_l, x_{i+1}, \dots, x_n)}.$$

Then the second result of the second chapter is:

Theorem 0.0.3. [Helali 2018] *The second largest eigenvalue of the Gibbs sampler for the one-dimensional Ising model with N states satisfies:*

$$\beta_1 < 1 - N \times n^{-2} \frac{e^{-\frac{4}{T}}}{1 + (N-1)e^{-\frac{4}{T}}}. \quad (5)$$

A theorem proved by Ingrassia (see [20], Theorem 5.3) gives the following lower bound for the smallest eigenvalue:

$$\beta_{|\chi|-1} \geq -1 + \frac{2}{1 + (c-1)e^{\frac{\Delta}{T}}}$$

where c is the number of orientations that a spin can have and Δ is the size of maximal discontinuity of the Hamiltonian.

The application of this theorem to the case of the 1–D Ising model with three states and more gives an upper bound for $|\beta_{|\chi|-1}|$ which is smaller than the bound we obtain for β_1 . Then the upper bound of β^* is given in the following corollary:

Corollary 0.0.4. *The upper bounds for β_1 given in theorem 0.0.2 and theorem 0.0.3 are also upper bounds for the absolute value of all eigenvalues $\{\beta_1, \dots, \beta_{|\chi|-1}\}$ of the Gibbs sampler for the one-dimensional Ising model with three states and more respectively.*

In the third chapter we study the Gibbs sampler for the 2–D Ising model with two states. The distribution of the two-dimensional square lattice Ising model is:

$$\pi(x) = \frac{1}{Z_T} \exp \left\{ \frac{1}{T} \left(\sum_{j=1}^n \sum_{i=1}^{n-1} x_i^j x_{i+1}^j + \sum_{i=1}^n \sum_{j=1}^{n-1} x_i^j x_i^{j+1} \right) \right\} \quad \forall x = (x_i^j)_{1 \leq i, j \leq n} \in \chi ,$$

where

Introduction

- $\chi = \{-1, 1\}^{n^2}$ is the state space,
- T is a positive real number representing the temperature,
- $Z_T = \sum_{x \in \chi} \exp \left\{ \frac{1}{T} \left(\sum_{j=1}^n \sum_{i=1}^{n-1} x_i^j x_{i+1}^j + \sum_{i=1}^n \sum_{j=1}^{n-1} x_i^j x_i^{j+1} \right) \right\}$ is the normalizing constant.

The transition between two neighbors x and y is:

$$\pi(y_i^j | x) = \frac{\pi(x_1^1, \dots, x_{i-1}^j, y_i^j, x_{i+1}^j, \dots, x_n^n)}{\pi(x_1^1, \dots, x_{i-1}^j, y_i^j, x_{i+1}^j, \dots, x_n^n) + \pi(x_1^1, \dots, x_{i-1}^j, -y_i^j, x_{i+1}^j, \dots, x_n^n)}.$$

Then we have the main result of the third chapter:

Theorem 0.0.5. [Franke and Helali 2018] *The second largest eigenvalue of the Gibbs sampler for the two-dimensional Ising model with two states satisfies:*

$$\beta_1 \leq 1 - n^{-4} \exp \left\{ -\frac{2}{T} (2n + 1) \right\}.$$

Due to the upper bound for $|\beta_{|\chi|-1}|$ presented in Theorem 5.3 by Ingrassia in [20] we obtain the following bound for β^* :

Corollary 0.0.6. *The second largest eigenvalue in absolute value of the Gibbs sampler for the two-dimensional Ising model with two states satisfies:*

$$\beta^* \leq 1 - n^{-4} \exp \left\{ -\frac{2}{T} (2n + 1) \right\}.$$

In the fourth chapter we study the case of some perturbed Gibbs sampler for the 1-D Ising model with two states. The distribution of the 1-D Ising model is:

$$\pi(x) = \frac{1}{Z_T} \exp \left\{ \frac{1}{T} \sum_{k=0}^{n-1} x_k x_{k+1} \right\} \quad \forall x = (x_1, \dots, x_n) \in \chi,$$

where

- $\chi = \{-1, 1\}^n$ is the state space,
- T is the temperature,
- $Z_T = \sum_{x \in \chi} \exp \left\{ \frac{1}{T} \sum_{k=1}^{n-1} x_k x_{k+1} \right\}$ is the normalizing constant.

One expects an acceleration of the convergence rate of the Gibbs sampler to its equilibrium once one adds some edges to the graph underlying the initial Markov chain. However only few edges should be added in order to keep the computational cost low.

Starting from this philosophy, we propose some perturbation of the Gibbs algorithm based on some permutation σ of χ .

Let σ be a two cyclic permutation defined for $x \in \chi$ as follows:

$$\begin{aligned} \sigma : \chi &\longrightarrow \chi \\ x &\longrightarrow -x. \end{aligned}$$

For some $0 < \alpha < 1$ we denote by P_σ the matrix which is defined for $x, y \in \chi$ as:

$$P_\sigma(x, y) = \begin{cases} \alpha & \text{if } y = \sigma(x) \\ 1 - \alpha & \text{if } x = y \\ 0 & \text{else.} \end{cases}$$

The idea behind this construction is that the permutation we want to introduce should not add many edges to the graph underlying the Markov chain of the sampler. Further, the fact that σ is a two-cyclic permutation guarantees the reversibility of the Markov chain which allows us thereafter to use the geometric bound introduced by Diaconis and Stroock (1991) in [5]. Furthermore, the computational effort to run a simulation for this simple permutation is relatively cheap since we add just one further option at each transition. For $0 < \delta < 1$, we introduce the perturbed Gibbs algorithm \tilde{P} as follows:

$$\tilde{P}(x, y) = \delta P(x, y) + (1 - \delta) P_\sigma(x, y) \quad \forall x, y \in \chi.$$

Remark 0.0.7. We remark that $\tilde{P}(x, y)$ is reversible as a convex combination of two reversible matrices. Their eigenvalues can be arranged as follows:

$$1 = \tilde{\beta}_0 > \tilde{\beta}_1 \geq \tilde{\beta}_2 \geq \dots \geq \tilde{\beta}_{|\chi|-1} > -1.$$

To be able to use the result of Diaconis and Stroock [5] and to compute the bound for the second largest eigenvalue, we should fix a collection of paths linking any two configurations x and y from χ .

For a given pair $(x, y) \in \chi$, let d_1, \dots, d_m be an increasing sequence satisfying that $x_i \neq y_i$ for $i \in \{d_1, \dots, d_m\}$ and $x_i = y_i$ otherwise.

The choice of the paths linking configurations will depend on the distance between x and y .

1. If $m \leq \frac{n}{2}$ we use the path introduced by Shiu and Chen in [25] then:

$$\begin{aligned} (x_1, \dots, x_n) &= (y_1, \dots, y_{d_1-1}, x_{d_1}, x_{d_1+1}, \dots, x_{d_2-1}, x_{d_2}, x_{d_2+1}, \dots, x_n) \\ &\rightarrow (y_1, \dots, y_{d_1-1}, y_{d_1}, x_{d_1+1}, \dots, x_{d_2-1}, x_{d_2}, x_{d_2+1}, \dots, x_n) \\ &= (y_1, \dots, y_{d_1-1}, y_{d_1}, y_{d_1+1}, \dots, y_{d_2-1}, x_{d_2}, x_{d_2+1}, \dots, x_n) \\ &\rightarrow (y_1, \dots, y_{d_1-1}, y_{d_1}, y_{d_1+1}, \dots, y_{d_2-1}, y_{d_2}, x_{d_2+1}, \dots, x_n) \\ &\vdots \\ &\rightarrow (y_1, \dots, y_n). \end{aligned}$$

2. If $m > \frac{n}{2}$, as the configurations differ by more than $\frac{n}{2}$ sites we change to $-x$ in the first step and we continue with the path introduced above:

$$\begin{aligned}
(x_1, \dots, x_n) &= (y_1, \dots, y_{d_1-1}, x_{d_1}, x_{d_1+1}, \dots, x_{d_2-1}, x_{d_2}, x_{d_2+1}, \dots, x_n) \\
&\rightarrow (-y_1, \dots, -y_{d_1-1}, -x_{d_1}, -x_{d_1+1}, \dots, -x_{d_2-1}, -x_{d_2}, -x_{d_2+1}, \dots, -x_n) \\
&= (-y_1, \dots, -y_{d_1-1}, y_{d_1}, -x_{d_1+1}, \dots, -x_{d_2-1}, y_{d_2}, -x_{d_2+1}, \dots, -x_n) \\
&\rightarrow (y_1, \dots, -y_{d_1-1}, y_{d_1}, -x_{d_1+1}, \dots, -x_{d_2-1}, y_{d_2}, -x_{d_2+1}, \dots, -x_n) \\
&\vdots \\
&\rightarrow (y_1, \dots, y_n).
\end{aligned}$$

Then the principle result of the fourth chapter for a suitable choice of α and δ is:

Theorem 0.0.8. [Franke and Helali 2018] *The second largest eigenvalue of the perturbed Gibbs sampler for the 1-D Ising model satisfies at high temperature:*

$$\beta_1 \leq 1 - 2n^{-2} e^{\frac{-2}{T}}.$$

Introduction

Le modèle d'Ising généralisé qui s'appelle aussi le modèle de Potts est un modèle constitué par des spins avec des interactions locales. Il peut être défini dans un graphe. Un atome est supposé se trouver sur chaque sommet du graphe et avoir des spins $x_i = c_1, \dots, c_N$ où N est le nombre des orientations que peut prendre chaque site; Le cas $N = 2$ est appelé le modèle d'Ising avec des spins qui prennent les valeurs $+1$ ou -1 . Les atomes sont supposés d'interagir seulement avec leurs proches voisins dans le graphe et avec un champ magnétique externe. La fonction Hamiltonienne H est définie sur l'espace des états $\chi = \{c_1, \dots, c_N\}^\Lambda$ comme suit:

$$\begin{aligned} H : \chi &\longrightarrow \mathbb{R} \\ x &\longrightarrow \sum_{i \sim j} (\mathbb{1}_{\{x_i = x_j\}} - \mathbb{1}_{\{x_i \neq x_j\}}) \end{aligned}$$

avec $i \sim j$ signifie que i et j sont des voisins et Λ est la taille du réseau. Avec cet Hamiltonien, aucun champ extérieur n'est supposé.

On définit la mesure de probabilité sur l'espace de toutes les configurations possibles. La probabilité dépend du niveau d'énergie de chaque configuration. D'après cela, la probabilité qu'une configuration $x \in \chi$ apparaît est:

$$\pi(x) = \frac{1}{Z_T} \exp \left\{ \frac{1}{T} H(x) \right\} \quad (6)$$

avec

- T est la température;
- $Z_T = \sum_{x \in \chi} \exp \left\{ \frac{1}{T} H(x) \right\}$ est la constante de normalisation.

De nombreuses quantités physiques pertinentes sont des intégrales de certaines fonctionnelles par rapport à la loi π . Cependant, cette intégration est difficile à réaliser explicitement en raison de la forte dimensionnalité de l'espace des états.

Une manière pour résoudre ce problème consiste à simuler la mesure de probabilité π à l'aide d'une chaîne de Markov qui se répartit selon la mesure π et qui est facile à simuler. Cette méthode est appelée la méthode de Monte Carlo par Chaîne de Markov (MCMC). La méthode MCMC est une technique pour échantillonner à partir du modèle d'Ising. L'échantillonneur de Gibbs introduit par Geman et Geman (voir [11]) et l'algorithme de Metropolis-Hastings introduit par Metropolis et ses collaborateurs (voir [21]) et Hastings (voir [13]) sont les méthodes MCMC les plus populaires, utilisées dans ce contexte. Ces deux algorithmes utilisent une chaîne de Markov apériodique et irréductible qui est

Introduction

réversible par rapport à la mesure π . La réversibilité est contenue dans l'équation suivante appelé *detailed balance equation*:

$$Q(x, y) := \pi(x)P(x, y) = \pi(y)P(y, x) := Q(y, x) \quad \forall x, y \in \chi.$$

Sous les conditions précédentes (apériodicité, irréductibilité et réversibilité) la mesure π est l'unique mesure invariante pour la matrice P . Les valeurs propres de la matrice P peuvent être arrangées comme suit:

$$1 = \beta_0 > \beta_1 \geq \beta_2 \geq \dots \geq \beta_{|\chi|-1} > -1.$$

Il existe plusieurs critères pour mesurer la vitesse de convergence d'une chaîne de Markov vers sa mesure d'équilibre π . Par exemple la variance asymptotique comme dans Sinclair (voir [26]), Frigessi et ses collaborateurs (voir [10]), Ingrassia (voir [20]), Chen et ses collaborateurs (voir [3]) et Chen et Hwang (voir [4]) et la distance de la variation totale comme dans Diaconis et Stroock (voir [5]). Diaconis and Stroock ont montré l'inégalité suivante:

$$4\|P^k(x, \cdot) - \pi\|_{var}^2 \leq \frac{1 - \pi(x)}{\pi(x)} (\beta^*)^{2k}$$

avec $\beta^* = \max\{\beta_1, |\beta_{|\chi|-1}|\}$.

Plusieurs auteurs comme Sinclair et Jerrum (voir [27]), Diaconis et Stroock (voir [5]) et Sinclair (voir [26]) introduisent des bornes pour les valeurs propres β_1 et $\beta_{|\chi|-1}$. Ingrassia (voir [20]) montre que le résultat de Diaconis et Stroock (voir [5]) conduit à la borne la plus étroite. Ceci nous motive à examiner de plus près les bornes explicites que l'on peut obtenir de l'approche de Diaconis et Stroock.

Pour ce faire, rappelons d'abord la borne supérieure de β_1 établie par Diaconis et Stroock (voir [5]) pour les chaînes de Markov réversible à nombre fini d'états dont la matrice de transition est P .

Soit $\mathcal{G}(P) = (\chi, E)$ le graphe construit à partir de la matrice de transition P où l'espace des états χ est l'ensemble des sommets et $E = \{(x, y) \mid P(x, y) > 0\}$ est l'ensemble des arêtes. Ensuite, pour chaque paire de configurations distinctes x et y on choisit une trajectoire γ_{xy} dans $\mathcal{G}(P)$ qui lie x à y . L'irréductibilité de la matrice P garantit que chaque trajectoire existe. Finalement, on définit l'ensemble $\Gamma = \{\gamma_{xy} : x, y \in \chi\}$. Notons qu'une seule trajectoire γ_{xy} pour chaque paire de configurations x, y est choisie. La deuxième plus grande valeur propre β_1 est majorée comme suit:

$$\beta_1 \leq 1 - \frac{1}{\kappa} \tag{7}$$

avec

$$\kappa = \max_{e \in E} Q(e)^{-1} \sum_{\gamma_{xy} \ni e} |\gamma_{xy}| \pi(x) \pi(y), \tag{8}$$

et $|\gamma_{xy}|$ désigne la longueur de la trajectoire γ_{xy} (c'est à dire le nombre de sites qui diffèrent entre les configurations x et y).

Dans (8) le maximum est sur toutes les arêtes orientées dans le graphe $\mathcal{G}(P)$ et la somme est sur toutes les trajectoires dans Γ qui passent par l'arête e fixée. Il est clair que κ mesure les goulots (arêtes chargées) dans le graphe $\mathcal{G}(P)$. On remarque qu'un κ petit

donne un meilleur résultat. Donc, afin d'obtenir la meilleure borne supérieure de β_1 , et d'une part, nous devrions choisir les trajectoires les plus courtes pour relier une paire de configurations (x, y) et d'autre part, il faut éviter que de nombreuses trajectoires passent par la même arête .

Dans cette thèse, nous considérons la situation où la mesure de probabilité π est définie sur les configurations sur un ensemble fini. Nous étudions l'algorithme de l'échantillonneur de Gibbs qui choisit aléatoirement un site à changer en fonction des probabilités conditionnelles, compte tenu des autres sites. La matrice associée à l'algorithme de l'échantillonneur de Gibbs pour le modèle d'Ising généralisé est:

$$P(x, y) = \begin{cases} \frac{1}{\Lambda} \pi(y_i|x) & \text{si } x \sim y \\ 1 - \frac{1}{\Lambda} \sum_{i=1}^{\Lambda} \pi(y_i|x) & \text{si } x=y \\ 0 & \text{sinon} \end{cases}$$

où

$$\pi(y_i|x) = \frac{\pi(x_1, \dots, x_{i-1}, y_i, x_{i+1}, \dots, x_{\Lambda})}{\sum_{y_i \in \{c_1, \dots, c_N\}} \pi(x_1, \dots, x_{i-1}, y_i, x_{i+1}, \dots, x_{\Lambda})}$$

est la probabilité de se déplacer d'une configuration x vers sa proche voisine y . Notons que les configurations x et y diffèrent par un seul site.

Pour pouvoir utiliser le résultat de Diaconis et Stroock [5] afin de calculer la borne supérieure de la deuxième plus grande valeur propre, nous devrions fixer une collection de trajectoires. Pour une paire de configurations (x, y) nous sélectionons l'une parmi les trajectoires les plus courtes. Soit d_1, \dots, d_m une suite croissante tel que:

- $x_i \neq y_i$ si $i \in \{d_1, \dots, d_m\}$;
- $x_i = y_i$ sinon.

Nous utilisons la technique introduite par Shiu et Chen dans [25] pour définir une trajectoire de x à y comme suit:

$$\begin{aligned} (x_1, \dots, x_n) &= (y_1, \dots, y_{d_1-1}, x_{d_1}, x_{d_1+1}, \dots, x_{d_2-1}, x_{d_2}, x_{d_2+1}, \dots, x_{\Lambda}) \\ &\rightarrow (y_1, \dots, y_{d_1-1}, y_{d_1}, x_{d_1+1}, \dots, x_{d_2-1}, x_{d_2}, x_{d_2+1}, \dots, x_{\Lambda}) \\ &= (y_1, \dots, y_{d_1-1}, y_{d_1}, y_{d_1+1}, \dots, y_{d_2-1}, x_{d_2}, x_{d_2+1}, \dots, x_{\Lambda}) \\ &\rightarrow (y_1, \dots, y_{d_1-1}, y_{d_1}, y_{d_1+1}, \dots, y_{d_2-1}, y_{d_2}, x_{d_2+1}, \dots, x_{\Lambda}) \\ &\vdots \\ &\rightarrow (y_1, \dots, y_{\Lambda}). \end{aligned}$$

Après avoir fixé les trajectoires reliant chaque paire de configurations (x, y) dans l'espace des états χ nous nous intéressons au calcul de κ défini dans (8). Nous commençons par donner le résultat de Shiu et Chen (voir [25]) qui ont étudié le cas de l'algorithme de

Introduction

l'échantillonneur de Gibbs pour le modèle d'Ising unidimensionnel avec deux états. La mesure π donnée dans (6) devient:

$$\pi(x) = \frac{1}{Z_T} \exp \left\{ \frac{1}{T} \sum_{k=1}^{n-1} x_k x_{k+1} \right\} \quad \forall x = (x_1, \dots, x_n) \in \chi,$$

avec

- $\chi = \{-1, 1\}^n$ est l'espace des états avec $\Lambda = n$ et $N = 2$,
- T est la température,
- $Z_T = \sum_{x \in \chi} \exp \left\{ \frac{1}{T} \sum_{k=1}^{n-1} x_k x_{k+1} \right\}$ est la constante de normalisation.

La probabilité de se déplacer d'une configuration x vers sa configuration voisine y est:

$$\pi(y_i | x) = \frac{\pi(x_1, \dots, x_{i-1}, y_i, x_{i+1}, \dots, x_n)}{\pi(x_1, \dots, x_{i-1}, y_i, x_{i+1}, \dots, x_n) + \pi(x_1, \dots, x_{i-1}, -y_i, x_{i+1}, \dots, x_n)}.$$

Shiu et Chen (voir [25]) ont donné le résultat suivant:

Théorème 0.0.1. [Shiu and Chen 2015] *La deuxième plus grande valeur propre associée à l'algorithme de l'échantillonneur de Gibbs pour le modèle d'Ising unidimensionnel satisfait:*

$$\beta_1 \leq 1 - e^{-\frac{2}{T} n^{-2}}.$$

Nous présentons maintenant les résultats obtenus dans cette thèse. Dans le deuxième chapitre nous sommes intéressé à l'étude de l'algorithme de l'échantillonneur de Gibbs pour le modèle d'Ising unidimensionnel avec trois états appelé aussi modèle de Potts. La mesure π donnée dans (6) devient:

$$\pi(x) = \frac{1}{Z_T} \exp \left\{ \frac{1}{T} \sum_{k=1}^{n-1} (\mathbb{1}_{\{x_k=x_{k+1}\}} - \mathbb{1}_{\{x_k \neq x_{k+1}\}}) \right\} \quad \forall x = (x_1, \dots, x_n) \in \chi,$$

avec

- $\chi = \{c_1, c_2, c_3\}^n$ est l'espace des états,
- T est la température,
- $Z_T = \sum_{x \in \chi} \exp \left\{ \frac{1}{T} \sum_{k=1}^{n-1} (\mathbb{1}_{\{x_k=x_{k+1}\}} - \mathbb{1}_{\{x_k \neq x_{k+1}\}}) \right\}$ la constante de normalisation.

La probabilité de se déplacer entre deux configurations voisines (qui diffèrent par un seul site) x et y est:

$$\pi(y_i | x) = \frac{\pi(x_1, \dots, x_{i-1}, y_i, x_{i+1}, \dots, x_n)}{\sum_{l=1}^3 \pi(x_1, \dots, x_{i-1}, c_l, x_{i+1}, \dots, x_n)}.$$

Nous citons le premier résultat du deuxième chapitre:

Théorème 0.0.2. [Helali 2018] *La deuxième plus grande valeur propre associée à l'algorithme de l'échantillonneur de Gibbs pour le modèle d'Ising unidimensionnel avec trois états satisfait:*

$$\beta_1 < 1 - 3 \times n^{-2} \frac{e^{-\frac{4}{T}}}{1 + 2e^{-\frac{4}{T}}}. \quad (9)$$

Par la suite, nous généralisons le résultat précédent au cas du modèle d'Ising avec plusieurs couleurs appelé aussi modèle de Potts. La mesure π donnée dans (6) peut s'écrire comme suit:

$$\pi(x) = \frac{1}{Z_T} \exp \left\{ \frac{1}{T} \sum_{k=1}^{n-1} (\mathbb{1}_{\{x_k=x_{k+1}\}} - \mathbb{1}_{\{x_k \neq x_{k+1}\}}) \right\} \quad \forall x = (x_1, \dots, x_n) \in \chi,$$

avec

- $\chi = \{c_1, \dots, c_N\}^n$ est l'espace des états où N est le nombre de couleurs,
- T est la température,
- $Z_T = \sum_{x \in \chi} \exp \left\{ \frac{1}{T} \sum_{k=1}^{n-1} (\mathbb{1}_{\{x_k=x_{k+1}\}} - \mathbb{1}_{\{x_k \neq x_{k+1}\}}) \right\}$ est la constante de normalisation.

De même la probabilité conditionnelle $\pi(y_i|x)$ est égale à:

$$\pi(y_i|x) = \frac{\pi(x_1, \dots, x_{i-1}, y_i, x_{i+1}, \dots, x_n)}{\sum_{l=1}^N \pi(x_1, \dots, x_{i-1}, c_l, x_{i+1}, \dots, x_n)}.$$

Alors le deuxième résultat du deuxième chapitre est:

Théorème 0.0.3. [Helali 2018] *La deuxième plus grande valeur propre associée à l'algorithme de l'échantillonneur de Gibbs pour le modèle d'Ising unidimensionnel avec N états satisfait:*

$$\beta_1 < 1 - N \times n^{-2} \frac{e^{-\frac{4}{T}}}{1 + (N-1)e^{-\frac{4}{T}}}. \quad (10)$$

Un théorème prouvé par Ingrassia (voir [20], Théorème 5.3) donne le minorant suivant pour la plus petite valeur propre dans ce modèle:

$$\beta_{|\chi|-1} \geq -1 + \frac{2}{1 + (c-1)e^{\frac{\Delta}{T}}}$$

où c est le nombre d'orientations que peut prendre un spin et Δ est le saut maximal de l'Hamiltonien.

L'application de ce théorème au cas du modèle d'Ising avec trois ou plusieurs états donne une borne supérieure de $|\beta_{|\chi|-1}|$ qui est plus petite que la borne obtenue pour β_1 . La borne supérieure de β^* est donnée dans le corollaire suivant:

Introduction

Corollaire 0.0.9. *Les bornes supérieures de β_1 données dans le théorème 0.0.2 et le théorème 0.0.3 sont aussi des bornes supérieures pour la valeur absolue de toutes les valeurs propres $\{\beta_1, \dots, \beta_{|\chi|-1}\}$ de l'échantillonneur de Gibbs pour le modèle d'Ising unidimensionnel avec trois états et plus respectivement.*

Dans le troisième chapitre, nous étudions l'échantillonneur de Gibbs pour le modèle d'Ising deux-dimensionnel avec deux états. La loi du modèle d'Ising deux-dimensionnel est:

$$\pi(x) = \frac{1}{Z_T} \exp \left\{ \frac{1}{T} \left(\sum_{j=1}^n \sum_{i=1}^{n-1} x_i^j x_{i+1}^j + \sum_{i=1}^n \sum_{j=1}^{n-1} x_i^j x_i^{j+1} \right) \right\} \quad \forall x = (x_i^j)_{1 \leq i, j \leq n} \in \chi,$$

avec

- $\chi = \{-1, 1\}^{n^2}$ est l'espace des états,
- T est la température,
- $Z_T = \sum_{x \in \chi} \exp \left\{ \frac{1}{T} \left(\sum_{j=1}^n \sum_{i=1}^{n-1} x_i^j x_{i+1}^j + \sum_{i=1}^n \sum_{j=1}^{n-1} x_i^j x_i^{j+1} \right) \right\}$ est la constante de normalisation.

La loi de transition entre deux configurations voisines x et y est:

$$\pi(y_i^j | x) = \frac{\pi(x_1^1, \dots, x_{i-1}^j, y_i^j, x_{i+1}^j, \dots, x_n^n)}{\pi(x_1^1, \dots, x_{i-1}^j, y_i^j, x_{i+1}^j, \dots, x_n^n) + \pi(x_1^1, \dots, x_{i-1}^j, -y_i^j, x_{i+1}^j, \dots, x_n^n)}.$$

Ensuite, le résultat principal du troisième chapitre est:

Théorème 0.0.4. [Franke et Helali 2018] *La deuxième plus grande valeur propre associée à l'algorithme de l'échantillonneur de Gibbs pour le modèle d'Ising deux-dimensionnel avec deux états satisfait:*

$$\beta_1 \leq 1 - n^{-4} \exp \left\{ -\frac{2}{T} (2n + 1) \right\}.$$

En raison de la borne supérieure de $|\beta_{|\chi|-1}|$ présentée dans le théorème 5.3 introduit par Ingrassia dans [20] nous obtenons une borne supérieure de β^* comme suit:

Corollaire 0.0.10. *La deuxième plus grande valeur propre en valeur absolue associée à l'algorithme de l'échantillonneur de Gibbs pour le modèle d'Ising deux-dimensionnel avec deux états satisfait:*

$$\beta^* \leq 1 - n^{-4} \exp \left\{ -\frac{2}{T} (2n + 1) \right\}.$$

Dans le quatrième chapitre, nous étudions le cas de l'échantillonneur de Gibbs perturbé pour le modèle d'Ising unidimensionnel avec deux états. La distribution du modèle d'Ising unidimensionnel est:

$$\pi(x) = \frac{1}{Z_T} \exp \left\{ \frac{1}{T} \sum_{k=0}^{n-1} x_k x_{k+1} \right\} \quad \forall x = (x_1, \dots, x_n) \in \chi,$$

avec

- $\chi = \{-1, 1\}^n$ est l'espace des états,
- T est la température,
- $Z_T = \sum_{x \in \chi} \exp \left\{ \frac{1}{T} \sum_{k=1}^{n-1} x_k x_{k+1} \right\}$ la constante de normalisation.

On s'attend à une accélération de la vitesse de convergence de l'échantillonneur de Gibbs vers son état d'équilibre une fois que l'on ajoute des arêtes au graphe sous-jacent à la chaîne de Markov initiale. Cependant, seuls quelques arêtes doivent être ajoutées afin de limiter les coûts de calcul.

Partant de cette philosophie, nous proposons des perturbations de l'algorithme de Gibbs basées sur des permutations σ de χ .

Soit σ une permutation à cycle deux d'une configuration $x \in \chi$ comme suit:

$$\begin{aligned} \sigma : \chi &\longrightarrow \chi \\ x &\longrightarrow -x. \end{aligned}$$

Pour certains $0 < \alpha < 1$ on note par P_σ la matrice qui est définie pour $x, y \in \chi$ comme suit:

$$P_\sigma(x, y) = \begin{cases} \alpha & \text{si } y = \sigma(x) \\ 1 - \alpha & \text{si } x = y \\ 0 & \text{sinon.} \end{cases}$$

L'idée à la base de cette construction est que la permutation que nous voulons introduire ne doit pas ajouter de nombreuses arêtes au graphe sous-jacent à la chaîne de Markov de l'échantillonneur de Gibbs. Davantage, le fait que la permutation σ est à cycle deux garantit la réversibilité de la chaîne de Markov ce qui nous permet d'utiliser la borne géométrique introduite par Diaconis et Stroock (1991) dans [5]. De plus, l'effort de calcul pour exécuter une simulation pour cette simple permutation n'est pas trop coûteux. Pour $0 < \delta < 1$, on introduit l'algorithme de l'échantillonneur de Gibbs perturbé \tilde{P} comme suit:

$$\tilde{P}(x, y) = \delta P(x, y) + (1 - \delta) P_\sigma(x, y) \quad \forall x, y \in \chi.$$

Remarque 0.0.11. On remarque que $\tilde{P}(x, y)$ est réversible comme une combinaison convexe de deux matrices réversibles. Leurs valeurs propres peuvent être organisées comme suit:

$$1 = \tilde{\beta}_0 > \tilde{\beta}_1 \geq \tilde{\beta}_2 \geq \dots \geq \tilde{\beta}_{|\chi|-1} > -1.$$

Pour pouvoir utiliser le résultat de Diaconis et Stroock [5] pour calculer une borne supérieure de la deuxième plus grande valeur propre, il faut fixer une collection de trajectoires reliant toutes paires de configurations x et y de l'espace χ . Pour une paires de configurations (x, y) donnée, soit d_1, \dots, d_m une suite croissante satisfaisant $x_i \neq y_i$ pour $i \in \{d_1, \dots, d_m\}$ et $x_i = y_i$ ailleurs.

Le choix des trajectoires qui relient les configurations dépendra de la distance entre x et y .

1. Si $m \leq \frac{n}{2}$ on utilise le chemin introduit par Chen et Shiu dans [25] alors:

$$\begin{aligned}
 (x_1, \dots, x_n) &= (y_1, \dots, y_{d_1-1}, x_{d_1}, x_{d_1+1}, \dots, x_{d_2-1}, x_{d_2}, x_{d_2+1}, \dots, x_n) \\
 &\rightarrow (y_1, \dots, y_{d_1-1}, y_{d_1}, x_{d_1+1}, \dots, x_{d_2-1}, x_{d_2}, x_{d_2+1}, \dots, x_n) \\
 &= (y_1, \dots, y_{d_1-1}, y_{d_1}, y_{d_1+1}, \dots, y_{d_2-1}, x_{d_2}, x_{d_2+1}, \dots, x_n) \\
 &\rightarrow (y_1, \dots, y_{d_1-1}, y_{d_1}, y_{d_1+1}, \dots, y_{d_2-1}, y_{d_2}, x_{d_2+1}, \dots, x_n) \\
 &\vdots \\
 &\rightarrow (y_1, \dots, y_n).
 \end{aligned}$$

2. Si $m > \frac{n}{2}$, comme les configurations diffèrent de plus de $\frac{n}{2}$ sites, on change de x à $-x$ dans la première étape puis on continue avec le chemin introduit dans [25]:

$$\begin{aligned}
 (x_1, \dots, x_n) &= (y_1, \dots, y_{d_1-1}, x_{d_1}, x_{d_1+1}, \dots, x_{d_2-1}, x_{d_2}, x_{d_2+1}, \dots, x_n) \\
 &\rightarrow (-y_1, \dots, -y_{d_1-1}, -x_{d_1}, -x_{d_1+1}, \dots, -x_{d_2-1}, -x_{d_2}, -x_{d_2+1}, \dots, -x_n) \\
 &= (-y_1, \dots, -y_{d_1-1}, y_{d_1}, -x_{d_1+1}, \dots, -x_{d_2-1}, y_{d_2}, -x_{d_2+1}, \dots, -x_n) \\
 &\rightarrow (y_1, \dots, -y_{d_1-1}, y_{d_1}, -x_{d_1+1}, \dots, -x_{d_2-1}, y_{d_2}, -x_{d_2+1}, \dots, -x_n) \\
 &\vdots \\
 &\rightarrow (y_1, \dots, y_n).
 \end{aligned}$$

Ensuite, le résultat principal du quatrième chapitre pour un choix bien précis de α et δ est:

Théorème 0.0.5. [Franke et Helali 2018] *La deuxième plus grande valeur propre de l'échantillonneur de Gibbs perturbé associé au modèle Ising unidimensionnel satisfait à haute température:*

$$\beta_1 \leq 1 - 2n^{-2} e^{-\frac{2}{T}}.$$

Chapter 1

Preliminaries

In this chapter we present some basic definitions and properties used to prove our results. We indicate references for more details.

1.1 Markov chains and some properties

Definition 1.1.1. Let χ be a finite set. A family $\{P(x, \cdot) : x \in \chi\}$ of probability distributions on χ is called a transition probability or a Markov kernel.

The space χ is called the state space. A Markov kernel P can be represented by a matrix denoted by P as well: $P(x, y)$ is the element in the x -th row and the y -th column. Then the matrix P is a square matrix with probability vectors in its rows. Each number $P(x, y)$ present the probability that the process changes from the state x to the state y .

Definition 1.1.2. A kernel is said to be primitive if there exist an integer τ such that $P^\tau(x, y) > 0$ for all $x, y \in \chi$.

Definition 1.1.3. An homogeneous chain which reaches after a finite number of steps from each state to each other state with positive probability is called irreducible.

The following properties can be found in the book of Winkler (see [28]):

Properties 1.1.4. • A kernel is primitive if and only if it is irreducible and aperiodic.

• A Markov kernel P and a distribution π fulfill the detailed balance equation if

$$\pi(x)P(x, y) = \pi(y)P(y, x) \text{ for all } x, y \in \chi.$$

Then the Markov chain with initial distribution π and transition kernel P is called reversible in time, and P is called reversible with respect to π .

• The matrix P has 1 as its biggest eigenvalue which is associated to the constant function as associated eigenfunction. The spectrum of the matrix P can be arranged in the reversible case as follows:

$$1 = \beta_0 > \beta_1 \geq \dots \geq \beta_{|\chi|-1} \geq -1.$$

- If the Markov chain P is aperiodic then, the smallest eigenvalue $\beta_{|\chi|-1}$ is different from -1 .

Lemma 1.1.5. *If the matrix P is reversible with respect to the measure π then π is P -invariant.*

Proof. For all configurations $x, y \in \chi$ we have:

$$\begin{aligned} \pi P(x) &= \sum_{y \in \chi} \pi(y) P(y, x) = \sum_{y \in \chi} \pi(x) P(x, y) \\ &= \pi(x) \sum_{y \in \chi} P(x, y) = \pi(x) \\ &= \pi(x). \end{aligned}$$

□

1.2 Diaconis and Stroock's approach

Diaconis and Stroock (see [5]) prove that the convergence rate of a reversible Markov chain to its equilibrium with respect to the total variation distance depends on its second largest eigenvalue in absolute value denoted by β^* . Further, they give an upper bound for β_1 . More details are given in the following paragraph.

Theorem 1.2.1. *[Diaconis and Stroock 1991] If P is a reversible Markov chain and π its unique invariant measure if also P is irreducible then for all $x \in \chi$ and $k \in \mathbb{N}$ we have:*

$$4 \left\| P^k(x, \cdot) - \pi(\cdot) \right\|_{var}^2 = \left(\sum_{y \in \chi} | P^k(x, y) - \pi(y) | \right)^2 \leq \frac{1 - \pi(x)}{\pi(x)} (\beta^*)^{2k}.$$

Proof. The following proof can be found in [5].

According to the definition of the total variation distance we have:

$$\begin{aligned} 4 \left\| P^k(x, \cdot) - \pi(\cdot) \right\|_{var}^2 &= \left(\sum_{y \in \chi} | P^k(x, y) - \pi(y) | \right)^2 \\ &= \left(\sum_{y \in \chi} \sqrt{\frac{\pi(y)}{\pi(y)}} | P^k(x, y) - \pi(y) | \right)^2. \end{aligned}$$

Using Cauchy-Schwarz inequality we obtain:

$$4 \left\| P^k(x, \cdot) - \pi(\cdot) \right\|_{var}^2 \leq \sum_{y \in \chi} \frac{1}{\pi(y)} \left((P^k(x, y))^2 - 2P^k(x, y)\pi(y) + \pi(y)^2 \right).$$

From the fact that the Markov chain P is reversible we have the following identity:

$$\sum_{y \in \chi} \frac{1}{\pi(y)} (P^n(x, y))^2 = \frac{1}{\pi(x)} P^{2n}(x, y).$$

It follows that:

$$\begin{aligned} 4 \|\| P^k(x, \cdot) - \pi(\cdot) \|\|_{var}^2 &\leq \sum_{y \in \chi} \frac{1}{\pi(y)} (P^k(x, y))^2 - 1 \\ &= \frac{1}{\pi(x)} P^{2k}(x, y) - 1. \end{aligned}$$

Let D be a diagonal matrix where the x -th diagonal position takes the value $\sqrt{\pi(x)}$. Then the matrix DPD^{-1} takes the value $\sqrt{\frac{\pi(x)}{\pi(y)}}P(x, y)$ in position (x, y) . As a result the matrix DPD^{-1} is symmetric. There exists an orthogonal matrix Θ (i.e.: with $\Theta\Theta^t = I$) such that $DPD^{-1} = \Theta B \Theta^t$ where B is a diagonal matrix containing the eigenvalues of the matrix P on its diagonal. All entries of P are real numbers between -1 and 1 . Moreover the value -1 is never reached. Then, the transition matrix P^{2n} can be written in the following form:

$$P^{2n} = D^{-1} \Theta B^{2n} \Theta^t D.$$

We obtain in each entry (x, y) :

$$P^{2n}(x, y) = \sqrt{\frac{\pi(y)}{\pi(x)}} \sum_w \Gamma_{xw} B_{ww}^{2n} \Theta_{wy}.$$

By the fact that β^{*2n} is an upper bound of B_{ww}^{2n} and the orthogonality of Θ this finishes the proof. \square

In what follows we turn to give an upper bound for β_1 . Let $V = (\chi, E)$ be a graph where χ is the set of vertices and $E = \{(x, y) \mid P(x, y) > 0\}$ the set of edges. Let $\Gamma = \{\gamma_{xy}; x, y \in \chi\}$ be a collection of paths γ_{xy} joining x to y (only one path for each pair of configurations).

Proposition 1.2.2. [Diaconis and Stroock 1991] *The geometric bound of the second largest eigenvalue for reversible Markov chain is,*

$$\beta_1 \leq 1 - \frac{1}{\kappa} \tag{1.1}$$

with

$$\kappa = \max_{e \in E} Q(e)^{-1} \sum_{\gamma_{xy} \ni e} |\gamma_{xy}| \pi(x) \pi(y) \tag{1.2}$$

where $|\gamma_{xy}|$ designate the length of the path γ_{xy} and the sum is over all paths $\gamma_{xy} \in \Gamma$ passing through the edge e .

Remark 1.2.3. The previous proposition is a discrete version of the Poincaré method for estimating the spectral gap of the Laplacien on a domain. More details are given by Bandle in [2]. Related ideas were used by Landau and Odlyzko (see [17]), by Holley and Stroock (see [15]) and by Mohar (see [22, 23]).

Proof. We present the proof that can be found in [5].

Horn and Johnson (1985) (see [16], page 176) prove that

$$\lambda_1 = \inf \left\{ \frac{\mathcal{E}(\phi, \phi)}{\text{Var}(\phi)} : \phi \neq \text{constant} \right\} \quad (1.3)$$

where $\text{Var}(\phi)$ designate the variance of ϕ relative to the measure π . and \mathcal{E} is the quadratic form which is defined as follows:

$$\mathcal{E}(\phi, \phi) = \frac{1}{2} \sum_{x,y} (\phi(y) - \phi(x))^2 Q(x, y).$$

On the other hand we have:

$$\begin{aligned} \text{Var}(\phi) &= \frac{1}{2} \sum_{x,y \in \chi} (\phi(x) - \phi(y))^2 \pi(x)\pi(y) \\ &= \frac{1}{2} \sum_{x,y \in \chi} \left(\sum_{e \in \gamma_{xy}} \left(\frac{Q(e)}{Q(e)} \right)^{\frac{1}{2}} \phi(e) \right)^2 \pi(x)\pi(y). \end{aligned}$$

Here $\phi(e) = \phi(e^+) - \phi(e^-)$ is the oriented edge from e^- to e^+ in a given path. From Cauchy-Schwarz inequality we obtain:

$$\text{Var}(\phi) \leq \frac{1}{2} \sum_{x,y \in \chi} |\gamma_{xy}|_Q \pi(x)\pi(y) \sum_{e \in \gamma_{xy}} Q(e) \phi(e)^2. \quad (1.4)$$

Finally, for all edges in the graph we have:

$$\text{Var}(\phi) \leq \frac{1}{2} \sum_e Q(e) \phi(e)^2 \sum_{x,y \in \chi} |\gamma_{xy}|_Q \pi(x)\pi(y).$$

From the definition, κ is the upper bound for $\sum_{x,y \in \chi} |\gamma_{xy}|_Q \pi(x)\pi(y)$, then we obtain

$$\text{Var}(\phi) \leq \kappa \mathcal{E}(\phi, \phi).$$

The result is finally obtained by the variational characterization in (1.3). \square

1.3 Ingrassia's approach

The method of Ingrassia (see [20]) consists in giving an upper bound for the second largest eigenvalue denoted by β_1 and a lower bound for the smallest eigenvalue denoted by $\beta_{|\chi|-1}$ in order to bound β^* . This yields some informations about the speed of convergence of the Markov chain P to its equilibrium π .

Let us first introduce the notations used by Ingrassia: T is the temperature, Z_T is the normalizing constant, S is the lattice of sites, Γ the collection of paths, γ_Γ is the maximum length of each path $\gamma_{xy} \in \Gamma$, b_Γ is the maximum number of paths containing any edge Γ , c is the number of configurations that differ by only one site, m is the least total elevation gain of the Hamiltonian function in the sense which is described by Holley and Stroock (see [15]).

1.3.1 Upper bound for β_1

Ingrassia also uses the bound for β_1 from Diaconis and Stroock and he obtain the following result:

Theorem 1.3.1. [Ingrassia 1994] *Let P be the transition matrix for the Gibbs sampler. Then the second largest eigenvalue satisfies:*

$$\beta_1 \leq 1 - \frac{Z_T}{b_\Gamma \gamma_\Gamma c |S|} e^{-\frac{m}{T}}.$$

Proof. The following proof is from Ingrassia (see [20]).

Let Γ be the set of paths linking each pair of configurations and $e \in E$ some edge where $P(e) = P(e^-, e^+) > 0$. We obtain:

$$\begin{aligned} |\gamma_{xy}|_Q \pi(x) \pi(y) &= \sum_{e \in \gamma_{xy}} Q(e)^{-1} \pi(x) \pi(y) \\ &= \sum_{e \in \gamma_{xy}} (\pi(e^-) P(e))^{-1} \pi(x) \pi(y) \\ &= \sum_{e \in \gamma_{xy}} \left(\frac{1}{|S|} \sum_{s \in S} \pi(e^-) P_s(e) \mathbf{1}_s(e) \right)^{-1} \pi(x) \pi(y) \\ &= \sum_{e \in \gamma_{xy}} \left(\frac{1}{|S|} \sum_{s \in S} \frac{\exp\left\{\frac{-H(e^+)}{T}\right\}}{\sum_{\tilde{e} \in N_s(e^-)} \exp\left\{\frac{-H(\tilde{e})}{T}\right\}} \frac{\exp\left\{\frac{-H(e^-)}{T}\right\}}{Z_T} \mathbf{1}_s(e) \right)^{-1} \pi(x) \pi(y). \end{aligned}$$

For all edge e the sum in $(\sum_{s \in S} \dots)^{-1}$ contains only one term different from zero, which gives:

$$\begin{aligned} |\gamma_{xy}|_Q \pi(x) \pi(y) &= \frac{|S|}{Z_T} \sum_{e \in \gamma_{xy}} \sum_{s \in S} \sum_{\tilde{e} \in N_s(e^-)} \exp\left(\frac{H(e^-) - H(\tilde{e})}{T}\right) \\ &\quad \times \exp\left(\frac{H(e^-) - H(x) - H(y)}{T}\right) \mathbf{1}_s(e). \end{aligned}$$

Also, the choice of paths guarantee that for all edge $e \in \gamma_{xy}$ we have:

$$\exp\left(\frac{H(e^-) - H(\tilde{e})}{T}\right) \times \exp\left(\frac{H(e^-) - H(x) - H(y)}{T}\right) \leq \exp\frac{m}{T}.$$

Then

$$\sum_{s \in S} \sum_{\tilde{e} \in N_s(e^-)} \exp\left(\frac{H(e^-) - H(\tilde{e})}{T}\right) \times \exp\left(\frac{H(e^-) - H(x) - H(y)}{T}\right) \mathbf{1}_s(e) \leq c \exp\left(\frac{m}{T}\right),$$

which gives:

$$|\gamma_{xy}|_Q \pi(x) \pi(y) \leq \frac{|S|}{Z_T} \sum_{e \in \gamma_{xy}} c e^{\frac{m}{T}}.$$

From the definition of γ_Γ the previous definition becomes:

$$|\gamma_{xy}|_Q \pi(x) \pi(y) \leq \frac{|S|}{Z_T} c \gamma_\Gamma e^{\frac{m}{T}}.$$

Then we obtain:

$$\kappa \equiv \max_{e' \in E} \sum_{\gamma_{xy} \in e'} |\gamma_{xy}|_Q \pi(x) \pi(y) \leq \frac{b_\Gamma \gamma_\Gamma c |S|}{Z_T} e^{\frac{m}{T}}.$$

Using equation (1.1) we obtain the result:

$$\beta_1 \leq 1 - \frac{Z_T}{b_\Gamma \gamma_\Gamma c |S|} e^{-\frac{m}{T}}.$$

□

1.3.2 Lower bound for $\beta_{|\chi|-1}$

We turn now to give a lower bound for the smallest eigenvalue associated to the matrix P . In order to do, we define Δ the maximum jump of the function H for all transitions:

$$\Delta = \max_{x \in \chi} \max_{y \in N(x)} |H(y) - H(x)| \quad (1.5)$$

where $N(x) = \{y \in \chi : P(x, y) > 0\}$ is the set of neighbors of x .

Theorem 1.3.2. [Ingrassia 1994] *Let P be the transition matrix of the Gibbs sampler. Then the smallest eigenvalue satisfies:*

$$\beta_{|\chi|-1} \geq \frac{2}{1 + (c-1)e^{\frac{\Delta}{T}}},$$

where Δ is defined in (1.5).

Proof. The details of the proof can be found in Ingrassia (see [20], page 371). □

1.4 Shiu and Chen's result

In this section we present the result of Shiu and Chen (see [25]) for the Gibbs sampler in the case of the 1-D Ising model using the geometric bound proved by Diaconis and Stroock (see [5]) for reversible Markov chains.

1.4.1 Description of the model

The distribution associated to the one-dimensional Ising model is:

$$\pi(x) = \frac{1}{Z_T} \exp \left\{ \frac{1}{T} \sum_{k=1}^{n-1} x_k x_{k+1} \right\}$$

for all $x \in \chi$ where $\chi = \{-1, +1\}^n$ is the state space, T is the Temperature and Z_T is the normalizing constant. We consider the Gibbs sampler which chooses randomly a site to be updated with given conditional probabilities. Explicitly, for a given configuration $x = (x_1, \dots, x_i, \dots, x_n)$, we chooses randomly a site x_i who is to be updated to y_i with probability:

$$\pi(y_i|x) = \frac{\pi(x_1, \dots, x_{i-1}, y_i, x_{i+1}, \dots, x_n)}{\pi(x_1, \dots, x_{i-1}, y_i, x_{i+1}, \dots, x_n) + \pi(x_1, \dots, x_{i-1}, -y_i, x_{i+1}, \dots, x_n)}.$$

The matrix of the Gibbs sampler is defined as follows:

$$P(x, y) = \begin{cases} \frac{1}{n} \pi(y_i | x) & \text{if } x_j \neq y_j \text{ for all } j \neq i \\ 1 - \frac{1}{n} \sum_{i=1}^n \pi(y_i | x) & \text{if } x = y \\ 0 & \text{else.} \end{cases}$$

1.4.2 Main result

To be able to use the result of Diaconis and Stroock (see [5]) to give an upper bound of the second largest eigenvalue of the matrix P , we must first of all fix a collection of paths linking each pair of configurations $(x, y) \in \chi^2$. The best bound is obtained for small κ . For a given pair of configurations $x, y \in \chi$ there exists an increasing sequence d_1, \dots, d_m such that $x_i \neq y_i$ for $i \in \{d_1, \dots, d_m\}$ and $x_i = y_i$ else.

We define a path linking any pair of configurations (x, y) as follow:

$$\begin{aligned} (x_1, \dots, x_n) &= (y_1, \dots, y_{d_1-1}, x_{d_1}, x_{d_1+1}, \dots, x_{d_2-1}, x_{d_2}, x_{d_2+1}, \dots, x_n) \\ &\rightarrow (y_1, \dots, y_{d_1-1}, y_{d_1}, x_{d_1+1}, \dots, x_{d_2-1}, x_{d_2}, x_{d_2+1}, \dots, x_n) \\ &= (y_1, \dots, y_{d_1-1}, y_{d_1}, y_{d_1+1}, \dots, y_{d_2-1}, x_{d_2}, x_{d_2+1}, \dots, x_n) \\ &\rightarrow (y_1, \dots, y_{d_1-1}, y_{d_1}, y_{d_1+1}, \dots, y_{d_2-1}, y_{d_2}, x_{d_2+1}, \dots, x_n) \\ &\vdots \\ &\rightarrow (y_1, \dots, y_n). \end{aligned}$$

The principal result of Shiu and Chen (see [25]) is given in the following theorem:

Theorem 1.4.1. [Shiu and Chen 2015] *The second largest eigenvalue of the Gibbs sampler for the one-dimensional Ising model satisfies:*

$$\beta_1 \leq 1 - n^{-2} e^{-\frac{2}{T}}.$$

In what follows, we give a sketch of the proof of the main result of Shiu and Chen. One can find more details in [25].

Proof. We define the edge $e = (e^-, e^+)$ as follow:

$$e^- = (z_1, \dots, z_{i-1}, z_i, z_{i+1}, \dots, z_n) \text{ et } e^+ = (z_1, \dots, z_{i-1}, -z_i, z_{i+1}, \dots, z_n).$$

We consider at the beginning the case where $i \notin \{1, n\}$. Then

$$P(e^-, e^+) = \frac{1}{n(1 + \exp\{\frac{2}{T}(z_{i-1}z_i + z_i z_{i+1})\}}.$$

According to the selection of our paths, for a path γ_{xy} passing through e the configurations x and y take the following form:

$$x = (x_1, \dots, x_{i-1}, z_i, z_{i+1}, \dots, z_n) \text{ et } y = (z_1, \dots, z_{i-1}, -z_i, y_{i+1}, \dots, y_n).$$

Which gives

$$\begin{aligned} |\gamma_{xy}|Q(e)^{-1}\pi(x)\pi(y) &\leq \frac{n^2}{Z_T} \exp \left\{ \frac{1}{T} \left(\sum_{k=1}^{i-2} x_k x_{k+1} + x_{i-1} z_i + z_i y_{i+1} + \sum_{k=i+1}^{n-1} y_k y_{k+1} \right) \right\} \\ &\times \exp \left\{ -\frac{2}{T} (z_{i-1} z_i + z_i y_{i+1}) \right\} \\ &+ \frac{n^2}{Z_T} \exp \left\{ \frac{1}{T} \left(\sum_{k=1}^{i-2} x_k x_{k+1} - x_{i-1} z_i - z_i y_{i+1} + \sum_{k=i+1}^{n-1} y_k y_{k+1} \right) \right\} \\ &\times \exp \left\{ \frac{2}{T} (z_{i-1} z_i + z_i y_{i+1}) \right\}. \end{aligned} \quad (1.6)$$

We define the configurations $x \oplus y$ and $x \ominus y$ as follow:

$$x \oplus y = (x_1, \dots, x_{i-1}, z_i, y_{i+1}, \dots, y_n) \text{ and } x \ominus y = (x_1, \dots, x_{i-1}, -z_i, y_{i+1}, \dots, y_n).$$

We have:

$$\bigcup_{(x,y)|\gamma_{xy} \ni e} \{x \oplus y, x \ominus y\} = \chi.$$

Combining equation (1.6) and the previous notations we obtain:

$$\begin{aligned} Q(e)^{-1} \sum_{\gamma_{xy} \ni e} |\gamma_{xy}| \pi(x) \pi(y) &= n^2 \sum_{w \in \chi, w_i = z_i} \pi(w) e^{-\frac{2}{T}(z_{i-1} z_i + z_i w_{i+1})} \\ &+ n^2 \sum_{w \in \chi, w_i = -z_i} \pi(w) e^{\frac{2}{T}(z_{i-1} z_i + z_i w_{i+1})}. \end{aligned} \quad (1.7)$$

Note that for a given pair of configurations (x, y) , the max over all edges of equation (1.6) is reached for an edge e where $-z_{i-1} = z_i = z_{i+1}$. Suppose that $z_i = 1$ then equation (1.7) becomes:

$$Q(e)^{-1} \sum_{\gamma_{xy} \ni e} |\gamma_{xy}| \pi(x) \pi(y) \leq 2n^2 \sum_{w \in \chi, w_i = 1} \pi(w) e^{\frac{2}{T}(1-w_{i+1})}. \quad (1.8)$$

Remark 1.4.2. Equation (1.8) can also be obtained in the case where $z_i = -1$.

A symmetry consideration gives that:

$$\sum_{w_i=1} \pi(w) e^{\frac{2}{T}(1-w_{i+1})} = \frac{1}{2} e^{\frac{2}{T}}.$$

Then equation (1.8) becomes:

$$\max_{\gamma_{xy} \ni e} Q(e)^{-1} \sum_{\gamma_{xy} \ni e} |\gamma_{xy}| \pi(x) \pi(y) \leq n^2 e^{\frac{2}{T}}. \quad (1.9)$$

In the case of $i \in \{1, n\}$ we obtain similar results.

Finally, we put the result from (1.9) in equation (1.1) to finish the proof. \square

1.5 Alison Gibbs's approach

Alison Gibbs (see [12]) uses coupling in order to measure the convergence rate of the Gibbs sampler for the 1–D and the 2–D Ising models. We explain briefly in this section her approach. More details and explicit proofs can be found in [12].

1.5.1 Explanation of the approach

The convergence time of the Gibbs sampler is defined to be:

$$T_0 = \max_{x^0} \min\{t : D_{x^0}(t') \leq \epsilon \text{ for all } t' \geq t\}$$

where $D_{x^0}(t')$ is the total variation distance at time t' , x^0 is the initial configuration and ϵ is the pre-specified error tolerance. Let

$$T^{x^1, x^2} = \min\{t : X_t^1 = X_t^2 | X_0^1 = x^1, X_0^2 = x^2\}$$

the coupling time of two Markov chains X_t^1 and X_t^2 starting respectively in x^1 and x^2 . The maximum mean coupling time T is

$$T = \max_{x^1, x^2} \mathbb{E}(T^{x^1, x^2}).$$

Aldous (see [1]) proves that the convergence time depends on the *max mean coupling time* T and the pre-specified error tolerance ϵ . His result is given in the following theorem:

Theorem 1.5.1. [Aldous 1983] *The convergence time of a Markov chain depends on the maximum mean coupling time as follows:*

$$T_0 \leq 2eT(1 + \log \epsilon^{-1}).$$

Alison Gibbs used some coupling idea from Luby, Randal and Sinclair (see [18]) and Propp and Wilson (see [24]) to obtain explicit bounds for the coupling time in Aldous result.

Theorem 1.5.2. [Alison Gibbs 2000] *Suppose there exist two coupled realizations, X_t^1, X_t^2 , for a Markov chain where $X_0^1 = x^1$ and $X_0^2 = x^2$. Also suppose a constant $a > 0$ can be found such that $\mathbb{E}(\Delta\phi(t) | X_t^1, X_t^2) \leq -a$ for all t for which $X_t^1 \neq X_t^2$, and the distance between the initial states is $\phi(0) = N$. Then the following bound exists on the mean coupling time*

$$\mathbb{E}(T^{x^{\max}, x^{\min}}) \leq \frac{N}{a}$$

where $\phi(t)$ is the number of different sites between two configurations at time t and $\Delta\phi(t) = \phi(t+1) - \phi(t)$.

Proof. The following proof is from [12].

Define the stochastic process $Z_T = \phi(t) + at$. Then Z_t is a supermartingale up to time T^{x^1, x^2} since

$$\mathbb{E}(Z_{t+1} | X_t^1, X_t^2) - Z_t = \mathbb{E}(\Delta\phi | X_t^1, X_t^2) + a \leq 0.$$

1.5. Alison Gibbs's approach

Also, T^{x^1, x^2} is a stopping time. Since Z_t is nonnegative, we can apply the optional stopping theorem (see Theorem 7.6 page 274 in [6]) giving

$$\mathbb{E}(Z_{T^{x^1, x^2}}) = \mathbb{E}(Z_0),$$

which gives

$$a\mathbb{E}(T^{x^1, x^2}) \leq N.$$

which finishes the proof. □

To avoid confusion in the notation of T between the temperature and the mean coupling time we give the result in term of inverse temperature β .

Based on the previous theorem, Alison Gibbs proves the following results:

1.5.2 Main results

Theorem 1.5.3. [Alison Gibbs 2000] *The convergence time of the Gibbs sampler for the 1-D Ising model with N sites is*

$$t \leq 2eN^2 \frac{e^{2\beta} + e^{-2\beta}}{2e^{-2\beta}} (1 + \log \epsilon^{-1}).$$

Theorem 1.5.4. [Alison Gibbs 2000] *The convergence time of the Gibbs sampler for the 1-D Ising model with N sites, where each site is influenced by its n nearest neighbors, is*

$$t \leq 2eN^2 \frac{e^{n\beta} + e^{-n\beta}}{(n+2)e^{-n\beta} - ne^{n\beta}} (1 + \log \epsilon^{-1})$$

for all $0 \leq \beta \leq \frac{1}{2n} \log \binom{n+2}{n}$.

Chapter 2

The convergence rate of the Gibbs sampler for generalized 1–D Ising model

The rate of convergence of the Gibbs sampler for the generalized one-dimensional Ising model is determined by the second largest eigenvalue of its transition matrix in absolute value denoted by β^* . In this chapter we generalize a result from Shiu and Chen (2015) for the one-dimensional Ising model with two states which gives a bound for β^* . The method is based on Diaconis and Stroock bound for reversible Markov processes. The new bound presented in this chapter can be found in Helali (2018) and it improves Ingrassia's (1994) result.

La vitesse de convergence de l'algorithme de l'échantillonneur de Gibbs pour le modèle d'Ising unidimensionnel généralisé est déterminé par la deuxième plus grande valeur propre en valeur absolue de sa matrice de transition notée β^* . Dans ce chapitre nous avons généralisé un résultat obtenu par Shiu et Chen (2015) pour le modèle d'Ising unidimensionnel avec deux états qui donnent une borne de β^* . La méthode est basée sur la borne introduite par Diaconis et Stroock pour les chaînes de Markov réversibles. La nouvelle borne présentée dans ce chapitre est donnée dans Helali (2018) et elle améliore les résultats classiques de Ingrassia (1994)

2.1 Introduction

The Ising model is a crude model for ferromagnetism. It is the simplest model of statistical mechanics and it has been applied in many other fields like chemistry, molecular biology and image analysis. The distribution of the one-dimensional Ising model with three states is:

$$\pi(x) = \frac{1}{Z_T} \exp \left\{ \frac{1}{T} \sum_{k=1}^{n-1} (\mathbb{1}_{\{x_k=x_{k+1}\}} - \mathbb{1}_{\{x_k \neq x_{k+1}\}}) \right\} \quad \forall x = (x_1, \dots, x_n) \in \chi,$$

where $\chi = \{c_1, c_2, c_3\}^n$ is the state space, T is the temperature and

$$Z_T = \sum_{x \in \chi} \exp \left\{ \frac{1}{T} \sum_{k=1}^{n-1} (\mathbb{1}_{\{x_k=x_{k+1}\}} - \mathbb{1}_{\{x_k \neq x_{k+1}\}}) \right\}$$

is the normalizing constant. Monte Carlo Markov chain *MCMC* method is a very useful technique to draw samples from the Ising model. Suppose that the transition probability $P(x, y)$ for an irreducible Markov chain has π as its invariant measure. The pair (P, π) is said to be reversible if it verifies the detailed balance equation:

$$Q(x, y) := \pi(x)P(x, y) = \pi(y)P(y, x) := Q(y, x) \quad \forall x, y \in \chi.$$

The Gibbs sampler introduced by Geman and Geman and the Metropolis-Hastings algorithm introduced by Metropolis et al. and Hastings (see [11], [21] and [13]) are the most popular Monte Carlo Markov chain methods. The matrix P satisfies detailed balance and thus is symmetric with respect to the scalar product introduced by the measure π . Therefore its eigenvalues are real and can be arranged as follows:

$$1 = \beta_0 > \beta_1 \geq \beta_2 \geq \dots \geq \beta_{|\chi|-1} > -1.$$

Let $\beta^* = \max\{\beta_1, |\beta_{|\chi|-1}|\}$. By using the total variance distance, the second largest eigenvalue in absolute value determines the convergence rate of the Markov chain. Ingrassia gives a lower bound for $\beta_{|\chi|-1}$ and an upper bound for β_1 (see [20]).

This chapter deals with the Gibbs sampler for the one-dimensional Ising model with multiple states. It chooses a random coordinate which is updated according to the conditional probability given the other coordinates. The resulting Markov chain is reversible and the associated transition matrix has the form:

$$P(x, y) = \begin{cases} \frac{1}{n} \pi(y_i | x) & \text{if } x_j \neq y_j \text{ for all } j \neq i \\ 1 - \frac{1}{n} \sum_{i=1}^n \sum_{y_i \in \{c_1, c_2, c_3\}} \pi(y_i | x) & \text{if } x = y \\ 0 & \text{else} \end{cases}$$

where

$$\pi(y_i | x) = \frac{\pi(x_1, \dots, x_{i-1}, y_i, x_{i+1}, \dots, x_n)}{\sum_{l=1}^3 \pi(x_1, \dots, x_{i-1}, c_l, x_{i+1}, \dots, x_n)}.$$

In theorem 1.2.1 Diaconis and Stroock (see [5]) give a bound for the total variation distance to equilibrium in terms of β^* . Moreover, Diaconis and Stroock (see [5]) develop a method to calculate an upper bound for the second largest eigenvalue β_1 for a reversible Markov chain using geometric quantities such as the maximum degree, diameter and covering number of the associated graph. Consider the graph $\mathcal{G}(P) = (\chi, E)$ where χ is the vertex set and $E = \{(x, y) \mid P(x, y) > 0\}$ is the edge set. For each pair of distinct points $x, y \in \chi$ we choose a path γ_{xy} from x to y , such that each edge appears at most once in a given path. The fact that P is irreducible guarantees that such paths exist. Let Γ be the collection of all such paths γ_{xy} (one for each pair). The geometric bound given by Diaconis and Stroock is,

$$\beta_1 \leq 1 - \frac{1}{\kappa} \tag{2.1}$$

with

$$\kappa = \max_{e \in E} Q(e)^{-1} \sum_{\gamma_{xy} \ni e} |\gamma_{xy}| \pi(x) \pi(y) \tag{2.2}$$

where $|\gamma_{xy}|$ designates the length of the path γ_{xy} (see [5]).

In their paper Shiu and Chen (see [25]) present a method to explicitly compute the Diaconis and Stroock (see [5]) bound of the Gibbs sampler for a two state one-dimensional Ising model.

In this chapter we generalize the result of Shiu and Chen to the case of the one-dimensional Ising model with three and more states (see [25]).

Our method is based on the idea from [25] which consists of defining suitable paths γ_{xy} linking each pair (x, y) from the state space χ and then to explicitly compute κ defined in equation (2.2) with some suitable symmetry argument. In the discussion section of the chapter we compare our bound to results from the literature. It turns out that the result generalizes the bound given in [25] to the case of the Ising model with three states (see Theorem 2.2.1) and also to multiple states (see Theorem 2.2.3). It also improves the bound presented by Ingrassia in [20].

2.2 Main result

2.2.1 Selection of paths

To be able to use the result of Diaconis and Stroock (see [5]) and to calculate the bound of the second largest eigenvalue, we have to fix a collection of paths joining any configuration $x \in \chi$ to any $y \in \chi$. To get a small upper bound for β_1 , we seek a small value for κ and we should therefore use short paths γ_{xy} to link x with y . Moreover, we have to keep the number of paths passing through a given edge low.

For a pair of distinct configurations $x, y \in \chi$ there exist some increasing sequence d_1, \dots, d_m such that $x_i \neq y_i$ for $i \in \{d_1, \dots, d_m\}$ and $x_i = y_i$ otherwise. In the same way as Shiu

2.2. Main result

and Chen (see [25]) we define a path linking a given pair (x, y) as follows:

$$\begin{aligned}
(x_1, \dots, x_n) &= (y_1, \dots, y_{d_1-1}, x_{d_1}, x_{d_1+1}, \dots, x_{d_2-1}, x_{d_2}, x_{d_2+1}, \dots, x_n) \\
&\rightarrow (y_1, \dots, y_{d_1-1}, y_{d_1}, x_{d_1+1}, \dots, x_{d_2-1}, x_{d_2}, x_{d_2+1}, \dots, x_n) \\
&= (y_1, \dots, y_{d_1-1}, y_{d_1}, y_{d_1+1}, \dots, y_{d_2-1}, x_{d_2}, x_{d_2+1}, \dots, x_n) \\
&\rightarrow (y_1, \dots, y_{d_1-1}, y_{d_1}, y_{d_1+1}, \dots, y_{d_2-1}, y_{d_2}, x_{d_2+1}, \dots, x_n) \\
&\vdots \\
&\rightarrow (y_1, \dots, y_n).
\end{aligned}$$

We turn now to give an upper bound for the value of κ defined in equation (2.2).

2.2.2 Geometric bound for the second largest eigenvalue

In what follows we will essentially follow the arguments from Shiu and Chen (see [25]) to find an upper bound for

$$\kappa = \max_{e \in E} Q(e)^{-1} \sum_{\gamma_{xy} \ni e} |\gamma_{xy}| \pi(x) \pi(y).$$

Let $e = (e^-, e^+)$ be some edge from E where e^- and e^+ are two configurations from χ which differ by only one coordinate. It follows then that there exists $k, l \in \{1, 2, 3\}$ with $k \neq l$ such that:

$$e^- = (z_1, z_2, \dots, z_{i-1}, c_k, z_{i+1}, \dots, z_n) \quad \text{and} \quad e^+ = (z_1, z_2, \dots, z_{i-1}, c_l, z_{i+1}, \dots, z_n).$$

Without loss of generality we consider the case where $k = 1$ and $l = 2$. The transition probability for a transition from e^- to e^+ can then be computed. A short computation shows for $i = 1$:

$$\begin{aligned}
P(e^-, e^+) &= \frac{1}{n} \frac{\pi(c_2, z_2, \dots, z_n)}{\sum_{k=1}^3 \pi(c_k, z_2, \dots, z_n)} \\
&= \frac{1}{n} \frac{\frac{1}{Z_T} \exp \left\{ \frac{1}{T} \left(\mathbb{1}_{\{c_2=z_2\}} - \mathbb{1}_{\{c_2 \neq z_2\}} + \sum_{i=2}^{n-1} (\mathbb{1}_{\{z_i=z_{i+1}\}} - \mathbb{1}_{\{z_i \neq z_{i+1}\}}) \right) \right\}}{\sum_{k=1}^3 \frac{1}{Z_T} \exp \left\{ \frac{1}{T} \left(\mathbb{1}_{\{c_k=z_2\}} - \mathbb{1}_{\{c_k \neq z_2\}} + \sum_{i=2}^{n-1} (\mathbb{1}_{\{z_i=z_{i+1}\}} - \mathbb{1}_{\{z_i \neq z_{i+1}\}}) \right) \right\}} \\
&= \frac{1}{n} \frac{\exp \left\{ \frac{1}{T} (\mathbb{1}_{\{c_2=z_2\}} - \mathbb{1}_{\{c_2 \neq z_2\}}) \right\}}{\sum_{j=1}^3 \exp \left\{ \frac{1}{T} (\mathbb{1}_{\{c_j=z_2\}} - \mathbb{1}_{\{c_j \neq z_2\}}) \right\}}, \tag{2.3}
\end{aligned}$$

for $i = n$:

$$\begin{aligned}
 P(e^-, e^+) &= \frac{1}{n} \frac{\pi(z_1, z_2, \dots, z_{n-1}, c_2)}{\sum_{k=1}^3 \pi(z_1, z_2, \dots, z_{n-1}, c_j)} \\
 &= \frac{1}{n} \frac{\frac{1}{Z_T} \exp \left\{ \frac{1}{T} \left(\sum_{i=1}^{n-2} (\mathbb{1}_{\{z_i=z_{i+1}\}} - \mathbb{1}_{\{z_i \neq z_{i+1}\}} + \mathbb{1}_{\{z_{n-1}=c_2\}} - \mathbb{1}_{\{z_{n-1} \neq c_2\}}) \right) \right\}}{\sum_{k=1}^3 \frac{1}{Z_T} \exp \left\{ \frac{1}{T} \left(\sum_{i=1}^{n-2} (\mathbb{1}_{\{z_i=z_{i+1}\}} - \mathbb{1}_{\{z_i \neq z_{i+1}\}} + \mathbb{1}_{\{z_{n-1}=c_j\}} - \mathbb{1}_{\{z_{n-1} \neq c_j\}}) \right) \right\}} \\
 &= \frac{1}{n} \frac{\exp \left\{ \frac{1}{T} (\mathbb{1}_{\{z_{n-1}=c_2\}} - \mathbb{1}_{\{z_{n-1} \neq c_2\}}) \right\}}{\sum_{j=1}^3 \exp \left\{ \frac{1}{T} (\mathbb{1}_{\{z_{n-1}=c_j\}} - \mathbb{1}_{\{z_{n-1} \neq c_j\}}) \right\}} \tag{2.4}
 \end{aligned}$$

and for $i \in \{2, \dots, n-1\}$:

$$\begin{aligned}
 P(e^-, e^+) &= \frac{1}{n} \frac{\pi(z_1, z_2, \dots, z_{i-1}, c_2, z_{i+1}, \dots, z_n)}{\sum_{k=1}^3 \pi(z_1, z_2, \dots, z_{i-1}, c_j, z_{i+1}, \dots, z_n)} \\
 &= \frac{1}{n} \frac{\exp \left\{ \frac{1}{T} (\mathbb{1}_{\{z_{i-1}=c_2\}} - \mathbb{1}_{\{z_{i-1} \neq c_2\}} + \mathbb{1}_{\{c_2=z_{i+1}\}} - \mathbb{1}_{\{c_2 \neq z_{i+1}\}}) \right\}}{\sum_{j=1}^3 \exp \left\{ \frac{1}{T} (\mathbb{1}_{\{z_{i-1}=c_j\}} - \mathbb{1}_{\{z_{i-1} \neq c_j\}} + \mathbb{1}_{\{c_j=z_{i+1}\}} - \mathbb{1}_{\{c_j \neq z_{i+1}\}}) \right\}}. \tag{2.5}
 \end{aligned}$$

Then we turn to compute an upper bound of κ defined in equation (2.2) for each class of edges. The main conclusion of this chapter is given in the following theorem:

Theorem 2.2.1. [Helali 2018] *The second largest eigenvalue of the Gibbs sampler for the one-dimensional Ising model with three states satisfies :*

$$\beta_1 < 1 - 3 \times n^{-2} \frac{e^{-\frac{4}{T}}}{1 + 2e^{-\frac{4}{T}}} \tag{2.6}$$

The proof of this theorem is given in section 2.4. The above theorem can be generalized to the case of multiple colors where the state space is $\chi = \{c_1, \dots, c_N\}^n$ as follows:

Theorem 2.2.2. [Helali 2018] *The second largest eigenvalue of the Gibbs sampler for the one-dimensional Ising model with N states satisfies:*

$$\beta_1 < 1 - N \times n^{-2} \frac{e^{-\frac{4}{T}}}{1 + (N-1)e^{-\frac{4}{T}}}. \tag{2.7}$$

We give a sketch of the proof of this theorem in section 2.4.

To be able to quantify the convergence rate with Theorem 2.1.1 given by Diaconis and Stroock (see [5]) we must control the smallest eigenvalue in order to bound the second largest eigenvalue in absolute value. The answer is given in what follows:

2.3. Discussion

2.2.3 Bound for the absolute value of the second largest eigenvalue

A theorem proved by Ingrassia (see [20], Theorem 5.3) gives the following lower bound for the smallest eigenvalue:

$$\beta_{|\chi|-1} \geq -1 + \frac{2}{1 + (c-1)e^{\frac{\Delta}{T}}}.$$

where c is the number of nearest configuration - i.e.: configurations which differ by only one site - and Δ is defined in (1.5).

For the one-dimensional Ising model with three states, $c = 3$ and $\Delta = 2$. This yields for any natural number $n > 3/\sqrt{2}$

$$|\beta_{|\chi|-1}| \leq \left| -1 + \frac{2}{1 + 2e^{\frac{2}{T}}} \right| = 1 - \frac{2}{1 + 2e^{\frac{2}{T}}} < 1 - \frac{2}{3}e^{-\frac{2}{T}} < 1 - 3n^{-2} \frac{e^{-\frac{4}{T}}}{2e^{-\frac{4}{T}} + 1}.$$

In the general case where $\chi = \{c_1, \dots, c_N\}^n$, the parameter c is equal to N and Ingrassia's bound behaves for any natural number $n > N/\sqrt{2}$ as follows:

$$\begin{aligned} |\beta_{|\chi|-1}| &\leq \left| -1 + \frac{2}{1 + (N-1)e^{\frac{2}{T}}} \right| = 1 - \frac{2}{1 + (N-1)e^{\frac{2}{T}}} < 1 - \frac{2}{N}e^{-\frac{2}{T}} \\ &< 1 - Nn^{-2} \frac{e^{-\frac{4}{T}}}{1 + (N-1)e^{-\frac{4}{T}}}. \end{aligned}$$

The previous considerations prove the following corollary:

Corollary 2.2.3. *The upper bounds for β_1 given in Theorem 2.2.1 and Theorem 2.2.2 are also upper bounds for the absolute value of all eigenvalues $\{\beta_1, \dots, \beta_{|\chi|-1}\}$ of the Gibbs sampler for the one-dimensional Ising model with three states and more respectively.*

2.3 Discussion

Ingrassia (see [20]) gives the following upper bound for the second largest eigenvalue of the Gibbs sampler:

$$\beta_1 \leq 1 - \frac{Z_\Gamma}{b_\Gamma \gamma_\Gamma c |S|} e^{-\frac{m}{T}}.$$

In this expression Z_Γ is the normalizing constant, S is the lattice of sites, Γ is the collection of paths, γ_Γ is the maximum length of each path $\gamma_{xy} \in \Gamma$, b_Γ is the maximum number of paths containing any edge of Γ , c is the number of configurations that differ by only one site and m is the least total elevation gain of the Hamiltonian function in the sense which is described by Holley and Stroock (see [15]).

In our case, we have: $|S| = n$, $\gamma_\Gamma = n$, $b_\Gamma = 3^{n-1}$, $c = 3$, $Z_\Gamma \leq 3(1 + 2e^{-\frac{1}{2T}})^{n-1}$ and $m = 2$. It gives that:

$$\beta_1 \leq 1 - n^{-2} \left(\frac{1 + 2e^{-\frac{1}{2T}}}{3} \right)^{n-1} e^{-\frac{2}{T}}.$$

This upper bound differ to the result introduced in Theorem 2 by the multiplicative factor

$$\theta = \frac{e^{\frac{2}{T}} + 2e^{-\frac{2}{T}}}{3} \left(\frac{1 + 2e^{-\frac{1}{2T}}}{3} \right)^{n-1}.$$

By the fact that

$$1 + 2e^{-\frac{1}{2T}} < 3,$$

θ is small for large n .

Remark 2.3.1. The application of Ingrassia’s bound to the one-dimensional Ising model with multiple states gives:

$$\beta_1 \leq 1 - n^{-2} \left(\frac{1 + (N-1)e^{\frac{-1}{2T}}}{N} \right)^{n-1} e^{\frac{-2}{T}}$$

which differ to the result introduced in Theorem 3 by the factor $\tilde{\theta}$ defined as follow:

$$\tilde{\theta} = \frac{e^{\frac{2}{T}} + (N-1)e^{-\frac{2}{T}}}{N} \left(\frac{1 + (N-1)e^{-\frac{1}{2T}}}{N} \right)^{n-1}.$$

We notice that $\tilde{\theta}$ is small for large n because $1 + (N-1)e^{-\frac{1}{2T}} < N$.

2.4 Proofs of the main results

2.4.1 Proof of theorem 2.2.1

We have two principle cases:

a) If $i \neq \{1, n\}$: According to the selection of the paths in Section 2.2.1 any path γ_{xy} passing through the edge $e = (e^-, e^+)$ and connecting x with y , these extremities must have the following form:

$$x = (x_1, x_2, \dots, x_{i-1}, c_1, z_{i+1}, \dots, z_n) \quad \text{and} \quad y = (z_1, z_2, \dots, z_{i-1}, c_2, y_{i+1}, \dots, y_n).$$

This yields:

$$\begin{aligned} \pi(x) &= \pi(x_1, x_2, \dots, x_{i-1}, c_1, z_{i+1}, \dots, z_n) \\ &= \frac{1}{Z_T} \exp \left\{ \frac{1}{T} \left(\sum_{k'=1}^{i-2} (\mathbb{1}_{\{x_{k'}=x_{k'+1}\}} - \mathbb{1}_{\{x_{k'} \neq x_{k'+1}\}}) + \sum_{k'=i+1}^{n-1} (\mathbb{1}_{\{z_{k'}=z_{k'+1}\}} - \mathbb{1}_{\{z_{k'} \neq z_{k'+1}\}}) \right. \right. \\ &\quad \left. \left. + (\mathbb{1}_{\{x_{i-1}=c_1\}} - \mathbb{1}_{\{x_{i-1} \neq c_1\}} + \mathbb{1}_{\{c_1=z_{i+1}\}} - \mathbb{1}_{\{c_1 \neq z_{i+1}\}}) \right) \right\}, \end{aligned}$$

$$\begin{aligned} \pi(y) &= \pi(z_1, z_2, \dots, z_{i-1}, c_2, y_{i+1}, \dots, y_n) \\ &= \frac{1}{Z_T} \exp \left\{ \frac{1}{T} \left(\sum_{k'=1}^{i-2} (\mathbb{1}_{\{z_{k'}=z_{k'+1}\}} - \mathbb{1}_{\{z_{k'} \neq z_{k'+1}\}}) + \sum_{k'=i+1}^{n-1} (\mathbb{1}_{\{y_{k'}=y_{k'+1}\}} - \mathbb{1}_{\{y_{k'} \neq y_{k'+1}\}}) \right. \right. \\ &\quad \left. \left. + (\mathbb{1}_{\{z_{i-1}=c_2\}} - \mathbb{1}_{\{z_{i-1} \neq c_2\}} + \mathbb{1}_{\{c_2=y_{i+1}\}} - \mathbb{1}_{\{c_2 \neq y_{i+1}\}}) \right) \right\} \end{aligned}$$

2.4. Proofs of the main results

and

$$\begin{aligned}\pi(e^-) &= \pi(z_1, z_2, \dots, z_{i-1}, c_1, z_{i+1}, \dots, c_n) \\ &= \frac{1}{Z_T} \exp \left\{ \frac{1}{T} \left(\sum_{k'=1}^{i-2} (\mathbb{1}_{\{z_{k'}=z_{k'+1}\}} - \mathbb{1}_{\{z_{k'} \neq z_{k'+1}\}}) + \sum_{k'=i+1}^{n-1} (\mathbb{1}_{\{z_{k'}=z_{k'+1}\}} - \mathbb{1}_{\{z_{k'} \neq z_{k'+1}\}}) \right. \right. \\ &\quad \left. \left. + (\mathbb{1}_{\{z_{i-1}=c_1\}} - \mathbb{1}_{\{z_{i-1} \neq c_1\}} + \mathbb{1}_{\{c_1=z_{i+1}\}} - \mathbb{1}_{\{c_1 \neq z_{i+1}\}}) \right) \right\}.\end{aligned}$$

From the expressions given previously and equation (2.5), it follows that:

$$\begin{aligned}Q(e)^{-1}\pi(x)\pi(y) &= \frac{\pi(x)\pi(y)}{\pi(e^-)P(e^-, e^+)} \\ &= \frac{n}{Z_T} \left\{ 1 + \exp \left\{ \frac{1}{T} \left(-\mathbb{1}_{\{z_{i-1}=c_2\}} + \mathbb{1}_{\{z_{i-1} \neq c_2\}} - \mathbb{1}_{\{c_2=z_{i+1}\}} + \mathbb{1}_{\{c_2 \neq z_{i+1}\}} \right) \right\} \right. \\ &\quad \left. \times \sum_{j=1,3} \exp \left\{ \frac{1}{T} \left(\mathbb{1}_{\{z_{i-1}=c_j\}} - \mathbb{1}_{\{z_{i-1} \neq c_j\}} + \mathbb{1}_{\{c_j=z_{i+1}\}} - \mathbb{1}_{\{c_j \neq z_{i+1}\}} \right) \right\} \right\} \\ &\quad \times \exp \left\{ \frac{1}{T} \left(\sum_{k'=1}^{i-2} (\mathbb{1}_{\{x_{k'}=x_{k'+1}\}} - \mathbb{1}_{\{x_{k'} \neq x_{k'+1}\}}) + \sum_{k'=i+1}^{n-1} (\mathbb{1}_{\{y_{k'}=y_{k'+1}\}} - \mathbb{1}_{\{y_{k'} \neq y_{k'+1}\}}) \right) \right\} \\ &\quad \times \frac{\exp \left\{ \frac{1}{T} \left(\mathbb{1}_{\{x_{i-1}=c_1\}} - \mathbb{1}_{\{x_{i-1} \neq c_1\}} + \mathbb{1}_{\{z_{i-1}=c_2\}} - \mathbb{1}_{\{z_{i-1} \neq c_2\}} \right) \right\}}{\exp \left\{ \frac{1}{T} \left(\mathbb{1}_{\{z_{i-1}=c_1\}} - \mathbb{1}_{\{z_{i-1} \neq c_1\}} \right) \right\}} \\ &\quad \times \exp \left\{ \frac{1}{T} \left(\mathbb{1}_{\{c_2=y_{i+1}\}} - \mathbb{1}_{\{c_2 \neq y_{i+1}\}} \right) \right\}.\end{aligned}$$

We introduce the notation:

$$(x, c_l, y) := (x_1, x_2, \dots, x_{i-1}, c_l, y_{i+1}, \dots, y_{n-1}, y_n) \quad (2.8)$$

for $l \in \{1, 2, 3\}$. The previous expression becomes:

$$\begin{aligned}Q(e)^{-1}\pi(x)\pi(y) &= n\pi(x, c_1, y) \exp \left\{ \frac{1}{T} \left(-\mathbb{1}_{\{c_1=y_{i+1}\}} + \mathbb{1}_{\{c_1 \neq y_{i+1}\}} + \mathbb{1}_{\{z_{i-1}=c_2\}} - \mathbb{1}_{\{z_{i-1} \neq c_2\}} \right. \right. \\ &\quad \left. \left. + \mathbb{1}_{\{c_2=y_{i+1}\}} - \mathbb{1}_{\{c_2 \neq y_{i+1}\}} - \mathbb{1}_{\{z_{i-1}=c_1\}} + \mathbb{1}_{\{z_{i-1} \neq c_1\}} \right) \right\} \\ &\quad + n\pi(x, c_2, y) \exp \left\{ \frac{1}{T} \left(-\mathbb{1}_{\{x_{i-1}=c_2\}} + \mathbb{1}_{\{x_{i-1} \neq c_2\}} - \mathbb{1}_{\{z_{i-1}=c_1\}} + \mathbb{1}_{\{z_{i-1} \neq c_1\}} \right. \right. \\ &\quad \left. \left. + \mathbb{1}_{\{x_{i-1}=c_1\}} - \mathbb{1}_{\{x_{i-1} \neq c_1\}} - \mathbb{1}_{\{c_2=z_{i+1}\}} + \mathbb{1}_{\{c_2 \neq z_{i+1}\}} \right) \right\} \sum_{j=1,3} \exp \left\{ \frac{1}{T} \left(\mathbb{1}_{\{z_{i-1}=c_j\}} \right. \right. \\ &\quad \left. \left. - \mathbb{1}_{\{z_{i-1} \neq c_j\}} + \mathbb{1}_{\{c_j=z_{i+1}\}} - \mathbb{1}_{\{c_j \neq z_{i+1}\}} \right) \right\}. \\ &= n\alpha\pi(x, c_1, y) \exp \left\{ \frac{1}{T} \left(-\mathbb{1}_{\{c_1=y_{i+1}\}} + \mathbb{1}_{\{c_1 \neq y_{i+1}\}} + \mathbb{1}_{\{c_2=y_{i+1}\}} - \mathbb{1}_{\{c_2 \neq y_{i+1}\}} \right) \right\} \\ &\quad + n\beta\pi(x, c_2, y) \exp \left\{ \frac{1}{T} \left(-\mathbb{1}_{\{x_{i-1}=c_2\}} + \mathbb{1}_{\{x_{i-1} \neq c_2\}} + \mathbb{1}_{\{x_{i-1}=c_1\}} - \mathbb{1}_{\{x_{i-1} \neq c_1\}} \right) \right\}.\end{aligned} \quad (2.9)$$

where

1. $\alpha = \exp \left\{ \frac{1}{T} (\mathbb{1}_{\{z_{i-1}=c_2\}} - \mathbb{1}_{\{z_{i-1} \neq c_2\}} - \mathbb{1}_{\{z_{i-1}=c_1\}} + \mathbb{1}_{\{z_{i-1} \neq c_1\}}) \right\}$.
2. $\beta = \exp \left\{ \frac{1}{T} (-\mathbb{1}_{\{z_{i-1}=c_1\}} + \mathbb{1}_{\{z_{i-1} \neq c_1\}} - \mathbb{1}_{\{c_2=z_{i+1}\}} + \mathbb{1}_{\{c_2 \neq z_{i+1}\}}) \right\} \sum_{j=1,3} \exp \left\{ \frac{1}{T} (\mathbb{1}_{\{z_{i-1}=c_j\}} - \mathbb{1}_{\{z_{i-1} \neq c_j\}} + \mathbb{1}_{\{c_j=z_{i+1}\}} - \mathbb{1}_{\{c_j \neq z_{i+1}\}}) \right\}$.

Remark 2.4.1. From the notation in equation (2.8) we have

$$\bigcup_{(x,y): \gamma_{xy} \ni e} \left\{ (x, c_1, y), (x, c_2, y), (x, c_3, y) \right\} := \chi.$$

This yields

$$\begin{aligned} Q(e)^{-1} \sum_{(x,y): \gamma_{xy} \ni e} |\gamma_{xy}| \pi(x) \pi(y) &\leq \alpha n^2 \sum_{(x,y): \gamma_{xy} \ni e} \pi(x, c_1, y) \exp \left\{ \frac{1}{T} (-\mathbb{1}_{\{c_1=y_{i+1}\}} + \mathbb{1}_{\{c_1 \neq y_{i+1}\}} \right. \\ &\quad \left. + \mathbb{1}_{\{c_2=y_{i+1}\}} - \mathbb{1}_{\{c_2 \neq y_{i+1}\}}) \right\} + \beta n^2 \sum_{(x,y): \gamma_{xy} \ni e} \pi(x, c_2, y) \\ &\quad \times \exp \left\{ \frac{1}{T} (-\mathbb{1}_{\{x_{i-1}=c_2\}} + \mathbb{1}_{\{x_{i-1} \neq c_2\}} + \mathbb{1}_{\{x_{i-1}=c_1\}} - \mathbb{1}_{\{x_{i-1} \neq c_1\}}) \right\} \\ &= \alpha n^2 \sum_{w \in \chi: w_i=c_1} \pi(w) \exp \left\{ \frac{1}{T} (-\mathbb{1}_{\{c_1=w_{i+1}\}} + \mathbb{1}_{\{c_1 \neq w_{i+1}\}} \right. \\ &\quad \left. + \mathbb{1}_{\{c_2=w_{i+1}\}} - \mathbb{1}_{\{c_2 \neq w_{i+1}\}}) \right\} + \beta n^2 \sum_{w \in \chi: w_i=c_2} \pi(w) \\ &\quad \times \exp \left\{ \frac{1}{T} (-\mathbb{1}_{\{w_{i-1}=c_2\}} + \mathbb{1}_{\{w_{i-1} \neq c_2\}} + \mathbb{1}_{\{w_{i-1}=c_1\}} - \mathbb{1}_{\{w_{i-1} \neq c_1\}}) \right\} \\ &= \alpha n^2 A + \beta n^2 B. \end{aligned}$$

where α, β are defined previously and

- 1) $A = \sum_{w \in \chi: w_i=c_1} \pi(w) \exp \left\{ \frac{1}{T} (-\mathbb{1}_{\{c_1=w_{i+1}\}} + \mathbb{1}_{\{c_1 \neq w_{i+1}\}} + \mathbb{1}_{\{c_2=w_{i+1}\}} - \mathbb{1}_{\{c_2 \neq w_{i+1}\}}) \right\}$,
- 2) $B = \sum_{w \in \chi: w_i=c_2} \pi(w) \exp \left\{ \frac{1}{T} (-\mathbb{1}_{\{w_{i-1}=c_2\}} + \mathbb{1}_{\{w_{i-1} \neq c_2\}} + \mathbb{1}_{\{w_{i-1}=c_1\}} - \mathbb{1}_{\{w_{i-1} \neq c_1\}}) \right\}$.

We now turn to the computation of the two terms A and B on the right side of the previous equation separately:

In order to compute A we generalize some symmetry argument from Shiu and Chen (see [25]) to the three state case. In this situation we define three spaces $W_k = \{w \in \chi, w_i = c_1, w_{i+1} = c_k\}$ for $k \in \{1, 2, 3\}$. In order to compute their π measure we will establish some equations between those numbers $\pi(W_1), \pi(W_2)$ and $\pi(W_3)$.

We now establish some identification between the elements from W_1 with the elements of W_2 and W_3 .

For any vertex $\xi^1 \in W_1$, there exist a unique vertex $\xi^2 \in W_2$ such that:

- If $k < i$, $\xi_k^1 = \xi_k^2$.
- If $k > i + 1$, then:

2.4. Proofs of the main results

- If $\xi_k^1 = c_1$ then $\xi_k^2 = c_2$.
- If $\xi_k^1 = c_2$ then $\xi_k^2 = c_1$.
- If $\xi_k^1 = c_3$ then $\xi_k^2 = c_3$.

Let $\xi^1 \in W_1$ be a configuration of the form

$$\xi^1 = \{\xi_1^1, \dots, \xi_{i-1}^1, c_1, c_1, \xi_{i+2}^1, \dots, \xi_n^1\}$$

then there exist a unique configuration $\xi^2 \in W_2$ of the form

$$\xi^2 = \{\xi_1^1, \dots, \xi_{i-1}^1, c_1, c_2, \xi_{2+1}^2, \dots, \xi_n^2\}$$

From this argument we obtain:

$$\begin{aligned} \pi(\xi_1) &= \pi(\xi_1^1, \dots, \xi_{i-1}^1, c_1, c_1, \xi_{i+1}^1, \dots, \xi_n^1) \\ &= \frac{1}{Z_T} \exp \left\{ \frac{1}{T} \left(\sum_{k'=1}^{i-2} (\mathbb{1}_{\{\xi_{k'}^1 = \xi_{k'+1}^1\}} - \mathbb{1}_{\{\xi_{k'}^1 \neq \xi_{k'+1}^1\}}) + \sum_{k'=i+2}^{n-1} (\mathbb{1}_{\{\xi_{k'}^1 = \xi_{k'+1}^1\}} - \mathbb{1}_{\{\xi_{k'}^1 \neq z_{k'+1}^1\}}) \right) \right. \\ &\quad \left. + \mathbb{1}_{\{\xi_{i-1}^1 = c_1\}} - \mathbb{1}_{\{\xi_{i-1}^1 \neq c_1\}} + \mathbb{1}_{\{c_1 = c_1\}} - \mathbb{1}_{\{c_1 \neq c_1\}} + (\mathbb{1}_{\{c_1 = \xi_{i+2}^1\}} - \mathbb{1}_{\{c_1 \neq \xi_{i+2}^1\}}) \right\} \\ &= \frac{1}{Z_T} \exp \left\{ \frac{1}{T} \left(\sum_{k'=1}^{i-2} (\mathbb{1}_{\{\xi_{k'}^1 = \xi_{k'+1}^2\}} - \mathbb{1}_{\{\xi_{k'}^2 \neq \xi_{k'+1}^2\}}) + \sum_{k'=i+2}^{n-1} (\mathbb{1}_{\{\xi_{k'}^2 = \xi_{k'+1}^2\}} - \mathbb{1}_{\{\xi_{k'}^2 \neq z_{k'+1}^2\}}) \right) \right. \\ &\quad \left. + \mathbb{1}_{\{\xi_{i-1}^2 = c_1\}} - \mathbb{1}_{\{\xi_{i-1}^2 \neq c_1\}} + \mathbb{1}_{\{c_1 = c_2\}} - \mathbb{1}_{\{c_1 \neq c_2\}} + 2 + (\mathbb{1}_{\{c_2 = \xi_{i+2}^2\}} - \mathbb{1}_{\{c_2 \neq \xi_{i+2}^2\}}) \right\} \\ &= e^{\frac{2}{T}} \pi(\xi_1^2, \dots, \xi_{i-1}^2, c_1, c_2, \xi_{i+2}^2, \dots, \xi_n^2) \end{aligned} \quad (2.10)$$

Similarly, for any vertex $\xi^1 \in W_1$, there exist a unique vertex $\xi^3 \in W_3$ such that :

- If $k < i$, $\xi_k^1 = \xi_k^3$.
- If $k > i + 1$, then:
 - If $\xi_k^1 = c_1$ then $\xi_k^3 = c_3$.
 - If $\xi_k^1 = c_3$ then $\xi_k^3 = c_1$.
 - If $\xi_k^1 = c_2$ then $\xi_k^3 = c_2$.

With an argument like in equation (2.10) we obtain

$$\pi(\xi_1) = e^{\frac{2}{T}} \pi(\xi^3).$$

Those relations yield that $\pi(\xi^1) = e^{\frac{2}{T}} \pi(\xi^2) = e^{\frac{2}{T}} \pi(\xi^3)$. Therefore, we obtain:

$$\sum_{w \in W_1} \pi(w) = e^{\frac{2}{T}} \sum_{w \in W_2} \pi(w) = e^{\frac{2}{T}} \sum_{w \in W_3} \pi(w). \quad (2.11)$$

On the other hand we have also:

$$\sum_{w \in W_1} \pi(w) + \sum_{w \in W_2} \pi(w) + \sum_{w \in W_3} \pi(w) = \frac{1}{3}. \quad (2.12)$$

Chapter 2. The convergence rate of the Gibbs sampler for generalized 1–D Ising model

We put equation (2.11) in equation (2.12) then we obtain:

$$\sum_{w \in W_1} \pi(w) + e^{\frac{-2}{T}} \sum_{w \in W_1} \pi(w) + e^{\frac{-2}{T}} \sum_{w \in W_1} \pi(w) = \frac{1}{3}$$

then

$$\sum_{w \in W_1} \pi(w) \left(1 + e^{\frac{-2}{T}} + e^{\frac{-2}{T}}\right) = \frac{1}{3}.$$

An elementary computation gives

$$\sum_{w \in W_1} \pi(w) = \frac{1}{3(1 + 2e^{\frac{-2}{T}})}.$$

From equation (2.11) we deduce that

$$\sum_{w \in W_2} \pi(w) = \frac{e^{\frac{-2}{T}}}{3(1 + 2e^{\frac{-2}{T}})} \quad \text{and} \quad \sum_{w \in W_3} \pi(w) = \frac{e^{\frac{-2}{T}}}{3(1 + 2e^{\frac{-2}{T}})}.$$

The subdivision of the sum in A to three sums over the sets W_1 , W_2 and W_3 gives

$$A = A_{W_1} + A_{W_2} + A_{W_3} \tag{2.13}$$

where for $k = 1, 2, 3$ we have

$$A_{W_k} = \sum_{w \in \mathcal{X}: w_i = c_1, w_{i+1} = c_k} \pi(w) \exp \left\{ \frac{1}{T} \left(\mathbb{1}_{\{c_2 = w_{i+1}\}} - \mathbb{1}_{\{c_2 \neq w_{i+1}\}} - \mathbb{1}_{\{c_1 = w_{i+1}\}} + \mathbb{1}_{\{c_1 \neq w_{i+1}\}} \right) \right\}.$$

We turn now to compute the sums in equation (2.13) separately. For A_{W_1} we have:

$$\begin{aligned} A_{W_1} &= \sum_{w \in \mathcal{X}: w_i = c_1, w_{i+1} = c_1} \pi(w) \exp \left\{ \frac{1}{T} \left(\mathbb{1}_{\{c_2 = w_{i+1}\}} - \mathbb{1}_{\{c_2 \neq w_{i+1}\}} - \mathbb{1}_{\{c_1 = w_{i+1}\}} + \mathbb{1}_{\{c_1 \neq w_{i+1}\}} \right) \right\} \\ &= e^{\frac{-2}{T}} \sum_{w \in \mathcal{X}: w_i = c_1, w_{i+1} = c_1} \pi(w) \\ &= \frac{e^{\frac{-2}{T}}}{3(1 + 2e^{\frac{-2}{T}})}. \end{aligned} \tag{2.14}$$

Similarly, we have

$$\begin{aligned} A_{W_2} &= \sum_{w \in \mathcal{X}: w_i = c_1, w_{i+1} = c_2} \pi(w) \exp \left\{ \frac{1}{T} \left(\mathbb{1}_{\{c_2 = w_{i+1}\}} - \mathbb{1}_{\{c_2 \neq w_{i+1}\}} - \mathbb{1}_{\{c_1 = w_{i+1}\}} + \mathbb{1}_{\{c_1 \neq w_{i+1}\}} \right) \right\} \\ &= e^{\frac{2}{T}} \sum_{w \in \mathcal{X}: w_i = c_1, w_{i+1} = c_2} \pi(w) \\ &= \frac{1}{3(1 + 2e^{\frac{-2}{T}})} \end{aligned} \tag{2.15}$$

2.4. Proofs of the main results

and

$$\begin{aligned}
A_{W_3} &= \sum_{w \in \chi: w_i=c_1, w_{i+1}=c_3} \pi(w) \exp \left\{ \frac{1}{T} (\mathbb{1}_{\{c_2=w_{i+1}\}} - \mathbb{1}_{\{c_2 \neq w_{i+1}\}} - \mathbb{1}_{\{c_1=w_{i+1}\}} + \mathbb{1}_{\{c_1 \neq w_{i+1}\}}) \right\} \\
&= e^{\frac{-2}{T}} \sum_{w \in \chi: w_i=c_1, w_{i+1}=c_3} \pi(w) \\
&= \frac{e^{\frac{-2}{T}}}{3(1+2e^{\frac{-2}{T}})}. \tag{2.16}
\end{aligned}$$

By regrouping the results in equations (2.14), (2.15) and (2.16) we obtain

$$\begin{aligned}
A &= A_{W_1} + A_{W_2} + A_{W_3} \\
&= \frac{e^{\frac{-2}{T}}}{3(1+2e^{\frac{-2}{T}})} + \frac{1}{3(1+2e^{\frac{-2}{T}})} + \frac{e^{\frac{-2}{T}}}{3(1+2e^{\frac{-2}{T}})} \\
&= \frac{1+2e^{\frac{-2}{T}}}{3 \times (1+2e^{\frac{-2}{T}})} \\
&= \frac{1}{3}. \tag{2.17}
\end{aligned}$$

We turn now to the computation of the expression of B . In order to compute it, we generalize some symmetry argument from Shiu and Chen (see [25]) to the three state case. In this situation we define three spaces $W'_k = \{w \in \chi, w_i = c_2, w_{i-1} = c_k\}$ for $k \in \{1, 2, 3\}$. In order to compute their π measure we will establish some equations between those numbers $\pi(W'_1), \pi(W'_2)$ and $\pi(W'_3)$.

We now establish some identification between the elements from W'_1 with the elements of W'_2 and W'_3 .

For any vertex $\xi^1 \in W'_1$, there exists a unique vertex $\xi^2 \in W'_2$ such that:

- If $k > i$, $\xi_k^1 = \xi_k^2$.
- If $k < i - 1$, then:
 - If $\xi_k^1 = c_1$ then $\xi_k^2 = c_2$.
 - If $\xi_k^1 = c_2$ then $\xi_k^2 = c_1$.
 - If $\xi_k^1 = c_3$ then $\xi_k^2 = c_3$.

Let $\xi^1 \in W'_1$ be a configuration of the form

$$\xi^1 = \{\xi_1^1, \dots, \xi_{i-2}^1, c_1, c_2, \xi_{i+1}^1, \dots, \xi_n^1\}$$

then there exists a unique configuration $\xi^2 \in W'_2$ of the form

$$\xi^2 = \{\xi_1^2, \dots, \xi_{i-2}^2, c_2, c_2, \xi_{i+1}^1, \dots, \xi_n^1\}.$$

From this argument we obtain:

$$\begin{aligned}
\pi(\xi^1) &= \pi(\xi_1^1, \dots, \xi_{i-2}^1, c_1, c_2, \xi_{i+1}^1, \dots, \xi_n^1) \\
&= \frac{1}{Z_T} \exp \left\{ \frac{1}{T} \left(\sum_{k'=1}^{i-3} (\mathbb{1}_{\{\xi_{k'}^1 = \xi_{k'+1}^1\}} - \mathbb{1}_{\{\xi_{k'}^1 \neq \xi_{k'+1}^1\}}) + \sum_{k'=i+1}^{n-1} (\mathbb{1}_{\{\xi_{k'}^1 = \xi_{k'+1}^1\}} - \mathbb{1}_{\{\xi_{k'}^1 \neq \xi_{k'+1}^1\}}) \right) \right. \\
&\quad \left. + \mathbb{1}_{\{\xi_{i-2}^1 = c_1\}} - \mathbb{1}_{\{\xi_{i-2}^1 \neq c_1\}} + \mathbb{1}_{\{c_1 = c_2\}} - \mathbb{1}_{\{c_1 \neq c_2\}} + (\mathbb{1}_{\{c_2 = \xi_{i+1}^1\}} - \mathbb{1}_{\{c_2 \neq \xi_{i+1}^1\}}) \right\} \\
&= \frac{1}{Z_T} \exp \left\{ \frac{1}{T} \left(\sum_{k'=1}^{i-3} (\mathbb{1}_{\{\xi_{k'}^2 = \xi_{k'+1}^2\}} - \mathbb{1}_{\{\xi_{k'}^2 \neq \xi_{k'+1}^2\}}) + \sum_{k'=i+1}^{n-1} (\mathbb{1}_{\{\xi_{k'}^2 = \xi_{k'+1}^2\}} - \mathbb{1}_{\{\xi_{k'}^2 \neq \xi_{k'+1}^2\}}) \right) \right. \\
&\quad \left. + \mathbb{1}_{\{\xi_{i-2}^1 = c_2\}} - \mathbb{1}_{\{\xi_{i-2}^1 \neq c_2\}} + \mathbb{1}_{\{c_2 = c_2\}} - \mathbb{1}_{\{c_2 \neq c_2\}} - 2 + (\mathbb{1}_{\{c_2 = \xi_{i+1}^1\}} - \mathbb{1}_{\{c_2 \neq \xi_{i+1}^1\}}) \right\} \\
&= e^{-\frac{2}{T}} \pi(\xi_1^2, \dots, \xi_{i-2}^2, c_2, c_2, \xi_{i+1}^2, \dots, \xi_n^2). \tag{2.18}
\end{aligned}$$

Similarly, for any vertex $\xi^2 \in W_2$, there exists a unique vertex $\xi^3 \in W_3$ such that:

- If $k > i$, $\xi_k^2 = \xi_k^3$.
- If $k < i - 1$, then:
 - If $\xi_k^2 = c_2$ then $\xi_k^3 = c_3$.
 - If $\xi_k^2 = c_3$ then $\xi_k^3 = c_2$.
 - If $\xi_k^2 = c_1$ then $\xi_k^3 = c_1$.

With an argument like in equation (2.18) we obtain

$$\pi(\xi^3) = e^{-\frac{2}{T}} \pi(\xi^2).$$

Those relations yield that $e^{-\frac{2}{T}} \pi(\xi^1) = \pi(\xi^2) = e^{-\frac{2}{T}} \pi(\xi^3)$. Therefore, we obtain:

$$e^{-\frac{2}{T}} \sum_{w \in W'_1} \pi(w) = \sum_{w \in W'_2} \pi(w) = e^{-\frac{2}{T}} \sum_{w \in W'_3} \pi(w). \tag{2.19}$$

On the other hand we have also:

$$\sum_{w \in W_1} \pi(w) + \sum_{w \in W_2} \pi(w) + \sum_{w \in W_3} \pi(w) = \frac{1}{3}. \tag{2.20}$$

We put equation (2.19) in equation (2.20) then we obtain:

$$\sum_{w \in W'_2} \pi(w) + e^{-\frac{2}{T}} \sum_{w \in W'_2} \pi(w) + e^{-\frac{2}{T}} \sum_{w \in W'_2} \pi(w) = \frac{1}{3}$$

then

$$\sum_{w \in W'_2} \pi(w) \left(1 + e^{-\frac{2}{T}} + e^{-\frac{2}{T}} \right) = \frac{1}{3}.$$

2.4. Proofs of the main results

An elementary computation gives

$$\sum_{w \in W'_2} \pi(w) = \frac{1}{3(1 + 2e^{-\frac{2}{T}})}.$$

From equation (2.19) we deduce that

$$\sum_{w \in W'_1} \pi(w) = \frac{e^{-\frac{2}{T}}}{3(1 + 2e^{-\frac{2}{T}})} \quad \text{and} \quad \sum_{w \in W'_3} \pi(w) = \frac{e^{-\frac{2}{T}}}{3(1 + 2e^{-\frac{2}{T}})}.$$

The subdivision of the sum in B to three sums over the sets W'_1 , W'_2 and W'_3 gives

$$B = B_{W'_1} + B_{W'_2} + B_{W'_3} \quad (2.21)$$

where for $k = 1, 2, 3$ we have

$$B_{W'_k} = \sum_{w \in \mathcal{X}: w_i = c_1, w_{i-1} = c_k} \pi(w) \exp \left\{ \frac{1}{T} (\mathbb{1}_{\{w_{i-1} = c_1\}} - \mathbb{1}_{\{w_{i-1} \neq c_1\}} - \mathbb{1}_{\{w_{i-1} = c_2\}} + \mathbb{1}_{\{w_{i-1} \neq c_2\}}) \right\}$$

We turn now to compute the sums in equation (2.21) separately. For $B_{W'_1}$ we have:

$$\begin{aligned} B_{W'_1} &= \sum_{w \in \mathcal{X}: w_i = c_2, w_{i-1} = c_1} \pi(w) \exp \left\{ \frac{1}{T} (\mathbb{1}_{\{w_{i-1} = c_1\}} - \mathbb{1}_{\{w_{i-1} \neq c_1\}} - \mathbb{1}_{\{w_{i-1} = c_2\}} + \mathbb{1}_{\{w_{i-1} \neq c_2\}}) \right\} \\ &= e^{\frac{2}{T}} \sum_{w \in \mathcal{X}: w_i = c_2, w_{i-1} = c_1} \pi(w) \\ &= \frac{1}{3(1 + 2e^{-\frac{2}{T}})}. \end{aligned} \quad (2.22)$$

Similarly, we have

$$\begin{aligned} B_{W'_2} &= \sum_{w \in \mathcal{X}: w_i = c_2, w_{i-1} = c_2} \pi(w) \exp \left\{ \frac{1}{T} (\mathbb{1}_{\{w_{i-1} = c_1\}} - \mathbb{1}_{\{w_{i-1} \neq c_1\}} - \mathbb{1}_{\{w_{i-1} = c_2\}} + \mathbb{1}_{\{w_{i-1} \neq c_2\}}) \right\} \\ &= e^{-\frac{2}{T}} \sum_{w \in \mathcal{X}: w_i = c_2, w_{i-1} = c_2} \pi(w) \\ &= \frac{e^{-\frac{2}{T}}}{3(1 + 2e^{-\frac{2}{T}})} \end{aligned} \quad (2.23)$$

and

$$\begin{aligned} B_{W'_3} &= \sum_{w \in \mathcal{X}: w_i = c_2, w_{i-1} = c_3} \pi(w) \exp \left\{ \frac{1}{T} (\mathbb{1}_{\{w_{i-1} = c_1\}} - \mathbb{1}_{\{w_{i-1} \neq c_1\}} - \mathbb{1}_{\{w_{i-1} = c_2\}} + \mathbb{1}_{\{w_{i-1} \neq c_2\}}) \right\} \\ &= \sum_{w \in \mathcal{X}: w_i = c_2, w_{i-1} = c_3} \pi(w) \\ &= \frac{e^{-\frac{2}{T}}}{3(1 + 2e^{-\frac{2}{T}})}. \end{aligned} \quad (2.24)$$

By regrouping the results in equations (2.22), (2.23) and (2.24) we obtain

$$\begin{aligned}
 B &= B_{W'_1} + B_{W'_2} + B_{W'_3} \\
 &= \frac{1}{3(1 + 2e^{\frac{-2}{T}})} + \frac{e^{\frac{-2}{T}}}{3(1 + 2e^{\frac{-2}{T}})} + \frac{e^{\frac{-2}{T}}}{3(1 + 2e^{\frac{-2}{T}})} \\
 &= \frac{1 + 2e^{\frac{-2}{T}}}{3 \times (1 + 2e^{\frac{-2}{T}})} \\
 &= \frac{1}{3}.
 \end{aligned} \tag{2.25}$$

Using equations (2.17) and (2.25) we obtain:

$$Q(e)^{-1} \sum_{(x,y): \gamma_{xy} \ni e} |\gamma_{xy}| \pi(x) \pi(y) \leq \alpha n^2 A + \beta n^2 B = \frac{n^2}{3} (\alpha + \beta).$$

The sum $\alpha + \beta$ takes the following possible values:

- $e^{\frac{2}{T}} + 2e^{\frac{-2}{T}}$ if $z_{i-1} = c_1$ and $z_{i+1} = c_1$,
- $2e^{\frac{-2}{T}} + e^{\frac{-4}{T}}$ if $z_{i-1} = c_1$ and $z_{i+1} = c_2$,
- $2 + e^{\frac{-2}{T}}$ if $z_{i-1} = c_1$ and $z_{i+1} = c_3$,
- $e^{\frac{2}{T}} + 2$ if $z_{i-1} = c_2$ and $z_{i+1} = c_1$,
- $e^{\frac{2}{T}} + 2e^{\frac{-2}{T}}$ if $z_{i-1} = c_2$ and $z_{i+1} = c_2$,
- $2e^{\frac{2}{T}} + 1$ if $z_{i-1} = c_2$ and $z_{i+1} = c_3$,
- $2e^{\frac{2}{T}} + 1$ if $z_{i-1} = c_3$ and $z_{i+1} = c_1$,
- $2 + e^{\frac{-2}{T}}$ if $z_{i-1} = c_3$ and $z_{i+1} = c_2$,
- $2 + e^{\frac{4}{T}}$ if $z_{i-1} = c_3$ and $z_{i+1} = c_3$.

By the fact that $\cosh\left(\frac{2}{T}\right) \geq 1$ we have $2 + e^{\frac{4}{T}} \geq 2e^{\frac{2}{T}} + 1$. Then, the worst value of $\alpha + \beta$ is obtained when $z_{i-1} = c_3 = z_{i+1}$. In this case we have:

$$Q(e)^{-1} \sum_{(x,y): \gamma_{xy} \ni e} |\gamma_{xy}| \pi(x) \pi(y) \leq \frac{n^2}{3} (2 + e^{\frac{4}{T}}). \tag{2.26}$$

b) If $i = 1$ then the configurations x and e^- coincide, from equations (2.2) and (2.3) we obtain:

$$\begin{aligned}
 \kappa &= \max_{e \in E} Q(e)^{-1} \sum_{\gamma_{xy} \ni e} |\gamma_{xy}| \pi(x) \pi(y) \leq n \sum_{\gamma_{xy} \ni e} \frac{\pi(x) \pi(y)}{\frac{1}{n} \pi(e^-) P(e^-, e^+)} \\
 &\leq n^2 (1 + 2e^{\frac{2}{T}}) \sum_{y \in \mathcal{X}: y_1 = c_2} \pi(y) = \frac{n^2}{3} (2 + e^{\frac{2}{T}}).
 \end{aligned} \tag{2.27}$$

2.4. Proofs of the main results

For $i = n$, then the configurations y and e^+ coincide and some computation gives a similar result as in equation (2.27).

By regrouping the results in equations (2.26) and (2.27) we obtain an upper bound for κ defined in equation (2.2) as follows:

$$\kappa = \max_{e \in E} Q(e)^{-1} \sum_{\gamma_{xy} \ni e} |\gamma_{xy}| \pi(x) \pi(y) \leq \frac{n^2}{3} (2 + e^{\frac{4}{T}}). \quad (2.28)$$

Remark 2.4.2. The above computation was done for the situation where $x_i = c_1$ and $y_i = c_2$. Obviously we obtain the same result in the other cases, ie.: $x_i = c_1$ and $y_i = c_3$, etc \dots .

Finally, from inequalities (2.1) and (2.28) we obtain an upper bound for β_1 which finish the proof.

2.4.2 Proof of theorem 2.2.2

We follow the same approach as in the case where $\chi = \{c_1, c_2, c_3\}^n$. We define an edge $e = (e^-, e^+)$ as in section 2.2.1 and then distinguish two cases:

a) For $i \neq \{1, n\}$, we pass in equations (2.3), (2.8) and (2.9) from the case of three colors where $\chi = \{c_1, c_2, c_3\}^n$ to the case of multiple colors where $\chi = \{c_1, \dots, c_N\}^n$. This yields:

$$\begin{aligned} Q(e)^{-1} \sum_{(x,y): \gamma_{xy} \ni e} |\gamma_{xy}| \pi(x) \pi(y) &\leq \alpha n^2 \sum_{w \in \chi: w_i = c_1} \pi(w) \exp \left\{ \frac{1}{T} \left(- \mathbb{1}_{\{c_1 = w_{i+1}\}} + \mathbb{1}_{\{c_1 \neq w_{i+1}\}} \right. \right. \\ &\quad \left. \left. + \mathbb{1}_{\{c_2 = w_{i+1}\}} - \mathbb{1}_{\{c_2 \neq w_{i+1}\}} \right) \right\} \\ &\quad + \beta n^2 \sum_{w \in \chi: w_i = c_2} \pi(w) \exp \left\{ \frac{1}{T} \left(- \mathbb{1}_{\{w_{i-1} = c_2\}} + \mathbb{1}_{\{w_{i-1} \neq c_2\}} \right. \right. \\ &\quad \left. \left. + \mathbb{1}_{\{w_{i-1} = c_1\}} - \mathbb{1}_{\{w_{i-1} \neq c_1\}} \right) \right\} \\ &= \alpha n^2 A' + \beta n^2 B'. \end{aligned}$$

where

1. $\alpha = \exp \left\{ \frac{1}{T} \left(\mathbb{1}_{\{z_{i-1} = c_2\}} - \mathbb{1}_{\{z_{i-1} \neq c_2\}} - \mathbb{1}_{\{z_{i-1} = c_1\}} + \mathbb{1}_{\{z_{i-1} \neq c_1\}} \right) \right\}$,
2. $\beta = \exp \left\{ \frac{1}{T} \left(- \mathbb{1}_{\{z_{i-1} = c_1\}} + \mathbb{1}_{\{z_{i-1} \neq c_1\}} - \mathbb{1}_{\{c_2 = z_{i+1}\}} + \mathbb{1}_{\{c_2 \neq z_{i+1}\}} \right) \right\}$
 $\times \sum_{\substack{j=1 \\ j \neq 2}}^N \exp \left\{ \frac{1}{T} \left(\mathbb{1}_{\{z_{i-1} = c_j\}} - \mathbb{1}_{\{z_{i-1} \neq c_j\}} + \mathbb{1}_{\{c_j = z_{i+1}\}} - \mathbb{1}_{\{c_j \neq z_{i+1}\}} \right) \right\}$.

To compute the term A' we define for $k = \{1, \dots, N\}$ the spaces

$$W_k = \{w \in \chi : w_i = c_1, w_{i+1} = c_k\}$$

and we consider some symmetry arguments as above:

For any vertex $\xi^1 \in W_1$ there exists a unique $\xi^l \in W_l$ where $l \in \{2, \dots, N\}$ such that:

- If $k < i$, $\xi_k^1 = \xi_k^l$.
- If $k > i + 1$, then:
 - If $\xi_k^1 = c_1$ then $\xi_k^l = c_l$.
 - If $\xi_k^1 = c_l$ then $\xi_k^l = c_1$.
 - If $\xi_k^1 = c_{\tilde{l}}$ where $\tilde{l} \neq \{1, l\}$ then $\xi_k^l = c_{\tilde{l}}$.

Then equations (2.11) and (2.12) becomes in the case of N colors:

$$\sum_{w \in W_1} \pi(w) = e^{\frac{2}{T}} \sum_{w \in W_2} \pi(w) = \dots = e^{\frac{2}{T}} \sum_{w \in W_N} \pi(w), \quad (2.29)$$

$$\sum_{w \in W_1} \pi(w) + \dots + \sum_{w \in W_N} \pi(w) = \frac{1}{N}. \quad (2.30)$$

From equations (2.29) and (2.30) and with the same tricks used to obtain equation (2.17) we get:

$$A' = \sum_{w \in \mathcal{X}: w_i = c_1} \pi(w) \exp \left\{ \frac{1}{T} (\mathbb{1}_{\{c_2 = w_{i+1}\}} - \mathbb{1}_{\{c_2 \neq w_{i+1}\}} - \mathbb{1}_{\{c_1 = w_{i+1}\}} + \mathbb{1}_{\{c_1 \neq w_{i+1}\}}) \right\} = \frac{1}{N}. \quad (2.31)$$

With the same tricks, we obtain the same result for B'

$$B' = \sum_{w \in \mathcal{X}: w_i = c_2} \pi(w) \exp \left\{ \frac{1}{T} (\mathbb{1}_{\{w_{i-1} = c_1\}} - \mathbb{1}_{\{w_{i-1} \neq c_1\}} - \mathbb{1}_{\{w_{i-1} = c_2\}} + \mathbb{1}_{\{w_{i-1} \neq c_2\}}) \right\} = \frac{1}{N}. \quad (2.32)$$

The application of equations (2.31) and (2.32) gives:

$$Q(e)^{-1} \sum_{(x,y): \gamma_{xy} \ni e} |\gamma_{xy}| \pi(x) \pi(y) \leq \alpha n^2 A' + \beta n^2 B' = \frac{n^2}{N} (\alpha + \beta).$$

The worst value of $\alpha + \beta$ is obtained when $z_{i-1} = c_l = z_{i+1}$ for $l \in \{3, \dots, N\}$. In this case we have:

$$Q(e)^{-1} \sum_{(x,y): \gamma_{xy} \ni e} |\gamma_{xy}| \pi(x) \pi(y) \leq \frac{n^2}{N} (N - 1 + e^{\frac{4}{T}}). \quad (2.33)$$

b) For the boundary cases, when $i = 1$ equation (2.27) becomes

$$Q(e)^{-1} \sum_{\gamma_{xy} \ni e} |\gamma_{xy}| \pi(x) \pi(y) \leq \frac{n^2}{N} (N - 1 + e^{\frac{2}{T}}). \quad (2.34)$$

Also, we obtain a similar result for the case where $i = n$.

Equation (2.33) and (2.34) together give an upper bound of κ defined in (2.2) as follow:

$$\kappa = Q(e)^{-1} \sum_{(x,y): \gamma_{xy} \ni e} |\gamma_{xy}| \pi(x) \pi(y) \leq \frac{n^2}{N} (N - 1 + e^{\frac{4}{T}}). \quad (2.35)$$

Finally, we apply the upper bound given in (2.35) in equation (2.1) to finish the proof.

Chapter 3

The Convergence rate of the Gibbs sampler for the 2–D Ising model via a geometric bound

We study the geometric bound introduced by Diaconis and Stroock (1991) of the Gibbs sampler for the two-dimensional Ising model with free boundary condition. The obtained result generalizes the method proposed by Shiu and Chen (2015) from dimension one to dimension two. The new bound presented in this chapter can be found in Franke and Helali (2018) and it improves the result given by Ingrassia (1994).

Nous étudions la borne géométrique introduite par Diaconis et Stroock (1991) de l'échantillonneur de Gibbs pour le modèle d'Ising deux-dimensionnel. Le résultat obtenu généralise la méthode proposée par Shiu et Chen (2015) du dimension un à la dimension deux. La nouvelle borne présentée dans ce chapitre est donnée dans Franke et Helali (2018) et elle améliore le résultat donné par Ingrassia (1994).

3.1 Introduction

The Ising model is the most basic model in statistical mechanics having non trivial interaction. It has many applications in pattern analysis, molecular biology and image analysis. The distribution of the two-dimensional square lattice Ising model is:

$$\pi(x) = \frac{1}{Z_T} \exp \left\{ \frac{1}{T} \left(\sum_{j=1}^n \sum_{i=1}^{n-1} x_i^j x_{i+1}^j + \sum_{i=1}^n \sum_{j=1}^{n-1} x_i^j x_i^{j+1} \right) \right\} \quad \forall x = (x_i^j)_{1 \leq i, j \leq n} \in \chi,$$

where $\chi = \{-1, 1\}^{n^2}$ is the state space, T is a positive real representing the temperature and

$$Z_T = \sum_{x \in \chi} \exp \left\{ \frac{1}{T} \left(\sum_{j=1}^n \sum_{i=1}^{n-1} x_i^j x_{i+1}^j + \sum_{i=1}^n \sum_{j=1}^{n-1} x_i^j x_i^{j+1} \right) \right\}$$

is the normalizing constant. Monte Carlo Markov chain *MCMC* method is a very useful technique to draw samples from the Ising model. The Gibbs sampler introduced by Geman and Geman (see [11]) and the Metropolis-Hastings algorithm introduced by Metropolis et al. (see [21]) and Hastings (see [13]) are the most popular Monte Carlo Markov chain methods used in this context. Those two algorithms use an aperiodic and irreducible Markov chain which is reversible with respect to the measure π . The reversibility is contained in the following *detailed balance equation*:

$$Q(x, y) := \pi(x)P(x, y) = \pi(y)P(y, x) := Q(y, x) \quad \forall x, y \in \chi.$$

Under the previous conditions (aperiodicity, irreducibility and reversibility) the measure π is the unique invariant measure for the matrix P . The eigenvalues of P can be arranged as follows:

$$1 = \beta_0 > \beta_1 \geq \beta_2 \geq \dots \geq \beta_{|\chi|-1} > -1.$$

There are two known criteria to measure the convergence rate of a Markov chain: On one hand, many researcher as Sinclair (see [26]), Frigessi et al. (see [10]), Ingrassia (see [20]), Chen et al. (see [3]) and Chen and Hwang (see [4]) use the asymptotic variance to study the convergence rate of the MCMC algorithms. On the other hand, in theorem 1.2.1 Diaconis and Stroock (see [5]) use the total variation distance to quantify the convergence of the MCMC algorithms to their stationary distribution.

Many authors as Sinclair and Jerrum (see [27]), Diaconis and Stroock (see [5]) and Sinclair (see [26]) introduced bounds for the eigenvalues β_1 and $\beta_{|\chi|-1}$. Ingrassia (see [20]) shows that the result of Diaconis and Stroock (see [5]) leads to the tightest bound. This motivates us to have a closer look at explicit bounds that one can obtain from the approach of Diaconis and Stroock.

In order to do so, let us first remind the classical result of Diaconis and Stroock (see [5]): Let $\mathcal{G}(P) = (\chi, E)$ be the graph constructed with the dynamic P where the state space χ is the vertex set and $E = \{(x, y) \mid P(x, y) > 0\}$ is the set of edges. Then, for each pair of distinct configurations x and y we choose a path γ_{xy} in $\mathcal{G}(P)$ linking x to y . The irreducibility of the matrix P guarantees that such paths exist. Finally, we define the set $\Gamma = \{\gamma_{xy} : x, y \in \chi\}$. Note that only one path γ_{xy} for each pair of configurations x, y is chosen. The second largest eigenvalue β_1 is then bounded from above as follows:

$$\beta_1 \leq 1 - \frac{1}{\kappa} \tag{3.1}$$

3.1. Introduction

where

$$\kappa = \max_{e \in E} Q(e)^{-1} \sum_{\gamma_{xy} \ni e} |\gamma_{xy}| \pi(x) \pi(y), \quad (3.2)$$

and where $|\gamma_{xy}|$ designates the length of the path γ_{xy} .

In equation (3.2) the maximum is over all directed edges in the graph $\mathcal{G}(P)$ and the sum is over all paths from the set Γ passing through the fixed edge e . It is clear that κ measures the bottlenecks (charged edges) in the graph $\mathcal{G}(P)$. We notice that a small κ gives a better result. So, on one hand we should choose the shortest paths to link some pair of configurations x, y and on the other hand we must avoid that many paths pass through the same edge in order to obtain the tightest bound for β_1 .

In this chapter we study the Gibbs sampler which chooses a random coordinate to be updated according to the conditional probabilities given the other coordinates.

The associated matrix for the two-dimensional square lattice Ising model is

$$P(x, y) = \begin{cases} \frac{1}{n^2} \pi(y_i^j | x) & \text{if } d(x, y) = 1 \text{ and } x_i^j \neq y_i^j \\ 1 - \frac{1}{n^2} \sum_{i,j=1}^n \pi(y_i^j | x) & \text{if } d(x, y) = 0 \\ 0 & \text{else} \end{cases}$$

where

- $\pi(y_i^j | x) = \frac{\pi(x_1^1, \dots, x_{i-1}^j, y_i^j, x_{i+1}^j, \dots, x_n^n)}{\pi(x_1^1, \dots, x_{i-1}^j, y_i^j, x_{i+1}^j, \dots, x_n^n) + \pi(x_1^1, \dots, x_{i-1}^j, -y_i^j, x_{i+1}^j, \dots, x_n^n)}$;
- $d(x, y) = \#\{i, x_i^j \neq y_i^j \text{ for } i, j = 1, \dots, n\}$ designates the number of sites that differ between two configurations x and y .

We apply the bound for β_1 introduced by Diaconis and Stroock (see [5]) to the Gibbs sampler for the two-dimensional Ising model with two states. The computation is based on some method introduced by Shiu and Chen (see [25]). In their paper they treat the one-dimensional lattice case.

The one-dimensional lattice case with multiple states which is also called Potts model was investigated by Helali with the same techniques (see [14]).

In the first section, we first define a path for each pair of configurations (x, y) from the state space $\chi = \{-1, +1\}^{n^2}$. Then, we turn to compute explicitly the expression:

$$\max_{e \in E} Q(e)^{-1} \sum_{(x,y): \gamma_{xy} \ni e} |\gamma_{xy}| \pi(x) \pi(y). \quad (3.3)$$

In Proposition 3.2.2, we give an upper bound for the expression in (3.3) for different classes of edges (interior, corners, etc ...) in the square lattice $\{1, \dots, n\}^2$. We notice that it is difficult to complete exact computations for the bounds given in Proposition 3.2.2. In Theorem 3.2.3 we present a bound for β_1 which results from some rough estimation of the terms from Proposition 3.2.2. To be able to use Theorem 3.1.1 we are referring to

Ingrassia (see Theorem 5.3 in [20]) which gives a lower bound for the smallest eigenvalue $\beta_{|\chi|-1}$. The main theorem is obtained by the fact that $|\beta_{|\chi|-1}|$ is smaller than the upper bound of β_1 . In the third section we compare our result with existing bounds from the literature. We notice that the main result of this chapter generalizes the one introduced by Shiu and Chen in [25] to higher dimension. It also improves the result of Ingrassia (see [20]). The fifth section contains the proofs for the main results and in the sixth section we give an asymptotic study for the upper bound for β_1 .

3.2 Main result

3.2.1 Selection of paths

To be able to use the result of Diaconis and Stroock (see [5]) in the computation of the bound for the second largest eigenvalue, we should fix a collection of paths linking any x to any y from χ . To get the best possible result, we should use the shortest paths γ_{xy} to obtain smallest possible κ which gives smaller upper bound for β_1 . An edge of the graph $\mathcal{G}(P)$ takes the following form: $e = (e^-, e^+)$ where

$$e^- = \begin{pmatrix} z_1^1 & \cdots & \cdots & z_p^1 & \cdots & \cdots & z_n^1 \\ \vdots & & & \vdots & & & \vdots \\ z_1^{q-1} & \cdots & z_{p-1}^{q-1} & z_p^{q-1} & z_{p+1}^{q-1} & \cdots & z_n^{q-1} \\ z_1^q & \cdots & z_{p-1}^q & z_p^q & z_{p+1}^q & \cdots & z_n^q \\ z_1^{q+1} & \cdots & z_{p-1}^{q+1} & z_p^{q+1} & z_{p+1}^{q+1} & \cdots & z_n^{q+1} \\ \vdots & & & \vdots & & & \vdots \\ z_1^n & \cdots & \cdots & z_p^n & \cdots & \cdots & z_n^n \end{pmatrix}$$

and e^+ takes the same form except that we have $-z_p^q$ in the (p, q) –th position. Without loss of generality we rearrange the lattice as a vector $x = (x_1, \dots, x_{n^2})$ then we use the same kind of paths introduced by Shiu and Chen in [25]. For a given pair $(x, y) \in \chi^2$ there exist an increasing sequence $d_1, \dots, d_m \in \{1, \dots, n^2\}$ such that:

- $x_i \neq y_i$ for $i \in \{d_1, \dots, d_m\}$
- $x_i = y_i$ otherwise.

The path linking x to y is defined as follows:

$$\begin{aligned} (x_1, \dots, x_{n^2}) &= (y_1, \dots, y_{d_1-1}, x_{d_1}, x_{d_1+1}, \dots, x_{d_2-1}, x_{d_2}, x_{d_2+1}, \dots, x_{n^2}) \\ &\rightarrow (y_1, \dots, y_{d_1-1}, y_{d_1}, x_{d_1+1}, \dots, x_{d_2-1}, x_{d_2}, x_{d_2+1}, \dots, x_{n^2}) \\ &= (y_1, \dots, y_{d_1-1}, y_{d_1}, y_{d_1+1}, \dots, y_{d_2-1}, x_{d_2}, x_{d_2+1}, \dots, x_{n^2}) \\ &\rightarrow (y_1, \dots, y_{d_1-1}, y_{d_1}, y_{d_1+1}, \dots, y_{d_2-1}, y_{d_2}, x_{d_2+1}, \dots, x_{n^2}) \\ &\vdots \\ &\rightarrow (y_1, \dots, y_{n^2}). \end{aligned}$$

Those paths subsequently update the differing sites in the configurations x and y .

3.2. Main result

3.2.2 Geometric bound for the second largest eigenvalue

In what follows we will find an upper bound for

$$\kappa = \max_{e \in E} Q(e)^{-1} \sum_{\gamma_{xy} \ni e} |\gamma_{xy}| \pi(x) \pi(y).$$

First, let $e = (e^-, e^+)$ be a fixed edge. The flux in equilibrium associated to this edge is:

$$Q(e) = \pi(e^-) P(e^-, e^+)$$

and the transition matrix $P(e^-, e^+)$ is described in the following lemma:

Lemma 3.2.1. *Let x be a configuration. From the position of the site who is to be updated we distinguish three principale cases:*

1) If $(p, q) \in \{2, \dots, n-1\}^2$, each configuration has four neighbors and

$$P(e^-, e^+) = \frac{1}{n^2(1 + e^{\frac{2}{T}(z_{p-1}^q z_p^q + z_p^q z_{p+1}^q + z_p^{q-1} z_p^q + z_p^q z_p^{q+1})})}.$$

2) If $(p, q) \in \{2, \dots, n-1\} \times \{1\}$, each configuration has three neighbors and

$$P(e^-, e^+) = \frac{1}{n^2(1 + e^{\frac{2}{T}(z_1^{q-1} z_1^q + z_1^q z_1^{q+1} + z_1^q z_2^q)})}$$

and similar results hold for:

i) $(p, q) \in \{2, \dots, n-1\} \times \{n\}$:

$$P(e^-, e^+) = \frac{1}{n^2(1 + e^{\frac{2}{T}(z_{n-1}^q z_n^q + z_n^{q-1} z_n^q + z_n^q z_n^{q+1})})},$$

ii) $(p, q) \in \{1\} \times \{2, \dots, n-1\}$:

$$P(e^-, e^+) = \frac{1}{n^2(1 + e^{\frac{2}{T}(z_{p-1}^1 z_p^1 + z_p^1 z_{p+1}^1 + z_p^1 z_p^2)})},$$

iii) $(p, q) \in \{n\} \times \{2, \dots, n-1\}$:

$$P(e^-, e^+) = \frac{1}{n^2(1 + e^{\frac{2}{T}(z_{p-1}^n z_p^n + z_p^n z_{p+1}^n + z_p^n z_p^{n-1})})}.$$

3) If $(p, q) = (1, 1)$, each configuration has two neighbors and

$$P(e^-, e^+) = \frac{1}{n^2(1 + e^{\frac{2}{T}(z_1^1 z_2^1 + z_1^1 z_1^2)})}$$

and similar results hold for:

i) $(p, q) = (1, n)$:

$$P(e^-, e^+) = \frac{1}{n^2(1 + e^{\frac{2}{T}(z_1^n z_2^n + z_1^{n-1} z_1^n)})},$$

ii) $(p, q) = (n, 1)$:

$$P(e^-, e^+) = \frac{1}{n^2(1 + e^{\frac{2}{T}(z_{n-1}^1 z_n^1 + z_n^1 z_n^2)})},$$

iii) $(p, q) = (n, n)$:

$$P(e^-, e^+) = \frac{1}{n^2(1 + e^{\frac{2}{T}(z_{n-1}^n z_n^n + z_n^{n-1} z_n^n)})}.$$

Chapter 3. The Convergence rate of the Gibbs sampler for the 2-D Ising model via a geometric bound

The proof of this lemma is given in section 5.

For each class of edges, we give an upper bound for κ defined in equation (3.2) in the following proposition:

Proposition 3.2.2. *The second largest eigenvalue of the Gibbs sampler for the two-dimensional Ising model satisfies:*

$$\beta_1 \leq 1 - \frac{1}{\kappa} \quad \text{where } \kappa \text{ is bounded as follow :}$$

1) For $(p, q) = (1, 1)$ or $(p, q) = (n, n)$ we have:

$$Q(e)^{-1} \sum_{\gamma_{xy} \ni e} |\gamma_{xy}| \pi(x) \pi(y) \leq \frac{n^4}{2} (1 + e^{\frac{4}{T}}).$$

2) For $(p, q) = (1, n)$ we have:

$$Q(e)^{-1} \sum_{\gamma_{xy} \ni e} |\gamma_{xy}| \pi(x) \pi(y) \leq 2n^4 e^{\frac{n-1}{T}} \sum_{w \in \chi: w_n^1=1} \pi(w) \exp \left\{ \frac{1}{T} \left(2(1 - w_n^2) + \sum_{i=1}^{n-1} (w_i^1 - w_i^2 - w_i^1 w_i^2) \right) \right\}.$$

and a similar result holds for $(p, q) = (n, 1)$.

3) For $(p, q) \in \{2, \dots, n-1\} \times \{1\}$ we have:

$$Q(e)^{-1} \sum_{\gamma_{xy} \ni e} |\gamma_{xy}| \pi(x) \pi(y) = n^4 e^{\frac{n+1}{T}} \sum_{w \in \chi: w_p^1=1} \pi(w) \exp \left\{ \frac{1}{T} \left(-2(w_2^q + w_1^{q+1}) + \sum_{i=2}^n (w_i^{q-1} - w_i^q - w_i^{q-1} w_i^q) \right) \right\} + n^4 e^{\frac{n+1}{T}} \sum_{w \in \chi: w_p^1=-1} \pi(w) \exp \left\{ \frac{1}{T} \left(2(1 + w_1^{q-1}) + \sum_{i=2}^n (w_i^{q-1} - w_i^q - w_i^{q-1} w_i^q) \right) \right\}.$$

A similar result holds for $(p, q) \in \{2, \dots, n-1\} \times \{n\}$. Moreover with the same tricks we obtain an upper bound in the cases where

$$(p, q) \in \{1\} \times \{2, \dots, n-1\} \quad \text{or} \quad (p, q) \in \{n\} \times \{2, \dots, n-1\}.$$

4) For $(p, q) \in \{2, \dots, n-1\}^2$ we have:

$$Q(e)^{-1} \sum_{\gamma_{xy} \ni e} |\gamma_{xy}| \pi(x) \pi(y) \leq 2n^4 e^{\frac{1}{T}(n-1)} \sum_{w \in \chi: w_q^p=1} \pi(w) \exp \left\{ \frac{1}{T} \left(\sum_{i=1}^{p-1} (-w_i^q + w_i^{q+1} - w_i^q w_i^{q+1}) - 2(w_{p+1}^q + w_p^{q+1}) + \sum_{i=p+1}^n (w_i^{q-1} - w_i^q - w_i^{q-1} w_i^q) \right) \right\}.$$

3.3. Comparison

The proof of this proposition is given later in section 5.

After this step, we notice that it is difficult to give an exact value for the sums in the previous proposition. So, in what follows we give an upper bound for those sums in order to obtain an upper bound for κ defined in equation (3.2). The main result is given in the following theorem:

Theorem 3.2.3. [Franke and Helali 2018] *The second largest eigenvalue of the Gibbs sampler for the two-dimensional Ising model with two states satisfies:*

$$\beta_1 \leq 1 - n^{-4} \exp \left\{ -\frac{2}{T} (2n + 1) \right\}.$$

The proof of this theorem is given in section 5.

To be able to use Theorem 3.1.1 given by Diaconis and Stroock [5] we must give an upper bound for the second largest eigenvalue in absolute value. Until now, we have given an upper bound for the second largest eigenvalue. So we turn now to control the smallest eigenvalue of the Gibbs sampler for the two-dimensional Ising model.

3.2.3 Bound for the second largest eigenvalue in absolute value

Theorem 5.3 introduced by Ingrassia in [20] gives a lower bound for the smallest eigenvalue as:

$$\beta_{|\chi|-1} \geq -1 + \frac{2}{1 + (c-1)e^{\frac{\Delta}{T}}}.$$

For the two-dimensional Ising model one has $c = 2$ and $\Delta = 4$. Therefore for large n :

$$|\beta_{|\chi|-1}| \leq \left| -1 + \frac{2}{1 + e^{\frac{4}{T}}} \right| = 1 - \frac{2}{1 + e^{\frac{4}{T}}} < 1 - \frac{2}{e^{\frac{4}{T}} + e^{\frac{4}{T}}} < 1 - e^{-\frac{4}{T}} < 1 - n^{-4} e^{-\frac{2}{T}(2n+1)}.$$

An upper bound for β^* is given in the following corollary:

Corollary 3.2.4. *The second largest eigenvalue in absolute value of the Gibbs sampler for the two-dimensional Ising model with two states satisfies:*

$$\beta^* \leq 1 - n^{-4} \exp \left\{ -\frac{2}{T} (2n + 1) \right\}.$$

Proof. We give an upper bound for $|\beta_{|\chi|-1}|$ in the above computation. Then, combining this result with the result given in Theorem 3.2.3 finishes the proof. \square

3.3 Comparison

Ingrassia (see [20]) develop a method to give an upper bound for the second largest eigenvalue of the Gibbs sampler for the general Ising model and he obtains:

$$\beta_1 \leq 1 - \frac{Z_T}{b_\Gamma \Gamma_\gamma c |S|} e^{-\frac{m}{T}},$$

where Z_T is the normalizing constant, S is the lattice of sites, Γ the collection of paths, γ_Γ the maximum length of each path $\gamma_{xy} \in \Gamma$, b_Γ is the maximum number of paths containing any edge of Γ , c is the number of configurations that differ by only one site and m is the

Chapter 3. The Convergence rate of the Gibbs sampler for the 2-D Ising model via a geometric bound

least total elevation gain of the Hamiltonian function in the sense which is described by Holley and Stroock (see [15]).

For the two-dimensional square lattice Ising model with two states we have:

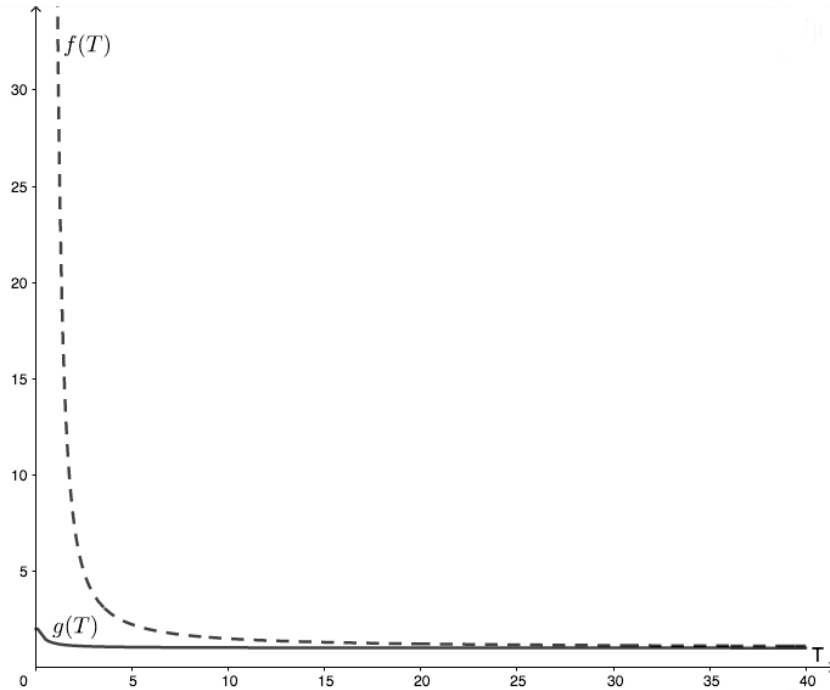
$$\Gamma_\gamma = n^2, \quad b_\gamma = 2^{n^2-1}, \quad c = 2, \quad m = 4, \quad |S| = n^2 \quad \text{and} \quad Z_T \leq 2(1 + e^{-\frac{1}{2T}})^{n^2-1}.$$

Which leads to:

$$\beta_1 \leq 1 - n^{-4} e^{-\frac{4}{T}} \left(\frac{1 + e^{-\frac{1}{2T}}}{2} \right)^{n^2-1}.$$

The comparison of the two results amounts to compare $f(T) = e^{\frac{4}{T}}$ and $g(T) = \frac{2}{1+e^{-\frac{1}{2T}}}$.

The following figure represents the two graphs.



We notice that for non-zero temperature, i.e.: $T > 0$ we have $e^{\frac{4}{T}} \geq \frac{2}{1+e^{-\frac{1}{2T}}}$. However for all sufficiently large natural numbers n holds:

$$\exp \left\{ \frac{2}{T}(2n-1) \right\} \leq \exp \left\{ \frac{4n}{T} \right\} = \left(e^{\frac{4}{T}} \right)^n \ll \left(\frac{2}{1+e^{-\frac{1}{2T}}} \right)^{n^2-1}.$$

This is quite natural to consider because MCMC method is used to give samples from probability measures defined on large spaces ($n \sim 10^{23}$). Then the main result of this chapter improves the one introduced by Ingrassia (see [20]) for large n .

Let t indicate the time. From Theorem 3.2.3 and the mathematical property: $1-x \leq e^{-x}$ we obtain:

$$\left(1 - n^{-4} e^{-\frac{2(2n+1)}{T}} \right)^{2t} \leq e^{-2tn^{-4} e^{-\frac{2(2n+1)}{T}}} \leq \epsilon$$

where ϵ is the pre-specified error tolerance. An elementary calculus gives that:

$$t \geq \frac{n^4}{2} e^{\frac{2(2n+1)}{T}} \log \epsilon^{-1}.$$

3.4. conclusion

Gibbs (see [12]) used coupling to bound the convergence time and he obtains the following bound:

$$2en^4(1 + \log \epsilon^{-1}) \frac{e^{\frac{4}{T}} + e^{-\frac{4}{T}}}{6e^{\frac{2}{T}} - 4e^{\frac{2}{T}}} \quad \text{for all } T \geq \frac{8}{\log\left(\frac{3}{2}\right)} \approx 20.$$

The result of Alison Gibbs (see [12]) improves the main result of this chapter but it is valid only for high temperature $T \geq 20$. On the contrary, we success to proof in this work a convergence result of the Gibbs sampler for the 2–D Ising model for all temperature T .

3.4 conclusion

This chapter deals with the bound of the second largest eigenvalue in absolute value of the Gibbs sampler for the two-dimensional Ising model with two states ± 1 . The main result generalizes some methods of Shiu and Chen (see [25]) obtained in dimension one to dimension two. It also improves Ingrassia's bound (see [20]).

3.5 Proof of the main result

3.5.1 Proof of Lemma 3.2.1

We concentrate on the case where $(p, q) \in \{2, \dots, n-1\}^2$. We have:

$$\begin{aligned} \pi(e^-) = & \frac{1}{Z_T} \exp \left\{ \frac{1}{T} \left(\sum_{j=1}^{q-1} \sum_{i=1}^{n-1} x_i^j x_{i+1}^j + \sum_{i=1}^{p-2} z_i^q z_{i+1}^q + z_{p-1}^q z_p^q + z_p^q z_{p+1}^q + \sum_{i=p+1}^{n-1} z_i^q z_{i+1}^q \right. \right. \\ & + \sum_{j=q+1}^n \sum_{i=1}^{n-1} z_i^j z_{i+1}^j + \sum_{i=1}^{p-1} \left(\sum_{j=1}^{q-2} z_i^j z_i^{j+1} + z_i^{q-1} z_i^q + z_i^q z_i^{q+1} + \sum_{j=q+1}^{n-1} z_i^j z_i^{j+1} \right) \\ & + \sum_{j=1}^{q-2} z_p^j z_p^{j+1} + z_p^{q-1} z_p^q + z_p^q z_p^{q+1} + \sum_{j=q+1}^{n-1} z_p^j z_p^{j+1} + \sum_{i=p+1}^n \left(\sum_{j=1}^{q-2} z_i^j z_i^{j+1} + z_i^{q-1} z_i^q \right. \\ & \left. \left. + z_i^q z_i^{q+1} + \sum_{j=q+1}^{n-1} z_i^j z_i^{j+1} \right) \right\}. \end{aligned}$$

and $\pi(e^+)$ takes a similar form except in the (p, q) -th position which is equal to $-z_p^q$. This yields:

$$\begin{aligned} P(e^-, e^+) &= \frac{1}{n^2} \frac{\pi(e^+)}{\pi(e^+) + \pi(e^-)} \\ &= \frac{e^{-\frac{1}{T}(z_{p-1}^q z_p^q + z_p^q z_{p+1}^q + z_p^{q-1} z_p^q + z_p^q z_p^{q+1})}}{e^{\frac{1}{T}(z_{p-1}^q z_p^q + z_p^q z_{p+1}^q + z_p^{q-1} z_p^q + z_p^q z_p^{q+1})} + e^{-\frac{1}{T}(z_{p-1}^q z_p^q + z_p^q z_{p+1}^q + z_p^{q-1} z_p^q + z_p^q z_p^{q+1})}} \\ &= \frac{1}{1 + e^{\frac{2}{T}(z_{p-1}^q z_p^q + z_p^q z_{p+1}^q + z_p^{q-1} z_p^q + z_p^q z_p^{q+1})}}. \end{aligned}$$

With similar techniques we can compute the transition matrix for each class of edges and the results are presented in lemma 3.2.1.

3.5.2 Proof of proposition 3.2.2

Depending on the position of the site who is to be updated we distinguish several cases.

Lets first assume that $(p, q) \in \{2, \dots, n-1\}^2$. Then the transition matrix is:

$$P(e^-, e^+) = \frac{1}{n^2(1 + e^{\frac{2}{T}(z_{p-1}^q z_p^q + z_p^q z_{p+1}^q + z_p^{q-1} z_p^q + z_p^q z_p^{q+1})})}.$$

According to the selection of the paths done in section 3.2.1., if a pair (x, y) satisfies $\gamma_{xy} \ni e$, then the configurations x and y must have the following form:

$$x = \begin{pmatrix} x_1^1 & \cdots & \cdots & x_p^1 & \cdots & \cdots & x_n^1 \\ \vdots & & & \vdots & & & \vdots \\ x_1^{q-1} & \cdots & x_{p-1}^{q-1} & x_p^{q-1} & x_{p+1}^{q-1} & \cdots & x_n^{q-1} \\ x_1^q & \cdots & x_{p-1}^q & z_p^q & z_{p+1}^q & \cdots & z_n^q \\ z_1^{q+1} & \cdots & z_{p-1}^{q+1} & z_p^{q+1} & z_{p+1}^{q+1} & \cdots & z_n^{q+1} \\ \vdots & & & \vdots & & & \vdots \\ z_1^n & \cdots & \cdots & z_p^n & \cdots & \cdots & z_n^n \end{pmatrix}$$

and

$$y = \begin{pmatrix} z_1^1 & \cdots & \cdots & z_p^1 & \cdots & \cdots & z_n^1 \\ \vdots & & & \vdots & & & \vdots \\ z_1^{q-1} & \cdots & z_{p-1}^{q-1} & z_p^{q-1} & z_{p+1}^{q-1} & \cdots & z_n^{q-1} \\ z_1^q & \cdots & z_{p-1}^q & -z_p^q & y_{p+1}^q & \cdots & y_n^q \\ y_1^{q+1} & \cdots & y_{p-1}^{q+1} & y_p^{q+1} & y_{p+1}^{q+1} & \cdots & y_n^{q+1} \\ \vdots & & & \vdots & & & \vdots \\ y_1^n & \cdots & \cdots & y_p^n & \cdots & \cdots & y_n^n \end{pmatrix}.$$

We then have:

$$\begin{aligned} \pi(x) = & \frac{1}{Z_T} \exp \left\{ \frac{1}{T} \left(\sum_{j=1}^{q-1} \sum_{i=1}^{n-1} x_i^j x_{i+1}^j + \sum_{i=1}^{p-2} x_i^q x_{i+1}^q + x_{p-1}^q z_p^q + z_p^q z_{p+1}^q + \sum_{i=p+1}^{n-1} z_i^q z_{i+1}^q \right. \right. \\ & + \sum_{j=q+1}^n \sum_{i=1}^{n-1} z_i^j z_{i+1}^j + \sum_{i=1}^{p-1} \left(\sum_{j=1}^{q-2} x_i^j x_{i+1}^{j+1} + x_i^{q-1} x_i^q + x_i^q z_i^{q+1} + \sum_{j=q+1}^{n-1} z_i^j z_i^{j+1} \right) \\ & + \sum_{j=1}^{q-2} x_p^j x_p^{j+1} + x_p^{q-1} z_p^q + z_p^q z_p^{q+1} + \sum_{j=q+1}^{n-1} z_p^j z_p^{j+1} + \sum_{i=p+1}^n \left(\sum_{j=1}^{q-2} x_i^j x_i^{j+1} + x_i^{q-1} z_i^q \right. \\ & \left. \left. + z_i^q z_i^{q+1} + \sum_{j=q+1}^{n-1} z_i^j z_i^{j+1} \right) \right\}. \end{aligned}$$

3.5. Proof of the main result

and

$$\begin{aligned}
\pi(y) = & \frac{1}{Z_T} \exp \left\{ \frac{1}{T} \left(\sum_{j=1}^{q-1} \sum_{i=1}^{n-1} z_i^j z_{i+1}^j + \sum_{i=1}^{p-2} z_i^q z_{i+1}^q - z_{p-1}^q z_p^q - z_p^q y_{p+1}^q + \sum_{i=p+1}^{n-1} y_i^q y_{i+1}^q \right. \right. \\
& + \sum_{j=q+1}^n \sum_{i=1}^{n-1} y_i^j y_{i+1}^j + \sum_{i=1}^{p-1} \left(\sum_{j=1}^{q-2} z_i^j z_i^{j+1} + z_i^{q-1} z_i^q + z_i^q y_i^{q+1} + \sum_{j=q+1}^{n-1} y_i^j y_i^{j+1} \right) \\
& + \sum_{j=1}^{q-2} z_p^j z_p^{j+1} - z_p^{q-1} z_p^q - z_p^q y_p^{q+1} + \sum_{j=q+1}^{n-1} y_p^j y_p^{j+1} + \sum_{i=p+1}^n \left(\sum_{j=1}^{q-2} z_i^j z_i^{j+1} + z_i^{q-1} y_i^q \right. \\
& \left. \left. + y_i^q y_i^{q+1} + \sum_{j=q+1}^{n-1} y_i^j y_i^{j+1} \right) \right\}.
\end{aligned}$$

Similarly, we can compute the expression of the measure $\pi(e^-)$. We then obtain:

$$\begin{aligned}
Q(e)^{-1} \pi(x) \pi(y) &= \frac{\pi(x) \pi(y)}{\pi(e^-) P(e^-, e^+)} \\
&= \frac{n^2}{Z_T} \left(\exp \left\{ \frac{-2}{T} (z_{p-1}^q z_p^q) \right\} + \exp \left\{ \frac{2}{T} (z_p^q z_{p+1}^q + z_p^{q-1} z_p^q + z_p^q z_p^{q+1}) \right\} \right) \\
&\times \exp \left\{ \frac{1}{T} \left(\sum_{j=1}^{q-1} \sum_{i=1}^{n-1} x_i^j x_{i+1}^j + \sum_{i=1}^{p-2} x_i^q x_{i+1}^q + x_{p-1}^q z_p^q - z_p^q y_{p+1}^q \right. \right. \\
&+ \sum_{i=p+1}^{n-1} y_i^q y_{i+1}^q + \sum_{j=q+1}^n \sum_{i=1}^{n-1} y_i^j y_{i+1}^j + \sum_{i=1}^{p-1} \left(\sum_{j=1}^{q-2} x_i^j x_i^{j+1} + x_i^{q-1} x_i^q \right. \\
&+ x_i^q z_i^{q+1} - z_i^q z_i^{q+1} + z_i^q y_i^{q+1} + \sum_{j=q+1}^{n-1} y_i^j y_i^{j+1} \left. \right) + \sum_{j=1}^{q-2} x_p^j x_p^{j+1} + x_p^{q-1} z_p^q \\
&- z_p^{q-1} z_p^q - z_p^{q-1} z_p^q - z_p^q y_p^{q+1} + \sum_{j=q+1}^{n-1} y_p^j y_p^{j+1} + \sum_{i=p+1}^n \left(\sum_{j=1}^{q-2} x_i^j x_i^{j+1} \right. \\
&\left. \left. + x_i^{q-1} z_i^q - z_i^{q-1} z_i^q + z_i^{q-1} y_i^q + y_i^q y_i^{q+1} + \sum_{j=q+1}^{n-1} y_i^j y_i^{j+1} \right) \right\}. \tag{3.4}
\end{aligned}$$

Following the line of arguments of Shiu and Chen (see [25]) we define the configurations

$$x \oplus y = \begin{pmatrix} x_1^1 & \cdots & \cdots & x_p^1 & \cdots & \cdots & x_n^1 \\ \vdots & & & \vdots & & & \vdots \\ x_1^{q-1} & \cdots & x_{p-1}^{q-1} & x_p^{q-1} & x_{p+1}^{q-1} & \cdots & x_n^{q-1} \\ x_1^q & \cdots & x_{p-1}^q & z_p^q & y_{p+1}^q & \cdots & y_n^q \\ y_1^{q+1} & \cdots & y_{p-1}^{q+1} & y_p^{q+1} & y_{p+1}^{q+1} & \cdots & y_n^{q+1} \\ \vdots & & & \vdots & & & \vdots \\ y_1^n & \cdots & \cdots & y_p^n & \cdots & \cdots & y_n^n \end{pmatrix}$$

Chapter 3. The Convergence rate of the Gibbs sampler for the 2-D Ising model via a geometric bound

and $x \ominus y$ with the same expression expect that in the position (p, q) we have $-z_p^q$. Equation (3.4) becomes:

$$\begin{aligned}
Q(e)^{-1} \pi(x) \pi(y) &= n^2 \pi(x \oplus y) \exp \left\{ \frac{1}{T} \left(-2(z_{p-1}^q z_p^q + z_p^{q-1} z_p^q) + \sum_{i=1}^{p-1} -z_i^q z_i^{q+1} + \sum_{i=p+1}^n -z_i^{q-1} z_i^q \right. \right. \\
&\quad \left. \left. + \sum_{i=1}^{p-1} (x_i^q z_i^{q+1} + z_i^q y_i^{q+1} - x_i^q y_i^{q+1}) - 2(z_p^q y_{p+1}^q + z_p^q y_p^{q+1}) + \sum_{i=p+1}^n (x_i^{q-1} z_i^q + z_i^{q-1} y_i^q \right. \right. \\
&\quad \left. \left. - x_i^{q-1} y_i^q) \right) \right\} + n^2 \pi(x \ominus y) \exp \left\{ \frac{1}{T} \left(2(z_p^q z_{p+1}^q + z_p^q z_p^{q+1}) + \sum_{i=1}^{p-1} -z_i^q z_i^{q+1} \right. \right. \\
&\quad \left. \left. - \sum_{i=p+1}^n z_i^{q-1} z_i^q + \sum_{i=1}^{p-1} (x_i^q z_i^{q+1} + z_i^q y_i^{q+1} - x_i^q y_i^{q+1}) + 2(x_{p-1}^q z_p^q + x_p^{q-1} z_p^q) \right. \right. \\
&\quad \left. \left. + \sum_{i=p+1}^n (x_i^{q-1} z_i^q + z_i^{q-1} y_i^q - x_i^{q-1} y_i^q) \right) \right\}. \tag{3.5}
\end{aligned}$$

Remark 3.5.1. We notice that the right side in equation (3.5) reaches its maximum value if

$$-z_i^{q-1} = z_i^q = -z_i^{q+1} \quad \text{for } i = 1, \dots, n.$$

Without loss of generality we can consider the situation where $z_p^q = 1$. This yields:

$$\begin{aligned}
Q(e)^{-1} \pi(x) \pi(y) &\leq n^2 e^{\frac{n-1}{T}} \pi(x \oplus y) \exp \left\{ \frac{1}{T} \left(\sum_{i=1}^{p-1} (-x_i^q + y_i^{q+1} - x_i^q y_i^{q+1}) - 2(y_{p+1}^q + y_p^{q+1}) \right. \right. \\
&\quad \left. \left. + \sum_{i=p+1}^n (x_i^{q-1} - y_i^q - x_i^{q-1} y_i^q) \right) \right\} + n^2 e^{\frac{n-1}{T}} \pi(x \ominus y) \exp \left\{ \frac{1}{T} \left(\sum_{i=1}^{p-1} (-x_i^q + y_i^{q+1} \right. \right. \\
&\quad \left. \left. - x_i^q y_i^{q+1}) + 2(x_{p-1}^q + x_p^{q-1}) + \sum_{i=p+1}^n (x_i^{q-1} - y_i^q - x_i^{q-1} y_i^q) \right) \right\}.
\end{aligned}$$

Then we obtain:

$$\begin{aligned}
Q(e)^{-1} \sum_{\gamma_{xy} \ni e} |\gamma_{xy}| \pi(x) \pi(y) &\leq n^4 e^{\frac{n-1}{T}} \sum_{\gamma_{xy} \ni e} \pi(x \oplus y) \exp \left\{ \frac{1}{T} \left(\sum_{i=1}^{p-1} (-x_i^q + y_i^{q+1} - x_i^q y_i^{q+1}) \right. \right. \\
&\quad \left. \left. - 2(y_{p+1}^q + y_p^{q+1}) + \sum_{i=p+1}^n (x_i^{q-1} - y_i^q - x_i^{q-1} y_i^q) \right) \right\} \\
&\quad + n^4 e^{\frac{n-1}{T}} \sum_{\gamma_{xy} \ni e} \pi(x \ominus y) \exp \left\{ \frac{1}{T} \left(\sum_{i=1}^{p-1} (-x_i^q + y_i^{q+1} - x_i^q y_i^{q+1}) \right. \right. \\
&\quad \left. \left. + 2(x_{p-1}^q + x_p^{q-1}) + \sum_{i=p+1}^n (x_i^{q-1} - y_i^q - x_i^{q-1} y_i^q) \right) \right\}.
\end{aligned}$$

3.5. Proof of the main result

Like in Shiu and Chen (see [25]) we notice that

$$\bigcup_{(x,y): \gamma_{xy} \ni e} \{x \oplus y, x \ominus y\} = \chi.$$

With new notation w for the elements from χ the previous expression becomes:

$$\begin{aligned} Q(e)^{-1} \sum_{\gamma_{xy} \ni e} |\gamma_{xy}| \pi(x) \pi(y) &\leq n^4 e^{\frac{n-1}{T}} \sum_{w \in \chi: w_p^q = 1} \pi(w) \exp \left\{ \frac{1}{T} \left(\sum_{i=1}^{p-1} (-w_i^q + w_i^{q+1} - w_i^q w_i^{q+1}) \right. \right. \\ &\quad \left. \left. - 2(w_{p+1}^q + w_p^{q+1}) + \sum_{i=p+1}^n (w_i^{q-1} - w_i^q - w_i^{q-1} w_i^q) \right) \right\} \\ &\quad + n^4 e^{\frac{n-1}{T}} \sum_{w \in \chi: w_p^q = -1} \pi(w) \exp \left\{ \frac{1}{T} \left(\sum_{i=1}^{p-1} (-w_i^q + w_i^{q+1} - w_i^q w_i^{q+1}) \right. \right. \\ &\quad \left. \left. + 2(w_{p-1}^q + w_p^{q-1}) + \sum_{i=p+1}^n (w_i^{q-1} - w_i^q - w_i^{q-1} w_i^q) \right) \right\}. \end{aligned}$$

If we look to the nearest neighbors of the site (p, q) , for each configuration $w_+ \in \{w \in \chi : w_p^q = 1\}$ there exist a unique configuration $w_- \in \{w \in \chi : w_p^q = -1\}$ such that:

$$\pi(w_+) = \pi(w_-) \exp \left\{ \frac{2}{T} (w_p^{q-1} + w_p^{q+1} + w_{p-1}^q + w_{p+1}^q) \right\}.$$

This yields the following equality:

$$\sum_{w \in \chi: w_p^q = 1} \pi(w) \exp \left\{ \frac{-2}{T} (w_{p+1}^q + w_p^{q+1}) \right\} = \sum_{w \in \chi: w_p^q = -1} \pi(w) \exp \left\{ \frac{2}{T} (w_{p-1}^q + w_p^{q-1}) \right\}. \quad (3.6)$$

Then, it follows that:

$$\begin{aligned} Q(e)^{-1} \sum_{\gamma_{xy} \ni e} |\gamma_{xy}| \pi(x) \pi(y) &\leq 2n^4 e^{\frac{1}{T}(n-1)} \sum_{w \in \chi: w_p^q = 1} \pi(w) \exp \left\{ \frac{1}{T} \left(\sum_{i=1}^{p-1} (-w_i^q + w_i^{q+1} - w_i^q w_i^{q+1}) \right. \right. \\ &\quad \left. \left. - 2(w_{p+1}^q + w_p^{q+1}) + \sum_{i=p+1}^n (w_i^{q-1} - w_i^q - w_i^{q-1} w_i^q) \right) \right\}. \quad (3.7) \end{aligned}$$

Remark 3.5.2. We obtain the same result by considering the case where $z_p^q = -1$.

We turn now to remaining cases where the site (p, q) lies on the boundary of the square:

i) For $(p, q) = (1, 1)$ we have that x and e^- coincide and

$$P(e^-, e^+) = \frac{1}{n^2 (1 + e^{\frac{2}{T}(z_1^1 z_2^1 + z_1^1 z_1^2)})}.$$

This yields:

$$\begin{aligned} Q(e)^{-1} \pi(x) \pi(y) &= n^2 (1 + e^{\frac{2}{T}(z_1^1 z_2^1 + z_1^1 z_1^2)}) \pi(y) \\ &\leq n^2 (1 + e^{\frac{4}{T}}) \pi(y) \end{aligned}$$

Chapter 3. The Convergence rate of the Gibbs sampler for the 2-D Ising model via a geometric bound

and it follows:

$$\begin{aligned} Q(e)^{-1} \sum_{\gamma_{xy} \ni e} |\gamma_{xy}| \pi(x) \pi(y) &\leq n^4 (1 + e^{\frac{4}{T}}) \sum_{y \in \mathcal{X}: y_1^1 = 1} \pi(y) \\ &= \frac{n^4}{2} (1 + e^{\frac{4}{T}}). \end{aligned} \quad (3.8)$$

For $(p, q) = (n, n)$ we notice that y and e^+ coincide and we obtain the same kind of result as for $(p, q) = (1, 1)$.

ii) For $(p, q) = (1, n)$, we have from lemma 3.2.1

$$P(e^-, e^+) = \frac{1}{n^2 (1 + e^{\frac{2}{T}(z_{n-1}^1 z_n^1 + z_n^1 z_n^2)})}.$$

According to the selection of paths in section 3.2.1 for any pair (x, y) where $\gamma_{xy} \ni e$ the configurations x and y must have the following form:

$$x = \begin{pmatrix} x_1^1 & \cdots & x_{n-1}^1 & z_n^1 \\ z_1^2 & \cdots & \cdots & z_n^2 \\ \vdots & & & \vdots \\ \vdots & & & \vdots \\ z_1^n & \cdots & \cdots & z_n^n \end{pmatrix} \quad \text{and} \quad y = \begin{pmatrix} z_1^1 & \cdots & z_{n-1}^1 & -z_n^1 \\ y_1^2 & \cdots & \cdots & y_n^2 \\ \vdots & & & \vdots \\ \vdots & & & \vdots \\ y_1^n & \cdots & \cdots & y_n^n \end{pmatrix}.$$

As in the case where $(p, q) = \{2, \dots, n-1\}^2$ we find:

$$\begin{aligned} Q(e)^{-1} \pi(x) \pi(y) &= \frac{n^2}{Z_T} (1 + e^{\frac{2}{T}(z_{n-1}^1 z_n^1 + z_n^1 z_n^2)}) \exp \left\{ \frac{1}{T} \left(\sum_{i=1}^{n-2} x_i^1 x_{i+1}^1 + x_{n-1}^1 z_n^1 \right. \right. \\ &\quad \left. \left. - 2z_{n-1}^1 z_n^1 + \sum_{j=2}^n \sum_{i=1}^{n-1} y_i^j y_{i+1}^j + \sum_{i=1}^{n-1} (x_i^1 z_i^2 + z_i^1 y_i^2 + \sum_{j=2}^{n-1} y_i^j y_i^{j+1} \right. \right. \\ &\quad \left. \left. - z_i^1 z_i^2) - z_n^1 y_n^2 + \sum_{j=2}^n y_n^j y_n^{j+1} \right) \right\}. \end{aligned} \quad (3.9)$$

Following the line of arguments from Shiu and Chen (see [25]) we define configurations:

$$x \oplus y = \begin{pmatrix} x_1^1 & \cdots & x_{n-1}^1 & z_n^1 \\ y_1^2 & \cdots & \cdots & y_n^2 \\ \vdots & & & \vdots \\ \vdots & & & \vdots \\ y_1^n & \cdots & \cdots & y_n^n \end{pmatrix} \quad \text{and} \quad x \ominus y = \begin{pmatrix} x_1^1 & \cdots & x_{n-1}^1 & -z_n^1 \\ y_1^2 & \cdots & \cdots & y_n^2 \\ \vdots & & & \vdots \\ \vdots & & & \vdots \\ y_1^n & \cdots & \cdots & y_n^n \end{pmatrix}.$$

3.5. Proof of the main result

Then equation (3.9) becomes:

$$\begin{aligned}
Q(e)^{-1}\pi(x)\pi(y) &= n^2\pi(x \oplus y) \exp \left\{ \frac{1}{T} \left(-2(z_{n-1}^1 z_n^1 + z_n^1 y_n^2) \right. \right. \\
&\quad \left. \left. + \sum_{i=1}^{n-1} (x_i^1 z_i^2 + z_i^1 y_i^2 - z_i^1 z_i^2 - x_i^1 y_i^2) \right) \right\} \\
&\quad + n^2\pi(x \ominus y) \exp \left\{ \frac{1}{T} \left(2(z_n^1 z_n^2 + x_{n-1}^1 z_n^1) + \sum_{i=1}^{n-1} (x_i^1 z_i^2 \right. \right. \\
&\quad \left. \left. + z_i^1 y_i^2 - z_i^1 z_i^2 - x_i^1 y_i^2) \right) \right\}. \tag{3.10}
\end{aligned}$$

Remark 3.5.3. We notice that equation (3.10) reaches its maximum if the following two points hold:

- $z_i^1 = -z_i^2 = -1$ for $i \leq n-1$.
- $z_n^1 = z_n^2 = +1$ and $z_{n-1}^1 = -z_{n-1}^2 = -1$.

This yields:

$$\begin{aligned}
Q(e)^{-1}\pi(x)\pi(y) &\leq n^2\pi(x \oplus y) \exp \left\{ \frac{1}{T} \left(2(1 - y_n^2) + \sum_{i=1}^{n-1} (x_i^1 - y_i^2 + 1 - x_i^1 y_i^2) \right) \right\} \\
&\quad + n^2\pi(x \ominus y) \exp \left\{ \frac{1}{T} \left(2(1 + x_{n-1}^1) + \sum_{i=1}^{n-1} (x_i^1 - y_i^2 + 1 - x_i^1 y_i^2) \right) \right\}. \tag{3.11}
\end{aligned}$$

Summing over all pairs (x, y) where the path γ_{xy} passes through the edge e of equation (3.11) gives:

$$\begin{aligned}
Q(e)^{-1} \sum_{\gamma_{xy} \ni e} |\gamma_{xy}| \pi(x)\pi(y) &\leq n^4 e^{\frac{n-1}{T}} \sum_{\gamma_{xy} \ni e} \pi(x \oplus y) \exp \left\{ \frac{1}{T} \left(2(1 - y_n^2) \right. \right. \\
&\quad \left. \left. + \sum_{i=1}^{n-1} (x_i^1 - y_i^2 - x_i^1 y_i^2) \right) \right\} \\
&\quad + n^4 e^{\frac{n-1}{T}} \sum_{\gamma_{xy} \ni e} \pi(x \ominus y) \exp \left\{ \frac{1}{T} \left(2(1 + x_{n-1}^1) \right. \right. \\
&\quad \left. \left. + \sum_{i=1}^{n-1} (x_i^1 - y_i^2 - x_i^1 y_i^2) \right) \right\}
\end{aligned}$$

Then

$$\begin{aligned}
 Q(e)^{-1}\pi(x)\pi(y) &\leq n^4 e^{\frac{n-1}{T}} \sum_{w \in \mathcal{X}: w_n^1=1} \pi(w) \exp \left\{ \frac{1}{T} \left(2(1-w_n^2) \right. \right. \\
 &\quad \left. \left. + \sum_{i=1}^{n-1} (w_i^1 - w_i^2 - w_i^1 w_i^2) \right) \right\} \\
 &\quad + n^4 e^{\frac{n-1}{T}} \sum_{w \in \mathcal{X}: w_n^1=-1} \pi(w) \exp \left\{ \frac{1}{T} \left(2(1+w_{n-1}^1) \right. \right. \\
 &\quad \left. \left. + \sum_{i=1}^{n-1} (w_i^1 - w_i^2 - w_i^1 w_i^2) \right) \right\} \\
 &= 2n^4 e^{\frac{n-1}{T}} \sum_{w \in \mathcal{X}: w_n^1=1} \pi(w) \exp \left\{ \frac{1}{T} \left(2(1-w_n^2) \right. \right. \\
 &\quad \left. \left. + \sum_{i=1}^{n-1} (w_i^1 - w_i^2 - w_i^1 w_i^2) \right) \right\}.
 \end{aligned}$$

The last equality is given by a symmetry argument similar to the one introduced in equation (3.6). For $(p, q) = (n, 1)$ we use the same technique and we obtain a similar result.

iii) For $(p, q) \in \{2 \cdots n-1\} \times \{1\}$, we have from lemma 3.2.1

$$P(e^-, e^+) = \frac{1}{n^2(1 + e^{\frac{2}{T}(z_1^{q-1}z_1^q + z_1^q z_1^{q+1} + z_1^q z_2^q)})}.$$

As before, a pair of configuration (x, y) where $\gamma_{xy} \ni e$ must have the following form:

$$x = \begin{pmatrix} x_1^1 & \cdots & \cdots & x_n^1 \\ \vdots & & & \vdots \\ x_1^{q-1} & \cdots & \cdots & x_n^{q-1} \\ z_1^q & \cdots & \cdots & z_n^q \\ z_1^{q+1} & \cdots & \cdots & z_n^{q+1} \\ \vdots & & & \vdots \\ z_1^n & \cdots & \cdots & z_n^n \end{pmatrix} \quad \text{and} \quad y = \begin{pmatrix} z_1^1 & \cdots & \cdots & z_n^1 \\ \vdots & & & \vdots \\ z_1^{q-1} & \cdots & \cdots & z_n^{q-1} \\ -z_1^q & y_2^q & \cdots & y_n^q \\ y_1^{q+1} & \cdots & \cdots & y_n^{q+1} \\ \vdots & & & \vdots \\ y_1^n & \cdots & \cdots & y_n^n \end{pmatrix}.$$

It follows that:

$$\begin{aligned}
 Q(e)^{-1}\pi(x)\pi(y) &= \frac{n^2}{Z_T} (1 + e^{\frac{2}{T}(z_1^{q-1}z_1^q + z_1^q z_1^{q+1} + z_1^q z_2^q)}) \\
 &\quad \times \exp \left\{ \frac{1}{T} \left(\sum_{j=1}^{q-1} \sum_{i=1}^{n-1} x_i^j x_{i+1}^j - z_1^q y_2^q + \sum_{i=2}^{n-1} y_i^q y_{i+1}^q + \sum_{j=q+1}^n \sum_{i=1}^{n-1} y_i^j y_{i+1}^j \right. \right. \\
 &\quad \left. \left. + \sum_{j=1}^{q-2} x_1^j x_1^{j+1} + x_1^{q-1} z_1^q - z_1^{q-1} z_1^q - z_1^q y_1^{q+1} + \sum_{j=q+1}^{n-1} y_1^j y_1^{j+1} - z_1^{q-1} z_1^q \right. \right. \\
 &\quad \left. \left. + \sum_{i=2}^n \left(\sum_{j=1}^{q-2} x_i^j x_i^{j+1} + x_i^{q-1} z_i^q + z_i^{q-1} y_i^q + y_i^q y_i^{q+1} + \sum_{j=q+1}^{n-1} y_i^j y_i^{j+1} - z_i^{q-1} z_i^q \right) \right) \right\}. \tag{3.12}
 \end{aligned}$$

3.5. Proof of the main result

Following the line of arguments of Shiu and Chen (see [25]) we define the configurations:

$$x \oplus y = \begin{pmatrix} x_1^1 & \cdots & \cdots & x_n^1 \\ \vdots & & & \vdots \\ x_1^{q-1} & \cdots & \cdots & x_n^{q-1} \\ z_1^q & y_2^q & \cdots & y_n^q \\ y_1^{q+1} & \cdots & \cdots & y_n^{q+1} \\ \vdots & & & \vdots \\ y_1^n & \cdots & \cdots & y_n^n \end{pmatrix} \quad \text{and} \quad x \ominus y = \begin{pmatrix} x_1^1 & \cdots & \cdots & x_n^1 \\ \vdots & & & \vdots \\ x_1^{q-1} & \cdots & \cdots & x_n^{q-1} \\ -z_1^q & y_2^q & \cdots & y_n^q \\ y_1^{q+1} & \cdots & \cdots & y_n^{q+1} \\ \vdots & & & \vdots \\ y_1^n & \cdots & \cdots & y_n^n \end{pmatrix}.$$

Equation (3.12) becomes:

$$\begin{aligned} \frac{\pi(x)\pi(y)}{Q(e)} &= n^2 \pi(x \oplus y) \exp \left\{ \frac{1}{T} \left(-2(z_1^q y_2^q + z_1^{q-1} z_1^q + z_1^q y_1^{q+1}) + \sum_{i=2}^n (x_i^{q-1} z_i^q + z_i^{q-1} y_i^q \right. \right. \\ &\quad \left. \left. - z_i^{q-1} z_i^q - x_i^{q-1} y_i^q) \right) \right\} + n^2 \pi(x \ominus y) \exp \left\{ \frac{1}{T} \left(2(z_1^q z_1^{q+1} + z_1^q z_2^q + x_1^{q-1} z_1^q) \right. \right. \\ &\quad \left. \left. + \sum_{i=2}^n (x_i^{q-1} z_i^q + z_i^{q-1} y_i^q - z_i^{q-1} z_i^q - x_i^{q-1} y_i^q) \right) \right\}. \end{aligned} \quad (3.13)$$

Remark 3.5.4. The maximum in equation (3.13) is reached once the following conditions hold:

- $-z_i^{q-1} = z_i^q = z_i^{q+1}$ for $i = 2, \dots, n$.
- $-z_1^{q-1} = z_1^q = z_1^{q+1} = 1$ and $z_1^q = z_2^q = 1$.

Then equation (3.13) becomes:

$$\begin{aligned} Q(e)^{-1} \pi(x)\pi(y) &\leq n^2 \pi(x \oplus y) \exp \left\{ \frac{1}{T} \left(-2(y_2^q - 1 + y_1^{q+1}) + \sum_{i=2}^n (x_i^{q-1} - y_i^q + 1 - x_i^{q-1} y_i^q) \right) \right\} \\ &\quad + n^2 \pi(x \ominus y) \exp \left\{ \frac{1}{T} \left(2(1 + 1 + x_1^{q-1}) + \sum_{i=2}^n (x_i^{q-1} - y_i^q + 1 - x_i^{q-1} y_i^q) \right) \right\}. \end{aligned}$$

As in Shiu and Chen (see [25]) we notice that

$$\bigcup_{(x,y): \gamma_{xy} \ni e} \{x \oplus y, x \ominus y\} = \chi.$$

In the previous expression we change the notation of an element $x \oplus y$ or $x \ominus y$ to w with

Chapter 3. The Convergence rate of the Gibbs sampler for the 2-D Ising model via a geometric bound

$w_p^1 = +1$ or $w_p^1 = -1$ respectively. This yields:

$$\begin{aligned}
Q(e)^{-1} \sum_{\gamma_{xy} \ni e} |\gamma_{xy}| \pi(x) \pi(y) &\leq n^4 \sum_{\gamma_{xy} \ni e} \pi(x \oplus y) \exp \left\{ \frac{1}{T} \left(2(1 - y_2^q - y_1^{q+1}) + \sum_{i=2}^n (x_i^{q-1} - y_i^q \right. \right. \\
&\quad \left. \left. + 1 - x_i^{q-1} y_i^q) \right) \right\} + n^4 \sum_{\gamma_{xy} \ni e} \pi(x \ominus y) \exp \left\{ \frac{1}{T} \left(2(1 + 1 + x_1^{q-1}) \right. \right. \\
&\quad \left. \left. + \sum_{i=2}^n (x_i^{q-1} - y_i^q + 1 - x_i^{q-1} y_i^q) \right) \right\} \\
&= n^4 e^{\frac{n+1}{T}} \sum_{w \in \mathcal{X}: w_p^1=1} \pi(w) \exp \left\{ \frac{1}{T} \left(-2(w_2^q + w_1^{q+1}) + \sum_{i=2}^n (w_i^{q-1} \right. \right. \\
&\quad \left. \left. - w_i^q - w_i^{q-1} w_i^q) \right) \right\} + n^4 e^{\frac{n+1}{T}} \sum_{w \in \mathcal{X}: w_p^1=-1} \pi(w) \exp \left\{ \frac{1}{T} \left(2(1 + w_1^{q-1}) \right. \right. \\
&\quad \left. \left. + \sum_{i=2}^n (w_i^{q-1} - w_i^q - w_i^{q-1} w_i^q) \right) \right\}.
\end{aligned}$$

We obtain the same result for $(p, q) \in \{2 \cdots n - 1\} \times \{n\}$. With similar techniques we can do the same calculus for $(p, q) \in \{1\} \times \{2 \cdots n - 1\}$ then we obtain:

$$\begin{aligned}
Q(e)^{-1} \sum_{\gamma_{xy} \ni e} |\gamma_{xy}| \pi(x) \pi(y) &\leq n^4 e^{\frac{n+1}{T}} \sum_{w \in \mathcal{X}: w_p^1=1} \pi(w) \exp \left\{ \frac{1}{T} \left(-2(w_{p+1}^1 + w_p^2) + \sum_{i=1}^{n-1} (w_i^1 \right. \right. \\
&\quad \left. \left. - w_i^2 - w_i^1 w_i^2) \right) \right\} + n^4 e^{\frac{n+1}{T}} \sum_{w \in \mathcal{X}: w_p^1=-1} \pi(w) \exp \left\{ \frac{1}{T} \left(2(1 + w_{p-1}^1) \right. \right. \\
&\quad \left. \left. + \sum_{i=1}^{n-1} (w_i^1 - w_i^2 - w_i^1 w_i^2) \right) \right\}.
\end{aligned}$$

A similar result is obtained in the case where $(p, q) \in \{n\} \times \{2 \cdots n - 1\}$.

We turn now to give the proof of the main theorem of this chapter:

3.5.3 Proof of Theorem 3.2.3

Since for all sites $(p, q) \in \{1, \dots, n\}^2$ we have $w_p^q = \pm 1$ it follows that

$$\max_{i \in \{1, \dots, n\}} \{w_i^{q-1} - w_i^q - w_i^{q-1} w_i^q\} = \max_{i \in \{1, \dots, n\}} \{-w_i^q + w_i^{q+1} - w_i^q w_i^{q+1}\} = 3.$$

Analyzing the worst cases yield the following inequalities

$$\text{i) } 2(1 - w_n^2) + \sum_{i=1}^{n-1} (w_i^1 - w_i^2 - w_i^1 w_i^2) \leq 3n + 1.$$

$$\text{ii) } -2(w_2^q + w_1^{q+1}) + \sum_{i=2}^n (w_i^{q-1} - w_i^q - w_i^{q-1} w_i^q) \leq 3n + 1.$$

3.6. Asymptotic study

$$\text{iii) } 2(1 + w_1^{q-1}) + \sum_{i=2}^n (w_i^{q-1} - w_i^q - w_i^{q-1}w_i^q) \leq 3n + 1.$$

$$\text{iv) } \sum_{i=1}^{p-1} (-w_i^q + w_i^{q+1} - w_i^q w_i^{q+1}) - 2(w_{p+1}^q + w_p^{q+1}) + \sum_{i=p+1}^n (w_i^{q-1} - w_i^q - w_i^{q-1}w_i^q) \leq 3n + 1.$$

By the fact that

$$\sum_{w \in \chi: w_p^q = +1} \pi(w) = \sum_{w \in \chi: w_p^q = -1} \pi(w) = \frac{1}{2}$$

we conclude that

$$\kappa \leq n^4 \exp \left\{ \frac{2}{T} (2n + 1) \right\}$$

which finishes the proof.

3.6 Asymptotic study

We now argue otherwise and the main goal is to obtain better bound of the second largest eigenvalue for the two-dimensional square lattice Ising model. We turn to study the asymptotic of κ for each class of edges. We gives a detailed calculus only for the case $(i, j) = \{2, \dots, n-1\} \times \{2, \dots, n-1\}$ then we do similar reasoning for the other class of edges.

3.6.1 For $(i, j) = \{2, \dots, n-1\} \times \{2, \dots, n-1\}$

First of all we will introduce some notations that are necessary for further calculus. Denote by:

- $\alpha_k = \sum_{i=1}^{p-1} (-w_i^q + w_i^{q+1} - w_i^q w_i^{q+1}) - 2(w_{p+1}^q + w_p^{q+1}) + \sum_{i=p+1}^n (w_i^{q-1} - w_i^q - w_i^{q-1}w_i^q)$

- $\theta(w)$ the number of signs changes in a given configuration, ie:

$$\theta(w) = \frac{1}{2} \# \{ (i, j) \in \{1, \dots, n\}^2 : \exists (k, l) \in \{1, \dots, n\}^2 : w_i^k = -w_k^l \text{ and } w_i^j \sim w_k^l \}$$

- θ_ℓ the number of sign changes of $w \in \chi$ where $w_p^q = 1$ and $\pi(w) = \ell$, ie: for $1 \leq i, j, k, l \leq n$

$$\theta_\ell = \# \frac{1}{2} \{ (i, j) : w_i^j = -w_k^l \text{ and } w_i^j \sim w_k^l \}.$$

- \mathbb{L} the different energy levels. ie:

$$\mathbb{L} = \{ \theta(w) : w \in \chi \text{ with } w_p^q = +1 \}.$$

- δ_ℓ the number of configurations of a given level of energy is:

$$\delta_\ell = \# \{ w \in \chi \text{ with } w_p^q = +1 : \text{ such that } \theta(w) = \ell \}.$$

Chapter 3. The Convergence rate of the Gibbs sampler for the 2-D Ising model via a geometric bound

Then the probability measure of a configuration takes the following form:

$$\pi(w) = \frac{\exp \frac{1}{T}(2n(n-1) - 2\theta(w))}{2 \sum_{\ell \in \mathbb{L}} \delta_\ell \exp \frac{1}{T}(2n(n-1) - 2\theta_\ell)}$$

For an asymptotic study, we need to know the configuration that achieves the maximum for

$$\pi(w) e^{\frac{1}{T}(n-1+\alpha_\ell)}$$

We notice that the number of signs changing between the neighboring sites of a configuration has a direct effect on the energy level of a given configuration. We look to the lines from $j = q - 2$ to $j = q + 2$ in the given configuration $w \in \{-1, +1\}^{n^2}$ and we call by Δ its energy level, then:

$$\begin{aligned} \Delta &= \sum_{i=1}^{n-1} w_i^{q-2} w_{i+1}^{q-2} + \sum_{i=1}^{n-1} w_i^{q-1} w_{i+1}^{q-1} + \sum_{i=1}^{n-1} w_i^{q+1} w_{i+1}^{q+1} + \sum_{i=1}^{n-1} w_i^{q+2} w_{i+1}^{q+2} \\ &+ \sum_{i=1}^{p-2} w_i^q w_{i+1}^q + w_{p-1}^q + w_{p+1}^q + \sum_{i=p+1}^{n-1} w_i^q w_{i+1}^q \\ &+ \sum_{i=1}^{p-1} (w_i^{q-2} w_i^{q-1} + w_i^{q-1} w_i^q + w_i^q w_i^{q+1} + w_i^{q+1} w_i^{q+2}) + (w_p^{q-2} w_p^{q-1} + w_p^{q-1} \\ &+ w_p^{q+1} + w_p^{q+1} w_p^{q+2}) + \sum_{i=p+1}^n (w_i^{q-2} w_i^{q-1} + w_i^{q-1} w_i^q + w_i^q w_i^{q+1} + w_i^{q+1} w_i^{q+2}) \\ &+ \sum_{i=1}^{p-1} (-w_i^q + w_i^{q+1} - w_i^q w_i^{q+1}) - 2(w_{p+1}^q + w_p^{q+1}) + \sum_{i=p+1}^n (w_i^{q-1} - w_i^q - w_i^{q-1} w_i^q) \\ &= \sum_{i=1}^{n-1} w_i^{q-2} w_{i+1}^{q-2} + \sum_{i=1}^{n-1} w_i^{q-1} w_{i+1}^{q-1} + \sum_{i=1}^{n-1} w_i^{q+1} w_{i+1}^{q+1} + \sum_{i=1}^{n-1} w_i^{q+2} w_{i+1}^{q+2} \\ &+ \sum_{i=1}^{p-2} w_i^q w_{i+1}^q + w_{p-1}^q - w_{p+1}^q + \sum_{i=p+1}^{n-1} w_i^q w_{i+1}^q \\ &+ \sum_{i=1}^{p-1} (w_i^{q-2} w_i^{q-1} + w_i^{q-1} w_i^q - w_i^q + w_i^{q+1} + w_i^{q+1} w_i^{q+2}) + (w_p^{q-2} w_p^{q-1} + w_p^{q-1} \\ &- w_p^{q+1} + w_p^{q+1} w_p^{q+2}) + \sum_{i=p+1}^n (w_i^{q-2} w_i^{q-1} + w_i^{q-1} - w_i^q + w_i^q w_i^{q+1} + w_i^{q+1} w_i^{q+2}) \end{aligned}$$

We will build the maximum configuration in three steps. In the first step we notice that a first observation gives that the previous expression reaches its maximum value if :

- $-w_i^q = w_i^{q+1} = 1$ for $i = 1, \dots, p - 1$
- $w_i^{q-1} = -w_i^q = 1$ for $i = p + 1, \dots, n$
- $w_{p+1}^q = w_p^{q+1} = -1$.

3.6. Asymptotic study

Then in the second step, the sums of the near neighbors is maximal if they have similar signs which is illustrated in the following three points:

- $w_p^q = w_p^{q-1} = \dots = w_p^1 = 1$ and $w_p^{q+1} = w_p^{q+2} = \dots = w_p^n = -1$
- $w_i^{q-1} = w_i^{q-2} = \dots = w_i^1 = 1$ for $i = p+1, \dots, n$
- $w_i^q = w_i^{q+1} = \dots = w_i^n = -1$ for $i = p+1, \dots, n$.

Finally, in the last step, we notice that the choice of small p gives the best result. By the fact that we study the case beyond the board, we consider that $p = 2$ and the max is reached for:

- $w_1^{q-1} = w_p^{q-2} = \dots = w_p^1 = 1$
- $w_p^{q+2} = w_p^{q+3} = \dots = w_p^n = -1$.

A natural candidate takes the following form:

$$\tilde{w} = \begin{pmatrix} +1 & +1 & +1 & \dots & \dots & +1 \\ \vdots & \vdots & \vdots & & & \vdots \\ +1 & +1 & +1 & \dots & \dots & +1 \\ +1 & +1 & +1 & \dots & \dots & +1 \\ -1 & +1 & -1 & \dots & \dots & -1 \\ +1 & -1 & -1 & \dots & \dots & -1 \\ -1 & -1 & -1 & \dots & \dots & -1 \\ \vdots & \vdots & \vdots & & & \vdots \\ -1 & -1 & -1 & \dots & \dots & -1 \end{pmatrix}$$

Let

$$\theta_{\max} = n + 5, \quad \alpha_{\max} = 3n + 1 \quad \text{and} \quad \delta_{\max} = \#\{w : \theta(w) = \theta_{\max}\}.$$

Then for all $w \in \chi$, $w_p^q = +1$, an elementary calculus gives:

$$\pi(w)e^{\frac{1}{T}(n-1+\alpha_\ell)} \leq \frac{1}{Z_T} \exp \frac{1}{T} (2n(n-1) + 2(n-5)).$$

$$\begin{aligned} Q(e)^{-1} \sum_{\gamma_{xy} \ni e} |\gamma_{xy}| \pi(x) \pi(y) &\leq 2n^4 \times \frac{\sum_{\ell \in \mathbb{L}} \delta_\ell \exp \left\{ \frac{1}{T} (2n(n-1) - 2\theta_\ell + n - 1 + \alpha_\ell) \right\}}{2 \sum_{\ell \in \mathbb{L}} \delta_\ell \exp \left\{ \frac{1}{T} (2n(n-1) - 2\theta_\ell) \right\}} \\ &= 2n^4 \times \frac{\sum_{\ell \in \mathbb{L}} \delta_\ell \exp \left\{ \frac{1}{T} (-2\theta_\ell + n - 1 + \alpha_\ell) \right\}}{2 \sum_{\ell \in \mathbb{L}} \delta_\ell \exp \left\{ \frac{1}{T} (-2\theta_\ell) \right\}}. \end{aligned}$$

Chapter 3. The Convergence rate of the Gibbs sampler for the 2-D Ising model via a geometric bound

By the fact that $-2\theta_\ell + n - 1 + \alpha_\ell \leq 2(n - 5)$ for all $\ell \in \mathbb{L}$ there exist a non-negative relative ζ_ℓ such that:

$$Q(e)^{-1} \sum_{\gamma_{xy} \ni e} |\gamma_{xy}| \pi(x) \pi(y) \leq 2n^4 e^{\frac{2}{T}(n-5)} \times \frac{\sum_{\ell \in \mathbb{L}} \delta_\ell e^{-\frac{\zeta_\ell}{T}}}{2 \sum_{\ell \in \mathbb{L}} \delta_\ell e^{-\frac{2\theta_\ell}{T}}}$$

in fact

$$\zeta_\ell = 2(n - 5) - (n - 1) + 2\theta_\ell - \alpha_\ell.$$

We notice that:

$$\frac{\sum_{\ell \in \mathbb{L}} \delta_\ell e^{-\frac{\zeta_\ell}{T}}}{2 \sum_{\ell \in \mathbb{L}} \delta_\ell e^{-\frac{2\theta_\ell}{T}}} = \begin{cases} \frac{\delta_{\max+} \sum_{\ell \in \mathbb{L} \setminus \{\max\}} \delta_\ell e^{-\frac{\zeta_\ell}{T}}}{2\delta_{\max+2} \sum_{\ell \in \mathbb{L} \setminus \{\max\}} \delta_\ell e^{-\frac{2\theta_\ell}{T}}} \longrightarrow \frac{1}{2} & \text{if } T \longrightarrow 0, \\ \frac{\sum_{\ell \in \mathbb{L}} \delta_\ell e^{-\frac{\zeta_\ell}{T}}}{2 \sum_{\ell \in \mathbb{L}} \delta_\ell e^{-\frac{2\theta_\ell}{T}}} \longrightarrow \frac{\sum_{\ell \in \mathbb{L}} \delta_\ell}{2 \times \sum_{\ell \in \mathbb{L}} \delta_\ell} = \frac{1}{2} & \text{if } T \longrightarrow +\infty. \end{cases}$$

Finally we obtain:

$$Q(e)^{-1} \sum_{\gamma_{xy} \ni e} |\gamma_{xy}| \pi(x) \pi(y) \leq n^4 e^{\frac{2}{T}(n-5)}. \quad (3.14)$$

3.6.2 For $(i, j) = \{2, \dots, n - 1\} \times \{1\}$

According to Proposition 3.2.2 we have:

$$Q(e)^{-1} \sum_{\gamma_{xy} \ni e} |\gamma_{xy}| \pi(x) \pi(y) \leq n^4 \left(e^{\frac{n+1}{T}} + e^{\frac{n+3}{T}} \right) \sum_{w \in \mathcal{X}: w_1^q = 1} \pi(w) \exp \left\{ \frac{1}{T} \left(-2(w_2^q + w_1^{q+1}) + \sum_{i=2}^n (w_i^{q-1} - w_i^q - w_i^{q-1} w_i^q) \right) \right\}.$$

As previously, a candidate that realizes the max of

$$\pi(w) \exp \left\{ \frac{1}{T} \left(-2(w_2^q + w_1^{q+1}) + \sum_{i=2}^n (w_i^{q-1} - w_i^q - w_i^{q-1} w_i^q) \right) \right\}$$

takes the following form:

$$\tilde{w} = \begin{pmatrix} +1 & +1 & +1 & \cdots & \cdots & +1 \\ \vdots & \vdots & \vdots & & & \vdots \\ +1 & +1 & +1 & \cdots & \cdots & +1 \\ +1 & +1 & +1 & \cdots & \cdots & +1 \\ -1 & -1 & -1 & \cdots & \cdots & -1 \\ -1 & -1 & -1 & \cdots & \cdots & -1 \\ \vdots & \vdots & \vdots & & & \vdots \\ -1 & -1 & -1 & \cdots & \cdots & -1 \end{pmatrix}$$

3.6. Asymptotic study

With the same tricks as previous, an asymptotic analysis gives for high and low temperature:

$$\begin{aligned} Q(e)^{-1} \sum_{\gamma_{xy} \ni e} |\gamma_{xy}| \pi(x) \pi(y) &\leq \frac{n^4}{2} (e^{\frac{2}{T}(n-1)} + e^{\frac{2}{T}n}) \\ &\leq n^4 e^{\frac{2n}{T}}. \end{aligned} \quad (3.15)$$

We obtain the same result for $(i, j) = (2 \cdots n - 1, n)$. Then with similar techniques we can do the same calculus for $(i, j) = (1, 2 \cdots n - 1)$ and $(i, j) = (n, 2 \cdots n - 1)$.

3.6.3 For $(i, j) = (1, n)$

From proposition 3.2.2 we have

$$\begin{aligned} Q(e)^{-1} \sum_{\gamma_{xy} \ni e} |\gamma_{xy}| \pi(x) \pi(y) &\leq 2n^4 \sum_{w \in \mathcal{X}: w_1^n = 1} \pi(w) \exp \left\{ \frac{1}{T} \left(2(1 - w_n^2) \right. \right. \\ &\quad \left. \left. + \sum_{i=1}^{n-1} (w_i^1 - w_i^2 + 1 - w_i^1 w_i^2) \right) \right\}. \end{aligned}$$

A natural candidate that realize the max of

$$\pi(w) \exp \left\{ \frac{1}{T} \left(2(1 - w_n^2) + \sum_{i=1}^{n-1} (w_i^1 - w_i^2 + 1 - w_i^1 w_i^2) \right) \right\}$$

takes the form:

$$\tilde{w} = \begin{pmatrix} +1 & +1 & +1 & \cdots & \cdots & +1 \\ -1 & -1 & -1 & \cdots & \cdots & -1 \\ +1 & +1 & +1 & \cdots & \cdots & +1 \\ \vdots & \vdots & \vdots & & & \vdots \\ \vdots & \vdots & \vdots & & & \vdots \\ -1 & -1 & -1 & \cdots & \cdots & -1 \end{pmatrix}$$

κ is bounded for high and low temperature as follow:

$$Q(e)^{-1} \sum_{\gamma_{xy} \ni e} |\gamma_{xy}| \pi(x) \pi(y) \leq n^4 e^{\frac{2n}{T}}. \quad (3.16)$$

We obtains the same result for $(i, j) = (n, 1)$.

By regrouping the results in (3.14), (3.8), (3.15) and (3.16) we obtains the following upper bound for β_1 :

Theorem 3.6.1. *The second largest eigenvalue of the Gibbs sampler for the 2-D Ising model satisfies at high and low temperature:*

$$\beta_1 < 1 - n^{-4} e^{-\frac{2n}{T}}.$$

Then, for all Temperature T , we have the following result:

Chapter 3. The Convergence rate of the Gibbs sampler for the 2-D Ising model via a geometric bound

Corollary 3.6.2. *There exists a real number τ_n such that the second largest eigenvalue of the Gibbs sampler for the 2-D Ising model satisfies:*

$$\beta_1 < 1 - n^{-4} \tau_n^{-1} e^{\frac{-2}{T}(n+1)}.$$

Remark 3.6.3. From Theorem 3.6.1 and Corollary 3.6.2 we deduce that τ_n is bounded as follows:

$$0 < \tau_n \leq e^{\frac{2}{T}(n+2)}.$$

Chapter 4

The convergence rate of some perturbed Gibbs sampler for the 1–D Ising model

In this chapter, We study the geometric bound introduced by Diaconis and Stroock (1991) of some perturbed Gibbs sampler for the one-dimensional Ising model. Shiu and Chen (2015) studies this problem for the classical Gibbs algorithm and obtained the best geometric bound in the literature for this case. We observe that the disruption of the Gibbs sampler improve the result of Chen and Shiu (2015) at high temperature. However, at low temperature we did not succeed to improve the geometric bound, although an improvement should be expected.

Dans ce chapitre, Nous étudions la borne géométrique introduite par Diaconis et Stroock (1991) pour l'algorithme de l'échantillonneur de Gibbs perturbé pour le modèle d'Ising unidimensionnel. Shiu et Chen (2015) ont étudié ce problème pour le cas de l'échantillonneur de Gibbs classique et ils ont obtenu la meilleure borne dans la littérature pour ce cas. On observe que la perturbation de l'algorithme de l'échantillonneur de Gibbs améliore le resultat de Shiu et Chen (2015) à haute temperature. Cependant, à basse température, nous n'avons pas réussi à améliorer la borne géométrique, même si une amélioration est à prévoir.

4.1 Introduction

The Ising model is a basic model for ferromagnetism. It is the simplest model of statistical mechanics and it has been applied in many other fields like chemistry, molecular biology and image analysis. The distribution of the 1-D Ising model is:

$$\pi(x) = \frac{1}{Z_T} \exp \left\{ \frac{1}{T} \sum_{k=0}^{n-1} x_k x_{k+1} \right\} \quad \forall x = (x_1, \dots, x_n) \in \chi,$$

where $\chi = \{-1, 1\}^n$ is the state space, T is the temperature and

$$Z_T = \sum_{x \in \chi} \exp \left\{ \frac{1}{T} \sum_{k=1}^{n-1} x_k x_{k+1} \right\}$$

is the normalizing constant. Monte Carlo Markov chain *MCMC* method is a very useful technique to draw samples from the Ising model.

Suppose that the transition probability $P(x, y)$ for an irreducible Markov chain has π as its invariant measure. The pair (P, π) is said to be reversible if it verifies the *detailed balance equation*:

$$Q(x, y) := \pi(x)P(x, y) = \pi(y)P(y, x) := Q(y, x) \quad \forall x, y \in \chi.$$

The Gibbs sampler introduced by Geman and Geman (1986) (see [11]) and the Metropolis-Hastings algorithm introduced by Metropolis, Rosenbluth, Rosenbluth, Teller and Teller (1953) (see [21]) and Hastings (1970) (see [13]) are both the most popular Monte Carlo Markov chain methods. The Gibbs sampler chooses a random coordinate which is updated according to the conditional probability given the other coordinates. The resulting Markov chain is reversible. The associated transition matrix has the form:

$$P(x, y) = \begin{cases} \frac{1}{n} \pi(y_i | x) & \text{if } d(x, y) = 1 \\ 1 - \frac{1}{n} \sum_{i=1}^n \pi(y_i | x) & \text{if } d(x, y) = 0 \\ 0 & \text{else} \end{cases}$$

where

- $\pi(y_i | x) = \frac{\pi(x_1, \dots, x_{i-1}, y_i, x_{i+1}, \dots, x_n)}{\pi(x_1, \dots, x_{i-1}, y_i, x_{i+1}, \dots, x_n) + \pi(x_1, \dots, x_{i-1}, -y_i, x_{i+1}, \dots, x_n)}$;
- $d(x, y) = \text{card}\{i, x_i \neq y_i \text{ for } i = 1, \dots, n\}$ designates the number of sites that differ between two configurations x and y .

One expects an acceleration of the convergence rate of the Gibbs sampler to its equilibrium once one adds some edges to the graph underlying the initial Markov chain. However only few edges should be added in order to keep the computational cost low.

Starting from this philosophy, we propose some perturbation of the Gibbs algorithm based

4.1. Introduction

on some permutation σ of χ .

Let σ be a two cyclic permutation defined for $x \in \chi$ as follow:

$$\begin{aligned}\sigma : \chi &\longrightarrow \chi \\ x &\longrightarrow -x.\end{aligned}$$

Remark 4.1.1. We remark that:

- $\sigma(\sigma(x)) = x$.
- π is σ -invariant.

For some $0 < \alpha < 1$ we denote by P_σ the matrix which is defined for $x, y \in \chi$ as:

$$P_\sigma(x, y) = \begin{cases} \alpha & \text{if } y = \sigma(x) \\ 1 - \alpha & \text{if } x = y \\ 0 & \text{else .} \end{cases}$$

The idea behind this construction is that the permutation we want to introduce should in one hand not add many edges to the graph underlying the Markov chain of the sampler and in the other hand the previous remark guarantees the reversibility of the Markov chain which allows us thereafter to use the geometric bound introduced by Diaconnis and Stroock (1991) in [5]. Furthermore, the computational effort to run a simulation for this simple permutation is very cheap.

For $0 < \delta < 1$, we introduce the perturbed Gibbs algorithm as follows:

$$\tilde{P}(x, y) = \delta P(x, y) + (1 - \delta) P_\sigma(x, y) \quad \forall x, y \in \chi.$$

We remark that $\tilde{P}(x, y)$ is reversible as a linear combination of two reversible matrices. Their eigenvalues can be arranged as follows:

$$1 = \tilde{\beta}_0 > \tilde{\beta}_1 \geq \tilde{\beta}_2 \geq \dots \geq \tilde{\beta}_{|\chi|-1} > -1.$$

Let $\tilde{\beta}^* = \max\{\tilde{\beta}_1, |\tilde{\beta}_{|\chi|-1}|\}$. Then the second largest eigenvalue in absolute value can be used to measure the distance between \tilde{P}^m and π . This follows from the Perron-Frobenius Theorem [19]. Suppose that $\Pi = (\pi, \dots, \pi)^t$ a $n \times n$ matrix where π is a $n \times 1$ vector then we have:

$$\tilde{P}^m - \Pi = O(m^{\gamma-1}(\beta^*)^m)$$

where γ is the multiplicity of $\tilde{\beta}^*$. If \tilde{P} is reversible we have a stronger result. In theorem 1.2.1 Diaconnis and Stroock [5] gives a bound on the variation distance to equilibrium in terms of $\tilde{\beta}^*$.

Moreover, Diaconnis and Stroock [5] develop a method to calculate an upper bound for the second largest eigenvalue $\tilde{\beta}_1$ for a reversible Markov chain using geometric quantities such as the maximum degree, diameter and covering number of the associated graph. Consider a graph with a vertex set χ and an edge set $E = \{(x, y) \mid \tilde{P}(x, y) > 0\}$. For each pair of distinct points $x, y \in \chi$ we choose a path γ_{xy} from x to y , such that each edge appears at most once in a given path. The fact that \tilde{P} is irreducible guarantees that

such paths exist. Let Γ be the collection of γ_{xy} (one for each pair). The geometric bound given by Diaconis and Stroock [5] is,

$$\tilde{\beta}_1 \leq 1 - \frac{1}{\tilde{\kappa}} \tag{4.1}$$

where

$$\tilde{\kappa} = \max_{e \in E} \tilde{Q}(e)^{-1} \sum_{\gamma_{xy} \ni e} |\gamma_{xy}| \pi(x) \pi(y) \tag{4.2}$$

and $|\gamma_{xy}|$ designates the length of the path γ_{xy} .

A suitable choice of the trajectories γ_{xy} linking x and y was proposed by Shiu and Chen in [25]. They obtained the following bound:

Theorem 4.1.2. [Shiu and Chen 2015] *The second largest eigenvalue of the Gibbs sampler on the 1-D Ising model satisfies:*

$$\beta_1 \leq 1 - e^{-\frac{2}{T} n^{-2}}.$$

In this chapter we discuss the improvement obtained by the passage from P to \tilde{P} in terms of the bound of Diaconis and Stroock for low and high temperature asymptotic. The chapter is organized as follow. At the beginning, we define a suitable path for each pair (x, y) of vertices in the state space χ ; we distinguish two cases:

- If the couple (x, y) differs by less or equal than $\frac{n}{2}$ sites we use the same paths as Shiu and Chen in [25].
- If $d(x, y) > \frac{n}{2}$ then γ first passes from x to $-x$ and then follows the trajectory proposed by Shiu and Chen as previous.

Then, we use the result introduced by Diaconis and Stroock in [5] to give a bound of the second largest eigenvalue associated to the perturbed Gibbs sampler for the 1-D Ising model; Indeed for each edge e we calculate separately the sum of the couples (x, y) where $d(x, y) \leq \frac{n}{2}$ and $d(x, y) > \frac{n}{2}$ passing through e . The main theorem is that at low temperature, we obtain the same result as Shiu and Chen in [25] but at high temperature, we double the convergence rate to equilibrium. Finally, by referring to [20] we give a bound for the second largest eigenvalue in absolute value. Some notations and elementary remarks which are used in the end of the proof of the main result are given in the appendix.

4.2 Main result

4.2.1 Selection of paths

To be able to use the result of Diaconis and Stroock [5] to calculate the bound of the second largest eigenvalue, we should fix a collection of paths. To get the best result, we should use the shortest paths γ_{xy} to obtain smallest κ which gives smaller upper bound of β_1 . Then, for each pair (x, y) we select one of the shorter paths.

We have two classes of edges:

4.2. Main result

- $e = (e^-, e^+)$ such that $d(e^-, e^+) = 1$.
- $e = (e^-, e^+)$ such that $d(e^-, e^+) = n$.

Let d_1, \dots, d_m be an increasing sequence, for any pair $(x, y) \in \chi^2$ suppose that $x_i \neq y_i$ for $i \in \{d_1, \dots, d_m\}$ and $x_i = y_i$ otherwise.

The choice of path linking configurations depends on the distance between them, so we have:

1. If $m \leq \frac{n}{2}$ we use the technique introduced by Shiu and Chen in [25] then:

$$\begin{aligned}
 (x_1, \dots, x_n) &= (y_1, \dots, y_{d_1-1}, x_{d_1}, x_{d_1+1}, \dots, x_{d_2-1}, x_{d_2}, x_{d_2+1}, \dots, x_n) \\
 &\rightarrow (y_1, \dots, y_{d_1-1}, y_{d_1}, x_{d_1+1}, \dots, x_{d_2-1}, x_{d_2}, x_{d_2+1}, \dots, x_n) \\
 &= (y_1, \dots, y_{d_1-1}, y_{d_1}, y_{d_1+1}, \dots, y_{d_2-1}, x_{d_2}, x_{d_2+1}, \dots, x_n) \\
 &\rightarrow (y_1, \dots, y_{d_1-1}, y_{d_1}, y_{d_1+1}, \dots, y_{d_2-1}, y_{d_2}, x_{d_2+1}, \dots, x_n) \\
 &\vdots \\
 &\rightarrow (y_1, \dots, y_n).
 \end{aligned}$$

2. If $m \geq \frac{n}{2}$, as the configurations differ by more than $\frac{n}{2}$ sites we change to $-x$ in the first step and we continue with the technique introduced in [25]:

$$\begin{aligned}
 (x_1, \dots, x_n) &= (y_1, \dots, y_{d_1-1}, x_{d_1}, x_{d_1+1}, \dots, x_{d_2-1}, x_{d_2}, x_{d_2+1}, \dots, x_n) \\
 &\rightarrow (-y_1, \dots, -y_{d_1-1}, -x_{d_1}, -x_{d_1+1}, \dots, -x_{d_2-1}, -x_{d_2}, -x_{d_2+1}, \dots, -x_n) \\
 &= (-y_1, \dots, -y_{d_1-1}, y_{d_1}, -x_{d_1+1}, \dots, -x_{d_2-1}, y_{d_2}, -x_{d_2+1}, \dots, -x_n) \\
 &\rightarrow (y_1, \dots, -y_{d_1-1}, y_{d_1}, -x_{d_1+1}, \dots, -x_{d_2-1}, y_{d_2}, -x_{d_2+1}, \dots, -x_n) \\
 &\vdots \\
 &\rightarrow (y_1, \dots, y_n).
 \end{aligned}$$

4.2.2 Geometric bound for the second largest eigenvalue

In what follows we will find an upper bound of

$$\tilde{\kappa} = \max_{e \in E} \tilde{Q}(e)^{-1} \sum_{\tilde{\gamma}_{xy} \ni e} |\gamma_{xy}| \pi(x) \pi(y).$$

Let $e = (e^-, e^+)$ defined as previously. Then we have:

$$\{(x, y) | \tilde{\gamma}_{xy} \ni e\} = \{(x, y) : d(x, y) \leq \frac{n}{2} \mid \tilde{\gamma}_{xy} \ni e\} \cup \{(x, y) : d(x, y) > \frac{n}{2} \mid \tilde{\gamma}_{xy} \ni e\}.$$

Classical edges $e = (e^-, e^+)$

Let $e = (e^-, e^+)$ with

$$e^- = (z_1, z_2, \dots, z_{i-1}, z_i, z_{i+1}, \dots, z_n) \text{ and } e^+ = (z_1, z_2, \dots, z_{i-1}, -z_i, z_{i+1}, \dots, z_n).$$

**Chapter 4. The convergence rate of some perturbed Gibbs sampler for the
1-D Ising model**

First we consider the case that $i \neq 1$ or n . The transition matrix is:

$$\tilde{Q}(e) = \pi(e^-) \tilde{P}(e^-, e^+) = \delta\pi(e^-) P(e^-, e^+),$$

and

$$\begin{aligned} P(e^-, e^+) &= \frac{1}{n} \frac{\pi(e^+)}{(\pi(e^-) + \pi(e^+))} \\ &= \frac{1}{n(1 + e^{\frac{2}{T}(z_{i-1}z_i + z_i z_{i+1})})}. \end{aligned}$$

We remark also that: $|\tilde{\gamma}_{xy}| \leq \frac{n}{2}$. Define

$x \oplus y = (x_1, \dots, x_{i-1}, z_i, y_{i+1}, \dots, y_n)$ and $x \ominus y = (x_1, \dots, x_{i-1}, -z_i, y_{i+1}, \dots, y_n)$.

According to the selection of paths in section 4.2.1(1). For any $\tilde{\gamma}_{xy} \ni e$, where $d(x, y) \leq \frac{n}{2}$, x and y must have the following form:

$$x = (x_1, x_2, \dots, x_{i-1}, z_i, z_{i+1}, \dots, z_n) \text{ and } y = (z_1, z_2, \dots, z_{i-1}, -z_i, y_{i+1}, \dots, y_n).$$

Then we obtain:

$$\begin{aligned} \tilde{Q}(e)^{-1} \pi(x) \pi(y) &= \frac{\pi(x) \pi(y)}{\delta\pi(e^-) P(e^-, e^+)} \\ &= \frac{n}{\delta Z_T} \left(1 + e^{\frac{2}{T}(z_{i-1}z_i + z_i z_{i+1})}\right) \times \frac{\exp\left(\frac{1}{T} \sum_{k=1}^{n-1} x_k x_{k+1}\right) + \exp\left(\frac{1}{T} \sum_{k=1}^{n-1} y_k y_{k+1}\right)}{\exp\left(\frac{1}{T} \sum_{k=1}^{n-1} e_k^- e_{k+1}^-\right)} \\ &= \frac{n}{\delta Z_T} \left\{ \exp \frac{1}{T} \left(\sum_{k=1}^{i-2} x_k x_{k+1} + x_{i-1} z_i + z_i y_{i+1} + \sum_{k=i+1}^{n-1} y_k y_{k+1} \right) e^{-\frac{2}{T}(z_{i-1}z_i + z_i y_{i+1})} \right. \\ &\quad \left. + \exp \frac{1}{T} \left(\sum_{k=1}^{i-2} x_k x_{k+1} - x_{i-1} z_i - z_i y_{i+1} + \sum_{k=i+1}^{n-1} y_k y_{k+1} \right) e^{\frac{2}{T}(z_i z_{i+1} + x_{i-1} z_i)} \right\}. \end{aligned}$$

It gives that:

$$\tilde{Q}(e)^{-1} \pi(x) \pi(y) = \frac{n}{\delta} \left\{ \pi(x \oplus y) \exp \left\{ -\frac{2}{T}(z_{i-1}z_i + z_i y_{i+1}) \right\} + \pi(x \ominus y) \exp \left\{ \frac{2}{T}(z_i z_{i+1} + x_{i-1} z_i) \right\} \right\}. \quad (4.3)$$

According to the selection of paths in section 4.2.1(2). For any $\tilde{\gamma}_{xy} \ni e$, where $d(x, y) \leq \frac{n}{2}$, x and y must have the following form:

$$x = (x_1, x_2, \dots, x_{i-1}, -z_i, -z_{i+1}, \dots, -z_n) \text{ and } y = (z_1, z_2, \dots, z_{i-1}, -z_i, y_{i+1}, \dots, y_n).$$

4.2. Main result

Then we obtain:

$$\begin{aligned}
\tilde{Q}(e)^{-1}\pi(x)\pi(y) &= \frac{\pi(x)\pi(y)}{\delta\pi(e^-)P(e^-, e^+)} \\
&= \frac{1}{\delta Z_T} \frac{\exp\left(\frac{1}{T}\sum_{k=1}^{n-1} x_k x_{k+1}\right) \times \exp\left(\frac{1}{T}\sum_{k=1}^{n-1} y_k y_{k+1}\right)}{\frac{1}{n\left(1+e^{\frac{2}{T}(z_{i-1}z_i+z_i z_{i+1})}\right)} \times \exp\left(\frac{1}{T}\sum_{k=1}^{n-1} e_k^- e_{k+1}^-\right)} \\
&= \frac{n}{\delta} \left(1 + e^{\frac{2}{T}(z_{i-1}z_i+z_i z_{i+1})}\right) e^{-\frac{2}{T}z_{i-1}z_i} \exp\left\{\frac{1}{T}\left(\sum_{k=1}^{i-2} x_k x_{k+1} \right. \right. \\
&\quad \left. \left. - x_{i-1}z_i - z_i y_{i+1} + \sum_{k=i+1}^{n-1} y_k y_{k+1}\right)\right\} \\
&= \frac{n}{\delta} \left(e^{-\frac{2}{T}z_{i-1}z_i} + e^{\frac{2}{T}z_i z_{i+1}}\right) \exp\left\{\frac{1}{T}\left(\sum_{k=1}^{i-2} x_k x_{k+1} - x_{i-1}z_i \right. \right. \\
&\quad \left. \left. - z_i y_{i+1} + \sum_{k=i+1}^{n-1} y_k y_{k+1}\right)\right\}
\end{aligned}$$

It gives that:

$$\tilde{Q}(e)^{-1}\pi(x)\pi(y) = \frac{n}{\delta} \left(e^{-\frac{2}{T}z_{i-1}z_i} + e^{\frac{2}{T}z_i z_{i+1}}\right) \pi(x \oplus y). \quad (4.4)$$

Let

$$\begin{aligned}
\chi_1 &= \bigcup_{\{(x,y):d(x,y)\leq\frac{n}{2}|\gamma_{xy}\ni e\}} \{x \oplus y, x \ominus y\}. \\
\chi_2 &= \bigcup_{\{(x,y):d(x,y)>\frac{n}{2}|\gamma_{xy}\ni e\}} \{x \oplus y, x \ominus y\}.
\end{aligned}$$

According to equation (4.3) we have:

$$\begin{aligned}
\sum_{\tilde{\gamma}_{xy}\ni e, d(x,y)\leq\frac{n}{2}} \tilde{Q}(e)^{-1}\pi(x)\pi(y) &= \frac{n}{\delta} \sum_{\{\tilde{\gamma}_{xy}\ni e, d(x,y)\leq\frac{n}{2}\}} \left(\pi(x \oplus y) e^{-\frac{2}{T}(z_{i-1}z_i+z_i y_{i+1})} \right. \\
&\quad \left. + \frac{n}{\delta} \pi(x \ominus y) e^{\frac{2}{T}(z_i z_{i+1}+x_{i-1}z_i)}\right) \\
&= \frac{n}{\delta} \left(\sum_{w\in\chi_1, w_i=z_i} \pi(w) e^{-\frac{2}{T}(z_{i-1}z_i+z_i w_{i+1})} \right. \\
&\quad \left. + \sum_{w\in\chi_1, w_i=-z_i} \pi(w) e^{\frac{2}{T}(z_i z_{i+1}+w_{i-1}z_i)} \right). \quad (4.5)
\end{aligned}$$

Also we have:

$$\begin{aligned}
\sum_{\tilde{\gamma}_{xy}\ni e, d(x,y)>\frac{n}{2}} \tilde{Q}(e)^{-1}\pi(x)\pi(y) &= \sum_{\tilde{\gamma}_{xy}\ni e, d(x,y)>\frac{n}{2}} \frac{n}{\delta} \times \left(e^{-\frac{2}{T}z_{i-1}z_i} + e^{\frac{2}{T}z_i z_{i+1}}\right) \times \pi(x \ominus y) \\
&= \frac{n}{\delta} \left(e^{-\frac{2}{T}z_{i-1}z_i} + e^{\frac{2}{T}z_i z_{i+1}}\right) \sum_{w\in\chi_2, w_i=-z_i} \pi(w). \quad (4.6)
\end{aligned}$$

By examining the equations (4.3) and (4.4) we notice the max is reached for

$$-z_{i-1} = z_i = z_{i+1}.$$

We assume that $-z_{i-1} = z_i = z_{i+1} = 1$. Using (4.5) and (4.6) we obtain:

$$\begin{aligned} \max_{e \in E} \tilde{Q}(e)^{-1} \sum_{\tilde{\gamma}_{xy} \ni e} |\tilde{\gamma}_{xy}| \pi(x) \pi(y) &\leq \frac{n}{2} \times \tilde{Q}(e)^{-1} \sum_{\tilde{\gamma}_{xy} \ni e} \pi(x) \pi(y) \\ &\leq \frac{n^2}{2\delta} \left\{ \sum_{w \in \chi_1, w_i=1} \pi(w) e^{\frac{2}{T}(1-w_{i+1})} + \sum_{w \in \chi_1, w_i=-1} \pi(w) e^{\frac{2}{T}(1+w_{i-1})} \right. \\ &\quad \left. + 2e^{\frac{2}{T}} \sum_{w \in \chi_2, w_i=-1} \pi(w) \right\} \\ &= \frac{n^2}{\delta} \left(\sum_{w \in \chi_1, w_i=1} \pi(w) e^{\frac{2}{T}(1-w_{i+1})} + e^{\frac{2}{T}} \sum_{w \in \chi_2, w_i=-1} \pi(w) \right). \end{aligned}$$

The last equality was obtained by the fact that:

$$\begin{aligned} \sum_{w \in \chi_1, w_i=1} \pi(w) e^{\frac{2}{T}(1-w_{i+1})} &= \sum_{w \in \chi_1, w_i=1} \frac{1}{Z_T} \exp \left\{ \frac{1}{T} \left(\sum_{k=1}^{i-2} w_k w_{k+1} + \sum_{k=i+1}^{n-1} w_k w_{k+1} \right) \right\} \\ &\quad \times \exp \left\{ \frac{1}{T} (2 + w_{i-1} - w_{i+1}) \right\} \\ &= \sum_{w \in \chi_1, w_i=-1} \frac{1}{Z_T} \exp \left\{ \frac{1}{T} \left(\sum_{k=1}^{i-2} w_k w_{k+1} + \sum_{k=i+1}^{n-1} w_k w_{k+1} \right) \right\} \\ &\quad \times \exp \left\{ \frac{1}{T} (2 + w_{i-1} - w_{i+1}) \right\} \\ &= \sum_{w \in \chi_1, w_i=-1} \pi(w) e^{\frac{2}{T}(1+w_{i-1})}. \end{aligned}$$

Remark 4.2.1. We obtain the same result by considering $-z_{i-1} = z_i = z_{i+1} = -1$.

It gives:

$$\kappa \leq \frac{n^2}{\delta} e^{\frac{2}{T}} \times \left(\sum_{w \in \chi_1, w_i=1} \pi(w) e^{\frac{-2}{T} w_{i+1}} + \sum_{w \in \chi_2, w_i=-1} \pi(w) \right).$$

Now, we will control the term:

$$\sum_{w \in \chi_1, w_i=1} \pi(w) e^{\frac{-2}{T} w_{i+1}} + \sum_{w \in \chi_2, w_i=-1} \pi(w).$$

Denote by δ_k^1 the number of states in $\{\chi_1 : w_i = 1\}$ where:

$$\pi(w) e^{-\frac{2}{T} w_{i+1}} = \frac{1}{Z_T} e^{\frac{n-1-2k}{T}}$$

and δ_k^2 the number of states in $\{\chi_2 : w_i = -1\}$ where:

$$\pi(w) = \frac{1}{Z_T} e^{\frac{n-1-2k}{T}}.$$

4.2. Main result

By referring to the Appendix we have:

$$\sum_{k=0}^{n-1} (\delta_k^1 + \delta_k^2) = |\chi_1^+| + |\chi_2^-| = 2^{n-1}$$

and we obtain:

$$\begin{aligned} \sum_{w \in \chi_1, w_i=1} \pi(w) e^{\frac{-2}{T} w_{i+1}} + \sum_{w \in \chi_2, w_i=-1} \pi(w) &= \frac{\sum_{k=0}^{n-1} (\delta_k^1 + \delta_k^2) e^{\frac{n-1-2k}{T}}}{2 \sum_{k=0}^{n-1} \binom{n-1}{k} e^{\frac{n-1-2k}{T}}} \\ &= \frac{\sum_{k=0}^{n-1} (\delta_k^1 + \delta_k^2) e^{\frac{-2k}{T}}}{2 \sum_{k=0}^{n-1} \binom{n-1}{k} e^{\frac{-2k}{T}}} \\ &= \frac{\delta_0^1 + \delta_0^2 + \sum_{k=1}^{n-1} (\delta_k^1 + \delta_k^2) e^{\frac{-2k}{T}}}{2 + 2 \sum_{k=1}^{n-1} \binom{n-1}{k} e^{\frac{-2k}{T}}}. \end{aligned}$$

Then we have

$$\frac{\delta_0^1 + \delta_0^2 + \sum_{k=1}^{n-1} (\delta_k^1 + \delta_k^2) e^{\frac{-2k}{T}}}{2 + 2 \sum_{k=1}^{n-1} \binom{n-1}{k} e^{\frac{-2k}{T}}} \rightarrow \begin{cases} \frac{\delta_0^1 + \delta_0^2}{2} & \text{if } T \rightarrow 0, \\ \frac{\sum_{k=0}^{n-1} (\delta_k^1 + \delta_k^2)}{\sum_{k=0}^{n-1} \binom{n-1}{k}} = \frac{2^{n-1}}{2^n} = \frac{1}{2} & \text{if } T \rightarrow +\infty. \end{cases}$$

At low temperature, there exist a unique configuration $w \in \chi_1^+$ that realizes the maximal energy, ie: $\pi(w) e^{-\frac{2}{T} w_{i+1}} = \frac{1}{Z_T} e^{\frac{n-1}{T}}$ and it takes the following form:

$$w = (+1, +1, \dots, +1, +1, -1) \text{ where } i = n - 1.$$

It gives that the edge $e = (e^-, e^+)$ have the form:

$$e = \begin{cases} e^- = (z_1, z_2, \dots, z_{n-3}, -1, +1, +1) \\ e^+ = (z_1, z_2, \dots, z_{n-3}, -1, -1, +1) \end{cases},$$

and the couple (x, y) such that $d(x, y) \leq \frac{n}{2}$ takes the following form:

$$\begin{cases} x = (+1, \dots, +1, +1, +1, +1) \\ y = (z_1, \dots, z_{n-3}, -1, -1, -1). \end{cases}$$

Chapter 4. The convergence rate of some perturbed Gibbs sampler for the 1-D Ising model

In the other hand, there exist a unique configuration $w \in \chi_2^-$ that realize the maximal energy, ie= $\pi(w) = \frac{1}{Z_T} e^{\frac{n-1}{T}}$ and it takes the following form:

$$w = (-1, -1, \dots, -1, -1, -1).$$

The corresponding couple (x, y) where $\gamma_{xy} \ni e$ and $d(x, y) > \frac{n}{2}$ takes the form:

$$\begin{cases} x = (-1, \dots, -1, -1, -1, -1) \\ y = (z_1, \dots, z_{n-3}, -1, -1, -1). \end{cases}$$

It gives that at low temperature we have the following result:

$$\frac{\delta_0^1 + \delta_0^2}{2} \longrightarrow 1.$$

It gives that:

- At low temperature:

$$\tilde{Q}(e)^{-1} \sum_{(x,y): \tilde{\gamma}_{xy} \ni e} |\gamma_{xy}| \pi(x) \pi(y) \leq \frac{n^2}{\delta} e^{\frac{2}{T}} \quad (4.7)$$

- At high temperature:

$$\tilde{Q}(e)^{-1} \sum_{(x,y): \tilde{\gamma}_{xy} \ni e} |\gamma_{xy}| \pi(x) \pi(y) \leq \frac{n^2}{2\delta} e^{\frac{2}{T}} \quad (4.8)$$

We turn now to the boundary cases.

- For $i = 1$, x and e^- coinsides and $y = (-z_1, y_2, \dots, y_n)$. The transition matrix is

$$\tilde{P}(e^-, e^+) = \delta \frac{\pi(e^+)}{n(\pi(e^-) + \pi(e^+))} = \frac{\delta}{n(1 + e^{\frac{2}{T} z_1 z_2})}.$$

Then

$$\tilde{Q}(e)^{-1} \pi(x) \pi(y) = \frac{n(1 + e^{\frac{2}{T} z_1 z_2}) \pi(y)}{\delta}.$$

The max is reached for $z_1 = z_2$ which gives for all $e \in E$:

$$\tilde{Q}(e)^{-1} \pi(x) \pi(y) \leq \frac{2n}{\delta} e^{\frac{2}{T}} \pi(y).$$

Finally, we have:

$$\begin{aligned} \tilde{Q}(e)^{-1} \sum_{\gamma_{xy} \ni e} |\gamma_{xy}| \pi(x) \pi(y) &\leq \frac{n}{2} \times \frac{2n}{\delta} e^{\frac{2}{T}} \sum_{\gamma_{xy} \ni e} \pi(y) \\ &< \frac{n^2}{\delta} e^{\frac{2}{T}} \sum_{\{y \in \chi, y_1 = -z_1, d(e^-, y) \leq \frac{n}{2}\}} \pi(y) \\ &\leq \frac{n^2}{2\delta} e^{\frac{2}{T}}. \end{aligned} \quad (4.9)$$

4.2. Main result

The last inequality was given by the fact that:

$$1 = \sum_{\gamma_{e^-y} \ni e} \pi(y) = \sum_{\{\gamma_{e^-y} \ni e: d(e^-, y) \leq \frac{n}{2}\}} \pi(y) + \sum_{\{\gamma_{e^-y} \ni e: d(e^-, y) > \frac{n}{2}\}} \pi(y)$$

and

$$\sum_{\{\gamma_{e^-y} \ni e: d(e^-, y) \leq \frac{n}{2}\}} \pi(y) = \sum_{\{\gamma_{e^-y} \ni e: d(e^-, y) \leq \frac{n}{2}, y_1 = -z_1\}} \pi(y) + \sum_{\{\gamma_{e^-y} \ni e: d(e^-, y) \leq \frac{n}{2}, y_1 = z_1\}} \pi(y).$$

Which leads to:

$$\sum_{\{\gamma_{e^-y} \ni e: d(e^-, y) \leq \frac{n}{2}, y_1 = -z_1\}} \pi(y) \leq \frac{1}{2}.$$

- For $i = n$, we have $x = (x_1, \dots, x_{n-1}, z_n)$ and $y = e^+$ and with the same techniques as previously we obtain:

$$\tilde{Q}(e)^{-1} \pi(x) \pi(y) = \frac{n}{\delta} (1 + e^{\frac{2}{T} z_{n-1} z_n}) \pi(x).$$

Then the max is obtained if $z_i = -z_{i-1}$, so we have:

$$\begin{aligned} \tilde{Q}(e)^{-1} \sum_{\gamma_{xy} \ni e} |\gamma_{xy}| \pi(x) \pi(y) &\leq \frac{n}{2\delta} (1 + e^{\frac{2}{T}}) \sum_{x \in \chi, x_n = z_n} \pi(x) \\ &\leq \frac{n^2}{2\delta} e^{\frac{2}{T}}. \end{aligned} \tag{4.10}$$

non classical edges $e = (e^-, \sigma(e^-) = e^+)$

The transition matrix is

$$\tilde{Q}(e) = \pi(e^-) \tilde{P}(e^-, e^+) = \alpha(1 - \delta) \pi(e^-).$$

As we apply the permutation only in the first step of each path γ_{xy} then all paths that passing through the edge e have the property that x and e^- coincides. It gives:

$$\begin{aligned} \frac{\pi(x) \pi(y)}{\tilde{Q}(e)} &= \frac{\pi(x) \pi(y)}{\alpha(1 - \delta) \pi(e^-)} \\ &= \frac{\pi(y)}{\alpha(1 - \delta)}. \end{aligned}$$

For e^- fixed, we have two disjoint sets such that:

$$\{y : d(e^-, y) \leq \frac{n}{2}\} \cup \{y : d(e^-, y) > \frac{n}{2}\} = \chi.$$

Noticing that for all $y \in \chi$ such that $d(e^-, y) \leq \frac{n}{2}$ we have $-y \in \chi$ and $d(e^-, -y) > \frac{n}{2}$ with the property $\pi(y) = \pi(-y)$. Then we have:

$$1 = \sum_{y \in \chi} \pi(y) = \sum_{\{y : d(e^-, y) \leq \frac{n}{2}\}} \pi(y) + \sum_{\{y : d(e^-, y) > \frac{n}{2}\}} \pi(y) = 2 \sum_{\{y : d(e^-, y) > \frac{n}{2}\}} \pi(y).$$

It gives:

$$\sum_{\gamma_{xy} \ni e} \pi(y) = \sum_{\{y \in \mathcal{X} : d(e^-, y) > \frac{n}{2}\}} \pi(y) = \frac{1}{2}.$$

Then the sum

$$\begin{aligned} \tilde{Q}(e)^{-1} \sum_{\gamma_{xy} \ni e} |\gamma_{xy}| \pi(x) \pi(y) &\leq \frac{1}{\alpha(1-\delta)} \times \frac{n}{2} \sum_{\gamma_{xy} \ni e} \pi(y) \\ &= \frac{n}{2\alpha(1-\delta)} \times \frac{1}{2} \\ &= \frac{n}{4\alpha(1-\delta)}. \end{aligned} \tag{4.11}$$

4.2.3 Summary

By regrouping the previous results in (4.7), (4.8), (4.9), (4.10) and (4.11) we obtain with a suitable choice of $\alpha = e^{-\frac{2}{T}}$ and $\delta = \frac{n}{n+1}$ that:

- $\kappa \leq n^2 e^{\frac{2}{T}}$ At low temperature;
- $\kappa \leq \frac{n^2}{2} e^{\frac{2}{T}}$ At high temperature.

Theorem 4.2.2. [Franke and Helali 2018] *The second largest eigenvalue of the perturbed Gibbs sampler for the 1–D Ising model satisfies at high temperature:*

$$\beta_1 < 1 - 2e^{-\frac{2}{T}} n^{-2}.$$

4.2.4 Bound for the second largest eigenvalue in absolute value

Theorem 5.3 introduced by Ingrassia (1994) in [20] gives a lower bound of the smallest eigenvalue as:

$$\beta_{|\mathcal{X}|-1} \geq -1 + \frac{2}{1 + (c-1)e^{\frac{\Delta}{T}}}.$$

For the 1–D Ising model, $c = 2$ and $\Delta = 2$. Therefore:

$$|\beta_{|\mathcal{X}|-1}| = \left| -1 + \frac{2}{1 + e^{\frac{2}{T}}} \right| = 1 - \frac{2}{1 + e^{\frac{2}{T}}} < 1 - \frac{2}{e^{\frac{2}{T}} + e^{\frac{2}{T}}} < 1 - e^{-\frac{2}{T}} < 1 - 2e^{-\frac{2}{T}} n^{-2}.$$

Combining this result with the result obtained in Theorem 4.2.2 we have the following proposition:

Proposition 4.2.3. *The second largest eigenvalue in absolute value of the Gibbs sampler on the 1–D Ising model satisfies at high temperature:*

$$\max\{\beta_1, |\beta_{|\mathcal{X}|-1}|\} < 1 - 2e^{-\frac{2}{T}} n^{-2}.$$

4.3 Comparison

Frigessi and his collaborators prove in [9] that the convergence rate of the MCMC method for the 1–D Ising model is $O(n \log(n))$. In this case the constant of proportionality isn't known which means that we can't calculate a precise bound. In this section we compare our result to existing result that include proportionality constant. The bound of Ingrassia [20] applied to the one-dimensional Ising model is:

$$\beta \leq 1 - e^{-\frac{2}{T}} \frac{1}{n^2} \left(\frac{1 + e^{-\frac{2}{T}}}{2} \right)^{n-1}.$$

We remark that at low temperature

$$\left(\frac{1 + e^{-\frac{2}{T}}}{2} \right)^{n-1} \longrightarrow \left(\frac{1}{2} \right)^{n-1}$$

which is very small for large n but in our case we obtain the same result as Shiu and Chen in [25].

But at high temperature

$$\left(\frac{1 + e^{-\frac{2}{T}}}{2} \right)^{n-1} \longrightarrow 1$$

and the given bound is similar to the result of Shiu and Chen [25] on the other hand we notice the appearance of a coefficient 2 in our case which doubles the convergence rate to equilibrium.

Let t indicate the time. Combining Proposition 4.2.3 and Theorem 4.2.2 and the mathematical property: $1 - x \leq e^{-x}$ we obtain:

$$(1 - 2n^{-2}e^{-\frac{2}{T}})^t \leq e^{-2tn^{-2}e^{-\frac{2}{T}}} \leq \epsilon$$

where ϵ is the pre-specified error tolerance.

An elementary calculus gives that:

$$t \geq \frac{n^2}{2} e^{\frac{2}{T}} \log \epsilon^{-1}. \tag{4.12}$$

Gibbs [12] used coupling to bound the convergence time and he obtains the following bound:

$$2en^2(1 + \log \epsilon^{-1}) \frac{e^{\frac{2}{T}} + e^{-\frac{2}{T}}}{e^{\frac{2}{T}}}.$$

Shiu and Chen proves in the comparison section in [25] that the convergence time is at most:

$$n^2 e^{\frac{2}{T}} \log \epsilon^{-1}$$

which improve the convergence time proposed by Gibbs by at least $\frac{1}{10}$.

In this chapter we gives a similar bound as Shiu and Chen at low temperature but at high temperature according to (4.12) we obtain that the convergence time is at most:

$$\frac{n^2}{2} e^{\frac{2}{T}} \log \epsilon^{-1}$$

which improve the convergence time to equilibrium by $\frac{1}{20}$ compared to Gibbs [12] and by $\frac{1}{2}$ compared to Shiu and Chen [25].

4.4 Overall conclusion

We applied the geometric bound derived by Diaconnis and Stroock [5] to the Gibbs sampler for the 1–D Ising model. Shiu and Chen in [25] gives a geometric bound that depends of the class of paths passing through each edge e and their length. They used this bound to proof that the convergence rate of the Gibbs sampler for the 1–D Ising model is $O(n^2)$ with a proportionality constant of $e^{\frac{2}{T}}$. Their result seems to be the best bound that we can found in the literature. Our main result is that the perturbation of the Gibbs algorithm with a suitable dynamic gives the same result at low temperature, but at high temperature it doubles the convergence rate to equilibrium.

4.5 Discussion of the convergence rate at low temperature

The following heuristic argument (at zero-temperature) seems to indicate that the improved sampler \tilde{P} should have a very fast convergence compared to the Gibbs sampler at low temperatures.

We remark that at low temperature, the evaluation of the convergence rate with the Diaconnis and Stroock bound gives the same result with or without perturbation of the Gibbs algorithm. Referring to proposition 5.2.1 in [28], the Gibbs measure behaves at low temperature as follows:

$$\lim_{T \rightarrow 0} \pi^T(x) = \begin{cases} \frac{1}{2} & \text{if } x = (+1, \dots \dots, +1) \\ \frac{1}{2} & \text{if } y = (-1, \dots \dots, -1) \\ 0 & \text{else.} \end{cases}$$

For $x = (+1, +1, \dots, +1, +1)$ and $y = (-1, -1, \dots, -1, -1)$ the transition probability is:

$$\tilde{P}(x, y) = (1 - \delta)P^\sigma(x, y) = \frac{1}{2(n+1)}.$$

which gives for $e = (x, y)$:

$$\tilde{Q}(e)^{-1} \sum_{\gamma_{xy} \ni e} |\gamma_{xy}| \pi(x) \pi(y) = 2(n+1) \sum_{\gamma_{xy} \ni e} \pi(y) = n+1.$$

Note that for $T = 0$ no other edges exists. Then the second largest eigenvalue has the following bound:

$$\beta_1 \leq 1 - (n+1)^{-1}.$$

4.6 Appendix

If n is even we have

$$\chi_1^+ = \{w \in \chi_1, w_i = +1\} = \bigcup_{\{(x,y): d(x,y) \leq \frac{n}{2} | \gamma_{xy} \ni e\}} \{x \oplus y\}.$$

4.6. Appendix

By the fact that: $|\chi_1| = 2 \times \sum_{k=0}^{\frac{n}{2}-1} \binom{n-1}{k}$ we obtain:

$$|\chi_1^+| = \sum_{k=0}^{\frac{n}{2}-1} \binom{n-1}{k}.$$

Also we have:

$$\chi_2^- = \{w \in \chi_2, w_i = -1\} = \bigcup_{\{(x,y):d(x,y)>\frac{n}{2}|\gamma_{xy}\ni e\}} \{x \ominus y\}.$$

By the fact that: $|\chi_2| = 2 \times \sum_{k=\frac{n}{2}}^{n-1} \binom{n-1}{k}$ we obtain:

$$|\chi_2^-| = \sum_{k=\frac{n}{2}}^{n-1} \binom{n-1}{k}.$$

It gives that:

$$|\chi_1^+| + |\chi_2^-| = \sum_{k=0}^{n-1} \binom{n-1}{k} = 2^{n-1}.$$

Similarly, if n is odd we have:

$$|\chi_1^+| = \sum_{k=0}^{\frac{n-1}{2}} \binom{n-1}{k} \quad \text{and} \quad |\chi_2^-| = \sum_{k=\frac{n+1}{2}}^{n-1} \binom{n-1}{k}$$

and it gives the same result as previously.

Discussion

In this paragraph, we give a summary of the results given in the thesis and we discuss some open questions.

Summary

Diaconis and Stroock (1991) use Poincaré inequality to bound the second largest eigenvalue for finite state reversible Markov chains. Then Shiu and Chen (2015) used this result to study the convergence rate of the Gibbs sampler for the 1–D Ising model. In the first part of this thesis, we generalize the result of Shiu and Chen to the case of the 1–D Potts model with three and more spin states (see Theorem 2.2.1 and Theorem 2.2.2). In the second part we generalize their method to the case of the 2–D Ising model with two spin states (see Theorem 3.2.3). In the third part, we prove that some perturbation of the Gibbs sampler improves the result of Shiu and Chen (2015) at high temperatures (see Theorem 4.2.2).

Open questions

We study in the second and third chapter of this thesis the rate of convergence of the Gibbs sampler algorithm for some model from statistical mechanics. We obtain in the case of the 1–D Potts model (see Theorem 2.2.1 and Theorem 2.2.2) the best bounds for the largest eigenvalue. Note that in the case of the 2–D Ising model (see Theorem 3.2.3) we did not succeed to compute the Diaconis and Stroock bound as explicitly in Theorem 3.2.3 as in the one dimensional case. We only obtained an upper bound for the bound they give for β_1 . This raises the following question: Is there a way to use some symmetry argument, like in the one dimensional case, to obtain an exact expression for the Diaconis Stroock bound?

It is also interesting to study other methods for obtaining the convergence rate of the Gibbs sampler.

One way of doing this is to use the Cheeger constant (see Section 3 in [5]). Diaconis and Stroock (see Example 3.2 in [5]) proved that the Cheeger constant is realized on one face of the hypercube $\{-1, 1\}^M$ in the case of the Gibbs sampler at infinite temperature, which is actually a random walk on $\{-1, 1\}^M$. The difficulty is to understand the Cheeger constant at finite temperature for Ising model and Potts model.

Yet another method is to use coupling to bound the convergence time. Alison Gibbs (see [12]) uses coupling to bound the convergence time of the Gibbs sampler for the 1–D and 2–D Ising model. We think that we can use this approach in the case of the Potts model.

This raises the following questions:

1) Can one compute the Cheeger constant in the case of the Gibbs sampler for those models?

2) What should be a good coupling in the case of the Gibbs sampler?

In the last chapter of this thesis we study the convergence rate of some perturbed Gibbs sampler for the 1–D Ising model (see Theorem 4.2.2). We observe that the disruption of the Gibbs sampler improves the result of Chen and Shiu (2015) at high temperatures. However, at low temperatures we did not succeed to improve the geometric bound, although an improvement should be expected. This prompts us to ask the following questions:

1) Is this due to a bad choice of the perturbation of the Gibbs sampler?

2) What is the perturbation which leads to the best improvement of the convergence time of the Gibbs algorithm for different models?

Bibliography

- [1] ALDOUS, D.(1983): Random walks on finite groups and rapidly mixing Markov chain. *In séminaire de probabilités XVII 1981/82, Lecture Notes in Mathematics, 986, Ed. J. Azema and M. Yor, pp 243-97. New York: Springer-Verlag.*
- [2] BUNDLE, C.(1980): Isoperimetric Inequalities and Applications, *Pitman, Boston.*
- [3] CHEN T.-L, CHEN, W.-K., HWANG, C.-R., PAI, H.-M.(2012): On the optimal transition matrix for Markov chain Monte Carlo sampling, *J. Control Optim. 59, 2743-2762..*
- [4] CHEN, T.-L., HWANG, C.-R.(2013): Accelerating reversible Markov chains. *Statist. Probab. Lett. 83 (9), 1956-1962.*
- [5] DIACONIS P. AND STROOCK D.(1991): Geometric bounds for eigenvalues of Markov chains. *Ann. Appl. Probab. 1 (1), 36-61.*
- [6] DURRETT R.(1996): Probability: Theory and examples, 2nd ed. Belmont, CA: *Wadsworth and Brooks/Cole.*
- [7] FRANKE B. AND HELALI A.(2018): The Convergence rate of the Gibbs sampler for the 2-D Ising model via a geometric bound. *arXiv preprint arXiv:1804.09198.*
- [8] FRANKE B. AND HELALI A.(2018): On the convergence rate of some perturbed Gibbs sampler for the 1-D Ising model. *(In preparation).*
- [9] FRIGESSI, A., HWANG, C.-R., SHEU, S.-J., DI STEFANO, P.(1993): Convergence rate of the Gibbs sampler, the Metropolis algorithm, and other single site update dynamics. *J. R. Stat. Soc. Ser. B 55, 205-219..*
- [10] FRIGESSI, A., HWANG, C.-R., YOUNES, L.(1992): Optimal spectral structure of reversible stochastic matrices, Monte Carlo methods and the simulation of Markov random fields. *Ann. Appl. Probab. 2, 610-628.*
- [11] GEMAN S. AND GEMAN D.(1984): Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images. *IEEE Trans. Pattern Anal. Mach. Intell. 6, 721-741.*
- [12] GIBBS AL(2000): Bounding the convergence time of the Gibbs sampler in Bayesian image restoration. *Biometrika 87 (4), 749-766.*

- [13] HASTINGS W.K.(1970): Monte Carlo sampling methods using Markov chains and their applications. *Biometrika* 57, 97-109.
- [14] HELALI A.(2018): The convergence rate of the Gibbs sampler for generalized 1-D Ising model. *arXiv preprint arXiv: 1804.04330*.
- [15] HOLLEY H. AND STROOCK D.(1988): Simulated Annealing via Sobolev Inequalities. *Communications in Mathematical Physics. Vol 115, N 4, 553-569*.
- [16] HORN, R. AND JOHNSON, C. (1985): Matrix Analysis. *Cambridge University Press*.
- [17] LANDAU, H. AND ODLYZKO A.(1981): Bounds for eigenvalues of certain stochastic matrices. *Linear Algebra Appl.* 38 5-15.
- [18] LUBY, M., RANDALL, D. AND SINCLAIR, A.(1995): Markov chain algorithms for planar Lattice structures (extended abstract) . *In Symposium on Foundation of Computer Science 36, pp 150-9. Long Beach, CA: IEEE Computer Soc.*
- [19] LOSIFESCU, M.: *Finite Markov processes and their applications*. Romania: John Wiley and Sons.
- [20] INGRASSIA S.(1994): On the rate of convergence of the metropolis algorithm and Gibbs sampler by geometric bounds. *Ann. Appl. Probab.* 4, 347-389.
- [21] METROPOLIS N., ROSENBLUTH A.W., ROSENBLUTH M.N., TELLER A.H., TELLER E.(1953): Equations of state calculations by fast computing machines. *J. Chem. Phys.* 21, 1087-1092.
- [22] MOHAR, B.(1989a): Eigenvalues, diameter, and mean distance in graphs. *Technical Report, Dept. Mathematics, Univ. E. K. Ljubljanos, Jadranska, Yugoslavia.*
- [23] MOHAR, B.(1989b): The Laplacian spectrum of graphs. *Technical Report, Dept. Mathematics, Univ. E. K. Ljubljanos, Jadranska, Yugoslavia..*
- [24] PROPP, J., G. AND WILSON, D., B.(1996): Exact sampling with coupled Markov chain and Applications to statistical mechanics. *Random Stuct. Algor.* 9, 223-52.
- [25] SHIU S.-Y. AND CHEN T.-L.(2015): On the rate of convergence of the Gibbs sampler for the 1-D Ising model by geometric bound. *Statistics and Probability Letters* 105 14-19.
- [26] SINCLAIR, A.(1991): Improved bounds for mixing rates of Markov chains and multicommodity flow. *Comb. Probab. Comput.* 1, 351-370.
- [27] SINCLAIR, A., JERRUM, M.(1989): Approximate counting, uniform generation and rapidly mixing Markov chain. *Inform. and Comput.* 82, 93-133.
- [28] WINKLER G.(2012): *Image Analysis, Random fields and Markov Chain Monte Carlo methods: A mathematical Introduction* (Vol. 27). Springer Science and Business Media.

Title: Vitesse de convergence de l'échantillonneur de Gibbs appliqué à des modèles de la physique statistique

Résumé Les méthodes de Monte Carlo par chaînes de Markov MCMC sont des outils mathématiques utilisées pour simuler des mesures de probabilités π définies sur des espaces de grandes dimensions. Le principe consiste à construire une chaîne de Markov P facile à simuler qui converge vers la mesure π . Une des questions les plus importantes dans ce contexte est de savoir à quelle vitesse converge la chaîne de Markov P vers la mesure invariante π .

Pour mesurer la vitesse de convergence de la chaîne de Markov P vers sa mesure invariante π certains auteurs utilisent la variance asymptotique alors que d'autres auteurs utilisent la distance de la variation totale. Il est bien connu que la vitesse de convergence d'une chaîne de Markov réversible P dépend de la deuxième plus grande valeur propre en valeur absolue de la matrice P notée β^* . Une partie importante dans l'estimation de β^* consiste à estimer la deuxième plus grande valeur propre de la matrice P , qui est notée β_1 . Diaconis et Stroock (1991) ont introduit une méthode basée sur l'inégalité de Poincaré pour estimer β_1 pour le cas général des chaînes de Markov réversibles avec un nombre fini d'état. Récemment Shiu et Chen (2015) ont réussi à obtenir une expression explicite de la borne de Diaconis et Stroock dans le cas d'un modèle d'Ising unidimensionnel avec deux états. Ce résultat améliore celui introduit par Ingrassia (1991) et il est considéré comme la meilleure borne existante dans la littérature pour ce modèle.

Dans le chapitre deux de cette thèse, nous utilisons la méthode de Shiu et Chen pour étudier le cas de l'algorithme de l'échantillonneur de Gibbs pour le modèle d'Ising unidimensionnel avec trois états ou plus appelé aussi modèle de Potts. Puis, nous généralisons dans le troisième chapitre le résultat de Shiu et Chen au cas du modèle d'Ising deux-dimensionnel avec deux états. Notre méthode est basée sur une généralisation des techniques introduites par Shiu et Chen (2015). Les résultats obtenus minorent ceux introduits par Ingrassia (1994).

Dans le dernier chapitre de la thèse nous avons pensé à perturber l'échantillonneur de Gibbs afin d'améliorer sa vitesse de convergence vers l'équilibre. En partant de cette idée, nous avons réussi à améliorer le résultat introduit par Shiu et Chen (2015) à haute température. Cependant, à basse température, nous n'avons pas réussi à améliorer la borne géométrique, même si une amélioration est à prévoir.

Mots-clés : Méthode de Monte Carlo par chaîne de Markov, Vitesse de convergence, Échantillonneur de Gibbs, modèle d'Ising, modèle de Potts, Systèmes de treillis, Permutation.

Title: The convergence rate of the Gibbs sampler for some statistical mechanics models

Abstract Monte Carlo Markov chain methods MCMC are mathematical tools used to simulate probability measures π defined on state spaces of high dimensions. The idea is to construct a Markov chain P easy to simulate which converges to the measure π . The speed of convergence of this Markov chain X to its invariant state π is a natural question to study in this context.

To measure the convergence rate of a Markov chain some authors use the asymptotic variance, while other authors use the total variation distance. It is well known that the convergence rate of a reversible Markov chain depends on its second largest eigenvalue in absolute value denoted by β^* . An important part in the estimation of β^* is the estimation of the second largest eigenvalue which is denoted by β_1 . Diaconis and Stroock (1991) introduced a method based on Poincaré inequality to obtain a bound for β_1 for general finite state reversible Markov chains.

Recently Shiu and Chen (2015) are able to obtain explicit value for the bound of Diaconis and Stroock in case of the 1–D Ising model with two states (spin down, spin up). This result improves the one introduced by Ingrassia (1994) and it is considered to be the best existing bound in the literature for this model.

In the second chapter of this thesis we use the Chen and Shiu approach to study the case of the Gibbs sampler for the 1–D Ising model with three and more states which is also called Potts model. Then, we generalize in the third chapter the result of Shiu and Chen to the case of the 2–D Ising model with two states. Our method is based on some generalization of the techniques introduced by Shiu and Chen (2015). The results we obtain improve the ones obtained by Ingrassia (1994).

In the last chapter of this thesis, we introduce some method to disrupt the Gibbs sampler in order to improve its convergence rate to equilibrium. From this idea we are able to improve the result introduced by Shiu and Chen (2015) at high temperatures. However, at low temperature we did not succeed to improve the geometric bound, although an improvement should be expected.

keywords : Markov chain Monte Carlo, Rate of convergence, Gibbs sampler, Ising model, Potts model, Lattice systems, Permutation.

Titre : Vitesse de convergence de l'échantillonneur de Gibbs appliqué à des modèles de physique statistique

Mots clés : Méthode de Monte Carlo par chaîne de Markov, Vitesse de convergence, échantillonneur de Gibbs, modèle d'Ising, modèle de Potts, Systèmes de treillis, Permutation.

Résumé : Les méthodes de Monte Carlo par chaînes de Markov MCMC sont des outils mathématiques utilisés pour simuler des mesures de probabilités π définies sur des espaces de grandes dimensions. Une des questions les plus importantes dans ce contexte est de savoir à quelle vitesse converge la chaîne de Markov P vers la mesure invariante π . Pour mesurer la vitesse de convergence de la chaîne de Markov P vers sa mesure invariante π nous utilisons la distance de la variation totale. Il est bien connu que la vitesse de convergence d'une chaîne de Markov réversible P dépend de la deuxième plus grande valeur propre en valeur absolue de la matrice P notée β^* . Une partie importante dans l'estimation de β^* consiste à estimer la deuxième plus grande valeur propre de la matrice P , qui est notée β_1 .

Diaconis et Stroock (1991) ont introduit une méthode basée sur l'inégalité de Poincaré pour estimer β_1 pour le cas général des chaînes de Markov réversibles avec un nombre fini d'état. Dans cette thèse, nous utilisons la méthode de Shiu et Chen (2015) pour étudier le cas de l'algorithme de l'échantillonneur de Gibbs pour le modèle d'Ising unidimensionnel avec trois états ou plus appelé aussi modèle de Potts. Puis, nous généralisons le résultat de Shiu et Chen au cas du modèle d'Ising deux- dimensionnel avec deux états.

Les résultats obtenus minorent ceux introduits par Ingrassia (1994). Puis nous avons pensé à perturber l'échantillonneur de Gibbs afin d'améliorer sa vitesse de convergence vers l'équilibre.

Title : The convergence rate of the Gibbs sampler for some statistical mechanics models

Keywords : Markov chain Monte Carlo, Rate of convergence, Gibbs sampler, Ising model, Potts model, Lattice systems, Permutation.

Abstract : Monte Carlo Markov chain methods MCMC are mathematical tools used to simulate probability measures π defined on state spaces of high dimensions. The speed of convergence of this Markov chain X to its invariant state π is a natural question to study in this context. To measure the convergence rate of a Markov chain we use the total variation distance. It is well known that the convergence rate of a reversible Markov chain depends on its second largest eigenvalue in absolute value denoted by β^* . An important part in the estimation of β^* is the estimation of the second largest eigenvalue which is denoted by β_1 .

Diaconis and Stroock (1991) introduced a method based on Poincaré inequality to obtain a bound for β_1 for general finite state reversible Markov chains.

In this thesis we use the Chen and Shiu approach to study the case of the Gibbs sampler for the 1-D Ising model with three and more states which is also called Potts model. Then, we generalize the result of Shiu and Chen (2015) to the case of the 2-D Ising model with two states. The results we obtain improve the ones obtained by Ingrassia (1994). Then, we introduce some method to disrupt the Gibbs sampler in order to improve its convergence rate to equilibrium.