Coupling, Concentration and Random Walks in Dynamic Random Environments

Proefschrift

ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft;
op gezag van de Rector Magnificus
prof. ir. K.C.A.M. Luyben;
voorzitter van het College voor Promoties
in het openbaar te verdedigen op
maandag 15 october 2012 om 10:00 uur

door Florian Maartenszoon VÖLLERING
Diplom-Mathematiker, Technische Universität Berlin
geboren te Berlijn, Duitsland.

Dit proefschrift is goedgekeurd door de promotoren:

Prof.dr. F.H.J. Redig
Prof.dr. W.Th.F. den Hollander

Samenstelling promotiecommissie:

Rector Magnificus, voorzitter
Prof.dr. F.H.J. Redig, Technische Universiteit Delft, promotor
Prof.dr. W.Th.F. den Hollander, Universiteit Leiden, promotor
Prof. C. Giardina, Università Degli Studi di Modena e Reggio Emilia
Prof. C. Maes, Katholieke Universiteit Leuven
Prof.dr. E.A. Cator, Radboud Universiteit Nijmegen
Prof.dr. J.M.A.M. van Neerven, Technische Universiteit Delft
Dr. C. Kraaikamp, Technische Universiteit Delft
Prof.dr.ir. G. Jongbloed, Technische Universiteit Delft, reservelid

Contents

Summary

In this thesis we discuss concentration inequalities, relaxation to equilibrium of stochastic dynamics, and random walks in dynamic random environments. In stochastic systems one is interested in macroscopic and/or asymptotic properties as well as in fluctuations around typical behaviour. But the dependence structure induced by the interaction between the components of the system makes the analysis challenging. In order to overcome this in different settings a variety of methods are employed.

Additive functionals of Markov processes play important roles in applications. In order to get exponential and moment estimates for their fluctuations a non-standard martingale approximation is used. The resulting general theorems do not require special properties like reversibility or a spectral gap. What is needed is some control on the expected evolution. That is, the difference of the evolution starting from two "adjacent" configurations has to be controlled. Coupling methods are well suited to do perform this comparison. In concrete examples couplings are used to prove the conditions of the theorems.

In statistical mechanics Gibbs measures and Markov random fields play important roles. The Poincaré inequality is an important property describing the regularity of the measure. We prove the Poincaré inequality via a martingale telescoping argument. To control the individual increments of the martingale we use a coupling method called disagreement percolation. If the clusters of this percolation are sufficiently small we obtain the Poincaré inequality.

When interacting spin systems and their dynamics have a delicate connection to their ergodic measure(s) one has to take more care. We carefully study the graphical construction of the dynamics to understand how the influence of the measure can be preserved. An assumption is made that one can control how fast the system in equilibrium can compensate for a single spin flip. Under this assumption we obtain relaxation speed estimates for general functions. In attractive spin systems the condition can be reduced to the decay of autocorrelation of the spin at the origin. An application where this is of use is the low-temperature Ising model.

Finally we look at random walks in dynamic random environments. Here a time-changing random environment drives the motion of a particle. The goal is to understand under which conditions the macroscopic behaviour of this random walk is like that of a Brownian motion. We use coupling to prove a law of large numbers as well as a functional central limit theorem for the position of the random walk. Only polynomial decay of correlations in time are needed for the environment, and the influence of the environment on the walk can be very general.

Koppeling, concentratie ongelijkheden en stochastische wandelingen in dynamische toevallige omgevingen

Samenvatting

In dit proefschrift worden concentratie ongelijkheden, de toenadering naar het evenwicht van stochastische dynamica en stochastische wandelingen in dynamische toevallige omgevingen besproken. In stochastische systemen is men geïnteresseerd in macroscopische alsook asymptotische eigenschappen en in de fluctuaties rond het typische gedrag. De afhankelijkheidsstructuur geïnduceerd uit de wisselwerking van de componenten van het systeem vormt de uitdaging van de analyse. Om dit te overkomen worden, afhankelijk van de situatie, verschillende methodes gebruikt.

Additieve functionalen van Markovprocessen spelen een belangrijke rol in toepassingen. Om exponentiële en regulaire momenten van hun fluctuaties te schatten wordt een niet-standaard martingaal voor de benadering gebruikt. De resulterende algemene stellingen vereisen geen speciale eigenschappen zoals reversibiliteit of een spectral gap. Wat nodig is, is enige controle op de te verwachte ontwikkeling van het proces. Dat wil zeggen het verschil van de ontwikkeling vanaf twee aangrenzende configuraties moet gecontroleerd worden. Koppelingsmethoden zijn geschikt om dit soort vergelijk uit te voeren. In concrete voorbeelden worden koppelingen gebruikt om de voorwaarden van de stellingen te bewijzen.

In de statistische mechanica spelen Gibbsmaten en Markov random fields een belangrijke rol. De Poincaré ongelijkheid is een belangrijke eigenschap die de regelmatigheid van een maat beschrijft. Bewezen wordt de Poincaré ongelijkheid door een telescopische som van martingale incrementen. Om het

individuele increment van de martingaal te controleren wordt van een koppelingsmethode genaamd disagreement percolation gebruik gemaakt. Als de clusters van deze percolatie voldoende klein zijn, krijgen we de Poincaré ongelijkheid.

Als interactieve spin systemen en hun dynamica een delicate verbinding met hun ergodische maat(en) hebben, moet men voorzichtig zijn. We bestuderen de grafische opbouw van de dynamiek zorgvuldig om er achter te komen hoe de invloed van de maat bewaard kan worden. Verondersteld wordt, dat men kan bepalen hoe snel het systeem in evenwicht een enkele spin-flip kan compenseren. Onder deze aanname krijgen we ramingen voor de relaxatie snelheid van algemene functies. In monotone spin-systemen kan dit teruggebracht worden tot het verval van autocorrelatie van de spin in de oorsprong. Een applicatie waar dit van gebruik is, is het lage temperatuur Ising-model.

Tot slot kijken we naar stochastische wandelingen in dynamisch toevallige omgevingen. Hier zorgt een in de tijd veranderende omgeving voor de beweging van een deeltje. Het doel is: te begrijpen onder welke voorwaarden het macroscopisch gedrag van deze stochastische wandeling gelijk is aan dat van een Brownse beweging. We gebruiken koppelingsmethoden om de wet van grote aantallen en een functionele centrale limietstelling voor de positie van de stochastische stochastische wandeling. Alleen polynoom verval van correlaties in de tijd is nodig voor de omgeving waarbij de invloed van de omgeving op de stochastische wandeling zeer algemeen kan zijn.

1 Introduction

This thesis has three main themes: concentration inequalities, relaxation to equilibrium of stochastic dynamics, and random walks in dynamic random environments. These three themes are intimately related. In this thesis the focus is then on dependent variables, where the dependence is Markovian in nature. In contrast to an i.i.d. context, which is well understood, many natural questions prove to be challenging because of the dependence. In this thesis a common approach to these themes are coupling techniques: in all chapters coupling methods play a role.

In the study of interacting stochastic systems one is usually interested in global and/or asymptotic behaviour. This means e.g. global changes in the process when changing parameters, such as survival/extinction, diffusive versus non-diffusive behaviour, exponential versus slow (power-law) decay of correlations in space and/or time.

In interacting particle systems the following questions typically emerge:

a) properties of the stationary measure(s), and the speed of relaxation to it;

b) fluctuations of macroscopic time-averages along trajectories;

c) tagged particles;

d) behaviour of a particle driven by the interacting particle system.

In the following sections we provide some more details on the different themes.

1.1 Concentration inequalities

It is well known that the average $X_n = \frac{1}{n}\sum_{i=1}^{n} Y_i$ of independent and identically distributed random variables $(Y_i)_{i\geq 1}$ converges to $\mathbb{E}Y_1$ (as long as the expectation exists). This fact, the law of large numbers, is perhaps the most well-known theorem in probability theory. It can be explained quite intuitively, by claiming that in the long run the fluctuations above and below the average balance out. It is, however, only an asymptotic statement. It does not include any information about how close an average of size ten, one hundred or a million is to the mean. The study of this question, the deviation probability from the mean, is the domain of concentration inequalities.

There are various approaches to the question of deviation from the mean. One key insight is the fact that changes in only one Y_i result in only small changes in X_n. This concept of small influence of each individual random variable can be used to obtain deviation probabilities for the average X_n [51]. The idea also generalizes from averages to other functions $f(Y_1, ..., Y_n)$. The corresponding concept of small influence of each Y_i is coordinate-wise Lipschitz continuity of f. Denoting the Lipschitz-constants in the individual coordinates by $(\delta_i f)$, one can look at the inequality

$$\mathbb{P}(f - \mathbb{E}f > r) \leq e^{-c\frac{r^2}{\sum_i (\delta_i f)^2}}. \tag{1.1}$$

This inequality is called the Gaussian concentration inequality, and can be obtained for example for i.i.d. bounded or Gaussian random variables. More general, instead of looking just at Gaussian concentration one can try to obtain inequalities of the type

$$\mathbb{P}(f - \mathbb{E}f > r) \leq \Psi(r, (\delta_i f)). \tag{1.2}$$

One aim is to obtain strong decay of Ψ as r increases. Besides Gaussian concentration especially exponential concentration ($\Psi(r) \leq c_1 e^{-c_2 r / \| \delta f \|}$) is studied much.

These types of inequalities have applications in many fields. Examples are statistics, computer science and various areas in probability theory. They can be used to judge the quality of an estimator or to calculate the required running time of Monte-Carlo simulations. When simplifying calculations by replacing a random value $f(Y_1, ..., Y_n)$ by its expectation concentration inequalities provide estimates on the error.

Besides the use in applications concentration inequalities have become a field of research in themselves. This field has connections not only to probability theory and statistics, but also to measure theory, functional analysis and geometry. In this field the relations between various types of inequalities of the underlying measure are studied. Among those inequalities are isoperimetric inequalities, transportation cost inequalities, functional inequalities and exponential and Gaussian concentration inequalities. Active research is done regarding conditions for obtaining these kinds of inequalities, as well as stability under perturbation or taking product measures.

Typical for concentration inequalities is some sort of size parameter, like volume, time, dimension, etc. How this size parameter effects improvements in the deviation estimate (1.2) is usually of interest in applications. The above mentioned stability under taking product measures which some inequalities satisfy is of great use in this regard.

In the literature there are quite a few distinct methods for obtaining concentration inequalities. Among those are the Chen-Stein method [15], transportation cost inequalities[28], log-Sobolev inequalities[49, 41, 31, 30], concentration of measure[36, 50].

In general the Gaussian and exponential concentration behaviour for Lipschitz functions on a product measure is well understood. It is going beyond this setting where there are still open questions. One can distinguish three directions.

a) Dependent random variables:
What forms of concentration inequalities can one obtain for dependent random variables? In many situations the i.i.d. assumption is not satisfied, for example for Gibbs random fields, interacting particle systems or trajectories of Markov processes.

b) Strong dependence on the function:
How can one characterize concentration for non-Lipschitz functions, or functions whose Lipschitz constants are very large but which do not depict the typical behaviour of the function. An example of this would be the maximal overlap of two i.i.d. binomial sequences of length n. Here a single flip can lead to a change of order n, but typically the change is at most of order $\log(n)$. Here the Poincaré inequality and other variants can provide some answers.

c) Beyond Gaussian and exponential estimates:
One is interested in forms of sub-exponential concentration. The need for those estimates can come from a variety of sources. It could be that the distribution simply has heavy tails, so one cannot hope for exponential estimates. Or strong dependence between the variables, like in low temperature spin systems. One could have a system which exhibits exponen-

tial concentration for "good" functions, but sub-exponential (polynomial, stretched exponential,...) for "bad" functions. This leads to the need for moment estimates or estimates non-uniform in nature.

On first glance these directions may seem rather different. However there are quite a few connections in between. For example, it is often possible to express a system of dependent random variables as a (complicated) function of independent random variables. So one can go from dependent to independent random variables at the price of more complicated functions.

Also, in many situations what one would consider "bad" functions are functions which are mostly benign but which are very sensitive for exceptional configurations. It is the interplay between "bad" configurations and the function which causes the complications.

1.2 Functional inequalities and relaxation to equilibrium

In the end one needs to understand the underlying measure and its interaction with functions better. Inequalities like the Poincaré inequality are naturally more sensitive to the interaction between function and measure. In the Poincaré inequality,

$$\mathrm{Var}(f) \leq c \sum_{i=1}^{d} \int_{\mathbb{R}^d} (\partial_{x_i} f)^2 \, \mu(dx), \tag{1.3}$$

one uses directional derivatives (or suitable analogues in other spaces) instead of the Lipschitz constants of the function f.

Directional derivatives together with the measure also provide naturally associated dynamics in many cases. These associated dynamics, like heat-bath Glauber dynamics for interacting spin systems, or diffusions in a potential, give additional insight. If the measure satisfies good mixing properties in space, then the dynamics have good mixing properties in time, and vice versa. For example, a measure satisfying the Poincaré inequality is equivalent to the dynamics relaxing exponentially fast to equilibrium in L^2.

This connection between properties of the measure and relaxation of the dynamics has many uses. It allows to obtain concentration along trajectories of the dynamics from understanding the equilibrium measure, and having control on the dynamics, for example via coupling, can provide concentration of the measure.

Most of the literature explores this connection in fast mixing regimes. In comparison there are a lot less methods which can be used for relaxation to equilibrium slower than exponential. The weak Poincaré inequality, which implies sub-exponential relaxation to equilibrium, is one of the few exceptions[46].

In the end there is still much to be explored in concentration for dependent systems. A direct proof of the Poincaré inequality using a method called disagreement percolation is done in chapter 3. In chapter 2 additive functionals of Markov processes are studied. The martingale method employed does not require any form of information on the equilibrium measure, instead a specific control on the relaxation of the dynamics is used. This has the advantage that also sub-exponential concentration can be obtained, and the concentration result can be specific to a function. In 4 a detailed analysis of the dynamics of Glauber dynamics of spin systems is used obtain relaxation to equilibrium of the semi-group based on coupling probabilities.

1.3 Random walks

Random walks belong to the most important Markov processes and have been heavily studied. The interest in random walks comes from the fact that they are the standard model to represent motion in a noisy or complex system.

Some of the most important properties of random walks are the law of large numbers and the invariance principle. The law of large numbers states that asymptotically the random walk will have a deterministic speed. This makes a line corresponding to that speed the best deterministic approximation to the trajectory of a random walk. The invariance principle then states that the random fluctuations around this line are asymptotically like a Brownian motion. This description of of the macroscopic behaviour of the elementary model of motion is key to explain diffusivity in a wide range of physical (and other) models.

In the mathematical model, the position of the random walk $(X_n)_{n\geq 0}$ is given by the sum of the increments, $X_n = \sum_{i=1}^{n} Y_i$. These increments are assumed to be independent and identically distributed. This assumption is of course an idealisation. In actual applications to model motion this assumption of i.i.d. increments cannot assumed to be satisfied. In more realistic situations, increments of the motion are produced by complex interactions with the environment. However, if the environment is rapidly changing in a chaotic manner and is homogeneous in space and time i.i.d. increments are a reasonable approximation. It is then a question how far the macroscopic behaviour of Brownian motion with drift can be generalized to models where the microscopic assumption of i.i.d. increments is violated.

In order to make the simple random walk more realistic various variants have been studied. When the environment's influence on the random walk is only slowly changing compared to the motion of the random walk, one can take the idealization of an unchanging environment. In this environment the random walk is exposed to inhomogeneities in space, which locally change the jump probabilities of the random walk. this setting is called random walk in (static) random environment.

In this thesis we focus on random walks in dynamic random environments. Random walks in dynamic random environments form the regime in between static environments and i.i.d. increments. The environment is assumed to evolve at a similar time scale as the random walk.

Examples where static environments are appropriate could be electron movement inside an impure conductor or migration of species with respect to geography. Examples for dynamic environments are protein movement inside cells or migration of species in competition with other species.

When studying a random walk (X_t) in a random environment, one wants to answer several questions:

a) The law of large numbers:
Does $X_t/t \to v$ a.s. as $t \to \infty$? If the asymptotic speed exists, is it possible to describe its value?

b) The (functional) central limit theorem:
Does $\frac{X_t - vt}{\sqrt{t}} \to \mathcal{N}(0, \sigma^2)$? Does the trajectory of (X_t) converge to a Brownian motion with drift? Can one quantify the variance?

c) Large deviation principle:
Is $\mathbb{P}(X_t/t \approx x) \approx e^{-tI(x)}$? How does I look like?

d) Concentration estimates:
How does $\mathbb{P}(|\, X_t - vt \,| > r)$ decrease in r in in non-asymptotic settings?

e) Recurrence and transience:
Under which conditions is the random walk recurrent or transient? How does that relate with the asymptotic speed, variance and the dimension?

Random walks in static random environments are well understood in 1 dimension. Going back to [48, 47], it is known that the macroscopic behaviour of the random walk can drastically change even under small inhomogeneities in space. Notably, recurrence is no longer equivalent to 0 speed. The random walk can have an asymptotic speed of 0 but be transient. That is, asymptotically the random walk will approach infinity (or minus infinity), but at a sub-linear speed. On the level of fluctuations the behaviour can be anomalous as well, with super-diffusivity and sub-diffusivity possible.

This non-standard behaviour of the random walk can be explained by looking at the environment as a potential landscape. Here valleys of the potential act as traps for the random walk. Once the random walk enters the valley it takes a long time to escape, with the time increasing exponentially in the depth of the valley. If those traps are frequent and strong enough they can slow down the random walk to such an extend that its asymptotic speed is 0 even though it is transient. Similarly the traps are responsible for the non-diffusive behaviour. But here also fluctuations in the frequency and depth of the traps play a role.

In dimensions 2 and higher the picture is a lot less clear. The reason for this is that geometric structures in the environment start to play a role as well. There are still many open questions in contrast to random walks in 1 dimensional random environments where understanding is much more complete. One focus of the literature are fluctuations of the random walk under the assumption that it is ballistic, i.e., has a positive speed in some direction. Another focus is understanding conditions for ballisticity better.

Random walks in dynamic random environments are very natural from the point of view of many applications. Here the environment evolves at a time-scale comparable to the movement of the random walk. The time- and space-inhomogeneous nature of the environment poses very different problems in comparison to the effect of disorder in the static case. On one hand diffusive behaviour is more easily achieved because of additional averaging over time. On the other hand the dependence structure of the increments is more complicated. While the random walk given the environment is a Markov process this is no longer true in dynamic environments because of the time-inhomogeneity.

For the behaviour of the random walk the mixing speed of the environment is very important. This is easy to see, as dynamic environments interpolate between static environments (limit of 0 mixing speed) and i.i.d. increments (limit of infinite mixing speed). This idea leads to two regimes, so called "slow mixing" and "fast mixing" environments.

For fast mixing environments the correlations decay fast enough so that on macroscopic scales the dependence is lost and the increments of the random walk become i.i.d. For slow mixing environments some dependence structure is retained even on macroscopic space and time scales. However, the existence of those macroscopic space-time dependence structures in the environment are not equivalent with non-standard behaviour of the random walk, as the nature of the influence of the environment on the random walk is relevant as well.

In fact, the slow mixing regime is very poorly understood. In the literature the only rigorous result is the existence of a flat piece in the large deviation rate function for a random walk on a 1-dimensional simple symmetric exclusion process [5]. Simulations for the same model suggest that depending on the interaction the random walk can exhibit both diffusive and non-diffusive behaviour [7].

As the slow mixing regime is very complicated much of the literature is focused on understanding and expanding the fast mixing regime. Diffusive behaviour has been proven when the environment is independent in time or space or uniformly exponentially fast mixing [11, 9, 12, 6, 23, 22, 13]. When the mixing is typically exponential but slower for exceptional starting configurations of the environment there are only a few results. In [10, 24] the supercritical contact process is considered, where for specific forms of interaction diffusivity is proven as well. In another work for non-uniformly mixing environments a random walk with a strong drift is considered on the simple symmetric exclusion process is studied [8], where the drift makes the effective environment exponentially fast mixing in a way to obtain diffusivity.

When looking at sub-exponential mixing speeds there are only few results available. In [2] environments with polynomial decay of correlations in space and time are studied. Diffusivity of the random walk is obtained using a renormalization argument under the additional assumption that the random walk is only weakly coupled to the environment. In [3] random conductance models are studied for environments with only weak polynomial mixing. In dimension 3 and higher only mixing faster than t^{-1} is needed. However in random conductance models the interaction between random walk and environment is restricted, which limits possible conclusions for other models.

The main contribution of chapter 5 to this field is using coupling to obtain diffusivity for random walks in environments which are only polynomially mixing. Here the influence of the environment on the random walk is completely general up to some continuity conditions.

One important idea in the study of random walks in random environments (both static and dynamic) is the environment-seen-from-the-particle process (EP). It is this process, the perceived environment, which determines the movement of the random walk. Just like the sequence $(Y_1, Y_2, ..)$ determines the classical random walk $X_n = \sum_{i=1}^{n} Y_i$, the EP determines the random walk in random environment. Studying the asymptotic properties of the EP allows to obtain much information about the random walk itself. In the case of Markovian dynamic environments the EP has another important property: in contrast to the random walk, the EP is a Markov process. The key theorem of chapter 5 then deals with strong ergodicity properties of the EP.

1.4 Overview of the individual chapters

Chapter 2: Concentration of Additive Functionals for Markov Processes and Applications to Interacting Particle Systems

In chapter 2 we develop concentration inequalities for (time-inhomogeneous) additive functionals of Markov processes on general state spaces. Both exponential and moment estimates are obtained. The method used is based on martingales and does not need any information about the stationary measure of the process. Instead a specific form of relaxation to equilibrium is used, which is well-adapted to coupling methods. The connection via coupling is explained via a general relation between the contractivity of the semigroup and bounds on the generalized coupling time.

Applications include diffusions, random walks, and interacting particle systems. In particular the simple symmetric exclusion process, which is far beyond the traditional methods.

Chapter 3: Poincaré inequality for Markov random fields via disagreement percolation

In chapter 3 we look at the Poincaré inequality for Markov random fields like the high-temperature Ising model. Even though the Poincaré inequality is already known to hold we obtain it in a new way. The method used is disagreement percolation. If the influence of a single flip in a configuration is small enough, the Poincaré inequality is obtained. If the influence is just finite, i.e., the disagreement percolation is sub-critical, then at least the weak Poincaré inequality holds.

Chapter 4: A Variance Inequality for Glauber dynamics with Application to Low Temperature Ising Model

Chapter 4 looks at the relaxation rate to equilibrium of Glauber dynamics. The influence of a single flip is tracked through space-time in a detailed manner as not to lose the influence of the ergodic measure. Because of this non-uniform nature the method can be applied even in low teperature regimes, as long as some control on the relaxation to equilibrium is available. In attractive systems the decay of the auto-correlation of the spin at the origin is sufficient for that.

Chapter 5: Random Walks in Dynamic Random Environments: A transference principle

We study a general class of random walks driven by a uniquely ergodic Markovian environment. Under a coupling condition on the environment we obtain strong ergodicity properties for the environment as seen from the position of the walker, i.e., the environment process. We can transfer the rate of mixing in time of the environment to the rate of mixing of the environment process with a loss of at most polynomial order. Therefore the method is applicable to environments with sufficiently fast polynomial mixing. We obtain unique ergodicity of the environment process. Moreover, the unique invariant measure of the environment process depends continuously on the jump rates of the walker.

As a consequence we obtain the law of large numbers and a central limit theorem with non-degenerate variance for the position of the walk.

2 Concentration of Additive Functionals for Markov Processes and Applications to Interacting Particle Systems[1]

2.0 Abstract

We consider additive functionals of Markov processes in continuous time with general (metric) state spaces. We derive concentration bounds for their exponential moments and moments of finite order. Applications include diffusions, interacting particle systems and random walks. In particular, for the symmetric exclusion process we generalize large deviation bounds for occupation times to general local functions. The method is based on coupling estimates and not spectral theory, hence reversibility is not needed. We bound the exponential moments(or the moments of finite order) in terms of a so-called coupled function difference, which in turn is estimated using the generalized coupling time. Along the way we prove a general relation between the contractivity of the semigroup and bounds on the generalized coupling time.

2.1 Introduction

The study of concentration properties of additive functionals of Markov processes is the subject of many recent publications, see e.g. [14], [55]. This subject is strongly connected to functional inequalities such as the Poincaré and log-Sobolev inequality, as well as to the concentration of measure phenomenon [36]. In this chapter we consider concentration properties of a general class

[1]Concentration of Additive Functionals for Markov Processes and Applications to Interacting Particle Systems
F. Redig, F. Völlering
http://arxiv.org/abs/1003.0006
submitted

of additive functionals of the form $\int_0^T f_t(X_t)\, dt$ in the context of continuous-time Markov processes on a Polish space. The simplest and classical case is where $f_t = f$ does not depend on time. However the fact that time-dependent functions f_t are allowed can be a significant advantage in applications.

Our approach is based on coupling ideas. More precisely, we estimate exponential moments or k-th order moments using the so-called coupled function difference which is estimated in terms of a so-called generalized coupling time, a generalization of the concept used in [20]. Because of this approach no knowledge about a possible stationary distribution is required.

Our method covers several cases such as diffusion processes, jump processes, random walks and interacting particle systems. The example of random walk shows that for unbounded state spaces, the concentration inequalities depend on which space the functions f_t belong to.

The main application to the exclusion process, which has slow relaxation to equilibrium and therefore does not satisfy any functional inequality such as e.g. log-Sobolev (in infinite volume), shows the full power of the method. Besides, we give a one-to-one correspondence between the exponential contraction of the semigroup and the fact that the generalized coupling time is bounded by the metric. For discrete state spaces, this means that the semigroup is exponentially contracting if and only if the generalized coupling time is bounded.

This chapter is organized as follows: in Section 2.2 we prove our concentration inequalities in the general context of a continuous-time Markov process on a metric space. We derive estimates for exponential moments and moments of finite order. In Section 2.3 we study the generalized coupling time and its relation to contractivity of the semigroup. Section 2.4 is devoted to examples. Section 2.5 deals with the symmetric exclusion process.

2.2 Concentration inequalities

Let $\mathbb{X} = (X_t)_{t\geq 0}$ be a Feller process in the Polish state space E. Denote by $\mathbb{P}_x$ its associated measure on the path space of cadlag trajectories $D_{[0,\infty[}(E)$ started in $x \in E$ and with

$$\mathfrak{F}_t := \sigma\{X_s; 0 \leq s \leq t\}, \quad t \geq 0,$$

the canonical filtration. We denote by $\mathbb{E}_x$ the expectation with respect to the measure $\mathbb{P}_x$. For ν a probability measure on E, we define $\mathbb{E}_\nu := \int \mathbb{E}_x \, \nu(dx)$, i.e. expectation in the process starting from ν. The associated semigroup we denote by $(S_t)_{t\geq 0}$ and with A its generator, both considered on a suitable space $(\mathcal{B}(E), \mathcal{C}(E), \mathcal{C}_0(E), ...)$.

The content of this section is to derive concentration inequalities for functionals of the form

$$F(\mathbb{X}) := \int_0^\infty f_t(X_t)\, dt, \quad f_t : E \to \mathbb{R}. \tag{2.1}$$

The most familiar case is when F is of the form

$$\int_0^T f(X_t)\, dt,$$

i.e. $f_t \equiv f$ for $t \leq T$ and $f_t \equiv 0$ for $t > T$. We first formulate conditions on the family of functions f_t which we will need later.

Definition 2.2.1. *We say the family of functions $\{f_t, t \geq 0\}$ is k-regular for $k \in \mathbb{N}$, if:*

a) The f_t are Borel measurable and $t \mapsto f_{t+s}(X_s)$ is Lebesgue-integrable $\mathbb{P}_x$-a.s. for every $x \in E, t \geq 0$, and $\mathbb{E}_x \int_0^\infty |\, f_{t+s}(X_s) \,|\, ds < \infty$;

b) $\mathbb{E}_x \sup_{0\leq s\leq \epsilon} |\, f_{t+s}(X_s) \,|^k$ is well-defined and finite for $t \geq 0$, $x \in E$ arbitrary and $\epsilon > 0$ small enough;

c) There exists a function $r : E \to \mathbb{R}$ and $\epsilon_0 > 0$ such that for $0 < \epsilon < \epsilon_0$ and $x \in E$

$$\sup_{t\geq 0} \mathbb{E}_x \int_0^\infty |\, f_{t+\epsilon+s}(X_s) - f_{t+s}(X_s) \,|\, ds \leq \epsilon r(x)$$

and $\mathbb{E}_x r(X_\epsilon)^k < \infty$.

Remark If $F(\mathbb{X}) = \int_0^T f(X_t)\, dt$, then $\mathbb{E}_x \sup_{0\leq t\leq T+\epsilon_0} |\, f(X_t) \,|^k < \infty$ for some $\epsilon_0 > 0$ implies conditions b) and c) of the k-regularity. In condition b) the statement of well-definedness can be replaced by the existence of a measurable upper bound.

The technique to obtain concentration inequalities for functionals of the form (2.1) is to use a telescoping approach where one conditions on $\mathfrak{F}_t$, i.e., where we average $F(\mathbb{X})$ under the knowledge of the path of the Markov process $\mathbb{X}$ up to time t.

Definition 2.2.2. *For $0 \le s \le t$, define the increments*

$$\Delta_{s,t} := \mathbb{E}[F(\mathbb{X})|\mathfrak{F}_t] - \mathbb{E}[F(\mathbb{X})|\mathfrak{F}_s]$$

and the initial increment

$$\Delta_{\star,0} := \mathbb{E}[F(\mathbb{X})|\mathfrak{F}_0] - \mathbb{E}_\nu[F(\mathbb{X})],$$

which depends on the initial distribution ν.

The basic property of the increments is the relation $\Delta_{s,u} = \Delta_{s,t} + \Delta_{t,u}$ for $s < t < u$. Also, we have

$$\mathbb{E}[F(\mathbb{X})|\mathfrak{F}_T] - \mathbb{E}_\nu[F(\mathbb{X})] = \Delta_{\star,0} + \Delta_{0,T},$$

where we have to use $\Delta_{\star,0}$ to accommodate for the initial distribution ν. To better work with the increment $\Delta_{s,t}$, we will rewrite it in a more complicated but also more useful way.

Definition 2.2.3. *Given the family of functions $\{f_t : t \ge 0\}$, the coupled function difference is defined as*

$$\Phi_t(x,y) := \int_0^\infty S_u f_{t+u}(x) - S_u f_{t+u}(y)\, du.$$

Remark We call Φ_t the coupled function difference because later we will see that we need estimates on $|\,\Phi_t\,|$, and for a coupling $\widehat{\mathbb{E}}$ of $\mathbb{X}$ starting in x and y we have the estimate

$$\Phi_t(x,y) \le \int_0^\infty \widehat{\mathbb{E}}_{x,y}\,|\, f_{t+u}(X_u) - f_{t+u}(Y_u)\,|\; du.$$

In the next lemma we express the increments $\Delta_{s,t}$ in terms of the coupled function difference Φ_t.

Lemma 2.2.4.

$$\Delta_{s,t} = \int_s^t f_u(X_u) - S_{u-s}f_u(X_s)\,du + [S_{t-s}\Phi_t(X_t,\cdot)](X_s).$$

Proof. First, we note that

$$\mathbb{E}[F(\mathbb{X})|\mathfrak{F}_t] = \int_0^t f_u(X_u)\,du + \int_t^\infty S_{u-t}f_u(X_t)\,du,$$

and

$$\mathbb{E}[F(\mathbb{X})|\mathfrak{F}_s] = \int_0^s f_u(X_u)\,du + \int_s^t S_{u-s}f_u(X_s)\,du + \left[S_{t-s}\int_t^\infty S_{u-t}f_u\,du\right](X_s).$$

Hence,

$$\begin{aligned}
\Delta_{s,t} &= \mathbb{E}[F(\mathbb{X})|\mathfrak{F}_t] - \mathbb{E}[F(\mathbb{X})|\mathfrak{F}_s] \\
&= \int_s^t f_u(X_u) - S_{u-s}f_u(X_s)\,du \\
&\quad + S_{t-s}\left[\int_t^\infty S_{u-t}f_u(X_t) - S_{u-t}f_u\,du\right](X_s) \\
&= \int_s^t f_u(X_u) - S_{u-s}f_u(X_s)\,du + [S_{t-s}\Phi_t(X_t,\cdot)](X_s). \qquad \square
\end{aligned}$$

The following lemma is crucial to obtain the concentration inequalities of Theorems 2.2.6 and 2.2.9 below. It expresses conditional moments of the increments in terms of the coupled function difference.

Lemma 2.2.5. *Fix $k \in \mathbb{N}$, $k \geq 2$. Assume that the family (f_t) is k-regular and suppose that $\Phi_t(\cdot,x)^k$ is in the domain of the generator A for all $x \in E$. Then*

$$\lim_{\epsilon\to 0}\frac{1}{\epsilon}\mathbb{E}\left[\Delta_{t,t+\epsilon}^k \,\middle|\, \mathfrak{F}_t\right] = (A(\Phi_t(\cdot,X_t)^k))(X_t).$$

Proof. We will use the following elementary fact repetitively. For $k \geq 2$, if $|\,b_\epsilon\,| \leq \epsilon\bar{b}_\epsilon$ and $\sup_{0\leq\epsilon\leq\epsilon_0}\mathbb{E}\bar{b}_\epsilon^k < \infty$, then

$$\lim_{\epsilon\to 0}\frac{1}{\epsilon}\mathbb{E}(a_\epsilon + b_\epsilon)^k = \lim_{\epsilon\to 0}\frac{1}{\epsilon}\mathbb{E}a_\epsilon^k. \tag{2.2}$$

By Lemma 2.2.4,

$$\Delta_{t,t+\epsilon} = \int_t^{t+\epsilon} f_u(X_u) - S_{u-t} f_u(X_t)\, du + [S_\epsilon \Phi_{t+\epsilon}(X_{t+\epsilon}, \cdot)](X_t).$$

First, we show that we can neglect the first term. Indeed,

$$\begin{aligned}
&\left| \int_t^{t+\epsilon} f_u(X_u) - S_{u-t} f_u(X_t)\, du \right| \\
&\quad \leq \epsilon \sup_{0\leq s\leq \epsilon} |\, f_{t+s}(X_{t+s})\, | + \epsilon \mathbb{E}^{\mathbb{Y}}_{X_t} \sup_{0\leq s\leq\epsilon} |\, f_{t+s}(Y_s)\, |,
\end{aligned}$$

we can use part b) of the k-regularity to apply fact (2.2) and get

$$\lim_{\epsilon\to 0} \frac{1}{\epsilon} \mathbb{E}\left[\Delta^k_{t,t+\epsilon} \,\middle|\, \mathfrak{F}_t\right] = \lim_{\epsilon\to 0} \frac{1}{\epsilon} \mathbb{E}\left[[S_\epsilon \Phi_{t+\epsilon}(X_{t+\epsilon}, \cdot)]^k (X_t) \,\middle|\, \mathfrak{F}_t \right].$$

Next, by writing $\Phi_{t+\epsilon} = \Phi_t + (\Phi_{t+\epsilon} - \Phi_t)$, we will show that the difference can be neglected in the limit $\epsilon \to 0$. To this end, we observe that

$$\begin{aligned}
|\, \Phi_{t+\epsilon}(x,y) - \Phi_t(x,y)\, | \leq & \int_0^\infty \mathbb{E}_x \,|\, f_{t+\epsilon+u}(X_u) - f_{t+u}(X_u)\, |\; du \\
& + \int_0^\infty \mathbb{E}^{\mathbb{X}}_y \,|\, f_{t+\epsilon+u}(X_u) - f_{t+u}(X_u)\, |\; du.
\end{aligned}$$

Part c) of the $k-$regularity condition allows us to invoke fact (2.2) again to obtain

$$\lim_{\epsilon\to 0} \frac{1}{\epsilon} \mathbb{E}\left[\Delta^k_{t,t+\epsilon} \,\middle|\, \mathfrak{F}_t\right] = \frac{1}{\epsilon} \mathbb{E}\left[[S_\epsilon \Phi_t(X_{t+\epsilon}, \cdot)]^k (X_t) \,\middle|\, \mathfrak{F}_t \right].$$

Finally, to replace $S_\epsilon \Phi_t(X_{t+\epsilon}, \cdot)$ by $\Phi_t(X_{t+\epsilon}, \cdot)$ by applying fact (2.2) for a third time, we estimate

$$\begin{aligned}
&|\, [S_\epsilon \Phi_t(y, \cdot)](x) - \Phi_t(y,x)\, | \\
&\leq \left| \int_0^\infty S_{u+\epsilon} f_{t+u+\epsilon}(x) - S_{u+\epsilon} f_{t+u}(x)\, du \right| + \left| \int_0^\epsilon S_u f_{t+u}(x)\, du \right| \\
&\leq \mathbb{E}_x \int_0^\infty |\, f_{t+u+\epsilon}(X_{u+\epsilon}) - f_{t+u}(X_{u+\epsilon})\, |\; du + \epsilon \mathbb{E}_x \sup_{0\leq u\leq\epsilon} f_{t+u}(X_u),
\end{aligned}$$

where parts b) and c) of the k-regularity then provide the necessary estimates. Now, the desired result is immediately achieved:

$$\begin{aligned}\lim_{\epsilon\to 0}\frac{1}{\epsilon}\mathbb{E}\left[\Delta^k_{t,t+\epsilon}\,\middle|\,\mathfrak{F}_t\right] &= \lim_{\epsilon\to 0}\frac{1}{\epsilon}\left[S_\epsilon\left(\Phi_t(\cdot,X_t)\right)^k\right](X_t)\\ &= A\Phi_t(\cdot,X_t)^k(X_t).\end{aligned}$$
□

We can now state our first main theorem, which is a bound of the exponential moment of $F(\mathbb{X})$ in terms of the coupled function difference Φ_t.

Theorem 2.2.6. *Assume that for all $k\in\mathbb{N}$, the f_t are k-regular and $\Phi_t(\cdot,x)^k\in\operatorname{dom}(A)$ for all $x\in E$. Then, for any distributions μ and ν on E,*

$$\log\mathbb{E}_\mu\left[e^{F(\mathbb{X})-\mathbb{E}_\nu F(\mathbb{X})}\right]\le\log(c_0)+\int_0^\infty\sup_{x\in E}\sum_{k=2}^\infty\frac{1}{k!}(A(\Phi_t^k(\cdot,x)))(x)\,dt,$$

$$\log\mathbb{E}_\mu\left[e^{F(\mathbb{X})-\mathbb{E}_\nu F(\mathbb{X})}\right]\ge\log(c_0)+\int_0^\infty\inf_{x\in E}\sum_{k=2}^\infty\frac{1}{k!}(A(\Phi_t^k(\cdot,x)))(x)\,dt,$$

where the influence of the distributions μ and ν is only present in the factor

$$c_0=\int e^{\nu(\Phi_0(x,\cdot))}\,\mu(dx).$$

Remark If $H_t : E\times E$ is an upper bound on $|\,\Phi_t\,|$ and $H_t(x,x)=0$ for all $x\in E$, then the upper bound of the theorem remains valid if Φ_t is replaced by H_t. In particular, if $f_t\equiv f\mathbb{1}_{t\le T}$, $H_t := |\,\Phi_0\,|\,\mathbb{1}_{t\le T}$ serves as a good initial estimate to obtain the upper bound

$$\log\mathbb{E}_\mu\left[e^{F(\mathbb{X})-\mathbb{E}_\nu F(\mathbb{X})}\right]\le\log(c_0)+T\sup_{x\in E}\sum_{k=2}^\infty\frac{1}{k!}A\,|\,\Phi_0\,|^k(\cdot,x)(x).$$

Further estimates on $|\,\Phi_0\,|$ specific to the particular process can then be used without the need to keep a dependence on t.

Proof. Define

$$\Psi(t):=\mathbb{E}_\mu\left[e^{\Delta_{\star,0}+\Delta_{0,t}}\right].$$

We see that for $\epsilon > 0$,

$$\begin{aligned}\Psi(t+\epsilon) - \Psi(t) &= \mathbb{E}_\mu \left(e^{\Delta_{\star,0}+\Delta_{0,t}} \mathbb{E}\left[e^{\Delta_{t,t+\epsilon}} - 1 \,\middle|\, \mathfrak{F}_t \right] \right) \\ &= \mathbb{E}_\mu \left(e^{\Delta_{\star,0}+\Delta_{0,t}} \mathbb{E}\left[e^{\Delta_{t,t+\epsilon}} - \Delta_{t,t+\epsilon} - 1 \,\middle|\, \mathfrak{F}_t \right] \right),\end{aligned}$$

where we used the fact that $\mathbb{E}[\Delta_{t,t+\epsilon}|\mathfrak{F}_t] = 0$. Hence, using Lemma 2.2.5, we can calculate the derivative of Ψ:

$$\Psi'(t) = \mathbb{E}_\mu \left(e^{\Delta_{\star,0}+\Delta_{0,t}} \sum_{k=2}^{\infty} \frac{1}{k!} (A(\Phi_t(\cdot, X_t)^k))(X_t) \right).$$

To get upper or lower bounds on Ψ', we move the sum out of the expectation as a supremum or infimum. Just continuing with the upper bound, as the lower bound is analogue,

$$\Psi'(t) \le \Psi(t) \sup_{x\in E} \sum_{k=2}^{\infty} \frac{1}{k!} (A(\Phi_t^k(\cdot, x)))(x).$$

After dividing by $\Psi(t)$ and integrating, we get

$$\ln \Psi(T) - \ln \Psi(0) \le \int_0^T \sup_{x\in E} \sum_{k=2}^{\infty} \frac{1}{k!} (A(\Phi_t^k(\cdot, x)))(x)\, dt,$$

which leads to

$$\lim_{T\to\infty} \Psi(T) = \mathbb{E}_\mu \left[e^{F(\mathbb{X}) - \mathbb{E}_\nu F(\mathbb{X})} \right] \le \Psi(0) e^{\int_0^\infty \sup_{x\in E} \sum_{k=2}^{\infty} \frac{1}{k!} (A(\Phi_t^k(\cdot,x)))(x)\, dt}.$$

The value of $c_0 = \Psi(0) = \mathbb{E}_\mu e^{\Delta_{\star,0}}$ is obtained from the identity

$$\Delta_{\star,0} = \nu\left(\Phi_0(X_0, \cdot)\right). \qquad \square$$

How the bound in Theorem 2.2.6 can be used to obtain a deviation probability in the most common case is shown by the following corollary.

Corollary 2.2.7. *Assume that* $F(\mathbb{X}) = \int_0^T f(X_t)\, dt$*, the conditions of Theorem 2.2.6 are satisfied, and* $\sup_{x\in E} A \,|\, \Phi_0 \,|^k (\cdot, x)(x) \le c_1 c_2^k$ *for some* $c_1, c_2 > 0$*.*

Then, for any initial condition $x \in E$,

$$\mathbb{P}_x(F(\mathbb{X}) - \mathbb{E}_x F(\mathbb{X}) > x) \leq e^{\frac{-\frac{1}{2}(\frac{x}{c_2})^2}{Tc_1 + \frac{1}{3}\frac{x}{c_2}}}.$$

Proof. By Markov's inequality,

$$\begin{aligned}\mathbb{P}_x(F(\mathbb{X}) - \mathbb{E}_x F(\mathbb{X}) > x) &\leq \mathbb{E}_x e^{\lambda F(\mathbb{X}) - \mathbb{E}_x \lambda F(\mathbb{X})} e^{-\lambda x} \\ &\leq e^{Tc_1 \sum_{k=2}^{\infty} \frac{1}{k!} \lambda^k c_2^k - \lambda x},\end{aligned}$$

where the last line is the result from Theorem 2.2.6. Through optimizing λ, the exponent becomes

$$\frac{x}{c_2} - (Tc_1 + \frac{x}{c_2}) \log(\frac{x}{Tc_1 c_2} + 1).$$

To show that this term is less than $\frac{-\frac{1}{2}(\frac{x}{c_2})^2}{Tc_1 + \frac{1}{3}\frac{x}{c_2}}$, we first rewrite it as the following inequality:

$$\log(\frac{x}{Tc_1 c_2} + 1) \geq \frac{\frac{\frac{1}{2}(\frac{x}{c_2})^2}{Tc_1 + \frac{1}{3}\frac{x}{c_2}} + \frac{x}{c_2}}{Tc_1 + \frac{x}{c_2}}.$$

Through comparing the derivatives, one concludes that the left hand side is indeed bigger than the right hand side. □

In applications one tries to find good estimates of Φ_t. When looking at the examples in Section 2.4, finding those estimates is where the actual work lies. In the case where the functions f_t are Lipschitz continuous with respect to a suitably chosen (semi)metric ρ, the problem can be reduced to questions about the generalized coupling time h, which is defined and discussed in detail in Section 2.3. In case that the exponential moment of $F(\mathbb{X}) - \mathbb{E}F(\mathbb{X})$ does not exist or the bound obtained from Theorem 2.2.6 is not useful, we turn to moment bounds. This is the content of the next theorem.

Lemma 2.2.8. *Assume that the f_t are 2-regular and $\Phi_t^2(\cdot, x)$ is in the domain of the generator A. Then the predictable quadratic variation of the martingale*

$(\Delta_{0,t})_{t\geq 0}$ *is*

$$\langle \Delta_{0,\cdot} \rangle_t = \int_0^t A\Phi_s^2(\cdot, X_s)(X_s)\, ds.$$

Proof. We have, using Lemma 2.2.5 for $k=2$,

$$\frac{d}{dt} \langle \Delta_{0,\cdot} \rangle_t = \lim_{\epsilon\to 0} \frac{1}{\epsilon} \mathbb{E}\left[\Delta_{t,t+\epsilon}^2 \,\middle|\, \mathfrak{F}_t\right] = A\Phi_t^2(\cdot, X_t)(X_t). \qquad \square$$

Theorem 2.2.9. *Let the functions f_t be 2-regular and $\Phi_t^2(\cdot, x)$ in the domain of the generator A. Then*

$$\left(\mathbb{E}_\mu \,|\, F(\mathbb{X}) - \mathbb{E}_\nu F(\mathbb{X}) \,|^p\right)^{\frac{1}{p}} \leq C_p \left[\left(\mathbb{E}_\mu \left(\int_0^\infty A\Phi_t^2(\cdot, X_t)(X_t)\, dt \right)^{\frac{p}{2}} \right)^{\frac{1}{p}} \right. \tag{2.3a}$$

$$\left. + \left(\mathbb{E}_\mu \left(\sup_{t\geq 0} |\, \Phi_t(X_t, X_{t-}) \,| \right)^p \right)^{\frac{1}{p}} \right] \tag{2.3b}$$

$$+ \left(\int |\, \nu\left(\Phi_0(x,\cdot)\right) |^p \;\mu(dx) \right)^{\frac{1}{p}} \tag{2.3c}$$

where the constant C_p only depends on p and behaves like $p/\log p$ as $p\to\infty$.

Proof. By the triangle inequality,

$$\left(\mathbb{E}_\mu \,|\, F(\mathbb{X}) - \mathbb{E}_\nu F(\mathbb{X}) \,|^p\right)^{\frac{1}{p}} \leq \left(\mathbb{E}_\mu \,|\, \Delta_{0,\infty} \,|^p\right)^{\frac{1}{p}} + \left(\mathbb{E}_\mu \,|\, \Delta_{\star,0} \,|^p\right)^{\frac{1}{p}}.$$

Since $(\Delta_{0,t})_{t\geq 0}$ is a square integrable martingale starting at 0, a version of Rosenthal's inequality([43], Theorem 1) implies

$$\left(\mathbb{E}_\mu \,|\, \Delta_{0,T} \,|^p\right)^{\frac{1}{p}} \leq C_p \left[\left(\mathbb{E}_\mu \langle \Delta_{0,\cdot} \rangle_T^{\frac{p}{2}} \right)^{\frac{1}{p}} + \left(\mathbb{E}_\mu \sup_{0\leq t\leq T} |\, \Delta_{0,t} - \Delta_{0,t-} \,|^p \right)^{\frac{1}{p}} \right].$$

Applying Lemma 2.2.8 to rewrite the predictable quadratic variation $\langle \Delta_{0,\cdot}\rangle_T$ and Lemma 2.2.4 to rewrite $\Delta_{t-,t}$, we end with the first two terms of our claim after letting $T \to \infty$. The last term is just a different way of writing $\Delta_{\star,0}$:

$$(\mathbb{E}_\mu \,|\, \Delta_{\star,0} \,|^p)^{\frac{1}{p}} = \left(\int |\, \nu\,(\Phi_0(x,\cdot))\,|^p \;\mu(dx) \right)^{\frac{1}{p}}. \qquad \square$$

Let us discuss the meaning of the three terms appearing on the right hand side in Theorem (2.2.9).

a) The first term gives the contribution, typically of order $T^{\frac{p}{2}}$, that one expects even in the simplest case of processes with independent increments. E.g. if μ is an invariant measure and $F(\mathbb{X}) = \int_0^T f(X_t)\,dt$, then

$$\mathbb{E}_\mu \left(\int_0^\infty A\Phi_t^2(\cdot, X_t)(X_t)\,dt \right)^{\frac{p}{2}} \le T^{\frac{p}{2}} \int \left(A\Phi_0^2(\cdot, x)(x) \right)^{\frac{p}{2}} \mu(dx).$$

In many cases (see examples below), $\int \left(A\Phi_0^2(\cdot, x)(x) \right)^{\frac{p}{2}} \mu(dx)$ can be treated as a constant, i.e., not depending on T. There are however relevant examples where this factor blows up as $T \to \infty$.

b) The second term measures rare events of possibly large jumps where it is very difficult to couple. If the process $\mathbb{X}$ has continuous paths, this term is not present. Usually this term is or bounded or is of lower order than the first term as $T \to \infty$.

c) The third term has only the hidden time dependence of Φ_0 on T. It measures the intrinsic variation given the starting measures μ and ν and it vanishes if and only if $\mu = \nu = \delta_x$.

It is also interesting to note that the estimate is sharp for small T: If one chooses $F(\mathbb{X}) = \frac{1}{T} \int_0^T f(X_t)\,dt$ and looks at the limit as $T \to 0$, the first two terms disappear and the third one becomes $(\int |\, f(x) - \nu(f)\,|^p \;\mu(dx))^{\frac{1}{p}}$, which is also the limit of the left hand side.

2.3 Generalized coupling time

In order to apply the results of Section 2.2 we need estimates on Φ_t. We can obtain these if we know more about the coupling behaviour of the underlying process $\mathbb{X}$. To characterize this coupling behaviour, we will look at how close we can get two versions of the process started at different points measured with respect to a distance.

Let $\rho : E \times E \to [0,\infty]$ be a lower semi-continuous semi-metric. With respect to this semi-metric, we define

$$\| f \|_{Lip} := \inf \{ r \geq 0 \mid f(x) - f(y) \leq r\rho(x,y) \; \forall\, x,y \in E \},$$

the Lipschitz-seminorm of f corresponding to ρ. Now we introduce the main objects of study in this section.

Definition 2.3.1. *a) The optimal coupling distance at time t is defined as*

$$\rho_t(x,y) := \inf_{\pi \in \mathfrak{P}(\delta_x S_t, \delta_y S_t)} \int \rho(x',y')\, \pi(dx'dy'),$$

where the infimum ranges over the set of all possible couplings with marginals $\delta_x S_t$ and $\delta_y S_t$, i.e., the distribution of X_t started from x or y.

b) The generalized coupling time is defined as

$$h(x,y) := \int_0^\infty \rho_t(x,y)\, dt.$$

Now that we have introduced the generalized coupling time, as first application we obtain, using the remark following Theorem 2.2.6:

Corollary 2.3.2. *Assume the functions f_t are Lipschitz continuous with respect to a semi-metric ρ, and that the conditions of Theorem 2.2.6 hold true. Then*

$$\mathbb{E}_\mu \left[e^{F(\mathbb{X}) - \mathbb{E}_\nu F(\mathbb{X})} \right] \leq c_0 e^{\sum_{k=2}^{\infty} \frac{c_k}{k!} \sup_{x \in E} (A(h^k(\cdot,x)))(x)},$$

where

$$c_0 = \int e^{\sup_{t\geq 0} \| f_t \|_{Lip} \nu(h(x,\cdot))} \mu(dx),$$
$$c_k = \int_0^\infty \sup_{t \geq 0} \| f_t \|_{Lip}^k \, dt.$$

In particular, if $f_t \equiv f$ for $t \leq T$ and $f_t \equiv 0$ for $t > T$, then

$$c_0 \leq \int e^{\| f \|_{Lip} \nu(h(x,\cdot))} \mu(dx),$$
$$c_k \leq T \| f \|_{Lip}^k .$$

Remark If $\overline{h}$ is an upper bound on the generalized coupling time h with $\overline{h}(x,x) = 0$, then the result holds true with h replaced by $\overline{h}$.

Proposition 2.3.3. *The optimal coupling distance ρ_t has the dual formulation*

$$\rho_t(x,y) = \sup_{\| f \|_{Lip}=1} (S_t f(x) - S_t f(y)).$$

Proof. By the Kantorovich-Rubinstein theorem ([53], Theorem 1.14), we have

$$\begin{aligned} \inf_{\pi \in \mathfrak{P}(\delta_x S_t, \delta_y S_t)} \int \rho \, d\pi &= \sup_{\| f \|_{Lip}=1} \left[\int f \, d(\delta_x S_t) - \int f \, d(\delta_y S_t) \right] \\ &= \sup_{\| f \|_{Lip}=1} \left[(S_t f)(x) - (S_t f)(y) \right]. \end{aligned}$$ □

Also, it is easy to see that the semi-metric properties of ρ translate to ρ_t and thereby to the generalized coupling time h.

Proposition 2.3.4. *Both the optimal coupling distance ρ_t and the generalized coupling time h are semi-metrics.*

Proof. We only have to prove the semi-metric properties of ρ_t, they translate naturally to h via integration.

Obviously, $\rho_t(x,x) = 0$ and $\rho_t(x,y) = \rho_t(y,x)$ is true for all $x, y \in E$ by definition of ρ_t. For the triangle inequality, we use the dual representation:

$$\begin{aligned} \rho_t(x,y) &= \sup_{\| f \|_{Lip}=1} (S_t f(x) - S_t f(y)) \\ &= \sup_{\| f \|_{Lip}=1} (S_t f(x) - S_t f(z) + S_t f(z) - S_t f(y)) \\ &\leq \rho_t(x,z) + \rho_t(y,z) \end{aligned}$$ □

A first result is a simple estimate on the decay of the semigroup S_t in terms of the optimal coupling distance.

Proposition 2.3.5. *Let μ be a stationary probability measure of the semigroup S_t. Then*

$$\| S_t f - \mu(f) \|_{L^p(\mu)} \leq \| f \|_{Lip} \left(\int \mu(dx) \left(\int \mu(dy) \rho_t(x,y) \right)^p \right)^{\frac{1}{p}}.$$

Remark When we choose the metric ρ to be the discrete metric $\mathbb{1}_{x \neq y}$ (a choice we can make even in a non-discrete setting), we can estimate $\rho_t(x,y)$ by $\widehat{\mathbb{P}}_{x,y}(\tau > t)$, the probability that the coupling time

$$\tau = \inf \left\{ t \geq 0 \mid X_s^1 = X_s^2 \ \forall s \geq t \right\}$$

is larger than t in an arbitrary coupling $\widehat{\mathbb{P}}_{x,y}$ of the Markov process started in x and y. In this case, the result of Proposition 2.3.5 reads

$$\| S_t f - \mu(f) \|_{L^p(\mu)} \leq \| f \|_{osc} \left(\int \mu(dx) \left(\int \mu(dy) \widehat{\mathbb{P}}_{x,y}(\tau > t) \right)^p \right)^{\frac{1}{p}},$$

where $\| f \|_{osc} = \sup_{x,y}(f(x) - f(y))$ is the oscillation norm. Note that this can also be gained from the well-known coupling inequality

$$\| \delta_x S_t - \delta_y S_t \|_{TVar} \leq 2\widehat{\mathbb{P}}_{x,y}(\tau > t).$$

Proof of Proposition 2.3.5. First,

$$\begin{aligned} | S_t f(x) - \mu(f) | &= | S_t f(x) - \mu(S_t f) | \\ &= \left| \mathbb{E}_x f(X_t) - \int \mu(dy) \mathbb{E}_y f(Y_t) \right| \\ &\leq \int \mu(dy) \, | \mathbb{E}_x f(X_t) - \mathbb{E}_y f(Y_t) | \\ &\leq \int \mu(dy) \, \| f \|_{Lip} \, \rho_t(x,y). \end{aligned}$$

This estimate can be applied directly to get the result:

$$\| S_t f - \mu(f) \|_{L^p(\mu)} = \left(\int \mu(dx) \, | S_t f(x) - \mu(f) |^p \right)^{\frac{1}{p}}$$

$$\leq \| f \|_{Lip} \left(\int \mu(dx) \left(\int \mu(dy) \rho_t(x,y) \right)^p \right)^{\frac{1}{p}}. \qquad \square$$

The above result did not use the semigroup property of S_t. When we use it we can improve estimates considerably. The price is that from now on, ρ has to be a metric, and this metric must be compatible with the Markov process, which we will define a little bit later under the notion of contraction with respect to this metric. The aim is to show how the uniform boundedness of the generalized coupling time implies an exponential decay of the semigroup (S_t) in the Lipschitz seminorm. To this end, we need the following lemma:

Lemma 2.3.6. *Under the condition that ρ is a metric,*

$$\sup_{x \neq y} \frac{\rho_t(x,y)}{\rho(x,y)} = \| S_t \|_{Lip}.$$

Proof. By the representation of the optimal coupling distance in Proposition 2.3.3,

$$\begin{aligned} \sup_{x \neq y} \frac{\rho_t(x,y)}{\rho(x,y)} &= \sup_{x \neq y} \sup_{\| f \|_{Lip} = 1} \frac{S_t f(x) - S_t f(y)}{\rho(x,y)} \\ &= \sup_{\| f \|_{Lip} = 1} \| S_t f \|_{Lip} = \| S_t \|_{Lip}. here \qquad \square \end{aligned}$$

Definition 2.3.7. *We say that the process $\mathbb{X}$ acts as a* contraction for the distance ρ *if*

$$\rho_t(x,y) \leq \rho(x,y) \quad \forall t \geq 0, \tag{2.4}$$

or equivalently,

$$\| S_t \|_{Lip} \leq 1 \quad \forall t \geq 0.$$

This property is sufficient to show that the process is contracting the distance monotonely:

Lemma 2.3.8. *Assume that the process $\mathbb{X}$ acts as a contraction for the distance. Then*

$$\rho_{t+s}(x,y) \leq \rho_t(x,y) \quad \forall x, y \in E, s, t \geq 0.$$

Proof. Using the dual representation,

$$\begin{aligned}\rho_{t+s}(x,y) &= \sup_{\| f \|_{Lip}=1} [S_{t+s}f(x) - S_{t+s}f(y)] \\ &= \sup_{\| f \|_{Lip}\leq 1} [S_t(S_s f)(x) - S_t(S_s f)(y)].\end{aligned}$$

By our assumption, the set of functions f with $\| f \|_{Lip} \leq 1$ are a subset of the set of functions f with $\| S_s f \|_{Lip} \leq 1$. Hence,

$$\begin{aligned}\rho_{t+s}(x,y) &\leq \sup_{f:\| S_s f \|_{Lip}\leq 1} [S_t(S_s f)(x) - S_t(S_s f)(y)] \\ &\leq \sup_{\| g \|\leq 1} [S_t g(x) - S_t g(y)] = \rho_t(x,y). \qquad \square\end{aligned}$$

With this property in mind, we can show the main theorem of this section.

Theorem 2.3.9. *Assume that ρ is a metric and that the process $\mathbb{X}$ acts as a contraction for the distance. Then the fact that the generalized coupling time h is bounded by the metric ρ is equivalent to the fact that the semigroup (S_t) is exponentially contracting. More precisely, for $\alpha > 1$ arbitrary,*

a) $\forall\, x,y \in E: \ h(x,y) \leq M\rho(x,y) \quad \Rightarrow \quad \forall\, t \geq M\alpha: \ \| S_t \|_{Lip} \leq \frac{1}{\alpha}$;

b) $\| S_T \|_{Lip} \leq \frac{1}{\alpha} \quad \Rightarrow \quad \forall\, x,y \in E: \ h(x,y) \leq \frac{\alpha T}{\alpha - 1}\rho(x,y)$.

Proof. a) For $x,y \in E$, set

$$T_{x,y} := \inf\left\{ t \geq 0 \,\middle|\, \rho_t(x,y) \leq \frac{1}{\alpha}\rho(x,y) \right\}.$$

Then,

$$M\rho(x,y) \geq h(x,y) = \int_0^\infty \rho_t(x,y)\,dt \geq \int_0^{T_{x,y}} \rho_t(x,y)\,dt \geq T_{x,y}\frac{1}{\alpha}\rho(x,y).$$

Therefore $T_{x,y}$ is bounded by $M\alpha$. By Lemma 2.3.8, $\rho_t(x,y) \leq \rho_{T_{x,y}}(x,y)$ for all $t \geq T_{x,y}$. Hence $\rho_{M\alpha}(x,y) \leq \frac{1}{\alpha}\rho(x,y)$ uniformly, which implies $\| S_{M\alpha} \|_{Lip} \leq \frac{1}{\alpha}$.

b) Since $\rho_t(x,y) \leq \rho(x,y) \| S_t \|_{Lip}$,

$$h(x,y) = \int_0^\infty \rho_t(x,y)\,dt \leq \rho(x,y) \int_0^\infty \| S_t \|_{Lip}\,dt$$

$$\leq \rho(x,y)T\sum_{k=0}^{\infty} \| S_T \|_{Lip}^k \leq \frac{\alpha T}{\alpha - 1}\rho(x,y). \qquad \square$$

When we apply this theorem to an arbitrary Markov process where we use the discrete distance, we get the following corollary:

Corollary 2.3.10. *The following two statements are equivalent:*

a) The generalized coupling time with respect to the discrete metric $\rho(x,y) = \mathbb{1}_{x\neq y}$ is uniformly bounded, i.e.

$$h(x,y) \leq M \quad \forall x,y \in E;$$

b) The semigroup is eventually contractive in the oscillation (semi)-norm, i.e., $\| S_T \|_{osc} < 1$ for some $T > 0$.

Remark Theorem 2.3.9 actually gives us more information, namely how the constants M and T can be related to each other.

Proof. Since obviously $\sup_{x\neq y} \rho_t(x,y) \leq 1$, the process $\mathbb{X}$ acts as a contraction for the discrete distance and the result follows from Theorem 2.3.9, where we also use the fact that in the case of the discrete metric, $\| \cdot \|_{Lip} = \| \cdot \|_{osc}$. $\square$

Since Theorem 2.3.9 part a) implies that $\| S_t \|_{Lip}$ decays exponentially fast, it is of interest to get the best estimate on the speed of decay, which is the content of the following proposition:

Proposition 2.3.11. *Assume that ρ is a metric, the process $\mathbb{X}$ acts as a contraction for the distance and the generalized coupling time h satisfies $h(x,y) \leq M\rho(x,y)$. Then*

$$\lim_{t\to\infty} \frac{1}{t} \log \| S_t \|_{Lip} \leq -\frac{1}{M}.$$

Proof. The proof uses the same structure as the proof of part a) in Theorem 2.3.9. First, fix ϵ between 0 and $\frac{1}{M}$. Then define

$$T_{x,y} = \inf\left\{ t > 0 \;\middle|\; \rho_t(x,y) \leq \rho(x,y) e^{-(\frac{1}{M}-\epsilon)t} \right\}.$$

By our assumption,

$$M\rho(x,y) \geq h(x,y) \geq \rho(x,y) \int_0^{T_{x,y}} e^{-(\frac{1}{M}-\epsilon)t}\, dt = M\rho(x,y) \frac{1-e^{-(\frac{1}{M}-\epsilon)T_{x,y}}}{1-M\epsilon}.$$

Since the fraction on the right hand side becomes bigger than 1 if $T_{x,y}$ is too large, there exists an uniform upper bound $T(\epsilon)$ on $T_{x,y}$. Hence, for all $t \geq T(\epsilon)$, $\| S_t \|_{Lip} \leq e^{-(\frac{1}{M}-\epsilon)t}$, which of course implies $\lim\limits_{t\to\infty} \frac{1}{t} \| S_t \|_{Lip} \leq -\frac{1}{M} + \epsilon$. By sending ϵ to 0, we finish our proof. □

Again, we apply this result to the discrete metric to see what it contains.

Corollary 2.3.12. *Let $\widehat{\mathbb{P}}_{x,y}$ be a coupling of the process $\mathbb{X}$ started in x resp. y, and denote with $\tau := \inf\left\{t \geq 0 \,\middle|\, X_s^1 = X_s^2 \ \forall s \geq t\right\}$ the coupling time. Set $M := \sup\limits_{x,y\in E} \widehat{\mathbb{E}}_{x,y}\tau$. Then*

$$\lim_{t\to\infty} \frac{1}{t} \log \| S_t \|_{osc} \leq -\frac{1}{M}.$$

Equivalently, for $f \in L^\infty$,

$$\lim_{t\to\infty} \frac{1}{t} \log \| S_t f - \mu(f) \|_\infty \leq -\frac{1}{M},$$

where μ is the unique stationary distribution of $\mathbb{X}$.

Remarks a) If the the Markov process $\mathbb{X}$ is also reversible, then the above result extends to L^1 and hence to any L^p, where the spectral gap is then also at least of size $\frac{1}{M}$.

b) As an additional consequence, when a Markov process can be uniformly coupled, i.e. $\sup\limits_{x,y\in E} \widehat{\mathbb{E}}_{x,y}\tau \leq M < \infty$ for a coupling $\widehat{\mathbb{E}}$, then there exists (a possibly different) coupling $\widetilde{\mathbb{E}}_{x,y}$, so that $\sup\limits_{x,y\in E} \widetilde{\mathbb{E}}_{x,y} e^{\lambda\tau} < \infty$ for all $\lambda < \frac{1}{M}$. Note that without Corollary 2.3.12 this property is obvious only for Markovian couplings.

2.4 Examples

2.4.1 Diffusions with a strictly convex potential

Let V be a twice continuously differentiable function on the real line with $V'' \geq c > 0$ and $\int e^{-V(x)}dx = Z_V < \infty$. To the potential V is associated the Gibbs measure

$$\mu_V(dx) = \frac{1}{Z_V} e^{-V(x)} dx$$

and a Markovian diffusion

$$dX_t = -V'(X_t) + \sqrt{2} dW_t$$

with μ_V as reversible measure.

To estimate the optimal coupling distance ρ_t at time t(see Definition 2.3.1), we couple two versions of the diffusion, X_t^x started in x and X_t^y started in y, by using the same Brownian motion $(W_t)_{t \geq 0}$. Then the difference process $X_t^x - X_t^y$ is deterministic, $x < y$ implies $X_t^x < X_t^y$ and by the convexity assumption

$$d(X_t^y - X_t^x) = -(V'(X_t^y) - V'(X_t^x)) \leq -c(X_t^y - X_t^x).$$

Using Gronwall's Lemma, we obtain the estimate

$$\rho_t(x, y) \leq |\, x - y \,|\, e^{-ct}$$

on the optimal coupling distance. By integration, the generalized coupling time h has the estimate $h(x, y) \leq \frac{1}{c} |\, x - y \,|$. As a consequence, Proposition 2.3.11 implies

$$\lim_{t \to \infty} \log \|\, S_t \,\|_{Lip} \leq -c.$$

Since the generator A of the diffusion is

$$A = \frac{d^2}{dx^2} - V' \cdot \frac{d}{dx},$$

we have

$$A\left(\frac{1}{c}|\cdot - x|\right)^k (x) = \begin{cases} \frac{2}{c^2}, & k = 2, \\ 0, & k > 2. \end{cases}$$

Therefore, for $f : \mathbb{R} \to \mathbb{R}$ be Lipschitz-continuous, we can use Corollary 2.3.2 to get the estimate

$$\mathbb{E}_{\nu_1}\left[e^{\int_0^T f(X_t)\,dt - \mathbb{E}_{\nu_2}\int_0^T f(X_t)\,dt}\right] \le c_{\nu_1,\nu_2} e^{T\frac{\|f\|_{Lip}^2}{c^2}}, \tag{2.5}$$

with the dependence on the distributions ν_1 and ν_2 given by

$$c_{\nu_1,\nu_2} = \mathbb{E}_{\nu_1}^x e^{\mathbb{E}_{\nu_2}^y \frac{\|f\|_{Lip}}{c}|x-y|}.$$

Remark a) An alternative proof that strict convexity is sufficient for (2.5) to be true can be found in [54]. A proof via the log-Sobolev inequality can be found in [36]. Hence the result is of no surprise, but the method of obtaining it is new.

b) This example demonstrates nicely how in the case of diffusions the higher moments of $Ah^k(\cdot, x)(x)$ can disappear because the generalized coupling time is bounded by a multiple of the initial distance.

c) The generalization to higher dimensions under strict convexity is straightforward.

2.4.2 Interacting particle systems

Let $E = \{0,1\}^{\mathbb{Z}^d}$ be the state space of the interacting particle system with a generator L given by

$$Lf(\eta) = \sum_x \sum_{\Delta \subset \mathbb{Z}^d} c(\eta, x + \Delta)[f(\eta^{x+\Delta}) - f(\eta)],$$

where η^Δ denotes the configuration η with all spins in Δ flipped. This kind of particle system is extensively treated in [37]. For $f : E \to \mathbb{R}$, we denote with $\delta_f(x) := \sup_{\eta \in E} f(\eta^x) - f(\eta)$ the maximal influence of a single flip at site x, and with $\delta_f = (\delta_f(x))_{x \in E}$ the vector of all those influences.

If there is a way to limit how flips in the configuration affect the system as time progresses, then we can obtain a concentration estimate. Again, denote with $F(\eta_\cdot) = \int_0^T f(\eta_t)\,dt$ the additive functional of the function f and the particle system $\eta_\cdot$.

Theorem 2.4.1. *Assume there exists a family of operators A_t so that $\delta_{S_t f} \leq A_t \delta_f$, and write*

$$G := \int_0^\infty A_t\,dt,$$

which is assumed to exist. Denote with

$$c_k := \sup_{\eta \in E, x \in \mathbb{Z}^d} \sum_{\Delta \subset \mathbb{Z}^d} c(\eta, x + \Delta) \,|\, \Delta \,|^k$$

the weighted maximal rate of spin flips. If $\|\, G \,\|_{p \to 2} < \infty$ for some $p \geq 1$, then for any f with $\delta_f \in \ell^p$ and any initial condition $\eta \in E$,

$$\mathbb{E}_\eta e^{F(\eta_\cdot) - \mathbb{E}_\eta F(\eta_\cdot)} \leq \exp\left[T \sum_{k=2}^\infty \frac{c_k \,\|\, G \,\|_{p\to 2}^k \,\|\, \delta_f \,\|_p^k}{k!} \right].$$

If additionally $\|\, G \,\|_1 < \infty$ and $|||\, f \,||| := \|\, \delta_f \,\|_1 < \infty$, then for any two probability distributions ν_1, ν_2,

$$\mathbb{E}_{\nu_1} e^{F(\eta_\cdot) - \mathbb{E}_{\nu_2} F(\eta_\cdot)} \leq \exp\left[\|\, G \,\|_1 \;|||\; f \;|||\; + T \sum_{k=2}^\infty \frac{c_k \,\|\, G \,\|_{p\to 2}^k \,\|\, \delta_f \,\|_p^k}{k!} \right].$$

Applications of this Theorem are for example spin flip dynamics in the so-called $M < \epsilon$ regime, where there exists an operator Γ with $\|\, \Gamma \,\|_1 = M$, so that

$$\delta_{S_t f} \leq e^{-t(\epsilon - \Gamma)} \delta_f$$

holds. Since $G = \int_0^\infty e^{-t(\epsilon-\Gamma)}\,dt = (\epsilon - \Gamma)^{-1}$, $\|\, G \,\|_1 \leq (\epsilon - M)^{-1}$. Hence $\|\, G \,\|_{1\to 2} \leq (\epsilon - M)^{-1}$ for a first application of the Theorem. If the process is reversible as well, $\|\, G \,\|_\infty = \|\, G \,\|_1$, and by Riesz-Thorin's Theorem, we have $\|\, G \,\|_2 \leq (\epsilon - M)^{-1}$, hence we get the result for functions f with $\|\, \delta_f \,\|_2 < \infty$.

Another example is the exclusion process. As a single discrepancy is preserved and moves like a random walk, $A_t(x,y) = p_t(x,y)$, the transition probability of the random walk. In high dimensions, $G(x,y) = \int_0^\infty p_t(x,y)\,dt$ has bounded $\ell^1 \to \ell^2$-norm:

$$\begin{aligned}
\| G \|_{1\to 2} &= \sup_{\| g \|_1 = 1} \sum_x \Big(\sum_y G(x,y) g(y)\Big)^2 \\
&\le \sup_{\| g \|_1 = 1} \sum_x \sum_y |\, g(y) \,|\, G(x,y)^2 \le \sum_x G(x,0)^2 \infty \\
&= \int_0^\infty \int_0^\infty \sum_x p_t(0,x) p_s(0,x)\, ds\, dt \\
&= \int_0^\infty \int_0^\infty p_{s+t}(0,0)\, ds\, dt < \infty
\end{aligned}$$

in dimension 5 and higher. As the exclusion process switches two sites, $c_k \le 2^k$, and hence

$$\mathbb{E}_\eta e^{F(\eta_\cdot) - \mathbb{E}_\eta F(\eta_\cdot)} \le \exp\left[T \sum_{k=2}^\infty \frac{2^k \| G \|_{1\to 2}^k \, \|\!| \, f \, |\!\|^k}{k!} \right].$$

However, this is only a quick result exploiting the strong diffusive behaviour in high dimensions. In the last section we will deal with the exclusion process in much more detail to obtain results for lower dimensions as well.

Proof of Theorem 2.4.1. First, we notice that the coupled function difference Φ_t for a single flip can be bounded by

$$\begin{aligned}
\Phi_t(\eta^x, \eta) &\le \int_0^\infty |\, S_t f(\eta^x) - S_t f(\eta) \,|\, dt \\
&\le \int_0^\infty \delta_{S_t f}(x)\, dt \le \int_0^\infty (A_t \delta_f)(x)\, dt \\
&\le (G\delta_f)(x)
\end{aligned}$$

uniformly in η. To estimate the coupled function difference Φ_t we telescope over single site flips,

$$\Phi_t^k(\eta^{x+\Delta}, \eta) \le |\, \Delta \,|^k \, ((G\delta_f)(x))^k,$$

and therefore

$$\begin{aligned} L\Phi_t^k(\cdot,\eta)(\eta) &= \sum_x \sum_{\Delta\subset\mathbb{Z}^d} c(\eta, x+\Delta)\Phi_t^k(\eta^{x+\Delta},\eta) \\ &\le \sum_x \sum_{\Delta\subset\mathbb{Z}^d} c(\eta, x+\Delta) \,|\, \Delta \,|^k \, (G\delta_f)^k(x) \\ &\le c_k \,\|\, G\delta_f \,\|_k^k \le c_k \,\|\, G\delta_f \,\|_2^k \le c_k \,\|\, G \,\|_{p\to 2}^k \,\|\, \delta_f \,\|_p^k \end{aligned}$$

Hence the first part is proven by applying these estimates to Theorem 2.2.6 for fixed and identical initial conditions. To prove the estimate for arbitrary initial distributions, we simply observe that, again by telescoping over single site flips,

$$\Phi_0(\eta,\xi) \le \sum_x \sup_\zeta \Phi_0(\zeta^x,\zeta) \le \sum_x (G\delta_f)(x) \le \|\, G \,\|_1 \,\|\, f \,\|_1 . \qquad \square$$

2.4.3 Simple symmetric random walk

The aim of this example is to show that we can get concentration estimates even if the process $\mathbb{X}$ - in this example a simple symmetric nearest neighbour random walk in $\mathbb{Z}^d$ - has no stationary distribution. We will consider three cases: $f \in \ell^1(\mathbb{Z}^d)$, $\ell^2(\mathbb{Z}^d)$ and $\ell^\infty(\mathbb{Z}^d)$, and $F(\mathbb{X}) = \int_0^T f(X_t)\,dt$. To apply Theorem 2.2.6, our task is to estimate $|\,\Phi_t(x,y)\,|$ where y is a neighbour of x. We will denote with $p_t(x,z)$ the transition probability from x to z in time t. We start with the estimate on the coupled function difference

$$\begin{aligned} |\,\Phi_t(x,y)\,| &= \left| \int_0^{T-t} \mathbb{E}_x f(X_s) - \mathbb{E}_y f(X_s)\, ds \right| \\ &= \left| \int_0^{T-t} \sum_{z\in\mathbb{Z}^d} f(z)(p_s(x,z) - p_s(y,z))\, ds \right| \\ &\le \sum_z |\, f(z) \,| \left| \int_0^{T-t} p_s(x,z) - p_s(y,z)\, ds \right| \\ &\le \sum_z |\, f(z) \,| \left| \int_0^{T} p_s(x,z) - p_s(y,z)\, ds \right| . \end{aligned}$$

Now, depending on the three cases of f, we proceed differently. First, let $f \in \ell^1$. Then,

$$\begin{aligned}
|\,\Phi_t(x,y)\,| &\leq \sum_z |\,f(z)\,| \left| \int_0^T p_s(x,z) - p_s(y,z)\,ds \right| \\
&\leq \|\,f\,\|_1 \sup_z \left| \int_0^T p_s(x,z) - p_s(y,z)\,ds \right| \\
&= \|\,f\,\|_1 \int_0^T p_s(0,0) - p_s(y-x,0)\,ds \leq C_1 \,\|\,f\,\|_1 \,.
\end{aligned}$$

Since $|\,x-y\,| = 1$, the constant $C_1 = \int_0^\infty p_s(0,0) - p_s(y-x,0)\,ds$ depends on the dimension but nothing else.

Second, let $f \in \ell^\infty$. Then,

$$\begin{aligned}
|\,\Phi_t(x,y)\,| &\leq \sum_z |\,f(z)\,| \left| \int_0^T p_s(x,z) - p_s(y,z)\,ds \right| \\
&\leq \|\,f\,\|_\infty \sum_z \left| \int_0^T p_s(x,z) - p_s(y,z)\,ds \right| \\
&= \|\,f\,\|_\infty \int_0^T \sum_z |\,p_s(x,z) - p_s(y,z)\,|\;ds \\
&= \|\,f\,\|_\infty \int_0^T \frac{1}{2}\,\|\,p_s(x,\cdot) - p_s(y,\cdot)\,\|_{TVar}\;ds \\
&\leq \|\,f\,\|_\infty \int_0^T \widehat{\mathbb{P}}_{x,y}(\tau > s)\,ds
\end{aligned}$$

In the last line, we used the coupling inequality. The coupling $\widehat{\mathbb{P}}_{x,y}$ is the Ornstein coupling, i.e., the different coordinates move independently until they meet. Since x and y are equal in all but one coordinate, the probability of not having succeeded at time t is of order $t^{-\frac{1}{2}}$. Hence we end up with

$$|\,\Phi_t(x,y)\,| \leq C_\infty \,\|\,f\,\|_\infty \sqrt{T}.$$

Third, let $f \in \ell^2$. This is the most interesting case.

Lemma 2.4.2. *Let $x, y \in \mathbb{Z}^d$ be neighbours. Then*

$$\sum_{z\in\mathbb{Z}^d}\left(\int_0^T p_t(x,z)-p_t(y,z)\,dt\right)^2 \leq \alpha(T)$$

with

$$\alpha(T)\in\begin{cases} O(\sqrt{T}), & d=1;\\ O(\log T), & d=2;\\ O(1), & d\geq 3.\end{cases}$$

Proof. By expanding the product and using the fact that $\sum_z p_t(a,z)p_s(b,z) = p_{t+s}(a,b) = p_{t+s}(a-b,0)$, we get

$$\begin{aligned}\sum_{z\in\mathbb{Z}^d}&\left(\int_0^T p_t(x,z)-p_t(y,z)\,dt\right)^2\\ &= 2\int_0^T\int_0^T p_{t+s}(0,0)-p_{t+s}(x-y,0)\,dt\,ds\\ &= 2\int_0^T\int_0^T (-\Delta)p_{t+s}(\cdot,0)(0)\,dt\,ds\\ &= 2\int_0^T p_s(0,0)-p_{T+s}(0,0)\,ds\\ &\leq 2\int_0^T p_s(0,0)\,ds =: \alpha(T).\end{aligned}$$

□

Using first the Cauchy-Schwarz inequality and then Lemma 2.4.2,

$$|\,\Phi_t(x,y)\,|^k \leq \|\,f\,\|_2^k\left(\sum_z\left(\int_0^T p_t(x,z)-p_t(y,z)\,dt\right)^2\right)^{\frac{k}{2}} \leq \|\,f\,\|_2^k\,\alpha(T)^{\frac{k}{2}}.$$

To conclude this example, we finally use the uniform estimates on Φ_t to apply Theorem 2.2.6 and obtain

$$\mathbb{E}_x\exp\left[\int_0^T f(X_t)\,dt-\mathbb{E}_x\int_0^T f(X_t)\,dt\right]\leq\exp\left[T\sum_{k=2}^{\infty}\frac{C_1^k\,\|\,f\,\|_1^k}{k!}\right],$$

$$\mathbb{E}_x \exp\left[\int_0^T f(X_t)\, dt - \mathbb{E}_x \int_0^T f(X_t)\, dt\right] \le \exp\left[T \sum_{k=2}^{\infty} \frac{\| f \|_2^k}{k!} \alpha(T)^{\frac{k}{2}}\right],$$

and

$$\mathbb{E}_x \exp\left[\int_0^T f(X_t)\, dt - \mathbb{E}_x \int_0^T f(X_t)\, dt\right] \le \exp\left[T \sum_{k=2}^{\infty} \frac{C_\infty^k \| f \|_\infty^k}{k!} T^{\frac{k}{2}}\right],$$

for $f \in \ell^1, f \in \ell^2$ and $f \in \ell^\infty$. Since the generator is $Af(x) = \frac{1}{2d} \sum_{y \sim x}(f(y) - f(x))$, we use the estimates $2d$ times and divide by $2d$, so no additional constants appear in the results.

2.5 Application: Simple symmetric exclusion process

This example is somewhat more involved(because of the conservation law), and shows the full power of our approach in the context where classical functional inequalities such as the log-Sobolev inequality do not hold.

The simple symmetric exclusion process is defined via its generator

$$Af(\eta) = \sum_{x \sim y} \frac{1}{2d}(f(\eta^{xy}) - f(\eta)).$$

It is known that the large deviation behaviour of the occupation time of the origin $\int_0^T \eta_t(0)\, dt$ is dependent on the dimension [35]. Its variance is of order $T^{\frac{3}{2}}$ in dimension $d = 1$, $T \log(T)$ in dimension $d = 2$ and T in dimensions $d \ge 3$ [4]. Here we will show the same kind of time dependence for the exponential moments, in dimension $d = 1$ for functionals of any quasi-local function f, and in dimension $d \ge 2$ for the occupation time of a finite set A.

Theorem 2.5.1. *Let $f : \{0,1\}^{\mathbb{Z}} \to \mathbb{R}$ be such that $|\!|\!| f |\!|\!| < \infty$, and fix an initial configuration $\eta_0 \in \{0,1\}^{\mathbb{Z}}$. Then*

$$\mathbb{E}_{\eta_0} \exp\left(\int_0^T f(\eta_t)\, dt - \mathbb{E}_{\eta_0} \int_0^T f(\eta_t)\, dt\right) \le \exp\left[T^{\frac{3}{2}} c_1 \sum_{k=2}^{\infty} \frac{(c_2 |\!|\!| f |\!|\!|)^k}{k!}\right],$$

and the constants $c_1, c_2 > 0$ are independent of f, η_0 and T.

While it is natural to assume the same kind of result in all dimensions(with a properly adjusted dependence on T), we can only prove it in high dimensions($d \geq 5$, see application of Theorem 2.4.1) or for a subset of the local functions, the occupation indicator $H_A(\eta) := \prod_{a \in A} \eta(a)$ of a finite set $A \subset \mathbb{Z}^d$, with a slightly worse dependence on the function(i.e. $|\,A\,|$).

Theorem 2.5.2. *Let $A \subset \mathbb{Z}^d$ be a finite, and fix an initial configuration $\eta_0 \in \{0,1\}^{\mathbb{Z}^d}$. Then, for all $\lambda > 0$,*

$$\mathbb{E}_{\eta_0} \exp\left(\int_0^T \lambda H_A(\eta_t)\, dt - \mathbb{E}_{\eta_0} \int_0^T \lambda H_A(\eta_t)\, dt \right) \leq e^{T\alpha(T) \sum_{k=2}^{\infty} \frac{(c\lambda |\,A\,|^3)^k}{k!}},$$

where $\alpha(T) \in O(T^{\frac{1}{2}}), O(\log T)$ or $O(1)$ in dimensions $d = 1$, $d = 2$ or $d \geq 3$. The constant $c > 0$ is independent of A, η_0 and T, but may depend on the dimension d.

The proofs of Theorems 2.5.1 and 2.5.2 are subject of the two subsections below. For Theorem 2.5.2, we will only look at $d \geq 2$, the case $d = 1$ is contained in Theorem 2.5.1.

2.5.1 Concentration of quasi-local functions in $d = 1$: Proof of Theorem 2.5.1

Let f be a quasi-local function. To derive an exponential estimate, we will create a coupling between the exclusion process started in η and started in η^{xy}:

Proposition 2.5.3. *There exists a coupling $\widehat{\mathbb{P}}_{\eta,\eta^{xy}}$ of $\mathbb{P}_\eta$ and $\mathbb{P}_{\eta^{xy}}$ for which*

$$\widehat{\mathbb{E}}_{\eta,\eta^{xy}} \mathbb{1}_{\eta_t^1(z) \neq \eta_t^2(z)} \leq C\,|\,p_t(x,z) - p_t(y,z)\,|$$

holds for some constant $C > 0$.

Proof. To couple two exclusion processes with almost identical initial conditions, we use a variation of the graphical representation to describe their development, which is the following: at each edge between two consecutive integer numbers, we put an independent Poissonian clock of rate 1, and whenever this clock rings we exchange the occupation status of the sites which are connected by the edge associated to the clock, which is represented by a double sided arrow. Now, to couple $\mathbb{P}_\eta$ with $\mathbb{P}_{\eta^{xy}}$, we instead take Poissonian clocks of rate 2, and additionally a sequence of independent fair coin flips associated to the arrows. For both η^1 and η^2, which use the same arrow configuration, if the coin flip corresponding to an arrow is tails, that arrow is ignored, with one exception explained a bit later. First, we notice that this leads to effective rates of 1. Second, since we start with just two discrepancies(one at x and one at y), those remain the only discrepancies, and they perform independent random walk movements until they encounter the same arrow, which leads us to the only exception of the mechanics described above: When there is an arrow connecting the two discrepancies, the exchange of process η^1 is suppressed if the coin flip is tails, but then η^2 performs the exchange, and if the coin flip is heads, η^1 performs the exchange and η^2 does not. After this event, η^1 and η^2 are identical.

If we denote the position of the discrepancies by X_t and Y_t, those perform independent random walks of rate 1 until they meet, then they stay together. Hence

$$\widehat{\mathbb{E}}_{\eta,\eta^{xy}} \mathbb{1}_{\eta_t^1(z)\neq\eta_t^2(z)} = \widehat{\mathbb{E}}_{x,y} \mathbb{1}_{X_t\neq Y_t, z\in\{X_t,Y_t\}} \leq C\,|\,p_t(x,z) - p_t(y,z)\,|\,,$$

where we used the fact that in dimension 1, the independent coupling of two random walks is optimal and hence

$$\widehat{\mathbb{P}}_{x,y}(X_t = Y_t = z) = p_t(x,z) \wedge p_t(y,z). \qquad \square$$

To apply Theorem 2.2.6, we have to estimate

$$\begin{aligned} L\Phi^k(\cdot,\eta)(\eta) &\leq \sum_{x\in\mathbb{Z}} \sum_{j=\pm 1} \left| \int_0^T S_t f(\eta^{x,x+j}) - S_t f(\eta)\, dt \right|^k \\ &\leq \sum_{x\in\mathbb{Z}} \sum_{j=\pm 1} \left(\int_0^T \sum_z \delta_f(z) \mathbb{E}_{\eta,\eta^{x,x+j}} \mathbb{1}_{\eta_t^1(z)\neq\eta_t^2(z)}\, dt \right)^k \\ &\leq C^k \sum_{x\in\mathbb{Z}} \sum_{j=\pm 1} \left(\int_0^T \sum_z \delta_f(z)\,|\,p_t(x,z) - p_t(y,z)\,|\, dt \right)^k \end{aligned}$$

$$\leq C^k \sum_{x\in\mathbb{Z}} \sum_{j=\pm 1} \int_0^T \sum_z \delta_f(z) \,|\, p_t(x,z) - p_t(y,z) \,|\; dt$$
$$\cdot \left(\sup_{x\in\mathbb{Z}} \sup_{j=\pm 1} \int_0^T \sum_z \delta_f(z) \,|\, p_t(x,z) - p_t(y,z) \,|\; dt \right)^{k-1},$$

where we used Proposition 2.5.3 to obtain the third line. To continue, we calculate

$$\begin{aligned}
&\sum_{x\in\mathbb{Z}} \sum_{j=\pm 1} \int_0^T \sum_z \delta_f(z) \,|\, p_t(x,z) - p_t(x+j,z) \,|\; dt \\
&= \sum_{j=\pm 1} \sum_z \delta_f(z) \int_0^T \|\, p_t(0,\cdot) - p_t(j,\cdot) \,\|_{TVar}\; dt \\
&\leq \tilde{C} \,|||\; f \;|||\, \sqrt{T}.
\end{aligned}$$

Next,

$$\begin{aligned}
&\sup_{x\in\mathbb{Z}} \sup_{j=\pm 1} \int_0^T \sum_z \delta_f(z) \,|\, p_t(x,z) - p_t(x+j,z) \,|\; dt \\
&\leq |||\, f \;|||\, \sup_{j=\pm 1} \sup_z \int_0^T |\, p_t(0,z) - p_t(j,z) \,|\; dt \\
&= |||\, f \;|||\, \sup_{z\geq 0} \int_0^T p_t(0,z) - p_t(-1,z)\, dt.
\end{aligned}$$

In order to control the supremum over z on the right hand side of the last line, let τ_0 denote the first time a simple symmetric random walk $(X_t)_{t\geq 0}$ hits 0. Then

$$\begin{aligned}
&\int_0^T p_t(0,z) - p_t(-1,z)\, dt \\
&\quad = \int_0^T p_t(0,z)\, dt - \mathbb{E}\left[\int_{\tau_0 \wedge T}^T p_{t-\tau_0}(0,z)\, dt \,\middle|\, X_0 = -1 \right] \\
&\quad = \mathbb{E}\left[\int_{T-\tau_0 \wedge T}^T p_t(0,z)\, dt \,\middle|\, X_0 = -1 \right]
\end{aligned}$$

$$\leq \mathbb{E}\left[\int_{T-\tau_0\wedge T}^{T} p_t(0,0)\,dt \middle| X_0 = -1\right]$$
$$= \int_0^T p_t(0,0) - p_t(-1,0)\,dt =: \overline{C} < \infty.$$

Hence

$$L\Phi^k(\cdot,\eta)(\eta) \leq \,|||\, f \,|||^k \sqrt{T} C_1 C_2^k$$

for suitable constants C_1 and C_2, and Theorem 2.2.6 implies

$$\mathbb{E}_\eta \exp\left(\int_0^T f(\eta_t)\,dt - \mathbb{E}_\eta \int_0^T f(\eta_t)\,dt\right) \leq e^{T^{\frac{3}{2}} C_1 \sum_{k=2}^{\infty} \frac{(C_2 |||\, f \,|||)^k}{k!}}$$

for any initial configuration η.

2.5.2 Concentration of the occupation time of a finite set in $d \geq 2$: Proof of Theorem 2.5.2

Now, we want to show that the occupation time of a finite set $A \subset \mathbb{Z}^d$,

$$\int_0^T H_A(\eta_t)\,dt, H_A(\eta) := \prod_{a\in A} \eta(a),$$

has the same time asymptotic behaviour as the occupation time of a single site. As a stating point to estimate $L \mid \Phi_t \mid^k (\cdot,\eta)$, we use the following result of [26]:

Theorem 2.5.4. *[26], Theorem 2.2*

$$\mathbb{E}_\eta \prod_{a\in A} \eta_t(a) - \prod_{a\in A} \rho_t^\eta(a)$$
$$= -\frac{1}{2}\int_0^t ds \sum_{\substack{Z\subset\mathbb{Z}^d \\ |\, Z \,|=|\, A \,|}} \mathbb{P}_A(X_s = Z)$$
$$\cdot \sum_{\substack{z_1,z_2\in Z \\ z_1\neq z_2}} p(z_1,z_2)(\rho_{t-s}^\eta(z_1) - \rho_{t-s}^\eta(z_2))^2 \prod_{\substack{z_3\in Z \\ z_3\neq z_1,z_2}} \rho_{t-s}^\eta(z_3)$$

Here $\mathbb{P}_A(X_s = Z)$ is the probability of exclusion walkers started in A occupying the set Z at time s, and $\rho_t^\eta(z) = \mathbb{E}_\eta \eta_t(z)$ is the occupation probability of z at time t given the initial configuration η.

By using this comparison of exclusion dynamics with independent random walkers, we get

$$\begin{aligned}
&\mathbb{E}_{\eta^{xy}} \prod_{a\in A} \eta_t(a) - \mathbb{E}_\eta \prod_{a\in A} \eta_t(a) \\
&= \mathbb{E}_{\eta^{xy}} \prod_{a\in A} \eta_t(a) - \prod_{a\in A} \rho_t^{\eta^{xy}}(a) + \prod_{a\in A} \rho_t^{\eta^{xy}}(a) \\
&\qquad - \prod_{a\in A} \rho_t^{\eta}(a) + \prod_{a\in A} \rho_t^{\eta}(a) - \mathbb{E}_\eta \prod_{a\in A} \eta_t(a) \\
&= \left(\prod_{a\in A} \rho_t^{\eta^{xy}}(a) - \prod_{a\in A} \rho_t^{\eta}(a) \right) - \frac{1}{2} \int_0^t ds \sum_{\substack{Z\subset \mathbb{Z}^d \\ |Z|=|A|}} \mathbb{P}_A(X_s = Z) \\
&\quad \cdot \sum_{\substack{z_1,z_2\in Z \\ z_1\neq z_2}} p(z_1,z_2) \left[(\rho_{t-s}^{\eta^{xy}}(z_1) - \rho_{t-s}^{\eta^{xy}}(z_2))^2 \prod_{\substack{z_3\in Z \\ z_3\neq z_1,z_2}} \rho_{t-s}^{\eta^{xy}}(z_3) \right. \\
&\quad \left. -(\rho_{t-s}^{\eta}(z_1) - \rho_{t-s}^{\eta}(z_2))^2 \prod_{\substack{z_3\in Z \\ z_3\neq z_1,z_2}} \rho_{t-s}^{\eta}(z_3) \right] .
\end{aligned}$$

Taking absolute values, we start to estimate the first difference:

$$\begin{aligned}
\left| \prod_{a\in A} \rho_t^{\eta^{xy}}(a) - \prod_{a\in A} \rho_t^{\eta}(a) \right| &\leq \sum_{a\in A} \left| \rho_t^{\eta^{xy}}(a) - \rho_t^{\eta}(a) \right| \\
&= \sum_{a\in A} | p_t(x,a) - p_t(y,a) | .
\end{aligned}$$

The next part is the big difference inside the integral. It is estimated by

$$\left| (\rho_{t-s}^{\eta^{xy}}(z_1) - \rho_{t-s}^{\eta^{xy}}(z_2))^2 - (\rho_{t-s}^{\eta}(z_1) - \rho_{t-s}^{\eta}(z_2))^2 \right|$$

$$+ \sum_{\substack{z_3 \in Z \\ z_3 \neq z_1, z_2}} \left| \rho^{\eta^{xy}}_{t-s}(z_3) - \rho^{\eta}_{t-s}(z_3) \right| (\rho^{\eta}_{t-s}(z_1) - \rho^{\eta}_{t-s}(z_2))^2$$

Now we come back to the original task of estimating $L \,|\, \Phi_t \,|^k (\cdot, \eta)$. From now on, multiplicative constants are ignored on a regular basis, which results in an omitted factor of the form $c_1 c_2^k$. However warning is given by using $\lesssim$ instead of $\leq$. By using the above estimates, we obtain the upper bound

$$\sum_{x \in \mathbb{Z}^d} \sum_{y \in \mathbb{Z}^d} p(x,y) \left(\int_0^T \sum_{a \in A} |\, p_t(x,a) - p_t(y,a) \,|\; dt \right)^k \tag{2.6}$$

$$+ \sum_{x \in \mathbb{Z}^d} \sum_{y \in \mathbb{Z}^d} p(x,y) \left(\int_0^T dt \int_0^t ds \sum_{\substack{Z \subset \mathbb{Z}^d \\ |\, Z \,| = |\, A \,|}} \mathbb{P}_A(X_s = Z) \sum_{\substack{z_1, z_2 \in Z \\ z_1 \neq z_2}} p(z_1, z_2) \cdot \right.$$
$$\left. \cdot \left| (\rho^{\eta^{xy}}_{t-s}(z_1) - \rho^{\eta^{xy}}_{t-s}(z_2))^2 - (\rho^{\eta}_{t-s}(z_1) - \rho^{\eta}_{t-s}(z_2))^2 \right| \right)^k \tag{2.7}$$

$$+ \sum_{x \in \mathbb{Z}^d} \sum_{y \in \mathbb{Z}^d} p(x,y) \left(\int_0^T dt \int_0^t ds \sum_{\substack{Z \subset \mathbb{Z}^d \\ |\, Z \,| = |\, A \,|}} \mathbb{P}_A(X_s = Z) \sum_{\substack{z_1, z_2 \in Z \\ z_1 \neq z_2}} p(z_1, z_2) \cdot \right.$$
$$\left. \cdot \sum_{\substack{z_3 \in Z \\ z_3 \neq z_1, z_2}} \left| \rho^{\eta^{xy}}_{t-s}(z_3) - \rho^{\eta}_{t-s}(z_3) \right| (\rho^{\eta}_{t-s}(z_1) - \rho^{\eta}_{t-s}(z_2))^2 \right)^k , \tag{2.8}$$

which we will treat individually.

For term (2.6), we estimate sum over A by the maximum times $|\, A \,|$. Hence

$$(2.6) \leq |\, A \,|^k \sum_{x \in \mathbb{Z}^d} \sum_{y \in \mathbb{Z}^d} p(x,y) \left(\int_0^T |\, p_t(x, a_0) - p_t(y, a_0) \,|\; dt \right)^k .$$

We note that

$$\begin{aligned}
&\sup_{x\in\mathbb{Z},y\sim x}\int_0^T |\, p_t(x,a_0)-p_t(y,a_0)\,|\; dt \\
&\quad\leq \sup_{|\,j\,|=1}\int_0^\infty p_t(0,0)-p_t(j,0)\,dt<\infty
\end{aligned}$$

and

$$\begin{aligned}
&\sum_{x\in\mathbb{Z}^d}\sum_{y\in\mathbb{Z}^d} p(x,y)\left(\int_0^T |\, p_t(x,a_0)-p_t(y,a_0)\,|\; dt\right)^2 \\
&\quad=\frac{1}{2d}\sum_{|\,j\,|=1}\sum_{x\in\mathbb{Z}^d}\left(\int_0^T p_t(x,a_0)-p_t(x,a_0+j)\,dt\right)^2 \\
&\quad\leq \alpha(T)
\end{aligned}$$

by Lemma 2.4.2. Hence

$$(2.6)\lesssim |\,A\,|^k\,\alpha(T).$$

Next, we must treat (2.7). In the case $k=1$,

$$(2.7)\lesssim \int_0^T dt\int_0^t ds\sum_{x\in\mathbb{Z}^d}\sum_{y\sim x}\sum_{z_1\in\mathbb{Z}^d}\sum_{z_2\sim z_1}\left(\sum_{Z:z_1,z_2\in Z}\mathbb{P}_A(X_s=Z)\right) \tag{2.9a}$$

$$\cdot\left|\rho_{t-s}^{\eta^{xy}}(z_1)-\rho_{t-s}^{\eta^{xy}}(z_2)-\rho_{t-s}^{\eta}(z_1)+\rho_{t-s}^{\eta}(z_2)\right| \tag{2.9b}$$

$$\cdot\left|\rho_{t-s}^{\eta^{xy}}(z_1)-\rho_{t-s}^{\eta^{xy}}(z_2)+\rho_{t-s}^{\eta}(z_1)-\rho_{t-s}^{\eta}(z_2)\right|. \tag{2.9c}$$

Regarding the exclusion walkers X_s in (2.9a), we can simplify by using Liggett's correlation inequality ([37], chapter 8):

$$\begin{aligned}
\sum_{Z:z_1,z_2\in Z}\mathbb{P}_A(X_s=Z)&=\mathbb{P}_A(z_1,z_2\in X_s)\leq\mathbb{P}_A(z_1\in X_s)\mathbb{P}_A(z_2\in X_s) \\
&=\left(\sum_{a\in A}p_s(z_1,a)\right)\left(\sum_{a\in A}p_s(z_2,a)\right).
\end{aligned}$$

Lemma 2.5.5. *For* $|\,i\,|, |\,j\,| = 1$,

a) For any η,

$$\left|\,\rho_t^{\eta^{x,x+j}}(z) - \rho_t^{\eta^{x,x+j}}(z+i) - \rho_t^{\eta}(z) + \rho_t^{\eta}(z+i)\,\right| \\ \leq |\,p_t(x,z) - p_t(x+j,z) - p_t(x,z+i) + p_t(x+j,z+i)\,|,$$

b) $\sum_{x\in\mathbb{Z}^d} |\,p_t(x,z) - p_t(x+j,z) - p_t(x,z+i) + p_t(x+j,z+i)\,| \lesssim (1+t)^{-1}$.

Part b) holds as well when we sum over z *instead of* x.

Proof. First we notice that

$$\rho_t^{\eta^{xy}}(z) - \rho_t^{\eta}(z) = \begin{cases} p_t(y,z) - p_t(x,z), & \eta(x)=1, \eta(y)=0; \\ p_t(x,z) - p_t(y,z), & \eta(x)=0, \eta(y)=1; \\ 0, & otherwise, \end{cases}$$

which immediately proves a). To show b),

$$\begin{aligned}
&\sum_{x\in\mathbb{Z}^d} |\,p_t(x,z) - p_t(x+j,z) - p_t(x,z+i) + p_t(x+j,z+i)\,| \\
&= \sum_x \Big|\sum_u p_{t/2}(x,u)p_{t/2}(u,z) - p_{t/2}(x+j,u)p_{t/2}(u,z) \\
&\qquad - p_{t/2}(x,u)p_{t/2}(u,z+i) + p_{t/2}(x+j,u)p_{t/2}(u,z+i)\Big| \\
&\leq \sum_x \sum_u |\,(p_{t/2}(x,u) - p_{t/2}(x+j,u))(p_{t/2}(u,z) - p_{t/2}(u,z+i))\,| \\
&= \sum_u |\,p_{t/2}(u,z) - p_{t/2}(u,z+i)\,| \sum_x |\,p_{t/2}(x,u) - p_{t/2}(x+j,u)\,| \\
&= 4\,\|\,p_{t/2}(0,\cdot) - p_{t/2}(i,\cdot)\,\|_{TVar}\,\|\,p_{t/2}(0,\cdot) - p_{t/2}(j,\cdot)\,\|_{TVar} \\
&\lesssim (1+t/2)^{-\frac{1}{2}}(1+t/2)^{-\frac{1}{2}} \leq 2(1+t)^{-1},
\end{aligned}$$

where the last line relies on optimal coupling of two random walks, see for example [38]. □

As a third observation,

$$\begin{aligned} |\, \rho_t^\eta(z_1) - \rho_t^\eta(z_2)\,| &= \left| \sum_x (p_t(z_1,x) - p_t(z_2,x))\eta(x) \right| \\ &\leq \|\, p_t(z_1,\cdot) - p_t(z_2,\cdot)\,\|_{TVar}\,, \end{aligned} \tag{2.10}$$

which leads to the estimate

$$(2.9\text{c}) \leq 2\,\|\, p_{t-s}(z_1,\cdot) - p_{t-s}(z_2,\cdot)\,\|_{TVar} \lesssim (1+t-s)^{-\frac{1}{2}}.$$

Applying the estimates for (2.9a) to (2.9c), we have (for $k=1$)

$$\begin{aligned} (2.7) &\lesssim \int_0^T dt \int_0^t ds \sum_{z_1\in\mathbb{Z}^d} \sum_{z_2\sim z_1} \sum_{a\in A} p_s(z_1,a) \sum_{a\in A} p_s(z_2,a)(1+t-s)^{-\frac{3}{2}} \\ &\leq 2d\,|\,A\,|^2 \int_0^T dt \int_0^t ds\; p_s(0,0)(1+t-s)^{-\frac{3}{2}} \\ &\lesssim |\,A\,|^2 \int_0^T dt \int_0^t ds\; (1+s)^{-\frac{d}{2}}(1+t-s)^{-\frac{3}{2}} \\ &\lesssim |\,A\,|^2\, \alpha(T), \end{aligned}$$

where the last line is due to the following lemma:

Lemma 2.5.6.

$$\int_0^T dt \int_0^t ds\; (1+s)^{-\frac{n}{2}}(1+t-s)^{-\frac{3}{2}} \lesssim \begin{cases} \sqrt{T}, & n=1; \\ \log(1+T), & n=2; \\ 1, & n\geq 3. \end{cases}$$

Proof. Write

$$f(m,n) := \int_0^t (1+s)^{-\frac{m}{2}}(1+t-s)^{-\frac{n}{2}}\, dt.$$

Then f satisfies $f(m,n) \leq (1+t)^{-\frac{1}{2}}(f(m-1,n)+f(m,n-1))$ for $m,n\geq 1$:

$$f(m,n) = (1+t)^{-\frac{1}{2}} \int_0^t \frac{(1+t)^{\frac{1}{2}}}{(1+s)^{\frac{m}{2}}(1+t-s)^{\frac{n}{2}}}\, dt$$

$$\leq (1+t)^{-\frac{1}{2}} \int_0^t \frac{(1+s)^{\frac{1}{2}} + (1+t-s)^{\frac{1}{2}}}{(1+s)^{\frac{m}{2}}(1+t-s)^{\frac{n}{2}}}\, dt$$
$$= (1+t)^{-\frac{1}{2}}(f(m-1,n) + f(m,n-1)).$$

Also, $f(n,0) = f(0,n) \lesssim (1+t)^{\frac{1}{2}}, \log(1+t)$ or 1 for $n = 1, n = 2$ or $n \geq 3$. Using these two rules we obtain the given estimates. $\square$

As we have already dealt with (2.7) when $k = 1$, we use the simple fact

$$\sum_x h(x)^k \leq (\sum_x h(x))(\sup_x h(x))^{k-1}, \quad h \geq 0,$$

to generalize to any k. However, we must show that (2.7) is bounded by a constant when we replace the sum by the supremum. When we use the same initial estimates as above, we get

$$\sup_{x\in\mathbb{Z}^d} \sup_{y\sim x} \int_0^T dt \int_0^t ds \sum_{z_1\in\mathbb{Z}^d} \sum_{z_2\sim z_1} \left(\sum_{a\in A} p_s(a,z_1)\right)\left(\sum_{a\in A} p_s(a,z_2)\right)$$
$$\cdot \,|\, p_{t-s}(x,z_1) - p_{t-s}(x,z_2) - p_{t-s}(y,z_1) + p_{t-s}(y,z_2)\,|$$
$$\cdot \,\|\, p_{t-s}(z_1,\cdot) - p_{t-s}(z_2,\cdot)\,\|_{TVar},$$

and by taking the sum over z_1 over the p_{t-s} differences,

$$\lesssim \int_0^T dt \int_0^t ds\ (|\,A\,|\, p_s(0,0))^2 (1+t-s)^{-\frac{3}{2}}$$
$$\lesssim |\,A\,|^2 \int_0^T dt \int_0^t ds\ (1+s)^{-d}(1+t-s)^{-\frac{3}{2}} \lesssim |\,A\,|^2 \quad \text{if } d \geq 2.$$

Hence, finally, we have obtained the estimate

$$(2.7) \lesssim |\,A\,|^{2k}\, \alpha(T).$$

Part (2.8) is treated in a similar way:

$$\sum_{\substack{Z\subset\mathbb{Z}^d\\ |\,Z\,|=|\,A\,|}} \mathbb{P}_A(X_s = Z) \sum_{\substack{z_1,z_2\in Z\\ z_1\neq z_2}} p(z_1,z_2) \sum_{\substack{z_3\in Z\\ z_3\neq z_1,z_2}} \left|\rho^{\eta^{xy}}_{t-s}(z_3) - \rho^{\eta}_{t-s}(z_3)\right|$$
$$\cdot\, (\rho^{\eta}_{t-s}(z_1) - \rho^{\eta}_{t-s}(z_2))^2$$

$$\lesssim \sum_{z_1} \sum_{z_2 \sim z_1} \sum_{z_3} \prod_{i=1}^{3} \left(\sum_{a \in A} p_s(a, z_i) \right) |\, p_{t-s}(y, z_3) - p_{t-s}(x, z_3) \,| \cdot \|\, p_{t-s}(z_1, \cdot) - p_{t-s}(z_2, \cdot) \,\|_{TVar}^2$$

By using the fact that

$$\sum_x |\, p_{t-s}(x + j, z) - p_{t-s}(x, z) \,| = 2 \,\|\, p_{t-s}(j, \cdot) - p_{t-s}(0, \cdot) \,\|_{TVar}$$

we can sum over x to obtain another power of the total variation distance. Also,

$$\sum_{z_1} \sum_{z_2 \sim z_1} \sum_{z_3} \prod_{i=1}^{3} \left(\sum_{a \in A} p_s(a, z_i) \right) \leq 2d \,|\, A \,|^3\, p_s(0, 0),$$

hence we obtain the compound estimate

$$|\, A \,|^3\, p_s(0, 0)(1 + t - s)^{-3/2},$$

which after integrating over s and t is again of order $\alpha(T)$. When we take the supremum over x, we can instead take the sum over z_3 on the middle term. Hence we keep another $p_s(0, 0)$ and we get and get

$$|\, A \,|^3\, p_s(0, 0)^2(1 + t - s)^{-3/2},$$

which after integration is of order 1 if $d \geq 2$. Hence,

$$(2.8) \lesssim |\, A \,|^{3k}\, \alpha(T).$$

Returning to the original question,

$$L \,|\, \Phi_t \,|^k\, (\cdot, \eta)(\eta) \lesssim |\, A \,|^k\, \alpha(T) + |\, A \,|^{2k}\, \alpha(T) + |\, A \,|^{3k}\, \alpha(T) \lesssim |\, A \,|^{3k}\, \alpha(T),$$

and after replacing $\lesssim$ with $\leq$,

$$L \,|\, \Phi_t \,|^k\, (\cdot, \eta)(\eta) \leq c_1 c_2^k \,|\, A \,|^{3k}\, \alpha(T).$$

Now that we have this estimate, Theorem 2.2.6 gives us the estimate

$$\mathbb{E}_\eta \exp\left(\int_0^T \lambda H_A(\eta_t)\,dt - \mathbb{E}_\eta \int_0^T \lambda H_A(\eta_t)\,dt\right)$$
$$\leq \exp\left(T\alpha(T)c_1 \sum_{k=2}^{\infty} \frac{(c_2\lambda \mid A \mid^3)^k}{k!}\right),$$

where the constants c_1 and c_2 do not depend on T or A. □

3 Poincaré inequality for Markov random fields via disagreement percolation[1]

3.0 Abstract

We consider Markov random fields of discrete spins on the lattice $\mathbb{Z}^d$. We use a technique of coupling of conditional distributions. If under the coupling the disagreement cluster is "sufficiently" subcritical, then we prove the Poincaré inequality. In the whole subcritical regime, we have a weak Poincaré inequality and corresponding polynomial upper bound for the relaxation of the associated Glauber dynamics.

3.1 Introduction

Concentration inequalities is an active field of research in probability, with applications in other areas of mathematics such as functional analysis, geometry of metric spaces, as well as in more applied areas such as combinatorics, optimization and computer science [36], [42], [25].

Gibbsian random fields on lattice spin systems provide examples of interacting random systems having at the same time non-trivial and natural (e.g. Markovian) dependence structure. They provide a good class of examples where the validity of concentration inequalities in the context of dependent random fields can be tested.

The relation between good mixing properties of Gibbs measures and exponential relaxation to equilibrium of the associated reversible Glauber dynamics

[1] Poincaré inequality for Markov random fields via disagreement percolation
J.-R. Chazottes, F. Redig, F.Völlering
Indagationes Mathematicae 22(3-4), 149-164 (2011)
doi:10.1016/j.indag.2011.09.003

is a thoroughly studied subject. Well-known results in this area were obtained by Aizenman and Holley, [1], Zegarliński [57], Stroock and Zegarliński [49], Martinelli and Olivieri [40]. One of the main results in this area is the equivalence between the log-Sobolev inequality (implying exponential relaxation of the dynamics in L^∞) and the Dobrushin-Shlosman complete analyticity condition.

More recently, a direct relation between the Dobrushin uniqueness condition and Gaussian concentration estimates was proved in [34], and a more general relation between the existence of a coupling of a system of conditional distributions and Gaussian and moment inequalities in [16]. Besides the Dobrushin uniqueness condition, disagreement percolation technique appears here as a basic tool in constructing a good coupling of conditional distributions. The deviation of a function from its expectation is estimated in terms of the sum of the squares of the maximal variation, via martingale difference approach combined with coupling.

So far, no relation has been established between Gaussian concentration estimates or moment estimates (such as the variance inequality) of a Gibbs measure and relaxation properties of the associated reversible Glauber dynamics.

In this chapter we show the correspondence between the existence of a good coupling of conditional distributions and the Poincaré inequality in the context of lattice Ising spin systems. In [18] this was proved in dimension one for a large class of Gibbs measures in the uniqueness regime. The extension to higher dimension which we deal with here (for finite-range potentials) presents new challenges. The Poincaré inequality estimates the variance of a function in terms of the sum of its expected quadratic variations (instead of maximal variation). Therefore, the Poincaré inequality gives much more information. In particular it is equivalent with relaxation of the corresponding reversible Glauber dynamics in L^2. The Poincaré inequality is strictly weaker than the log-Sobolev inequality. So in the complete analyticity regime, the Poincaré inequality is satisfied. A direct proof of the Poincaré inequality in the Dobrushin uniqueness regime can be found in [56].

Our result gives a direct road between "good" coupling of conditional distributions and the Poincaré inequality. By good coupling we mean that if in some region of the space we condition on two configurations that differ only in a single point, then we can couple the unconditioned spins such that the set of sites where we have a discrepancy in the coupling is small. Small here means:

behaving as a subcritical percolation cluster, uniformly in the conditioning. The size of this region of discrepancies can be thought of as the analogue of the "coupling time" for processes. In order to derive the Poincaré inequality, we need the existence of an exponential moment of the disagreement cluster. which corresponds to a non-optimal high-temperature condition (which is e.g. stronger than Dobrushin uniqueness, for the ferromagnetic case).

We want to stress however that the main message of the chapter is the direct link between coupling of conditional distributions and the Poincaré inequality, rather than finding an optimal region of β where the inequality holds.

In case the required exponential moment of the disagreement cluster does not exist, we still obtain the so-called weak Poincaré inequality which gives a polynomial upper bound for the relaxation of the corresponding Glauber dynamics.

This chapter is organized as follows: in section 3.2 we introduce the basic ingredients and discuss coupling via disagreement percolation. In section 3.4 we prove the Poincaré inequality for small β and h close to zero, in section 3.5 we treat the case h large, in section 3.6 we prove the weak Poincaré inequality in the whole subcritical regime.

Acknowledgment. We thank Pierre Collet for fruitful discussions.

3.2 Setting

3.2.1 Configurations

We work in the context of Ising spin systems on a lattice, i.e., with state space $\Omega = \{-1,+1\}^{\mathbb{Z}^d}$ $(d \geq 2)$. Elements of Ω are denoted σ, η, ξ, and are called spin configurations. We fix a "spiraling" enumeration of $\mathbb{Z}^d$

$$\mathbb{Z}^d = \{x_1, x_2, \ldots, x_n, \ldots\}.$$

such that x_{i+1} lies in the exterior boundary of $\{x_1, \ldots, x_i\}$. This enumeration induces an order and lattice intervals like

$$[1,i] = \{x_k, 1 \leq k \leq i\}.$$

We use the notation ξ_i^j, $1 \leq i \leq j \leq \infty$, for a configuration supported on the set $\{x_k, i \leq k \leq j\}$. We denote by $\xi_1^{i-1}+_i$ the concatenation of ξ_1^{i-1} with a 'plus' spin at site x_i. More generally, we write $\xi_V \xi_W$ for the concatenation of a configuration ξ_V supported on V with a configuration ξ_W supported on W.

3.2.2 Functions

For a function $f : \Omega \to \mathbb{R}$ we define the "discrete derivative" in the direction η_x at the configuration η to be

$$\nabla_x f(\eta) = f(\eta^x) - f(\eta),$$

where η^x denotes the configuration obtained from η by "flipping" the spin at site x, i.e., $\eta^x_y = \eta_y$ for all $y \neq x$ and $\eta^x_x = 1 - \eta_x$. For a finite subset $A \subset \mathbb{Z}^d$ we denote by σ^A the configuration obtained from σ by flipping all the spins in A, and

$$\nabla_A f(\sigma) = f(\sigma^A) - f(\sigma).$$

For an enumeration $A = \{y_1, \ldots, y_n\}$ of A, and $x \in A$, we denote by $A_{<x}$ the set of those elements in A preceding x (x not included). For the minimal element $x^* \in A$, in the chosen order of enumeration of A, $A_{<x^*} = \emptyset$ by definition. Elementary telescoping yields the estimate

$$|\nabla_A f(\sigma)| \leq \sum_{x \in A} \left|\nabla_x f(\sigma^{A_{<x}})\right|.$$

Notice that if $A \subset B$ then we have the inequality

$$\sum_{x \in A} |\nabla_x f(\sigma^{A_{<x}})| \leq \sum_{x \in B} |\nabla_x f(\sigma^{B_{<x}})|$$

in an order where we enumerate B by first enumerating A and then the elements of $B \backslash A$.
The variation in direction σ_x is defined as

$$\delta_x f = \sup_{\eta \in \Omega} (f(\eta^x) - f(\eta)).$$

The collection $\{\delta_x f : x \in \mathbb{Z}^d\}$ is denoted by δf, and

$$\|\delta f\|_2^2 = \sum_{x \in \mathbb{Z}^d} (\delta_x f)^2 .$$

3.2.3 Markov random fields

Let $\mathbb{X} = \{X_x, x \in \mathbb{Z}^d\}$ be a Markov random field of "Ising spins", i.e., X_x takes values in $\{-1, +1\}$. In accordance with the previous section, we use the notations X_1^i, X_V, $X_V\xi_W$, etc.
The conditional probabilities of $\mathbb{X}$ are thus given by

$$\mathbb{P}\big(X_x = +1 | X_{\mathbb{Z}^d\setminus x} = \sigma_{\mathbb{Z}^d\setminus x}\big) = \frac{e^{\beta h} e^{\beta J \sum_{y\sim x} \sigma_y}}{2\cosh\left(\beta h + \beta J \sum_{y\sim x} \sigma_y\right)}. \tag{3.1}$$

In this formula $x \sim y$ means that x and y are nearest neighbors, $J \in \mathbb{R}$ is the coupling strength and $h \geq 0$ is interpreted as a uniform magnetic field. Without loss of generality we can assume that $|J| = 1$. The case $J = 1$ is the Ising ferromagnet whereas the case $J = -1$ is the Ising anti-ferromagnet.

An easy consequence of (3.1) is the following uniform bound on the Radon-Nikodym derivative w.r.t. spin-flip:

$$\left\| \frac{\mathrm{d}\mathbb{P}^x}{\mathrm{d}\mathbb{P}} \right\|_\infty \leq e^{2\beta h + 4\beta d} =: e^c \tag{3.2}$$

where $\mathbb{P}^x$ denotes the image measure of $\mathbb{P}$ under spin-flip at lattice site x.
From the previous estimate we deduce that, for a finite subset $A \subset \mathbb{Z}^d$,

$$\left\| \frac{\mathrm{d}\mathbb{P}^A}{\mathrm{d}\mathbb{P}} \right\|_\infty \leq e^{|A|(2\beta h + 4\beta d)}, \tag{3.3}$$

where $\mathbb{P}^A$ is the image measure of $\mathbb{P}$ under simultaneous flips of all the spins in A.

3.2.4 Glauber dynamics

In this section we review some well-known facts about Glauber dynamics. Much more information can be found in [37], chapter 3.

Given a random field $\mathbb{X}$ with distribution $\mathbb{P}$, the natural Glauber dynamics associated to it is a Markovian spin-flip dynamics that flips the spin configuration σ with rate $c(x, \sigma)$ at lattice site x. This is the Markov process $\{\sigma_t : t \geq 0\}$

with generator acting on the core of local functions given by

$$Lf(\sigma) = \sum_{x \in \mathbb{Z}^d} c(x,\sigma) \nabla_x f(\sigma). \tag{3.4}$$

We denote by S_t the associated semigroup generated by L, i.e.,

$$S_t f(\sigma) = \mathbb{E}_\sigma(f(\sigma_t)).$$

The rates $c(x,\sigma)$ are assumed to be local, uniformly bounded away from zero and uniformly bounded from above, i.e., there exist $0 < \delta < M < \infty$ such

$$\delta < c(x,\sigma) < M. \tag{3.5}$$

Moreover, we assume the so-called detailed balance relation between $c(x,\sigma)$ and $\mathbb{P}$ which reads, informally,

$$c(x,\sigma)\mathbb{P}(\sigma) = c(x,\sigma^x)\mathbb{P}(\sigma^x).$$

This is formally rewritten as

$$\frac{c(x,\sigma)}{c(x,\sigma^x)} = \frac{\mathrm{d}\mathbb{P}^x}{\mathrm{d}\mathbb{P}}(\sigma) \tag{3.6}$$

i.e., the lhs of (3.6) is a (and hence the unique) continuous (as a function of σ) version of the Radon-Nikodym derivative of $\mathbb{P}$ w.r.t. spin-flip at site x (i.e., the rhs).

Several choices for the rates are possible, one common choice is the heat-bath dynamics where

$$c(x,\sigma) = \mathbb{P}\big(X_x = -\sigma_x | X_{\mathbb{Z}^d \setminus x} = \sigma_{\mathbb{Z}^d \setminus x}\big).$$

The condition (3.6) ensures that $\mathbb{P}$ is a reversible measure for the Markov process with generator (3.4), i.e., the closure of L is a self-adjoint operator on $L^2(\mathbb{P})$.

The Dirichlet form associated to the rates $c(x,\sigma)$ is given by

$$\mathscr{E}_c(f,f) = 2\langle f(-L)f\rangle = \sum_{x \in \mathbb{Z}^d} \int c(x,\sigma)(\nabla_x f)^2 \, \mathbb{P}(\mathrm{d}\sigma). \tag{3.7}$$

where $\langle\cdot\rangle$ denotes inner product in $L^2(\mathbb{P})$. We say that the Glauber dynamics has a spectral gap if for all f local functions with $\int f \mathrm{d}\mathbb{P} = 0$,

$$\mathscr{E}_c(f,f) \geq \kappa \|f\|_2^2.$$

This implies that the $(-L)$ has simple eigenvalue zero and that the $L^2(\mathbb{P})$ spectrum has κ as a lower bound. This in turn implies the estimate

$$\mathrm{Var}(S_t f) \leq e^{-\kappa t} \|f\|_2^2$$

i.e., exponential relaxation to equilibrium in $L^2(\mathbb{P})$-sense.

Defining the quadratic form

$$\mathscr{E}(f,f) = \sum_{x \in \mathbb{Z}^d} \int (\nabla_x f)^2 \mathrm{d}\mathbb{P}.$$

we have by (3.5) the estimate

$$\delta \mathscr{E}(f,f) \leq \mathscr{E}_c(f,f) \leq M \mathscr{E}(f,f).$$

Hence, estimating the variance of a function in terms of the quadratic form $\mathscr{E}(f,f)$ is equivalent with estimating the variance in terms of the Dirichlet form (3.7) and therefore gives relevant information about the presence of a spectral gap and hence L^2-relaxation properties of the associated Glauber dynamics.

3.2.5 Coupling of conditional probabilities

We write $\mathbb{P}_{\xi_1^i}$ for the conditional distribution of $X_{[i+1,\infty)}$ given $X_1^i = \xi_1^i$.

Remark Notice that we have the same bound (3.3) for the measure $\mathbb{P}_{\xi_1^i}$, when $A \subset [1,i]^c$, uniformly in ξ.

We denote by $\widehat{\mathbb{P}}_{\xi_1^{i-1}+_i, \xi_1^{i-1}-_i}$ a coupling of the distributions $\mathbb{P}_{\xi_1^{i-1}+_i}$ and $\mathbb{P}_{\xi_1^{i-1}-_i}$. This coupling is a distribution of a random field

$$\{(Y_x, Z_x), x \in [i+1,\infty)\} \quad \text{on} \quad \big(\{-1,+1\} \times \{-1,+1\}\big)^{[i+1,\infty)}.$$

Similarly we write $\widehat{\mathbb{P}}_{X_1^{i-1}+_i, X_1^{i-1}-_i}$. We define the random set of discrepancies

$$\mathscr{C}_i = \{x_k : k \geq i, Y_{x_k} \neq Z_{x_k}\}.$$

The distribution of this set depends of course on the choice of the coupling.

The coupling $\widehat{\mathbb{P}}_{\xi_1^{i-1}+_i,\xi_1^{i-1}-_i}$ which we will use throughout this chapter is the one used in [52]. For the sake of self-consistency, we explain here the construction of this coupling.

First we pick a site x_{i+1}^1, with index higher than i, which is a neighbor of x_i. The couple $(Y_{x_{i+1}^1}, Z_{x_{i+1}^1})$ is generated according to the optimal coupling of $\mathbb{P}_{\xi_1^{i-1}+_i}(X_{x_{i+1}^1} = \cdot)$ and $\mathbb{P}_{\xi_1^{i-1}-_i}(X_{x_{i+1}^1} = \cdot)$, i.e., the coupling that maximizes the probability of agreement.

Having generated $(Y_{x_k^{k-i}}, Z_{x_k^{k-i}})$ for $i+1 \leq k \leq j$, either we choose a new lattice point x_{j+1}^{j+1-i} that has a neighbor in the previously generated sites where Y and Z disagree, or if such a point does not exist, then we choose an arbitrary neighbor higher in the order than the previously generated sites, and generate the couple $(Y_{x_{j+1}^{j+1-i}}, Z_{x_{j+1}^{j+1-i}})$ according to the optimal coupling of

$$\mathbb{P}_{\xi_1^{i-1}Y_<}(X_{x_{j+1}^{j+1-i}} = \cdot) \quad \text{and} \quad \mathbb{P}_{\xi_1^{i-1}Z_<}(X_{x_{j+1}} = \cdot).$$

where $Y_<, Z_<$ denote the values already generated before.

By the Markov character of the random field $\mathbb{X}$, the sets of discrepancies $\mathscr{C}_i$ are almost-surely (nearest-neighbor) *connected*. So we can think of the $\mathscr{C}_i$'s as *"percolation clusters"* containing for sure the lattice site x_i, where we have by the conditioning a disagreement. If these clusters behave as sub-critical percolation clusters, then we say that we are in the "good coupling regime", see [52, 27]. We then expect to obtain corresponding good relaxation properties of the natural Glauber dynamics associated to $\mathbb{P}$. The reason to expect this is that in the entire subcritical regime for the disagreement clusters, the corresponding Gibbs measure is unique. In the case of the Ising model in $d = 2$ it is known that in the entire uniqueness regime we have the log-Sobolev inequality, which implies the Poincaré inequality. It is therefore natural to expect that also in higher dimensions, and for arbitrary Markov fields, being in the uniqueness regime implies at least exponential relaxation of the Glauber dynamics in L^2.

3.2.6 Subcritical disagreement percolation

We suppose that, under the coupling $\widehat{\mathbb{P}}_{\xi_1^{i-1}+_i,\xi_1^{i-1}-_i}$, the disagreement clusters $\mathscr{C}_i$ are dominated by independent subcritical site-percolation clusters, uniformly in the conditioning ξ. In fact, we shall need more than subcriticality.

We believe that it is an artefact of our method and that the Poincaré inequality holds in the entire subcritical regime.

We denote by $\mathbb{P}_p$ the distribution of independent site-percolation with parameter $0 \leq p < 1$ and by p_c the corresponding critical value. Let $\mathfrak{C}_i$ be the open cluster containing x_i. In our model (3.1), by the construction of the coupling, we have domination by independent clusters, i.e., for any finite subset $A \subset \mathbb{Z}^d$

$$\sup_i \sup_\xi \widehat{\mathbb{P}}_{\xi_1^{i-1}+_i, \xi_1^{i-1}-_i} (\mathscr{C}_i \supset A) \leq \mathbb{P}_p(\mathfrak{C}_0 \supset A), \tag{3.8}$$

with

$$p = p(\beta, h) = e^{-2\beta h} \ \left(e^{4\beta d} - e^{-4\beta d}\right). \tag{3.9}$$

In particular,

$$\sup_i \sup_\xi \widehat{\mathbb{P}}_{\xi_1^{i-1}+_i, \xi_1^{i-1}-_i} (|\mathscr{C}_i| \geq n) \leq \mathbb{P}_p(|\mathfrak{C}| \geq n),$$

where $\mathfrak{C} = \mathfrak{C}_0$. Our subcriticality assumption reads as follows:

$$\mathbb{E}_p \left(|\mathfrak{C}| e^{c|\mathfrak{C}|}\right) < \infty, \tag{3.10}$$

where c is defined in (3.2). This condition is satisfied for β sufficiently small or h sufficiently large; see below for the precise region of (β, h).
By the uniform bound (3.8), the coupling $\widehat{\mathbb{P}}_{\xi_1^{i-1}+_i, \xi_1^{i-1}-_i}$ can be realized in two stages. Having generated Y_{x_k}, Z_{x_k} for $k = i+1, \ldots, i+n$, we first generate $Y_{x_{i+n+1}}$. Then we flip an independent coin with success probability $1-p$ (corresponding to certain agreement) given by (3.9). Given that we have success, we put $Z_{x_{i+n+1}} = Y_{x_{i+n+1}}$. If we do not have success, then we possibly choose $Z_{x_{i+n+1}} = Y_{x_{i+n+1}}$ or $Z_{x_{i+n+1}} \neq Y_{x_{i+n+1}}$ in order to obtain the correct marginal distributions of the coupling. The crucial point here is that the cluster of failures (=no success), which we denote $\tilde{\mathscr{C}}_i$, is a cluster that, is *independent of* Y and contains the cluster of disagreement $\mathscr{C}_i$. Therefore, in events that depend in a monotone way on the cluster of disagreements $\mathscr{C}_i$, we can replace it by $\mathfrak{C}_i$, the cluster of failures.

3.2.7 Sufficient conditions on β

A sufficient condition for (3.10) to hold is that

$$\sum_{n=0}^{\infty} n\, p^n \big(2d-1\big)^n e^{cn} < \infty,$$

where c is the constant appearing in (3.2) and p is defined in (3.9), and where the factor $n\big(2d-1\big)^n$ arrizes from counting self-avoiding paths. In turn, the above series is finite if

$$e^{4\beta d} - e^{-4\beta d} < \frac{e^{-4\beta d}}{2d-1},$$

which gives

$$\beta < \frac{1}{8d} \log\Big(\frac{2d}{2d-1}\Big). \tag{3.11}$$

Notice that this condition is independent of h and of the sign of J i.e., holds both in the ferromagnetic and the antiferromagnetic case.
For the ferromagnetic case $J = 1$, however, the Dobrushin uniqueness condition reads

$$2d \tanh(\beta) < 1$$

which is weaker. See [27] for more details and a comparison between uniqueness based on disagreement percolation versus Dobrushin uniqueness.

3.3 The Poincaré inequality and related variance inequalities

The general idea of concentration inequalities is to give an estimate of the probability of a deviation event $\{|f - \mathbb{E}(f)| > a\}$, in terms of a quantity that measures the influence on f of variations of the spin configuration at different sites. Usually, such estimates are obtained via Chebychev's inequality, by estimating moments of $|f - \mathbb{E}(f)|$, such as the variance of f, or higher order moments, exponential moments etc., in terms of a norm measuring the variability of f. In this section we concentrate on estimates of the variance.

3.3.1 Uniform variance estimate

The semi-norm

$$\|\delta f\|_2^2 = \sum_{x\in\mathbb{Z}^d} (\delta_x f)^2$$

measures the influence of spin-flips on f in a uniform way, i.e., for each x the worst influence is computed.

The first inequality measures the variance in terms of $\|\delta f\|_2^2$.

Definition 3.3.1. *We say that a random field $\mathbb{X}$ satisfies the uniform variance inequality if there exists $C > 0$, such for all $f : \Omega \to \mathbb{R}$, $f \in L^2(\mathbb{P})$, we have*

$$\mathbb{E}((f - \mathbb{E}(f))^2) \leq C\|\delta f\|_2^2 \tag{3.12}$$

The uniform variance inequality estimates the variance in terms of the rather "rough" norm $\|\delta f\|_2^2$. Surprisingly, it is still a powerful inequality with many useful applications, such as almost-sure central limit theorems, convergence of the empirical distribution in a strong (Kantorovich) distance, etc. See [17] for a list of applications.

Examples where the uniform variance inequality is satisfied include high-temperature Gibbsian random fields (where it follows from the much stronger log-Sobolev inequality) and plus phase of the Ising model at low enough temperatures, see [16].

3.3.2 Poincaré inequality

The quadratic form

$$\mathscr{E}(f,f) = \sum_{x\in\mathbb{Z}^d} \int (\nabla_x f)^2 \mathrm{d}\mathbb{P}$$

measures the influence of spin-flips on f, taking into account the distribution of the spin-configuration, i.e., large differences between $f(\sigma^x)$ and $f(\sigma)$ are weighted less if they correspond to exceptional configurations (in the sense of the measure $\mathbb{P}$). We have the obvious inequality $\mathscr{E}(f,f) \leq \|\delta f\|_2^2$, therefore, estimating the variance in terms of $\mathscr{E}(f,f)$ is clearly better, and, as we will see in examples below, this difference can be substantial.

Definition 3.3.2. *We say that the random field $\mathbb{X}$ satisfies the Poincaré inequality if there exists a constant $C_P > 0$ such that for all $f \in L^2(\mathbb{P})$*

$$\int \big(f - \mathbb{E}(f)\big)^2 \mathrm{d}\mathbb{P} \leq C_P \ \mathscr{E}(f,f) \ . \tag{3.13}$$

The Poincaré inequality is strictly stronger than the uniform variance inequality. Moreover, contrary to the uniform variance estimate, the Poincaré inequality gives exponentially fast decay to equilibrium for the associated Glauber dynamics in $L^2(\mathbb{P})$. Indeed, (3.13) implies

$$\mathrm{Var}(f) \leq \frac{1}{\delta} C_P \mathscr{E}_c(f,f) = 2\langle f, (-L)f\rangle$$

from which one easily sees that $(-L)$ has a spectral gap in $L^2(\mathbb{P})$ of at least $\kappa = 2\delta/C_P$, which implies the relaxation estimate

$$\mathrm{Var}(S_t f) \leq e^{-\kappa t}\|f\|_2^2$$

3.3.3 Weak Poincaré inequality

Finally, the variance can be estimated in terms of a combination of $\mathscr{E}(f,f)$ and another term $\Phi(f)$, where Φ is homogeneous of degree 2, i.e., $\Phi(\lambda f) = \lambda^2\Phi(f)$. Examples are $\Phi(f) = \|f\|_\infty^2$, or $\Phi(f) = \|\delta f\|_2^2$. The idea here is that if the Poincaré inequality does not hold, it can be due to "bad events" which have relatively small probability (e.g. large disagreement clusters). The idea is then to estimate the variance by $\mathscr{E}(f,f)$ on the good configurations and by $\Phi(f)$ on the bad configurations. This leads to the weak Poincaré inequality, initially introduced by Röckner and Wang [46]. This inequality contains enough information to conclude relaxation properties of the associated Glauber dynamics, but now with $\mathrm{Var}(S_t f)$ estimated with a stronger norm than the $L^2(\mathbb{P})$-norm.

Definition 3.3.3. *The measure $\mathbb{P}$ satisfies the weak Poincaré inequality if there exists a decreasing function $\alpha : (0,\infty) \to (0,\infty)$ such that for all bounded measurable functions $f : \Omega \to \mathbb{R}$ we have, for all $r > 0$*

$$\int \big(f - \mathbb{E}(f)\big)^2 \mathrm{d}\mathbb{P} \leq \alpha(r) \ \mathscr{E}(f,f) + r\Phi(f) \ .$$

If we have

$$\Phi(S_t f) \le \Phi(f) \tag{3.14}$$

i.e., if S_t contracts $\Phi(\cdot)$, then we obtain a relaxation estimate for the dynamics from the weak Poincaré inequality. More precisely, in that case, for bounded measurable functions f with $\int f d\mathbb{P} = 0$, the weak Poincaré inequality implies the estimate

$$\mathrm{Var}(S_t f) \le \xi(t) \left(\|f\|_2^2 + \Phi(f)\right)$$

where $\xi(t) \to 0$ as $t \to \infty$ is determined by α:

$$\xi(t) = \inf\left\{ r > 0 : -\frac{1}{\delta}\alpha(r)\log r \le 2t \right\},\ t > 0.$$

where $\delta > 0$ is the lower-bound on the spin-flip rates. In the case when $\alpha(r) \le Cr^{-\kappa}$ for $C, \kappa > 0$, we get $\xi(t) \le \left(1 + \frac{1}{\kappa}\right)^{1+\frac{1}{\kappa}} \left(\frac{2t\delta}{C}\right)^{-\frac{1}{\kappa}}$. We refer the reader to [46] for more background and details.

3.3.4 Examples

Here we illustrate with some simple examples that the Poincaré inequality is much stronger than the uniform variance inequality. The examples are representants of a whole class of functions for which the effect of spin-flip is only "typically small", which gives a good estimate of $\mathscr{E}(f,f)$, but where the uniform variation $\delta_i f$ is always of order one.

Let $d = 1$ and $\mathbb{P}$ be a translation invariant probability measure on configurations $\sigma \in \Omega = \{-1,+1\}^{\mathbb{Z}}$ such that there exists $0 < \theta < 1$ with

$$\mathbb{P}(\sigma_1 = \alpha_1, \ldots, \sigma_n = \alpha_n) \le \theta^n \tag{3.15}$$

for all $n \in \mathbb{N}$, $\alpha_1, \ldots, \alpha_n \in \{-1, 1\}$. Examples of such $\mathbb{P}$ are translation-invariant Gibbs measures.

Consider for $n \in \mathbb{N}, k < n$

$$f_k(\sigma_1, \ldots, \sigma_n) = |\{i \in \{1, \ldots, n-k\} : \sigma_i = \sigma_{i+1} = \cdots = \sigma_{i+k} = +1\}|$$

i.e., the number of lattice intervals of size k, contained in $[1, n]$ and filled with plus spins.

We have

$$\nabla_r f_k(\sigma) = \sum_{\substack{j\in[1,n-k]:\\ r\in[j,j+k]}} (\mathbb{1}\{\sigma_r = -1\} - \mathbb{1}\{\sigma_r = +1\}) \prod_{i\in[j,j+k], i\neq r} \mathbb{1}\{\sigma_i = +1\}$$

which gives

$$\int (\nabla_r f_k)^2 \mathrm{d}\mathbb{P} \leq 2k\theta^k$$

and hence

$$\mathscr{E}(f_k, f_k) \leq 2k(n-k)\theta^k.$$

Therefore, if $\mathbb{P}$ satisfies the Poincaré inequality (e.g. for a large class of Gibbs measures in one dimension in the uniqueness regime, [18]) then

$$\mathrm{Var}(f_k) \leq C_P 2k(n-k)\theta^k$$

Choosing now $k = c\log(n)$, and putting $\theta = e^{-\alpha}$ we find that

$$\mathrm{Var}(f_{c\log n}) \leq 2c\log(n)(n - c\log(n))n^{-\alpha c}.$$

Hence if $\alpha c > 1$, $\mathrm{Var}(f_{c\log n})$ goes to zero as $n \to \infty$. It is immediate from (3.15) that $\alpha > c$ the first moment $\mathbb{E}(f_{c\log(n)})$ converges to zero as $n \to \infty$. Therefore, $\alpha c > 1$ implies that $f_{c\log n}$ converges to zero in $L^2(\mathbb{P})$ (and hence in probability) as $n \to \infty$.

On the other hand, it is clear that $\delta_i(f) = 1$ for all $i = 1, \ldots, n$, therefore the uniform variance estimate gives $\mathrm{Var}(f_k) \leq Cn$, which is not useful here.

One can consider similar quantities like the number of clusters of size k of plus-spins, the number of self-overlaps of size k, etc. Such quantities will have small $\mathscr{E}(f,f)$ (for measures satisfying (3.15)) and large $\|\delta f\|_2^2$.

3.4 Poincaré inequality for the case $h = 0$

We start with the following result.

Theorem 3.4.1. *Consider the Markov random defined in* (3.1) *with* $h = 0$.

For β chosen such that

$$\mathbb{E}_p\left(|\mathfrak{C}|e^{c|\mathfrak{C}|}\right) < \infty,$$

the Poincaré inequality (3.13) *holds.*

In section 3.5 below (Theorem 3.5.1), we will give a complementary result which covers the case of large β and (correspondingly) large h.

Proof. The proof is divided in four steps.
Step 1 (Martingale decomposition).
Let $f : \Omega \to \mathbb{R}$ be a bounded measurable function. Define

$$\Delta_i = \Delta_i(X_1^i) = \mathbb{E}(f|\mathcal{F}_i) - \mathbb{E}(f|\mathcal{F}_{i-1})$$

where $\mathcal{F}_i$ is the sigma-field generated by $\{X_{x_k} : 1 \le k \le i\}$ for $i \ge 1$ and where $\mathcal{F}_0$ is the trivial sigma-field $\{\emptyset, \Omega\}$. Then we have

$$\mathrm{Var}(f) = \sum_{i\in\mathbb{N}} \mathbb{E}(\Delta_i^2).$$

Step 2 (Coupling representation of Δ_i)
We have (using that spins can take only two values)

$$\begin{aligned}
|\Delta_i| &= \left|\int \mathrm{d}\mathbb{P}_{X_1^{i-1}}(\xi_i) \int \mathrm{d}\widehat{\mathbb{P}}_{X_1^i, X_1^{i-1}\xi_i}(\sigma_{i+1}^\infty, \eta_{i+1}^\infty)\right. \\
&\qquad \left.\cdot \left(f(X_1^{i-1}X_i\sigma_{i+1}^\infty) - f(X_1^{i-1}\xi_i\eta_{i+1}^\infty)\right)\right| \qquad (3.16) \\
&\le \int \left|f(X_1^{i-1} +_i \sigma_{i+1}^\infty) - f(X_1^{i-1} -_i \eta_{i+1}^\infty)\right| \mathrm{d}\widehat{\mathbb{P}}_{X_1^{i-1}+_i, X_1^{i-1}-_i}(\sigma_{i+1}^\infty, \eta_{i+1}^\infty) \\
&= \int \left|f(X_1^{i-1} +_i \sigma_{i+1}^\infty) - f(X_1^{i-1} -_i \eta_{i+1}^\infty)\right| \mathrm{d}\widehat{\mathbb{P}}_{X_1^{i-1}+_i, X_1^{i-1}-_i}(\sigma_i^\infty, \eta_i^\infty) \\
&= \sum_{A\ni x_i} \int \mathrm{d}\widehat{\mathbb{P}}_{X_1^{i-1}+_i, X_1^{i-1}-_i}(\sigma_i^\infty, \eta_i^\infty) \times \\
&\qquad \mathbb{1}\{\mathscr{C}_i = A\} \left|f(X_1^{i-1}\eta_A\sigma_{(A\cup[1,i-1])^c}) - f(X_1^{i-1}\sigma_A\sigma_{(A\cup[1,i-1])^c})\right|, \qquad (3.17)
\end{aligned}$$

where $\widehat{\mathbb{P}}_{X_1^{i-1}+_i, X_1^{i-1}-_i}$ is the coupling of conditional probabilities defined in subsection 3.2.5. Notice that the sum over A runs over *finite* connected subsets of $\mathbb{Z}^d$ containing x_i since $\mathscr{C}_i$ is dominated by a subcritical percolation cluster.

In the sequel, we simply write $\sigma_V \xi_W \eta$ for $\sigma_V \xi \eta_{(V\cup W)^c}$ to alleviate notations.

Step 3 (Telescoping and domination by independent clusters).
Start again from (3.17) and telescope the disagreement cluster:

$$\begin{aligned}
|\Delta_i| &\le \int \left|\nabla_{\mathscr{C}_i} f(X_1^{i-1}\sigma^{\mathscr{C}_i})\right| \, \mathrm{d}\widehat{\mathbb{P}}_{X_1^{i-1}+_i, X_1^{i-1}-_i}(\sigma,\eta) \\
&\le \int \sum_{x\in\mathscr{C}_i} \left|\nabla_x f(X_1^{i-1}\sigma^{(\mathscr{C}_i)_{<x}})\right| \, \mathrm{d}\widehat{\mathbb{P}}_{X_1^{i-1}+_i, X_1^{i-1}-_i}(\sigma,\eta) \\
&\le \int \sum_{x\in\tilde{\mathscr{C}}_i} \left|\nabla_x f(X_1^{i-1}\sigma^{(\tilde{\mathscr{C}}_i)_{<x}})\right| \, \mathrm{d}\widehat{\mathbb{P}}_{X_1^{i-1}+_i, X_1^{i-1}-_i}(\sigma,\eta) \\
&= \widetilde{\mathbb{E}} \int \sum_{x\in\tilde{\mathscr{C}}_i} \left|\nabla_x f(X_1^{i-1}\sigma^{(\tilde{\mathscr{C}}_i)_{<x}})\right| \, \mathrm{d}\mathbb{P}_{X_1^{i-1}+_i}(\sigma) \\
&= \sum_{A\ni x_i}\sum_{x\in A} \mathbb{P}_p(\mathfrak{C}_i = A) \int \left|\nabla_x f(X_1^{i-1}\sigma^{A_{<x}})\right| \, \mathrm{d}\mathbb{P}_{X_1^{i-1}+_i}(\sigma).
\end{aligned}$$

In the third inequality the expectation is over the "failure cluster" $\tilde{\mathscr{C}}_i$ only, which is independent of σ. This independence gives the factorization in the last equality, by decomposing over the realization of this cluster (which is finite with $\widehat{\mathbb{P}}_{X_1^{i-1}+_i, X_1^{i-1}-_i}$ probability one under the subcriticality assumption).

Step 4 (Change of measure).
Using now the bound (3.3) and the remark in the beginning of subsection 3.2.5, we further estimate, using

$$|\Delta_i| \le \sum_{A\ni x_i}\sum_{x\in A} \mathbb{P}_p(\mathfrak{C}_i = A)\; e^{c|A|} \int |\nabla_x f(X_1^{i-1} +_i \sigma)| \, \mathrm{d}\mathbb{P}_{X_1^{i-1}+_i}(\sigma)$$

where c is defined in (3.2).
Define the finite number (by the subcriticality assumption (3.10))

$$K := \sum_{A\ni 0} |A| \; \mathbb{P}_p(\mathfrak{C} = A)\; e^{c|A|} = \mathbb{E}_p\big(|\mathfrak{C}| e^{c|\mathfrak{C}|}\big).$$

Then, using the elementary inequality

$$\left(\sum_k a_k b_k\right)^2 \leq \sum_k a_k \sum_k a_k b_k^2 \tag{3.18}$$

for $a_k, b_k \geq 0$, we obtain

$$\begin{aligned}\sum_{i\in\mathbb{N}} \mathbb{E}(\Delta_i^2) &\leq Ke^{2c} \sum_{i\in\mathbb{N}} \sum_{A\ni x_i} \sum_{x\in A} e^{c|A|} \mathbb{P}_p(\mathfrak{C}_i = A) \int (\nabla_x f)^2 \mathrm{d}\mathbb{P} \\ &= K^2 e^{2c} \mathscr{E}(f,f),\end{aligned}$$

where the extra factor e^c arises from removing the plus in the conditioning in $\mathbb{P}_{X_1^{i-1}+_i}$. This finishes the proof of Theorem 3.4.1 □

3.5 Non-zero magnetic field

In this section we show how to prove the Poincaré inequality under a subcricality condition different from Theorem 3.4.1. It is strictly worse in the case $h = 0$ (since it uses Cauchy-Schwarz to seperate the realization of the disagreement cluster from the gradient of f) but can be used for β large and h large, where the condition (3.10) fails.

Theorem 3.5.1. *Suppose that p given in* (3.9) *is such that*

$$\sum_n n(2d-1)^n e^{c'n} \mathbb{P}_p(|\mathfrak{C}| \geq n)^{1/2} < \infty, \tag{3.19}$$

where

$$c' = 4\beta d. \tag{3.20}$$

Then the Poincaré inequality holds.

For (3.19) to hold, it is sufficient that

$$(2d-1)p^{\frac{1}{2}} e^{c'} < 1$$

which gives

$$(2d-1)^2 e^{-2\beta h}(e^{12\beta d} - e^{4\beta d}) < 1.$$

This is satisfied for β small enough or h large enough.

Proof. The telescoping and coupling steps are the same as in the proof of Theorem 1. So we arrive at

$$|\Delta_i| \leq \sum_{A \ni x_i} \sum_{x \in A} \int d\widehat{\mathbb{P}}_{X_1^{i-1}+_i, X_1^{i-1}-_i}(\sigma_i^\infty, \eta_i^\infty) \mathbb{1}\{\mathscr{C}_i = A\} \big|\nabla_x f(X_1^{i-1}\sigma_{A_{<x}}\eta)\big|.$$

Now we use Cauchy-Schwarz inequality to obtain

$$|\Delta_i| \leq \sum_{A \ni x_i} \sum_{x \in A} \Big(\widehat{\mathbb{P}}_{X_1^{i-1}+_i, X_1^{i-1}-_i}\big(\mathscr{C}_i = A\big)\Big)^{1/2} \times \\ \left(\int d\widehat{\mathbb{P}}_{X_1^{i-1}+_i, X_1^{i-1}-_i}(\sigma_i^\infty, \eta_i^\infty)\Big(\nabla_x f(X_1^{i-1}\sigma_{A_{<x}}\eta)\Big)^2\right)^{1/2}. \tag{3.21}$$

Step 4 (Change of measure). In the r.h.s. of (3.21) we integrate over the "composite" configuration $\sigma_{A_{<x}}\eta$ under the coupling $\widehat{\mathbb{P}}_{X_1^{i-1}+_i, X_1^{i-1}-_i}$. To recover the measure $\mathbb{P}$ (see later) we need to replace $\sigma_{A_{<x}}$ by $\eta_{A_{<x}}$. The cost of this replacement is independent of h and is estimated in the following lemma where $\widehat{\mathbb{P}}_{\xi_1^{i-1}+_i, \xi_1^{i-1}-_i}$ is the coupling introduced above.

Lemma 3.5.2. *Let A be a finite subset of $\mathbb{Z}^d$ containing x_i and let $x \in A$. Let $\mathbb{P}_1$ be the distribution of $Z_{A_{<x}} Y_{(A_{<x})^c}$ and $\mathbb{P}_2$ be the distribution of $\{Y_x, x \in \mathbb{Z}^d\}$. Then $\mathbb{P}_1$ is absolutely continuous with respect to $\mathbb{P}_2$ and*

$$\left\|\frac{d\mathbb{P}_1}{d\mathbb{P}_2}\right\|_\infty \leq e^{c'|A|}$$

where c' is defined in (3.20).

Proof. Let $\Lambda \subset \mathbb{Z}^d$ finite, large enough to contain A. We have by construction of the coupling $\widehat{\mathbb{P}}_{\xi_1^{i-1}+_i, \xi_1^{i-1}-_i}$ (see subsection 3.2.5):

$$\frac{\widehat{\mathbb{P}}_{\xi_1^{i-1}+_i, \xi_1^{i-1}-_i}\big(Z_{A_{<x}} = \sigma_{A_{<x}}, Y_{\Lambda \setminus A_{<x}} = \eta_{\Lambda \setminus A_{<x}}\big)}{\widehat{\mathbb{P}}_{\xi_1^{i-1}+_i, \xi_1^{i-1}-_i}\big(Y_{A_{<x}} = \sigma_{A_{<x}}, Y_{\Lambda \setminus A_{<x}} = \eta_{\Lambda \setminus A_{<x}}\big)}$$

$$
\begin{aligned}
&= \sum_{\zeta_{A_{<x}}} \mathbb{P}_{\xi_1^{i-1}-_i}(\sigma_{A_{<x}})\ \mathbb{P}_{\xi_1^{i-1}+_i\zeta_{A_{<x}}}(\eta_{\Lambda\setminus A_{<x}})\times \\
&\qquad\qquad \frac{\widehat{\mathbb{P}}_{\xi_1^{i-1}+_i,\xi_1^{i-1}-_i}\big(Z_{A_{<x}}=\zeta_{A_{<x}}\big|Y_{A_{<x}}=\sigma_{A_{<x}}\big)}{\mathbb{P}_{\xi_1^{i-1}+_i\sigma_{A_{<x}}}\big(\eta_{\Lambda\setminus A_{<x}}\big)} \\
&\le \sup_{\zeta} \frac{\mathbb{P}_{\xi_1^{i-1}+_i\zeta_{A_{<x}}}\big(\eta_{\Lambda\setminus A_{<x}}\big)}{\mathbb{P}_{\xi_1^{i-1}+_i\sigma_{A_{<x}}}\big(\eta_{\Lambda\setminus A_{<x}}\big)} \\
&\le e^{c'|\partial A_{<x}|} \le e^{c'|A|}.
\end{aligned}
$$

We conclude by letting $\Lambda\uparrow\mathbb{Z}^d$. $\square$

Returning to (3.21) and using the preceding lemma we get

$$
\begin{aligned}
|\Delta_i| &\le \sum_{A\ni x_i}\sum_{x\in A}\Big(\widehat{\mathbb{P}}_{X_1^{i-1}+_i,X_1^{i-1}-_i}\big(\mathscr{C}_i=A\big)\Big)^{1/2}\ e^{c'|A|}\ \times \\
&\qquad\qquad\qquad \left(\int \mathrm{d}\mathbb{P}_{X_1^{i-1}-_i}(\eta)\big(\nabla_x f(X_1^{i-1}\eta)\big)^2\right)^{1/2} \\
&\le e^{c}\ \sum_{A\ni x_i}\sum_{x\in A}\Big(\widehat{\mathbb{P}}_{X_1^{i-1}+_i,X_1^{i-1}-_i}\big(\mathscr{C}_i=A\big)\Big)^{1/2}\ e^{c'|A|}\ \times \\
&\qquad\qquad\qquad \left(\int \mathrm{d}\mathbb{P}_{X_1^{i}}(\eta)\big(\nabla_x f(X_1^{i}\eta)\big)^2\right)^{1/2},
\end{aligned}
\tag{3.22}
$$

where for the second inequality we used that, under the measure $\mathbb{P}$, the cost of flip at a single site is bounded by e^c (see (3.2)).

Step 5 (Domination by independent clusters). Using (3.8) we get from (3.22)

$$
\begin{aligned}
|\Delta_i| &\le e^{c}\ \sum_{A\ni x_i}\sum_{x\in A}\big(\mathbb{P}_p(|\mathfrak{C}|\ge|A|)\big)^{1/2}\ e^{c'|A|}\ \times \\
&\qquad\qquad\qquad \left(\int \mathrm{d}\mathbb{P}_{X_1^{i}}(\eta)\big(\nabla_x f(X_1^{i}\eta)\big)^2\right)^{1/2}.
\end{aligned}
\tag{3.23}
$$

Now let

$$
K'=\sum_{A\ni x_i}\sum_{x\in A}\mathbb{P}_p(|\mathfrak{C}|\ge|A|)^{1/2}\ e^{c'|A|}=\sum_{A\ni 0}\ |A|\ \mathbb{P}_p(|\mathfrak{C}|\ge|A|)^{1/2}\ e^{c'|A|}.
$$

By assumption (3.19) K' is finite. Using once more the elementary inequality (3.18) we deduce from (3.23) that

$$\begin{aligned}\sum_i \mathbb{E}(\Delta_i^2) &\leq e^{2c} K' \sum_i \sum_{A \ni x_i} \sum_{x \in A} \mathbb{P}_p\big(|\mathfrak{C}| \geq |A|\big)^{1/2} e^{c'|A|} \int (\nabla_x f)^2 \mathrm{d}\mathbb{P} \\ &= e^{2c} K' \sum_x \left(\int (\nabla_x f)^2 \mathrm{d}\mathbb{P} \right) \sum_{A \ni x} |A| \, \mathbb{P}_p\big(|\mathfrak{C}| \geq |A|\big)^{1/2} e^{c'|A|} \\ &= C_P \sum_x \int (\nabla_x f)^2 \, \mathrm{d}\mathbb{P}\end{aligned}$$

where

$$C_P := e^{2c} K'^2.$$

This finishes the proof of Theorem 3.5.1. □

3.6 Weak Poincaré inequality

If the assumption (3.10) fails, but $p < p_c$ (where p_c denotes the critical value for independent site percolation) then we are still in the uniqueness regime (i.e., the conditional probabilities (3.1) admit a unique Gibbs measure) and expect suitable decay properties of the Glauber dynamics.

We show that in this regime the weak Poincaré inequality holds, which gives polynomial relaxation to equilibrium.

Theorem 3.6.1. *Suppose that p (defined in (3.9)) satisfies $p < p_c$. Then the weak Poincaré inequality is satisfied. Moreover, there exists $C, \kappa > 0$ such that*

$$\alpha(r) \leq C r^{-\kappa}.$$

As a consequence,

$$\mathrm{Var}(S_t f) \leq \left(1 + \frac{1}{\kappa}\right)^{1+\frac{1}{\kappa}} \left(\frac{2t\delta}{C}\right)^{-\frac{1}{\kappa}} \left(\|f\|_2^2 + 4\|f\|_\infty^2\right)$$

where δ is defined in (3.5)*.*

Proof. The proof follows the lines of the proof of Theorem 1, so we sketch where we start to deviate from it: In the estimation of the variance, the contribution involving $\|f\|_\infty^2$ will arise by cutting the cluster of disagreement at some order of magnitude N.

The sum in (3.10) is now possibly infinite, so we define

$$K_N = \sum_{n=0}^{N} n\ e^{cn}\ \mathbb{P}_p\left(|\mathfrak{C}| \geq n\right).$$

Following the line of proof of Theorem 3.4.1, we follow the change of measure road for realizations of the cluster $\mathfrak{C}_i = A$ of cardinality less than or equal to N, and for A with $|A| > N$ we use the uniform estimate

$$\sup_\eta |f(\eta^A) - f(\eta)| \leq \sum_{x\in A} \delta_x f \leq 2|A|\|f\|_\infty$$

Next estimate, using Jensen and the elementary inequality (3.18),

$$\begin{aligned} &\sum_{i\in\mathbb{N}}\Big(\sum_{A\ni x_i, |A|>N} \mathbb{P}_p(\mathfrak{C}_i = A)\sum_{x\in A}(\delta_x f)\Big)^2 \\ \leq\ & 4\left(\mathbb{E}_p(|\mathfrak{C}|^2 \mathbb{1}\{|\mathfrak{C}| > N\})\right)^2 \|f\|_\infty^2. \end{aligned}$$

This gives the inequality

$$\mathrm{Var}(f) \leq 2e^c K_N^2 \mathscr{E}(f,f) + 8\left(\mathbb{E}_p(|\mathfrak{C}|^2 \mathbb{1}\{|\mathfrak{C}| > N\})\right)^2 \|f\|_\infty^2.$$

The constant in front of $\mathscr{E}(f,f)$ blows up at most exponentially in N, i.e., we have the estimate

$$2e^c K_N^2 \leq C_1 e^{aN}$$

where C_1, a are strictly positive and (β, h)-dependent. The constant in front of $\|f\|_\infty^2$ is exponentially small in the whole subcritical regime, by the exponential decay of the cluster size, [29] i.e., we have the estimate

$$2\left(\mathbb{E}_p(|\mathfrak{C}|^2 \mathbb{1}\{|\mathfrak{C}| > N\})\right)^2 \leq C_2 e^{-bN}$$

where C_2, b are strictly positive and (β, h)-dependent. Therefore we can take

$$\alpha(r) \leq C_1 \left(\frac{r}{C_2}\right)^{-\frac{a}{b}}$$

and $\kappa = a/b$. $\square$

4 A Variance Inequality for Glauber dynamics with Application to Low Temperature Ising Model

4.0 Abstract

A variance inequality for spin-flip systems is obtained using comparatively weaker knowledge of relaxation to equilibrium based on a coupling estimate for a single site disturbance. We obtain a variance inequality interpolating between the Poincaré inequality and the uniform variance inequality, as well as a general weak Poincaré inequality. In the case of monotone dynamics the variance inequality is obtained from the decay of the autocorrelation of the spin at the origin, i.e. from that decay we conclude decay for general functions. This method is then applied to the low temperature Ising model, where the time-decay of the auto-correlation of the origin is extended to arbitrary quasi-local functions.

4.1 Introduction

Variance estimates and related inequalities have a long history in the study of interacting particle systems. Classical inequalities are the log-Sobolev inequality or Poincaré's inequality. A basic distinction between various types of estimates is whether they deal with the mixing structure in space, with respect to some measure, or in time, with respect to some dynamics. It is well-established that strong mixing properties in space imply strong mixing properties in time, and vice versa[40, 31]. Often this connection is made via tensorization arguments of the corresponding inequalities.

In [19] it is shown how a different method, disagreement percolation[52], can be used to obtain a Poincaré inequality. The idea used is to track how the influence of a single spin-flip possibly percolates through space, and then use subcriticality of the percolation to obtain results.

However the picture is a lot less clear when only weaker mixing properties hold. One of the few general tools available are weak Poincaré inequalities, which allow to translate a weaker type of spatial mixing to a form of mixing in time.

In this chapter, we approach the problem of mixing in another direction. We go from a restricted form of decay of correlations in time to general decay of correlations in time. The idea is to track the influence of a single spin-flip through time and space. In systems with weak mixing properties typically the influence of such a single flip is limited, but there is the possibility of a big influence, which leads to moment conditions on certain coupling times.

Given that an interacting particle system with nearest-neighbour Glauber dynamics satisfies those coupling conditions we obtain variance estimates for the ergodic measures as well as the relaxation of the dynamics. In the case of attractive dynamics, the coupling condition can be relaxed to a condition on the auto-correlation of the spin at the origin. Using the recent progress in [39] on the low-temperature Ising model we can extend the results to obtain quasi-polynomial relaxation to equilibrium of the Glauber dynamics.

4.2 Definitions and Notation

4.2.1 Setting

We consider the state space $\Omega = \{-1,+1\}^{\mathbb{Z}^d}$. For a function $f : \Omega \to \mathbb{R}$, which is generally assumed to be bounded and measurable, define

$$\nabla_x f(\eta) := f(\eta^x) - f(\eta), \quad \eta \in \Omega, x \in \mathbb{Z}^d,$$

where η^x is the configuration η flipped at x, i.e., $\eta^x(x) = -\eta(x)$ and $\eta^x(y) = \eta(y)$ for $y \neq x$. We call f local if $\nabla_x f = 0$ for all but finitely many $x \in \mathbb{Z}^d$. In addition, we define a family of semi-norms for functions on Ω,

$$|||\, f \,|||_p := \left(\sum_{x \in \mathbb{Z}^d} \sup_{\eta \in \Omega} \left(\nabla_x f(\eta)\right)^p \right)^{\frac{1}{p}}, \quad p \geq 1.$$

A probability measure μ on the space Ω is a called a Markov random field if the probability of observing a plus-spin(or minus-spin) given the spin of all other sites depends only the spin of the nearest neighbours. In terms of a random variable ξ on Ω that means

$$\begin{aligned}&\mu\big(\xi(x) = +1 \,\big|\, \forall\, y \neq x : \xi(y) = \eta(y)\big)\\ &\quad = \mu\big(\xi(x) = +1 \,\big|\, \forall\, y, |\, y - x \,| = 1 : \xi(y) = \eta(y)\big)\end{aligned}$$

for any $\eta \in \Omega$. With this fact in mind, define

$$\begin{aligned}c_+(x, \eta) &= \mu\left(\xi(x) = +1 \,|\, \forall\, y \neq x : \xi(y) = \eta(y)\right);\\ c_-(x, \eta) &= \mu\left(\xi(x) = -1 \,|\, \forall\, y \neq x : \xi(y) = \eta(y)\right) = 1 - c_+(x, \eta).\end{aligned}$$

The conditional probabilities are called translation invariant if $c_+(x, \eta) = c_+(0, \tau_x \eta)$, where $\tau_x \eta(y) = \eta(x + y)$.

A natural dynamics with respect to μ is the Glauber dynamics, where spins at site x flip individually according to some rates $c(x, \eta)$. Here we choose the heat-bath Glauber dynamics, where the flip rates are given by the conditional probabilities c_+, c_-:

$$c(x, \eta) := \begin{cases} c_+(x, \eta), & \eta(x) = -1;\\ c_-(x, \eta), & \eta(x) = +1.\end{cases}$$

The associated Markov process $(\eta_t)_{t \geq 0}$ is then defined via its generator L acting on the core of local functions,

$$Lf(\eta) = \sum_{x \in \mathbb{Z}^d} c(x, \eta) \nabla_x f(\eta).$$

Let $\mathbb{P}_\eta, \eta \in \Omega$, be the path measures on the space of cadlag trajectories and $S_t f(\eta) = \mathbb{E}_\eta f(\eta_t)$ the corresponding semi-group.

4.2.2 Poincaré and uniform variance inequalities

The Dirichlet form $\mathcal{E}$ associated to L is given by

$$\mathcal{E}(f, f) = -2 \int f(\eta) Lf(\eta)\, \mu(d\eta) = \sum_{x \in \mathbb{Z}^d} \int c(x, \eta) (\nabla_x f)^2(\eta)\, \mu(d\eta).$$

A Poincaré inequality is said to hold if for some $K > 0$

$$\mathrm{Var}_\mu(f) \leq K\mathcal{E}(f,f) = K \sum_{x\in\mathbb{Z}^d} \int c(x,\eta)(\nabla_x f)^2(\eta)\,\mu(d\eta) \tag{4.1}$$

hold for all $f \in L^2(\mu)$. The Poincaré inequality is equivalent to a spectral gap of the (self-adjoined) generator L in $L^2(\mu)$ and implies exponential relaxation of the semi-group in $L^2(\mu)$. Under the assumption that $\inf_{\eta\in\Omega} c(\eta,0) > 0$ (4.1) is equivalent to

$$\mathrm{Var}_\mu(f) \leq K' \sum_{x\in\mathbb{Z}^d} \int (\nabla_x f)^2(\eta)\,\mu(d\eta) = K' \sum_{x\in\mathbb{Z}^d} \|\,(\nabla_x f)^2\,\|_{L^2(\mu)}\,. \tag{4.2}$$

A much weaker inequality is the uniform variance inequality

$$\mathrm{Var}_\mu(f) \leq K'' \,|||\, f \,|||_2^2 = K'' \sum_{x\in\mathbb{Z}^d} \|\,(\nabla_x f)^2\,\|_\infty\,. \tag{4.3}$$

To the authors knowledge this inequality is not related to any form of relaxation of the semi-group.

4.2.3 Weak Poincaré inequality

When the Poincaré inequality does not hold ($K = K' = \infty$) but (4.3) is too weak because one still wants to obtain some information about the relaxation speed to equilibrium one can go to other inequalities. One is the so-called weak Poincaré inequality, usually formulated as

$$\mathrm{Var}_\mu(f) \leq \alpha(r)\mathcal{E}(f,f) + r\Phi(f), \quad \mu(f) = 0, r > 0, \tag{4.4}$$

where $\Phi(\lambda f) = \lambda^2\Phi(f), \Phi(f) \in [0,\infty]$, and α is a function decreasing to 0. This implies the following relaxation to equilibrium:

$$\mathrm{Var}_\mu(S_T f) \leq \xi(T)\left(\sup_{t\geq 0}\Phi(S_t f) + \mathrm{Var}_\mu(f)\right)$$

with $\xi(T) = \inf\{r \geq 0 : -\frac{1}{2}\alpha(r)\log(r) \leq T\}$ (see [46]).

4.3 Results and discussion

4.3.1 Main results

Let $\widehat{\mathbb{P}}_{\eta,\xi}$ be the basic coupling (based on the graphical construction, see Section 4.4. See also for example [37]) between two copies of the dynamics starting from the configurations $\eta, \xi \in \Omega$. Set

$$\theta_t(\eta) = c(0,\eta)\widehat{\mathbb{P}}_{\eta^0,\eta}(\eta_t^1 \neq \eta_t^2), \quad t \geq 0. \tag{4.5}$$

For $p \in [1,\infty]$ define the function $D_p : [0,\infty[\to [0,\infty]$ as

$$D_p(T) = \int_T^\infty (t+1)^{2d+2} \parallel \theta_t \parallel_{L^q(\mu)} dt,$$

where $\frac{1}{p} + \frac{1}{q} = 1$.

The function D_p is going to determine the relaxation speed of $S_t f$ for general functions. Note that by definition D_p is decreasing.

Theorem 4.3.1. *Let μ be a translation invariant Markov random field, and S_t the associated heat-bath semi-group. Fix $p \in [1,\infty]$ and assume $D_p(0) < \infty$. For all $f : \Omega \to \mathbb{R}$ with $|||\, f \,|||_2 < \infty$ the following inequality holds:*

$$\mathrm{Var}_\mu(S_T f) \leq C_d D_p(T) \sum_{x\in\mathbb{Z}^d} \left\| (\nabla_x f)^2 \right\|_{L^p(\mu)}. \tag{4.6}$$

Here C_d is a universal constant depending only on the dimension d.

Remark For $T = 0$ we obtain the variance inequality

$$\mathrm{Var}_\mu(f) \leq D_p(0) \sum_{x\in\mathbb{Z}^d} \left\| (\nabla_x f)^2 \right\|_{L^p(\mu)},$$

which interpolates between the Poincaré inequality ($p = 1$) and the uniform variance inequality($p = \infty$).

Remark The factor $t^{2d+2} = t^{2(d+1)}$ appears for two reasons. First the volume of a space-time cone in $\mathbb{Z}^d$ starting at the origin and growing as time progresses is of order t^{d+1}. Then an application of Cauchy-Schwarz's inequality produces the square.

The need to obtain the necessary estimate for Theorem 4.3.1 in terms of a specific coupling instead of some other measure of suitable decay of correlations in time can be an obstacle to applications. If the spin-system is attractive that obstacle can be avoided. By exploiting that attractive spin-systems are well-adapted to the coupling $\widehat{\mathbb{P}}$ we reduce the dependence on the coupling to the auto-correlation of the origin:

Theorem 4.3.2. *Assume that the spin-system is attractive. Let* $\varphi(t) := \mathrm{Var}_\mu(S_t g)$, $g(\eta) = \eta(0)$, *be the auto-correlation of the spin at the origin. Then the function* D_p *can be estimated by*

$$D_p(T) \leq C'_d \int_T^\infty (t+1)^{3d+2} \left(\varphi(t)\right)^{\frac{p-1}{4p}} \, dt,$$

with a dimension dependent constant $C'_d > 0$.

A good example where this result can be applied is the two-dimensional low-temperature Ising model. Recently in [39] the estimate

$$\mathrm{Var}_{\mu^+}(S_t g) \leq \exp\left(-e^{c(\beta)\sqrt{\log(t+1)}}\right)$$

was obtained, with $c(\beta)$ some temperature dependent constant. Combining this with Theorem 4.3.2 gives a variance estimate for general functions.

Corollary 4.3.3. *Fix* $p > 1$. *Let* $\widetilde{D}_p : [0,\infty[\to [0,\infty[$ *be given by*

$$\widetilde{D}_p(T) = c_p(\beta) \int_T^\infty \exp\left(8\log(t+1) - \frac{p-1}{4p} e^{c(\beta)\sqrt{\log(t+1)}}\right) dt.$$

For all $f : \Omega \to \mathbb{R}$ *the relaxation of the semi-group in the plus-phase is estimated by,*

$$\mathrm{Var}_{\mu^+}(S_T f) \leq \widetilde{D}_p(T) \sum_{x \in \mathbb{Z}^d} \left\| (\nabla_x f)^2 \right\|_{L^p(\mu)}.$$

4.3.2 Discussion

If $D_1(0) < \infty$ Theorem 4.3.1 implies a Poincaré inequality, and hence exponentially fast convergence to equilibrium. If $\| \theta_t \|_{L^q(\mu)}$ decays exponentially fast Theorem 4.3.1 still implies exponentially fast decay of the variance, but with respect to a stronger norm. This, however, is sufficient to prove a spectral gap of the generator L or, equivalently, a Poincaré inequality.

Proposition 4.3.4. *Suppose* $\int_0^\infty \| \theta_t \|_{L^q(\mu)} e^{\lambda t}\, dt < \infty$ *for some* $\lambda > 0$ *and* $1 \le q \le \infty$. *Then* $]-\lambda/2, 0[$ *belongs to the resolvent set of* L.

In fact, we can say even more about the connection between $\| \theta_t \|_{L^q(\mu)}$ and the Poincaré inequality.

Proposition 4.3.5. *Suppose the spin system is attractive and* $\inf_{\eta\in\Omega} c(\eta, 0) > 0$. *If the spin system satisfies the Poincaré inequality, then* $\| \theta_t \|_{L^q(\mu)}$ *decays exponentially fast for any* $1 \le q < \infty$.

This shows that for attractive spin systems equivalence between exponential decay of $\| \theta_t \|_{L^q(\mu)}, 1 \le < \infty$, and the existence of a spectral gap.

It is then natural to ask if even stronger inequalities are implied by fast decay of $\| \theta_t \|_{L^\infty(\mu)}$. This is indeed the case:

Proposition 4.3.6. *Set* $\widehat{D}(T) := \int_T^\infty (t+1)^{d+1} \| \theta_t \|_{L^\infty(\mu)}^{\frac{1}{2}}\, dt$ *and suppose* $\widehat{D}(0) < \infty$. *Then for all* $f : \Omega \to \mathbb{R}$ *with* $|||\, f \,|||_1 < \infty$ *and all* $T \ge 0$

$$\| S_T f - \mu(f) \|_\infty \le C_d^{\frac{1}{2}} \widehat{D}(T)\, |||\ f\ |||_1 .$$

If $D_1(0) = \infty$, but $D_p(0) < \infty$ for some $p > 1$ (but with sub-exponential decay), it is natural to compare Theorem 4.3.1 with a weak Poincaré inequality. Under essentially the same conditions, we can prove the following weak Poincaré inequality.

Proposition 4.3.7. *Assume the conditions of Theorem 4.3.1 and* $c(\eta, x) \ge \delta > 0$. *Then for all* $f : \Omega \to \mathbb{R}$ *with* $\sum_{x\in\mathbb{Z}} \left\| (\nabla_x f)^2 \right\|_{L^p(\mu)}$ *and all* $t \ge 0$,

$$\mathrm{Var}_\mu(S_t f) \le C_d \delta^{-1} D_1(0, R)\mathcal{E}(f, f) + D_p(R)\Phi_R(f),$$

where

$$D_p(0, R) = \int_0^R (t+1)^{2d+2} \| \theta_t \|_{L^p(\mu)},$$

$$\Phi_R(f) = \frac{\mathrm{Var}_\mu(S_R f)}{D_P(R)} \le C_d \sum_{x\in\mathbb{Z}^d} \left\| (\nabla_x f)^2 \right\|_{L^p(\mu)}.$$

This weak Poincaré inequality leads (with a minor modification of the proof in [46]) to

$$\mathrm{Var}_\mu(S_T f) \leq \xi(T) \left(C_d \sum_{x \in \mathbb{Z}^d} \| (\nabla_x f)^2 \|_{L^p(\mu)} + \mathrm{Var}_\mu(f) \right).$$

The decay $\xi(T)$ is of order $D_p(T^{\frac{1}{2d+3}})$, which is worse than the one from Theorem 4.3.1. The reason for that is that in the weak Poincaré inequality the diverging $D_1(0,R)$ is partially used, while in Theorem 4.3.1 only the converging $D_p(0,R)$ is used.

4.4 Graphical construction

The graphical construction of the Glauber heat bath dynamics is the encoding of the random evolution of the process η_t into basic random components and a deterministic function of this randomness and the initial configuration. It is a well-known tool in the study of spin and particle systems.

Let $\overline{N}$ be a Poisson point process on $\mathbb{Z}^d \times [0,\infty[$ with intensity one(wrt. the counting measure on $\mathbb{Z}^d$ and the Lebesgue measure on $[0,\infty[$). A point $(x,t) \in \overline{N}$ represents a *chance* of flipping the spin at site x and time t. To realize this chance let $\overline{U} = (\overline{U}_n)_{n \in \mathbb{N}}$ be a countable iid. collection of $[0,1]$-uniform random variables independent of $\overline{N}$. We assume that to each $(x,t) \in \overline{N}$ there is an associated U from $\overline{U}$ (which can be realized by a bijection from $\overline{N}$ to $\mathbb{N}$, and we simply write $\overline{U} : \overline{N} \to [0,1]$). We denote the expectation with respect to $\overline{N}$ and $\overline{U}$ by $\int d\overline{N}$ and $\int d\overline{U}$.

The elementary step is then as follows. Given the configuration η_{t-} before a possible flip at $(x,t) \in \overline{N}$ and the to (x,t) associated random variable $U = \overline{U}((x,t))$ we determine the configuration η_t after the possible flip deterministically. All sites $y \in \mathbb{Z}^d, y \neq x$, are unchanged, i.e., $\eta_t(y) = \eta_{t-}(y)$. If $U < c_+(x, \eta_{t-})$, then $\eta_t(x) = +1$, otherwise $\eta_t(x) = -1$. Since we ignore the original spin at x and simply replace it with a new one drawn according to conditional probability given the other spins we call this a *resampling event*.

The configuration η_t is then given by the successive application of all resampling events to the initial configuration η_0. As those are infinitely many steps one has to take care that this is indeed well-defined. The goal is to define a deterministic function Ψ which will output the configuration at time t, η_t, given the inputs $\overline{N}, \overline{U}$ and η_0. We now focus on the precise construction of the graphical representation and its properties.

For a single resampling event the definition of Ψ is simple. Let $\Psi : \Omega \times (\mathbb{Z}^d \times [0,\infty[\times[0,1]) \to \Omega$ be given by

$$\Psi(\eta,(x,t,u))(y) := \begin{cases} +1, & y = x, c_+(x,\eta) \leq u; \\ -1, & y = x, c_-(x,\eta) > u; \\ \eta(y), & y \neq x. \end{cases}$$

This definition is directly extended recursively to a finite number of resampling events. For $(x_n,t_n,u_n)_{1\leq n\leq N} \subset \mathbb{Z}^d \times [0,\infty[\times[0,1]$ with $t_1 < t_2 < ... < t_N$,

$$\Psi\left(\eta,(x_n,t_n,u_n)_{1\leq n\leq N}\right) := \Psi\left(\Psi\left(\eta,(x_1,t_1,u_1)\right),(x_n,t_n,u_n)_{2\leq n\leq N}\right),$$

and $\Psi(\eta,\emptyset) = \eta$.

Definition 4.4.1. *Let G be a countable subset of $\mathbb{Z}^d \times [0,\infty[$.*

a) A partial order $<_G$ on $\mathbb{Z}^d \times [0,\infty[$ is defined as follows: $(x,t) <_G (y,s)$ iff either $x = y$ and $t < s$ or there exists a finite subset $\{(x_1,t_1),\ldots(x_K,t_K)\} \subset G$ such that $t < t_1 < t_2 < \ldots < t_K \leq s$ and $|\,x_m - x_{m-1}\,| = 1$, $2 \leq m \leq K$, as well as $|\,x_1 - x\,| = 1$ and $x_K = y$.

b) Write $T_x := \sup\{t : (x,t) \in G\}, x \in \mathbb{Z}^d$, and $G_{<x} := \{(y,t) \in G : (y,t) \leq_G (x,T_x)\}$. We call G locally finite*, if $|\,G_{<x}\,| < \infty$ for all $x \in \mathbb{Z}^d$.*

c) For G^U a countable subset of $\mathbb{Z}^d \times [0,\infty[\times[0,1]$ the definitions a) and b) are copied in the canonical way(projection of G^U onto $\mathbb{Z}^d \times [0,\infty[$).

The purpose of this definition becomes transparent by the following fact.

Lemma 4.4.2. *For any $G^U \subset \mathbb{Z}^d \times [0,\infty[\times[0,1]$ finite, $x \in \mathbb{Z}^d$ and $\eta \in \Omega$,*

$$\Psi(\eta,G^U)(x) = \Psi(\eta,G^U_{<x})(x).$$

Proof. The nearest-neighbour property of c_+ means that to determine the new spin after a resampling event (x,t) it is sufficient to know the spin value of the neighbours of x. Those might depend on earlier resampling events, which have again nearest neighbour dependencies, and all resampling events (y, s) which have an influence on (x, t) satisfy $(y, s) <_G (x, t)$. □

This leads is to the final definition of Ψ. For G^U a locally finite subset of $\mathbb{Z}^d \times [0, \infty[\times[0, 1]$ (or $G \subset \mathbb{Z}^d \times [0, \infty[, U : G \to [0, 1], G^U := \{(x, t, U(x, t)) : (x, t) \in G\}$),

$$\Psi(\eta, G^U)(x) := \Psi(\eta, G^U_{<x})(x), \quad x \in \mathbb{Z}^d.$$

An important property of the graphical construction evident here is that Ψ is tolerant to certain changes in the order of resampling events. Intuitively, a resampling event (x, t) is influenced only by resampling events which happen before t and are not too distant from x. This intuition can be formalized via the ordering $>_G$, which we now do.

Lemma 4.4.3. *Let $G^U \subset \mathbb{Z}^d \times [0, \infty[\times[0, 1]$ be locally finite and $A, B \subset G^U$ a decomposition of G^U such that $\forall\, (x_1, t_1, u_1) \in A, (x_2, t_2, u_2) \in B : (x_1, t_1) \not>_G (x_2, t_2)$. In words, A does not happen after B. Then*

$$\Psi(\eta, G^U) = \Psi\left(\Psi(\eta, A), B\right).$$

Proof. Assume G^U is finite. If not, restrict to $G^U_{<x}$.

The proof is a consequence from the following basic fact. For $(x_i, t_i, u_i) \in \mathbb{Z}^d \times [0, \infty[\times[0, 1], i = 1, 2$, with $|\, x_1 - x_2 \,| > 1$,

$$\Psi(\eta, \{(x_1, t_1, u_1), (x_2, t_2, u_2)\}) = \Psi(\Psi(\eta, (x_1, t_1, u_1)), (x_2, t_2, u_2)). \tag{4.7}$$

By the property of the decomposition for each $(x_1, t_1, u_1) \in A, (x_2, t_2, u_2) \in B$, either $t_1 < t_2$ or $|\, x_1 - x_2 \,| > 1$. The proof of the lemma is an iterative application of fact (4.7). Let $a_i, i = 1..|\, A \,|$ be the elements of A ordered in increasing time. Starting from $\Psi(\eta, A \cup B) = \Psi(\Psi(\eta, \emptyset), \{a_i : i = 1, ..., |\, A \,|\} \cup B)$, we can use fact (4.7) to move a_1 past all resampling events in B and perform this resampling event first:

$$\Psi(\eta, A \cup B) = \Psi(\Psi(\eta, \{a_1\}), \{a_i : i = 2, ..., |\, A \,|\} \cup B).$$

Repeating this procedure for all other elements of A in their time-order then proves the claim of the lemma. □

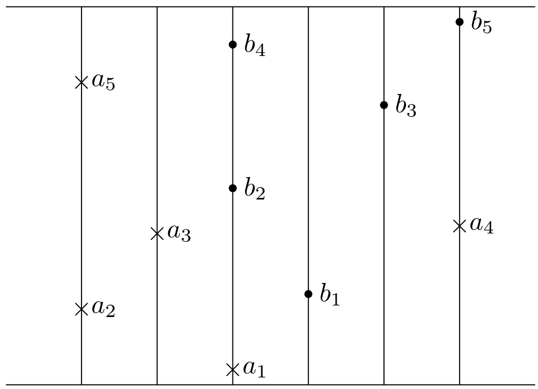

Figure 4.1: Resampling events $a_1, ..., a_5$ do not depend on $b_1, ..., b_5$.

The final proposition of this section sums up the properties of the graphical representation.

Proposition 4.4.4. *Let $f : \Omega \to \mathbb{R}$ be quasi-local. The function Ψ has the following properties:*

a) $\int \int f \left(\Psi(\eta, \overline{N}_t^{\overline{U}}) \right) d\overline{U} d\overline{N} = S_t f(\eta)$ *, where* $\overline{N}_t^{\overline{U}} = \{(x,s,u) \in \overline{N}^{\overline{U}} : s \leq t\}$*;*

b) For any locally finite $G \subset \mathbb{Z}^d \times [0, \infty[$,

$$\int \int f(\Psi(\eta, G^{\overline{U}}))\, d\overline{U} \mu(d\eta) = \int f(\eta)\, \mu(d\eta);$$

c) For $\eta^1, \eta^2 \in \Omega$ the coupling $\widehat{\mathbb{P}}_{\eta^1,\eta^2}$ of $\mathbb{P}_{\eta^1}$ and $\mathbb{P}_{\eta^2}$ is defined via

$$\widehat{\mathbb{E}}_{\eta^1,\eta^2} f(\eta_t^1, \eta_t^2) = \int \int f \left(\Psi(\eta^1, \overline{N}_t^{\overline{U}}), \Psi(\eta^2, \overline{N}_t^{\overline{U}}) \right) d\overline{U} d\overline{N}.$$

Proof. a) The point process $(\overline{N}_t^{\overline{U}})_{t \geq 0}$ is a Markov process on the subsets of $\mathbb{Z}^d \times [0, \infty[\times [0,1]$ under $d\overline{U}\, d\overline{N}$ and with respect to the canonical filtration.

The image process

$$\tilde{\eta}_t := \Psi(\eta, \overline{N}_t^{\overline{U}})$$

is also a Markov process since Ψ preserves the Markov property:

$$\tilde{\eta}_t = \Psi\left(\eta, \overline{N}_t^{\overline{U}}\right) = \Psi\left(\Psi(\eta, \overline{N}_s^{\overline{U}}), \overline{N}_t^{\overline{U}} \backslash \overline{N}_s^{\overline{U}}\right) = \Psi\left(\eta_s, \overline{N}_t^{\overline{U}} \backslash \overline{N}_s^{\overline{U}}\right), \quad t > s \geq 0.$$

The generator of $\tilde{\eta}_t$ is

$$\widetilde{L} f(\eta) = \sum_{x \in \mathbb{Z}^d} \int_0^1 f\left(\Psi(\eta, (x, 0, u))\right) - f(\eta)\, du. \tag{4.8}$$

Since $\Psi(\eta, (x, 0, u))$ is either η or η^x, after integrating over u we obtain $\widetilde{L} f = Lf$ on the core of local functions $f : \Omega \to \mathbb{R}$.

b) The proof follows the construction of Ψ. Let $G = \{(x, t)\}$ and write $\eta_+^x(x) = +1$, $\eta_+^x(y) = \eta(y)$ for $y \neq x$ (η_-^x analogue). Then

$$\begin{aligned}
\int \int f\left(\Psi(\eta, G^{\overline{U}})\right) d\overline{U}\, \mu(d\eta) &= \int \int_0^1 f\left(\Psi(\eta, (x, t, u))\right) du\, \mu(d\eta) \\
&= \int c_+(x, \eta) f(\eta_+^x) + c_-(x, \eta) f(\eta_-^x)\, \mu(d\eta) \\
&= \int f(\eta)\, \mu(d\eta).
\end{aligned}$$

For G a finite set the result is true by the iterative construction. For G countable but locally finite we observe that for local f only finitely many resampling steps have to be performed to determine the expectation of f.

c) By part a) $\widehat{\mathbb{E}}_{\eta^1, \eta^2} f(\eta_t^1) = S_t f(\eta^1)$ and $\widehat{\mathbb{E}}_{\eta^1, \eta^2} f(\eta_t^2) = S_t f(\eta^2)$, so $\widehat{\mathbb{P}}_{\eta^1, \eta^2}$ is indeed a coupling.

□

4.5 Proofs of the results

The first step is to rewrite the variance. As the following formula holds fairly generally and not just in this setting we formulate the lemma with more abstract conditions.

Lemma 4.5.1. *Let μ be an ergodic measure wrt. S_t and $f : \Omega \to \mathbb{R}$ such that $S_t f, (S_t f)^2 \in dom(L)$. Then, for $0 \leq T < S \leq \infty$,*

$$\begin{aligned}
&\mathrm{Var}_\mu(S_T f) - \mathrm{Var}_\mu(S_S f) && (4.9)\\
&= \int_T^S \int \left[L(S_t f - S_t f(\eta))^2\right](\eta)\,\mu(d\eta)\,dt && (4.10)\\
&= \int_T^S \int \sum_{x\in\mathbb{Z}^d} c(x,\eta)\left(S_t f(\eta^x) - S_t f(\eta)\right)^2\,\mu(d\eta)\,dt. && (4.11)
\end{aligned}$$

Note that by ergodicity $\lim_{S\to\infty} \mathrm{Var}_\mu(S_S f) = 0$.

Proof. Since

$$\frac{d}{dt}\mathrm{Var}_\mu(S_t f) = \int 2 S_t f(\eta) L S_t f(\eta)\,\mu(d\eta),$$

we can express the variance as

$$\mathrm{Var}_\mu(S_T f) - \mathrm{Var}_\mu(S_S f) = \int_T^S \int -2 S_t f(\eta) L S_t f(\eta)\,\mu(d\eta)\,dt.$$

By stationarity, $\int [L(S_t f)^2](\eta)\,\mu(d\eta) = 0$, hence

$$\begin{aligned}
&\mathrm{Var}_\mu(S_T f) - \mathrm{Var}_\mu(S_S f)\\
&= \int_T^S \int [L(S_t f)^2](\eta) - 2 S_t f(\eta) L S_t f(\eta)\,\mu(d\eta)\,dt\\
&= \int_T^S \int [L(S_t f - S_t f(\eta))^2](\eta)\,\mu(d\eta)\,dt.
\end{aligned}$$

□

Note that $|||\, f \,|||_1 < \infty$ implies both $S_t f \in dom(L)$ and $(S_t f)^2 \in dom(L)$ in the setting of Glauber dynamics([37]).

The idea of the proof of Theorem 4.3.1 is to rewrite (4.9) using the graphical representation to describe the semi-group S_t. Then various applications of Hölder's inequality are used to separate different parts contributing to the variance formulation (4.9). However the calculation is fairly sensitive to the order in which different aspects are treated, and has one crucial non-trivial use of the graphical construction on the infinite volume.

We start by looking how the graphical construction can be used in light of Lemma 4.5.1. Let, by slight abuse of notation, $\overline{N} \subset \mathbb{Z}^d \times [0, \infty[$ be a fixed realization of the Poisson point process on $\mathbb{Z}^d \times [0, \infty[$, the set of resampling events. Almost surely this is a locally finite subset of $\mathbb{Z}^d \times]0, \infty[$. We denote all resampling events up to time t by $\overline{N}_t := \{(y, s) \in \overline{N} \, : \, s \leq t\}$.

To determine what influence a flip at site x has on the configuration at time t we use the graphical construction, particularly the partial order introduced in definition 4.4.1. Given the fixed realization $\overline{N}$, the cone

$$C_{t,x} := \{(y, s) \in \overline{N}_t : (y, s) >_{\overline{N}} (x, 0)\}$$

contains all resampling events which depend on the value of the initial configuration at site x, see also figure 4.2. Motivated by (4.11) we also introduce the

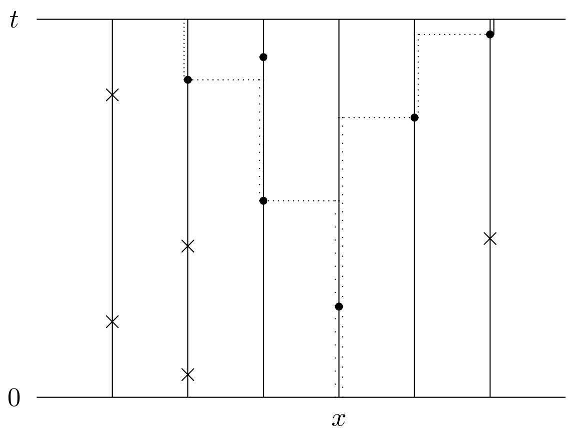

Figure 4.2: The cone $C_{t,x}$ containing all resampling events depending on $(0, x)$.

same cone with another resampling event added at site x and time 0:

$$\widetilde{C}_{t,x} := C_{t,x} \cup \{x, 0\}.$$

Given a realization of the independent uniform $[0, 1]$ variables associated to the resampling events, $\overline{U} : \overline{N} \to [0, 1]$, we extend the above sets to

$$\overline{N}_t^{\overline{U}} := \{(y, s, \overline{U}((y, s))) : (y, s) \in \overline{N}_t\};$$

$$C_{t,x}^{\overline{U}} := \{(y,s,\overline{U}((y,s))) : (y,s) \in C_{t,x}\}.$$

In the case of the added resampling event at $(x,0)$ we assume a given $u \in [0,1]$ to extend the event to $(x,0,u)$. This leads to

$$\begin{aligned}
\widetilde{\overline{N}}_t^{\overline{U}} &:= \overline{N}_t^{\overline{U}} \cup \{(x,0,u)\},\\
\widetilde{C}_{t,x}^{\overline{U}} &:= C_{t,x}^{\overline{U}} \cup \{(x,0,u)\},
\end{aligned}$$

and, from $\eta \in \Omega$,

$$\widetilde{\eta} := \Psi(\eta, (x,0,u)).$$

Now we are ready to formulate the crucial idea. We want to compare the evolution of two configurations η_t^1, η_t^2 under the graphical construction coupling when started from two initial configurations $\eta, \widetilde{\eta}$. By the graphical construction,

$$\begin{aligned}
\eta_t^1 &= \Psi(\eta, \overline{N}_t^{\overline{U}}),\\
\eta_t^2 &= \Psi(\widetilde{\eta}, \overline{N}_t^{\overline{U}}) = \Psi(\eta, \widetilde{\overline{N}}_t^{\overline{U}}).
\end{aligned}$$

By the reordering principle of the graphical construction in Lemma 4.4.3,

$$\begin{aligned}
\eta_t^1 &= \Psi(\xi, C_{t,x}^{\overline{U}}), \qquad (4.12)\\
\xi &= \Psi(\eta, \overline{N}_t^{\overline{U}} \backslash C_{t,x}^{\overline{U}}).
\end{aligned}$$

Similarly,

$$\eta_t^2 = \Psi(\xi, \widetilde{C}_{t,x}^{\overline{U}}). \qquad (4.13)$$

So we can see ξ as a common ancestor of η_t^1 and η_t^2 in terms of the graphical construction (it is not an ancestor in time). This is very important, as both configurations only differ from ξ by a finite number of resampling events, namely those in $C_{t,x}^{\overline{U}}$ or $\widetilde{C}_{t,x}^{\overline{U}}$ respectively. The proof of Theorem 4.3.1 is based on this observation, with Lemma 4.5.1 as a starting point.

To further facilitate the comparison of η_t^1, η_t^2 with ξ, write $C_{t,x}$ as the enumeration $\{(x_k, t_k, U_k), 1 \leq k \leq |\, C_{t,x}\,|\}$ with $t_k \geq t_{k-1}$ and $(x_0, t_0, U_0) = (x, 0, u)$.

With this,

$$\begin{aligned}
\xi_k &:= \Psi(\xi_{k-1}, (x_k, t_k, U_k)), \quad 1 \le k \le |\, C_{t,x} \,|, && (4.14)\\
\xi_0 &:= \xi,\\
\widetilde{\xi}_k &:= \Psi(\widetilde{\xi}_{k-1}, (x_k, t_k, U_k)), \quad 1 \le k \le |\, C_{t,x} \,|, && (4.15)\\
\widetilde{\xi}_0 &:= \Psi(\xi, (x, 0, u)).
\end{aligned}$$

By Proposition 4.4.4 ξ_k, $\widetilde{\xi}_k$ are μ-distributed since they are obtained via resampling steps. So we can describe η_t^1 and η_t^2 via finitely many flips from a common ancestor ξ, and each step in between is μ-distributed.

With the observations above we can rewrite part of (4.11) using the graphical representation.

Lemma 4.5.2. *Using above notation,*

$$\begin{aligned}
&(S_t f(\widetilde{\eta}) - S_t f(\eta))^2\\
&\quad \le \widehat{\mathbb{P}}_{\widetilde{\eta},\eta}\left(\eta_t^1 \neq \eta_t^2\right) \int d\overline{N}\, (2\,|\, C_{t,x} \,| + 1) \int d\overline{U}\\
&\qquad \left[\sum_{k=1}^{|\, C_{t,x} \,|} \left(\nabla_{x_k} f(\widetilde{\xi}_{k-1})\right)^2 + \sum_{k=1}^{|\, C_{t,x} \,|} \left(\nabla_{x_k} f(\xi_{k-1})\right)^2 + \left(\nabla_x f(\xi)\right)^2 \right].
\end{aligned}$$

Proof. Start with

$$\begin{aligned}
(S_t(\tilde{\eta}) - S_t f(\eta))^2 &= \left(\widehat{\mathbb{E}}_{\tilde{\eta},\eta}(f(\eta_t^1) - f(\eta_t^2))\mathbb{1}_{\eta_t^1 \neq \eta_t^2}\right)^2\\
&\le \widehat{\mathbb{E}}_{\tilde{\eta},\eta}\left(f(\eta_t^1) - f(\eta_t^2)\right)^2 \widehat{\mathbb{P}}_{\tilde{\eta},\eta}(\eta_t^1 \neq \eta_t^2).
\end{aligned}$$

Now let $\widehat{\mathbb{P}}$ be the graphical construction coupling, then, in the notation of Section 4.4,

$$\widehat{\mathbb{E}}_{\tilde{\eta},\eta}(f(\eta_t^1) - f(\eta_t^2))^2 = \int d\overline{N} \int d\overline{U} \left[f\left(\Psi(\tilde{\eta}, \overline{N}_t^{\overline{U}})\right) - f\left(\Psi(\eta, \overline{N}_t^{\overline{U}})\right)\right]^2.$$

Using (4.12) and (4.13),

$$\left[f\left(\Psi(\tilde{\eta}, \overline{N}_t^{\overline{U}})\right) - f\left(\Psi(\eta, \overline{N}_t^{\overline{U}})\right)\right]^2 = \left[f\left(\Psi(\xi, \widetilde{C}_{t,x}^{\overline{U}})\right) - f\left(\Psi(\xi, C_{t,x}^{\overline{U}})\right)\right]^2. \tag{4.16}$$

This can be rewritten using the telescopic sum over the individual resampling steps (4.14),(4.15):

$$f\left(\Psi\left(\xi, C_{t,x}^{\overline{U}}\right)\right) - f(\xi_0) = \sum_{k=1}^{|\, C_{t,x}\,|} f(\xi_k) - f(\xi_{k-1}),$$

$$f\left(\Psi\left(\xi, \widetilde{C}_{t,x}^{\overline{U}}\right)\right) - f(\tilde{\xi}_0) = \sum_{k=1}^{|\, C_{t,x}\,|} f(\tilde{\xi}_k) - f(\tilde{\xi}_{k-1}).$$

Putting the telescopic sums into (4.16) and using the inequality $(\sum_{i=1}^n a_i)^2 \leq n\sum_{i=1}^n a_i^2$ leads to the upper bound

$$(2\,|\, C_{t,x}\,| + 1)\left[\sum_{k=1}^{|\, C_{t,x}\,|}\left(f(\tilde{\xi}_k) - f(\tilde{\xi}_{k-1})\right)^2 \right.$$
$$\left. + \sum_{k=1}^{|\, C_{t,x}\,|}\left(f(\xi_k) - f(\xi_{k-1})\right)^2 + \left(f(\tilde{\xi}_0) - f(\xi_0)\right)^2\right].$$

Notice that by construction, ξ_k and ξ_{k-1} are identical except for a possible flip at site x_k. Consequently, we can further estimate by

$$(2\,|\, C_{t,x}\,| + 1)\left[\sum_{k=1}^{|\, C_{t,x}\,|}(\nabla_{x_k} f(\tilde{\xi}_{k-1}))^2 + \sum_{k=1}^{|\, C_{t,x}\,|}(\nabla_{x_k} f(\xi_{k-1}))^2 + (\nabla_x f(\xi))^2\right].$$

□

The next lemma deals with rearranging and separating integrals as well as condensing the individual terms as much as possible, continuing where Lemma 4.5.2 left off.

Lemma 4.5.3. *For* $1 \le p \le q \le \infty$ *with* $\frac{1}{p} + \frac{1}{q} = 1$,

$$\sum_{x\in\mathbb{Z}^d} \int \mu(d\eta) \int_0^1 du \left(S_t f(\Psi(\eta,(x,0,u))) - S_t f(\eta)\right)^2$$
$$\le \left(\int c(0,\eta)\theta_t(\eta)^q \, \mu(d\eta)\right)^{\frac{1}{q}} \left(\int (2\,|\, C_{t,0} \,|+1)^2 \, d\overline{N}\right)$$
$$\cdot \left(\sum_{x\in\mathbb{Z}^d} \|\, (\nabla_x f)^2 \,\|_{L^p(\mu)}\right),$$

where

$$\theta_t(\eta) = \widehat{\mathbb{P}}_{\eta^0,\eta}(\eta_t^1 \neq \eta_t^2).$$

Proof. We start by using Lemma 4.5.2 to estimate the inner term of

$$\sum_{x\in\mathbb{Z}^d} \int \mu(d\eta) \int_0^1 du \left(S_t f(\Psi(\eta,(x,0,u))) - S_t f(\eta)\right)^2.$$

Upon reordering some of the integrals and sums, we obtain

$$\sum_{x\in\mathbb{Z}^d} \int d\overline{N}\, (2\,|\, C_{t,x} \,|+1)$$
$$\left[\sum_{k=1}^{|\, C_{t,x} \,|} \int \mu(d\eta) \int_0^1 du \int d\overline{U} \left(\nabla_{x_k} f(\widetilde{\xi}_{k-1})\right)^2 \widehat{\mathbb{P}}_{\widetilde{\eta},\eta}\left(\eta_t^1 \neq \eta_t^2\right) \right. \tag{4.17}$$
$$+ \sum_{k=1}^{|\, C_{t,x} \,|} \int \mu(d\eta) \int_0^1 du \int d\overline{U} \left(\nabla_{x_k} f(\xi_{k-1})\right)^2 \widehat{\mathbb{P}}_{\widetilde{\eta},\eta}\left(\eta_t^1 \neq \eta_t^2\right) \tag{4.18}$$
$$\left. + \int \mu(d\eta) \int_0^1 du \int d\overline{U} \left(\nabla_x f(\xi)\right)^2 \widehat{\mathbb{P}}_{\widetilde{\eta},\eta}\left(\eta_t^1 \neq \eta_t^2\right)\right]. \tag{4.19}$$

Now we use Hölder's inequality with respect to the integration $\int \mu(d\eta) \int_0^1 du \int d\overline{U}$. In all three cases this produces as second term

$$\left(\int \mu(d\eta) \int_0^1 du \int d\overline{U}\, \widehat{\mathbb{P}}_{\widetilde{\eta},\eta}\left(\eta_t^1 \neq \eta_t^2\right)^q\right)^{\frac{1}{q}}.$$

Note that, depending on u, $\tilde{\eta}$ is either η^x or η, in which case $\widehat{\mathbb{P}}_{\tilde{\eta},\eta}(\eta_t^1 \neq \eta_t^2) = 0$. Using this as well as translation invariance shows that the above term equals

$$\left(\int c(0,\eta)\theta_t(\eta)^q \, \mu(d\eta) \right)^{\frac{1}{q}} .$$

The other term of Hölder's inequality varies slightly from line to line, but as it is mostly the same we focus on line (4.17):

$$\left(\int \mu(d\eta) \int_0^1 du \int d\overline{U} \left(\nabla_{x_k} f(\widetilde{\xi}_{k-1}) \right)^{2p} \right)^{\frac{1}{p}}$$

Here we can finally use the fact that the configurations ξ_k,$\tilde{\xi}_k$ are μ-distributed. Because of this fact we have the following identity:

$$\begin{aligned}
&\left(\int \mu(d\eta) \int_0^1 du \int d\overline{U} \left(\nabla_{x_k} f(\widetilde{\xi}_{k-1}) \right)^{2p} \right)^{\frac{1}{p}} \\
&\quad = \left(\int \mu(d\eta) \, (\nabla_{x_k} f(\eta))^{2p} \right)^{\frac{1}{p}} = \| \, (\nabla_{x_k} f)^2 \, \|_{L^p(\mu)} .
\end{aligned}$$

Applying the same argument to (4.18) and (4.19),

$$\begin{aligned}
&\sum_{x \in \mathbb{Z}^d} \int \mu(d\eta) \int_0^1 du \, (S_t f(\Psi(\eta,(x,0,u))) - S_t f(\eta))^2 \\
&\quad \leq \left(\int c(0,\eta)\theta_t(\eta)^q \, \mu(d\eta) \right)^{\frac{1}{q}} \sum_{x \in \mathbb{Z}^d} \int d\overline{N} \, (2 \, | \, C_{t,x} \, | + 1) \\
&\qquad \cdot \left[2 \sum_{k=1}^{|\, C_{t,x} \,|} \| \, (\nabla_{x_k} f)^2 \, \|_{L^p(\mu)} + \| \, (\nabla_x f)^2 \, \|_{L^p(\mu)} \right] .
\end{aligned}$$

By translation invariance of the law of $\overline{N}$,

$$\sum_{x \in \mathbb{Z}^d} \int d\overline{N} \, (2 \, | \, C_{t,x} \, | + 1) \left[2 \sum_{k=1}^{|\, C_{t,x} \,|} \| \, (\nabla_{x_k} f)^2 \, \|_{L^p(\mu)} + \| \, (\nabla_x f)^2 \, \|_{L^p(\mu)} \right]$$

$$= \sum_{x\in\mathbb{Z}^d} \int d\overline{N}\,(2\,|\,C_{t,0}\,|+1)\left[2\sum_{k=1}^{|\,C_{t,0}\,|} \|\,(\nabla_{x_k+x} f)^2\,\|_{L^p(\mu)} + \|\,(\nabla_x f)^2\,\|_{L^p(\mu)}\right]$$
$$= \int d\overline{N}\,(2\,|\,C_{t,0}\,|+1)^2 \sum_{x\in\mathbb{Z}^d} \|\,(\nabla_x f)^2\,\|_{L^p(\mu)}\,. \qquad \square$$

In order to proceed we need estimates on the size of $C_{t,0}$. The following two lemmas provides us with those.

Lemma 4.5.4. *Denote by $B_t \subset \mathbb{Z}^d$ the set of sites which are represented in $C_{t,0}$, i.e.,*

$$B_t := \{x \in \mathbb{Z}^d \,|\, \exists\, s \in [0,t] : (x,s) \in C_{t,0}\} \cup \{0\}.$$

Then there exist dimension-dependent constants $c_1, c_2 > 0$ such that

a) $\int |\,B_t\,|^2 \, d\overline{N} \le c_1(t+1)^{2d}$;

b) $\sum_{x\in\mathbb{Z}^d} \left(\int \mathbb{1}_{x\in B_t}\, d\overline{N}\right)^{\frac{1}{2}} \le c_2(t+1)^d$.

Proof. The proof rests on the observation that B_t is strongly related to first passage percolation: Consider first passage percolation with iid. exponentially distributed edge weights(see for example [33]). Let E be the edge set of $\mathbb{Z}^d$, and $r_e, e \in E$ iid. the exponentially distributed edge weights. Then the first passage percolation distance is $T(0,x) = \inf\{\sum_{e\in\gamma} r_e \,|\, \gamma \text{ path from } 0 \text{ to } x\}$. Now we compare the ball $\widetilde{B}_t := \{x \in \mathbb{Z}^d : T(0,x) \le t\}$ of reachable sites within distance t to B_t in terms of growth. Denote the outer boundary of a finite subset A of $\mathbb{Z}^d$ by $\partial A = \{x \in \mathbb{Z}^d \backslash A \,|\, \exists\, y \in A : |\,x-y\,| = 1\}$. The rate at which a site $x \in \partial\widetilde{B}_t$ is encompassed by $\widetilde{B}_t$ is given by the number of edges connecting x to $\widetilde{B}_t$. On the other hand B_t grows to contain a site $x \in \partial B_t$ just at rate 1. Therefore $\widetilde{B}_t$ stochastically dominates B_t, and proving a) and b) for $\widetilde{B}_t$ suffices.

From the theory of first passage percolation(see [33], Theorems 3.10, 3.11) we use the following fact : There exist positive constants k_1, k_2, k_3 (possibly dimension-dependent) such that for all $x \in \mathbb{Z}^d$ with $|\,x\,| > k_1 t$:

$$\mathbb{P}(x \in \widetilde{B}_t) = \mathbb{P}(T(0,x) \le t) \le k_2 e^{-k_3 |\,x\,|}. \qquad (4.20)$$

To prove b),

$$\begin{aligned}
\sum_{x\in\mathbb{Z}^d} \mathbb{P}(x \in \widetilde{B}_t)^{\frac{1}{2}} &\leq \sum_{x:|\,x\,|\leq k_1 t} 1 + \sum_{x:|\,x\,|>k_1 t} k_2 e^{-k_3|\,x\,|}\\
&\leq (2k_1+1)^d t^d + \sum_{x\in\mathbb{Z}^d} k_2 e^{-k_3|\,x\,|}\\
&\leq c_2(t+1)^d
\end{aligned}$$

for a suitable constant c_2. To prove a), fix an integer $r > k_1 t$. Since $\left|\,\widetilde{B}_t\,\right| > (2r+1)^d$ implies that at least one site in $\widetilde{B}_t$ lies outside a cube of size $2r+1$. Hence

$$\mathbb{P}\left(\left|\,\widetilde{B}_t\,\right| > (2r+1)^d\right) \leq k_2 e^{-k_3(r+1)} 2d(2r+3)^{d-1},$$

which proves exponentially decaying tails for the volume of $\widetilde{B}_t$. □

Utilizing Lemma 4.5.4 we now prove the second moment estimate of $|\,C_{t,0}\,|$ needed for Lemma 4.5.3.

Lemma 4.5.5. *There exists a dimension-dependent constant $C_d >$ so that the following estimate holds:*

$$\int (2\,|\,C_{t,0}\,|+1)^2\, d\overline{N} \leq C_d(t+1)^{2d+2}.$$

Proof. Let B_t be as in Lemma 4.5.4. Then for each $x \in B_t$ we denote by t_x the time of first time of appearance of x in $C_{t,0}$,

$$t_x := \inf\{s \in [0,t]\,|\,(x,s) \in C_{t,0}\}.$$

We have

$$C_{t,0} = \overline{N} \cap \{(x,s) \in \mathbb{Z}^d \times [0,t]\,|\,x \in B_t, s \geq t_x\} \subset \overline{N} \cap \{(x,s) \in \mathbb{Z}^d \times [0,t]\,|\,x \in B_t\}.$$

Conditioned on B_t and t_x the last set is Poisson distributed with the addition of the points $(x, t_x), x \in B_t$. Because of this, conditioned on B_t, $|\,C_{t,0}\,| - |\,B_t\,|$ is stochastically dominated by a Poisson distributed with parameter $t\,|\,B_t\,|$. As a consequence,

$$\int (2\,|\,C_{t,0}\,|+1)^2\,d\overline{N} \leq 4\int (t+1)^2(|\,B_t\,|+1)^2\,d\overline{N}.$$

Finally the estimate from Lemma 4.5.4,a) completes the proof. □

With all ingredients present we can quickly prove the main result in form of a slightly more general lemma.

Lemma 4.5.6. *Let $f:\Omega\to\mathbb{R}$ with $|||\,f\,|||_2<\infty$ and $0\leq T\leq S$. Then*

$$\begin{aligned}&\mathrm{Var}_\mu(S_Tf)-\mathrm{Var}_\mu(S_Sf)\\&\quad\leq C_d\int_T^S (t+1)^{2d+2}\left(\int c(0,\eta)\theta_t(\eta)^q\,\mu(d\eta)\right)^{\frac{1}{q}}\,dt\sum_{x\in\mathbb{Z}^d}\|\,(\nabla_x f)^2\,\|_{L^p(\mu)}\,.\end{aligned}$$

C_d is a constant depending just on the dimension.

Proof. Assume that f satisfies $|||\,f\,|||_1<\infty$. This then implies that $S_tf,(S_tf)^2\in$ *$dom(L)$* and by Lemma 4.5.1,

$$\mathrm{Var}_\mu(S_Tf)-\mathrm{Var}_\mu(S_Sf)=\int_T^S\int\left[L(S_tf-S_tf(\eta))^2\right](\eta)\,\mu(d\eta)\,dt.$$

By using the formulation of the generator using the graphical construction (see (4.8)),

$$Lf(\eta)=\sum_{x\in\mathbb{Z}^d}\int_0^1 f\left(\Psi(\eta,(x,0,u))\right)-f(\eta)\,du,$$

we apply Lemma 4.5.3 and obtain

$$\begin{aligned}&\mathrm{Var}_\mu(S_Tf)-\mathrm{Var}_\mu(S_Sf)\\&\quad\leq\int_T^S\int(2\,|\,C_{t,0}\,|+1)^2\,d\overline{N}\,\|\,\theta_t\,\|_{L^q(\mu)}\,dt\sum_{x\in\mathbb{Z}^d}\|\,(\nabla_x f)^2\,\|_{L^p(\mu)}\,.\end{aligned}$$

Finally Lemma 4.5.5 gives us the estimate on $\int(2\,|\,C_{t,0}\,|+1)^2\,d\overline{N}$ to complete the proof.

If f only satisfies $|||\,f\,|||_2<\infty$ we then approximate f by local functions. □

Proof of Theorem 4.3.1. A direct consequence of Lemma 4.5.6 with $S = \infty$ and the estimate $\left(\int c(0,\eta)\theta_t(\eta)^q \, \mu(d\eta)\right)^{\frac{1}{q}} \leq \| \theta_t \|_{L^q(\mu)}$. □

We now prove Theorem 4.3.2, which is a modification of Theorem 4.3.1 for attractive spin-systems.

Proof of Theorem 4.3.2. This result is also based on Lemma 4.5.6. To estimate $\left(\int c(0,\eta)\theta_t(\eta)^q \, \mu(d\eta)\right)^{\frac{1}{q}}$ in terms of the auto-correlation, we start with the fact that in the coupling the spread of discrepancies is limited to B_t(as in Lemma 4.5.4):

$$\begin{aligned}\theta_t(\eta) = \widehat{\mathbb{P}}_{\eta^0,\eta}(\eta_t^1 \neq \eta_t^2) &\leq \widehat{\mathbb{E}}_{\eta^0,\eta} \sum_{x \in B_t} \mathbb{1}_{\eta_t^1(x) \neq \eta_t^2(x)} \\ &= \sum_{x \in \mathbb{Z}^d} \widehat{\mathbb{E}}_{\eta^0,\eta} \mathbb{1}_{x \in B_t} \mathbb{1}_{\eta_t^1(x) \neq \eta_t^2(x)}.\end{aligned}$$

Next, since $\theta_t \leq 1$,

$$\begin{aligned}\int c(0,\eta)\theta_t(\eta)^q \, \mu(d\eta) &\leq \int c(0,\eta)\theta_t(\eta) \, \mu(d\eta) \\ &\leq \sum_{x \in \mathbb{Z}^d} \int \widehat{\mathbb{E}}_{\eta^0,\eta} \mathbb{1}_{x \in B_t} \mathbb{1}_{\eta_t^1(x) \neq \eta_t^2(x)} c(0,\eta) \, \mu(d\eta).\end{aligned}$$

We can now use Cauchy-Schwarz to obtain

$$\sum_{x \in \mathbb{Z}^d} \left(\int \mathbb{1}_{x \in B_t} \, d\overline{N}\right)^{\frac{1}{2}} \left(\int \widehat{\mathbb{E}}_{\eta^0,\eta} \mathbb{1}_{\eta_t^1(x) \neq \eta_t^2(x)} c(0,\eta)^2 \, \mu(d\eta)\right)^{\frac{1}{2}}. \tag{4.21}$$

Since the model is attractive the coupling $\widehat{\mathbb{P}}$ preserves an initial ordering. Since either $\eta^0 < \eta$ or $\eta < \eta^0$,

$$\begin{aligned}\widehat{\mathbb{E}}_{\eta^0,\eta} \mathbb{1}_{\eta_t^1(x) \neq \eta_t^2(x)} &= \frac{1}{2} \widehat{\mathbb{E}}_{\eta^0,\eta} \left| \eta_t^1(x) - \eta_t^2(x) \right| = \frac{1}{2} \left| \widehat{\mathbb{E}}_{\eta^0,\eta} \left(\eta_t^1(x) - \eta_t^2(x) \right) \right| \\ &= \frac{1}{2} \left| \mathbb{E}_{\eta^0} \eta_t(x) - \mathbb{E}_\eta \eta_t(x) \right|.\end{aligned}$$

When we use the notation $g_x(\eta) := \eta(x)$ and $m = \mu(g_0) = \mu(g_x)$,

$$\begin{aligned}
&\int \widehat{\mathbb{E}}_{\eta^0,\eta} \mathbb{1}_{\eta_t^1(x)\neq\eta_t^2(x)} c(0,\eta)^2 \,\mu(d\eta) \\
&= \frac{1}{2}\int \big|\, S_t g_x(\eta^0) - S_t g_x(\eta) \,\big|\, c(0,\eta)^2 \,\mu(d\eta) \\
&\leq \frac{1}{2}\int \big|\, S_t g_x(\eta^0) - m \,\big|\, c(0,\eta)^2 \,\mu(d\eta) + \frac{1}{2}\int |\, S_t g_x(\eta) - m \,|\, c(0,\eta)^2 \,\mu(d\eta).
\end{aligned}$$

Using $c(0,\eta) \leq 1$, as well as

$$\begin{aligned}
\int |\, S_t g_x(\eta) - m \,| \;\mu(d\eta) &\leq \left(\int \left(S_t g_x(\eta) - m\right)^2 \,\mu(d\eta)\right)^{\frac{1}{2}} \\
&= \mathrm{Var}_\mu(S_t g_x)^{\frac{1}{2}} = \mathrm{Var}_\mu(S_t g_0)^{\frac{1}{2}}
\end{aligned}$$

and

$$\begin{aligned}
\int \big|\, S_t g_x(\eta^0) - m \,\big|\, c(0,\eta)\,\mu(d\eta) &= \int \big|\, S_t g_x(\eta^0) - m \,\big|\, c(0,\eta) \frac{\mu(d\eta)}{\mu(d\eta^0)}\,\mu(d\eta^0) \\
&\leq \int \big|\, S_t g_x(\eta^0) - m \,\big| \;\mu(d\eta^0) \leq \mathrm{Var}_\mu(S_t g_0)^{\frac{1}{2}}
\end{aligned}$$

we estimate (4.21) and obtain

$$\int c(0,\eta)\theta_t(\eta)^q \,\mu(d\eta) \leq \sum_{x\in\mathbb{Z}^d} \left(\int \mathbb{1}_{x\in B_t}\, d\overline{N}\right)^{\frac{1}{2}} \mathrm{Var}_\mu(S_t g_0)^{\frac{1}{4}}.$$

Furthermore Lemma 4.5.4 gives us an estimate for the sum, so that

$$\left(\int c(0,\eta)\theta_t(\eta)^q \,\mu(d\eta)\right)^{\frac{1}{q}} \leq c_2^{\frac{1}{q}} (t+1)^{\frac{d}{q}}\, \mathrm{Var}_\mu(S_t g_0)^{\frac{1}{4q}}.$$

Omitting the q-th root where convenient and with a constant $C_d' = C_d(1 \vee c_2)$ as well as writing $\frac{1}{q} = \frac{p-1}{p}$ we obtain the result

$$D_p(T) \leq C_d' \int_T^\infty (t+1)^{3d+2} \left(\mathrm{Var}_\mu(S_t g_0)\right)^{\frac{p-1}{4p}} \;dt. \qquad \square$$

Proof of Proposition 4.3.4. Let $f : \Omega \to \mathbb{R}$ be a local function with $\mu(f) = 0$.

By Theorem 4.3.1, for any $0 < \lambda' < \lambda$,

$$\| S_t f \|^2_{L^2(\mu)} \leq const \cdot e^{-\lambda' t},$$

and for any $0 < a < \lambda'$

$$\int_0^\infty e^{at} \| S_t f \|^2_{L^2(\mu)} \, dt < \infty. \tag{4.22}$$

Let $E_{f,f}$ be the associated measure wrt. to the spectral decomposition of $-L$. Then

$$\| S_t f \|^2_{L^2(\mu)} = \int_0^\infty e^{-2\gamma t} E_{f,f}(d\gamma).$$

By (4.22),

$$\int_0^\infty \int_0^\infty e^{at-2\gamma t} E_{f,f}(d\gamma) \, dt < \infty.$$

Therefore $E_{f,f}(]0, \lambda/2[) = 0$.

Let now $f \in L^2(\mu)$ and approximate it by local functions f_n. Assuming that $(f_n), f$ have norm 1 makes $E_{f_n,f_n}, E_{f,f}$ probability measures and E_{f_n,f_n} weakly converges to $E_{f,f}$. By the Portmanteau theorem $E_{f,f}(]0, \lambda/2[) = 0$, which completes the proof. $\square$

Proof of Proposition 4.3.5. By the Poincaré inequality, the auto-correlation of the spin at the origin, $\varphi(t) = \mathrm{Var}_\mu(S_t g), g(\eta) = \eta(0)$, decays exponentially fast. The proof of Theorem 4.3.2 contains the estimate of $\left(\int c(\eta, 0)\theta_t(\eta)^q \, \mu(d\eta)\right)^{\frac{1}{q}}$ in terms of φ. $\square$

Proof of Proposition 4.3.6. We have

$$\| S_T - \mu(f) \|_\infty = \left\| \int_T^\infty L S_t f \, dt \right\| \leq \sup_{\eta \in \Omega} \int_T^\infty \sum_{x \in \Omega} | \nabla_x f(\eta) | \, .$$

Write $\delta_x(f) := \| \nabla_x f \|_\infty$. Then

$$\sum_{x \in \mathbb{Z}^d} | \nabla_x S_t f(\eta) | \leq \sum_{x \in \mathbb{Z}^d} \widehat{\mathbb{E}}_{\eta^x, \eta} \left| f(\eta_t^1) - f(\eta_t^2) \right|$$

$$\leq \sum_{x\in\mathbb{Z}^d} \widehat{\mathbb{E}}_{\eta^x,\eta}\left(\sum_{y\in\mathbb{Z}^d} \mathbb{1}_{\eta_t^1(y)\neq\eta_t^2(y)}\delta_y(f)\right) = \widehat{\mathbb{E}}_{\eta^0,\eta}\sum_{y\in\mathbb{Z}^d} \mathbb{1}_{\eta_t^1(y)\neq\eta_t^2(y)} \,|||\, f \,|||_1$$
$$\leq \widehat{\mathbb{E}}_{\eta^0,\eta} \,|\, C_{t,0} \,|\, \mathbb{1}_{\eta_t^1\neq\eta_t^2} \,|||\, f \,|||_1$$

By the Cauchy-Schwarz inequality and Lemma 4.5.5,

$$\begin{aligned}\widehat{\mathbb{E}}_{\eta^0,\eta} \,|\, C_{t,0} \,|\, \mathbb{1}_{\eta_t^1\neq\eta_t^2} &\leq \widehat{\mathbb{P}}_{\eta^0,\eta}(\eta_t^1 \neq \eta_t^2)^{\frac{1}{2}} \left(\widehat{\mathbb{E}}_{\eta^0,\eta} \,|\, C_{t,0} \,|^2\right)^{\frac{1}{2}} \\ &\leq \|\, \theta_t \,\|_{L^\infty(\mu)}^{\frac{1}{2}} \, C_d^{\frac{1}{2}} (t+1)^{d+1}.\end{aligned}$$

Therefore

$$\|\, S_T - \mu(f) \,\|_\infty \leq \int_T^\infty C_d^{\frac{1}{2}} (t+1)^{d+1} \,\|\, \theta_t \,\|_{L^\infty(\mu)}^{\frac{1}{2}} \, dt \,|||\, f \,|||_1 \,. \qquad \square$$

Proof of Proposition 4.3.7. Fix $R > 0$. Then, by Lemma 4.5.6,

$$\begin{aligned}\mathrm{Var}_\mu(S_t f) &\leq C_d D_1(0,R) \sum_{x\in\mathbb{Z}^d} \left\| (\nabla_x S_t f)^2 \right\|_{L^1(\mu)} + \mathrm{Var}_\mu(S_R S_t f) \\ &\leq C_d \delta^{-1} D_1(0,R)\mathcal{E}(S_t f, S_t f) + \mathrm{Var}_\mu(S_R f) \\ &\leq C_d \delta^{-1} D_1(0,R)\mathcal{E}(f,f) + D_p(R)\Phi_R(f).\end{aligned}$$

The estimate $\Phi_R(S_t f) \leq C_d \sum_{x\in\mathbb{Z}^d} \left\| (\nabla_x S_t f)^2 \right\|_{L^p(\mu)}$ is a direct consequence of Theorem 4.3.1. $\square$

5 Random Walks in Dynamic Random Environments: A transference principle[1]

5.0 Abstract

We study a general class of random walks driven by a uniquely ergodic Markovian environment. Under a coupling condition on the environment we obtain strong ergodicity properties for the environment as seen from the position of the walker, i.e., the environment process. We can transfer the rate of mixing in time of the environment to the rate of mixing of the environment process with a loss of at most polynomial order. Therefore the method is applicable to environments with sufficiently fast polynomial mixing. We obtain unique ergodicity of the environment process. Moreover, the unique invariant measure of the environment process depends continuously on the jump rates of the walker.

As a consequence we obtain the law of large numbers and a central limit theorem with non-degenerate variance for the position of the walk.

5.1 Introduction

In recent days random walks in dynamic random environment have been studied by several authors. Motivation comes among others from non-equilibrium statistical mechanics -derivation of Fourier law- [21] and large deviation theory [45]. In principle random walk in dynamic random environment contains as a particular case random walk in static random environment. However, mostly, in

[1] revised version of "Limit theorems for random walks in dynamic random environment"
Frank Redig, Florian Völlering
http://arxiv.org/abs/1106.4181
submitted

turning to dynamic environments, authors concentrate more on environments with sufficient mixing properties. In that case the fact that the environment is dynamic helps to obtain self-averaging properties that ensure standard limiting behaviour of the walk, i.e., law of large numbers and central limit theorem.

In the study of the limiting behaviour of the walker, the environment process, i.e., the environment as seen from the position of the walker plays a crucial role. See also [32], [44] for the use of the environment process in related context. In a translation invariant setting the environment process is a Markov process and its ergodic properties fully determine corresponding ergodic properties of the walk, since the position of the walker equals an additive function of the environment process plus a controllable martingale.

The main theme of this chapter is the following natural question: if the environment is uniquely ergodic, with a sufficient speed of mixing, then the environment process shares similar properties. In several works ([11], [9], [12], [6]) this transfer of "good properties of the environment" to "similar properties of the environment process" is made via a perturbative argument, and therefore holds only in a regime where the environment and the walker are weakly coupled. Some non-perturbative results also exist, but those require strong mixing properties of the environment in space and time ([23], [22], [13]).

In this chapter we consider the context of general Markovian uniquely ergodic environments, which are such that the semigroup contracts at a minimal speed in a norm of variation type. Examples of such environments include interacting particle systems in "the $M < \epsilon$ regime" [37] and weakly interacting diffusion processes on a compact manifold. Our conditions on the environment are formulated in the language of coupling. More precisely, we impose that for the environment there exists a coupling such that the distance between every pair of initial configurations in this coupling decays fast enough so that multiplied with t^d it is still integrable in time. As a result we then obtain that for the environment process there exists a coupling such that the distance between every pair of initial configurations in this coupling decays at a speed which is at least integrable in time. In fact we show more, namely in going from the environment to the environment process we essentially loose a factor t^d in the rate of decay to equilibrium. E.g., if for the environment there is a coupling where the distance decays exponentially, then the same holds for the environment process (with possibly another rate).

Once we have controllable coupling properties of the environment process, we can draw strong conclusions for the position of the walker. For example a law of large numbers with an asymptotic speed that depends continuously on the rates, and a central limit theorem. We also prove recurrence in $d = 1$ under condition of zero speed.

This chapter is organized as follows. The model and necessary notation are introduced in Section 5.2. Section 5.3 is dedicated to lift properties of the environment to the environment process. The focus is on Theorem 5.3.1 and its refinements. Based on these results consequences for the walker are summarized in 5.3.5. In Section 5.4 we give examples for environments to which the results are applicable and present one example which has polynomial mixing in space and time. Section 5.5 is devoted to proofs.

5.2 The model

5.2.1 Environment

A random walk in dynamic random environment is a process $(X_t)_{t\geq 0}$ on the lattice $\mathbb{Z}^d$ which is driven by a second process $(\eta_t)_{t\geq 0}$ on $E^{\mathbb{Z}^d}$, the (dynamic) environment. This is interpreted as a random walk moving through the environment, with time-dependent transition rates being determined by the local environment around the random walk.

To become more precise, the *environment* $(\eta_t)_{t\geq 0}$ we assume to be a Feller Process on the state space $\Omega := E^{\mathbb{Z}^d}$, where (E, ρ) is a compact metric space (examples in mind are $E = \{0, 1\}$ or $E = [0, 1]$). We assume (without loss of generality) that the distance ρ on E is bounded from above by 1. The generator of the Markov process $(\eta_t)_{t\geq 0}$ is denoted by L^E and its semigroup by S_t^E, both considered on the space of continuous functions $\mathcal{C}(\Omega; \mathbb{R})$. We assume that the environment is translation invariant, i.e.,

$$\mathbb{P}_\eta^E(\theta_x \eta_t \in \cdot) = \mathbb{P}_{\theta_x \eta}^E(\eta_t \in \cdot)$$

with θ_x denoting the shift operator $\theta_x \eta(y) = \eta(y - x)$ and $\mathbb{P}_\eta^E$ the path space measure of the process $(\eta_t)_{t\geq 0}$ starting from η.

5.2.2 Lipschitz functions

Denote, for $x \in \mathbb{Z}^d$,

$$(\Omega\times\Omega)_x := \left\{(\eta,\xi) \in \Omega^2 : \eta(x) \neq \xi(x) \text{ and } \eta(y) = \xi(y)\ \forall\, y \in \mathbb{Z}^d\backslash\{x\}\right\}, \quad x \in \mathbb{Z}^d.$$

Definition 5.2.1. *For any $f : \Omega \to \mathbb{R}$, we denote by $\delta_f(x)$ the Lipschitz-constant of f w.r.t. the variable $\eta(x)$,*

$$\delta_f(x) := \sup_{(\eta,\xi)\in(\Omega\times\Omega)_x} \frac{f(\eta) - f(\xi)}{\rho(\eta(x),\xi(x))}.$$

We write

$$|||\, f \,||| := \sum_{x\in\mathbb{Z}^d} \delta_f(x). \tag{5.1}$$

Note that $|||\, f \,||| < \infty$ implies that f is bounded, continuous and the value of f is uniformly weakly dependent on sites far away. A weaker semi-norm we also use is the oscillation (semi)-norm

$$\|\, f \,\|_{osc} := \sup_{\eta,\xi\in\Omega} \left(f(\eta) - f(\xi)\right).$$

From telescoping over single site changes one sees $\|\, f \,\|_{osc} \leq |||\, f \,|||$.

5.2.3 The random walker and assumption on rates

The random walk X_t is a process on $\mathbb{Z}^d$, whose transition rates depend on the state of the environment as seen from the walker. More precisely, the rate to jump from site x to site $x + z$ given that the environment is in state η is $\alpha(\theta_{-x}\eta, z)$. We make two assumptions on these jump rates α. First, we guarantee that the position of the walker X_t has a first moment by assuming

$$\|\, \alpha \,\|_1 := \sum_{z\in\mathbb{Z}^d} \|\, z \,\| \sup_{\eta\in\Omega} |\, \alpha(\eta, z) \,| < \infty. \tag{5.2}$$

More generally, as sometimes higher moments are necessary, we write

$$\|\, \alpha \,\|_p^p := \sum_{z\in\mathbb{Z}^d} \|\, z \,\|^p \sup_{\eta\in\Omega} |\, \alpha(\eta, z) \,|, \quad p \geq 1.$$

Second, we limit the sensitivity of the rates to small changes in the environment by assuming that

$$||| \alpha ||| := \sum_{z\in\mathbb{Z}^d} ||| \alpha(\cdot, z) ||| < \infty. \tag{5.3}$$

Finally, sometimes we will have to assume the stronger estimate

$$||| \alpha |||_1 := \sum_{z\in\mathbb{Z}^d} | z | \; ||| \; \alpha(\cdot, z) ||| < \infty. \tag{5.4}$$

5.2.4 Environment process

While the random walker X_t itself is not a Markov process due to the dependence on the environment, the pair (η_t, X_t) is a Markov process with generator

$$Lf(\eta, x) = L^E f(\cdot, x)(\eta) + \sum_{z\in\mathbb{Z}^d} \alpha(\theta_{-x}\eta, z) \left[f(\eta, x+z) - f(\eta, x)\right],$$

corresponding semigroup S_t (considered on the space of functions continuous in $\eta \in \Omega$ and Lipschitz continuous in $x \in \mathbb{Z}^d$) and path space measure $\mathbb{P}_{\eta,x}$.

The environment as seen from the walker is of crucial importance to understand the asymptotic behaviour of the walker itself. This process, $(\theta_{-X_t}\eta_t)_{t\geq 0}$, is called the *environment process* (not to be confused with the environment η_t). It is a Markov process with generator

$$L^{EP} f(\eta) = L^E f(\eta) + \sum_{z\in\mathbb{Z}^d} \alpha(\eta, z) \left[f(\theta_{-z}\eta) - f(\eta)\right],$$

corresponding semigroup S_t^{EP} (on $\mathcal{C}(\Omega)$) and path space measure $\mathbb{P}_\eta^{EP}$. Notice that this process is meaningful only in the translation invariant context.

5.2.5 Coupling of the environment

In the remainder of this chapter we will need a coupling $\widehat{\mathbb{P}}^E_{\eta,\xi}$, $\eta, \xi \in \Omega$, of the environment. For η, ξ the coupled pair $(\eta_t^1, \eta_t^2)_{t\geq 0}$ consists of two copies of the environment, started in η and ξ. By definition of a coupling, it has the marginals $\widehat{\mathbb{P}}^E_{\eta,\xi}(\eta^1 \in \cdot) = \mathbb{P}_\eta(\eta_t \in \cdot)$ and $\widehat{\mathbb{P}}^E_{\eta,\xi}(\eta^2 \in \cdot) = \mathbb{P}_\xi(\eta_t \in \cdot)$. Let $(\mathcal{F}_t)_{t\geq 0}$ be the canonical filtration in the path space of coupled processes. We say such a coupling satisfies the marginal Markov property if, for any $f : \Omega \to \mathbb{R}$,

$$\widehat{\mathbb{E}}^E_{\eta,\xi}\left[f(\eta^i_t)\,\middle|\,\mathcal{F}_s\right] = S^E_{t-s}f(\eta^i_s), \quad i = 1,2; t \geq s \geq 0. \tag{5.5}$$

We say it satisfies the strong marginal Markov property if, for any $f : \Omega \to \mathbb{R}$ and any stopping time τ,

$$\widehat{\mathbb{E}}^E_{\eta,\xi}\left[\mathbb{1}_{t\geq\tau}f(\eta^i_t)\,\middle|\,\mathcal{F}_\tau\right] = \mathbb{1}_{t\geq\tau}S^E_{t-\tau}f(\eta^i_\tau), \quad i = 1,2. \tag{5.6}$$

Note that the (strong) Markov property for the coupling implies the (strong) marginal Markov property.

5.3 Ergodicity of the environment process

5.3.1 Assumptions on the environment

In order to conclude results for the random walk we need to have sufficient control on the environment. To this end we assume there exists a translation invariant coupling $\widehat{\mathbb{P}}^E_{\eta,\xi}$ of the environment, which satisfies the strong marginal Markov property (5.6). In this coupling we look at $\widehat{\mathbb{E}}^E_{\eta,\xi}\rho(\eta^1_t, \eta^2)$, measuring the distance of the states at the origin. If this decays sufficiently fast we will be able to obtain ergodicity properties of the environment.

Assumption 1a. *The coupling $\widehat{\mathbb{P}}^E$ satisfies*

$$\int_0^\infty t^d \sup_{\eta,\xi\in\Omega} \widehat{\mathbb{E}}^E_{\eta,\xi}\rho(\eta^1_t(0), \eta^2_t(0))\, dt < \infty.$$

This assumption is already sufficient to obtain the law of large numbers for the position of the walker and unique ergodicity of the environment process, but it does not give quite enough control on local fluctuations. The following stronger assumption remedies that.

Assumption 1b. *The coupling $\widehat{\mathbb{P}}^E$ satisfies*

$$\int_0^\infty t^d \sum_{x\in\mathbb{Z}^d} \sup_{(\eta,\xi)\in(\Omega\times\Omega)_0} \widehat{\mathbb{E}}^E_{\eta,\xi}\rho(\eta^1_t(x), \eta^2_t(x))\, dt < \infty.$$

Remark Typically, a coupling which satisfies Assumption 1b) also satisfies Assumption 1a). It is however not automatic. But given a translation invariant coupling $\widehat{\mathbb{P}}^E$ which satisfies 1b) it is possible to construct from $\widehat{\mathbb{P}}^E$ a new coupling $\widetilde{\mathbb{P}}^E$ via a telescoping argument so that $\widetilde{\mathbb{P}}$ satisfies both 1b) and 1a).

In Section 5.4 we will discuss some examples which satisfy those assumptions. Beside natural examples where $\widehat{\mathbb{E}}^E_{\eta,\xi}\rho(\eta^1_t(0),\eta^2_t(0))$ decays exponentially fast we give an example where other decay rates like polynomial decay are obtained.

5.3.2 Statement of the main theorem

The main result of this section is the following theorem, which tells us how the coupling property of the environment lifts to the environment process.

Theorem 5.3.1. *Let $f : \Omega \to \mathbb{R}$ with $|||\, f \,||| < \infty$.*

a) Under Assumption 1a, there exists a constant $C_a > 0$ so that

$$\sup_{\eta,\xi\in\Omega} \int_0^\infty \left| \, S^{EP}_t f(\eta) - S^{EP}_t f(\xi) \, \right| \, dt \le C_a \, |||\, f \,||| \, .$$

b) Under Assumption 1b, there exists a constant $C_b > 0$ so that

$$\sum_{x\in\mathbb{Z}^d} \sup_{(\eta,\xi)\in(\Omega\times\Omega)_x} \int_0^\infty \left| \, S^{EP}_t f(\eta) - S^{EP}_t f(\xi) \, \right| \, dt \le C_b \, |||\, f \,||| \, .$$

This theorem is the key to understand the limiting behaviour of the random walk, i.e., law of large numbers as well as for the central limit theorem. Section 5.5 is devoted the proof of Theorem 5.3.1. In Section 5.3.4 we generalize this result to give more information about decay in time. Here we continue with results we can obtain using Theorem 5.3.1. Most results about the environment process just use part a) of the theorem, part b) shows how more sophisticated properties lift from the environment to the environment process as well. Those can be necessary to obtain more precise results on the walker, like how likely atypical excursions from the expected trajectory are.

It is possible to lift other properties from the environment to the environment process as well. For example if Assumption 1b is modified to state

$$\int_0^\infty t^d \sum_{x\in\mathbb{Z}^d} \sup_{(\eta,\xi)\in(\Omega\times\Omega)_0} \widehat{\mathbb{E}}^E_{\eta,\xi} \frac{\rho(\eta_t^1(x),\eta_t^2(x))}{\rho(\eta(0),\xi(0))}\, dt < \infty,$$

then that implies for the environment process

$$\sum_{x\in\mathbb{Z}^d} \sup_{(\eta,\xi)\in(\Omega\times\Omega)_x} \int_0^\infty \frac{\left| S_t^{EP} f(\eta) - S_t^{EP} f(\xi) \right|}{\rho(\eta(x),\xi(x))}\, dt \le C_{b'} \,|||\, f \,|||\,.$$

This kind of condition can be relevant in the context of diffusive environments to show that small changes in the environment are causing only small changes in the environment process.

5.3.3 Existence of a unique ergodic measure and continuity in the rates

First of, the environment process, i.e., the environment as seen from the walker, is ergodic.

Lemma 5.3.2. *Under Assumption 1a the environment process has a unique ergodic probability measure μ^{EP}.*

Proof. As E is compact, so is Ω, and therefore the space of stationary measures is non-empty. So we must just prove uniqueness.

Assume μ,ν are both stationary measures. Choose an arbitrary $f:\Omega\to\mathbb{R}$ with $|||\, f\,||| < \infty$. By Theorem 5.3.1a), for any $T>0$,

$$\begin{aligned} T\,|\,\mu(f)-\nu(f)\,| &\le \int\int\int_0^T \left| S_t^{EP} f(\eta) - S_t^{EP} f(\xi) \right| \, dt\, \mu(d\eta)\, \nu(d\xi) \\ &\le \sup_{\eta,\xi\in\Omega} \int_0^\infty \left| S_t^{EP} f(\eta) - S_t^{EP} f(\xi) \right| \, dt < \infty. \end{aligned}$$

As T is arbitrary, $\mu(f)=\nu(f)$. As functions f with $|||\, f\,||| < \infty$ are dense in $\mathcal{C}(\Omega)$, there is at most one stationary probability measure. $\square$

It is of interest not only to know that the environment process has a unique ergodic measure μ^{EP}, but also to know how this measure depends on the rates α.

Theorem 5.3.3. *Under Assumption 1a, the unique ergodic measure μ_α^{EP} depends continuously on the rates α. For two transition rate functions α, α', we have the following estimate:*

$$\left| \mu_\alpha^{EP}(f) - \mu_{\alpha'}^{EP}(f) \right| \leq \frac{C(\alpha)}{p(\alpha)} \| \alpha - \alpha' \|_0 \, ||| \, f \, |||,$$

i.e.,

$$(\alpha, f) \mapsto \mu_\alpha^{EP}(f)$$

is continuous in $\|\cdot\|_0 \times |||\cdot|||$. The functions $C(\alpha), p(\alpha)$ satisfy $C(\alpha) > 0, p(\alpha) \in]0,1[$. In the case that the rates α do not depend on the environment, i.e., $\alpha(\eta, z) = \alpha(z)$, they are given by $p(\alpha) = 1$,

$$C(\alpha) = \int_0^\infty \sup_{\eta,\xi\in\Omega} \widehat{\mathbb{E}}_{\eta,\xi}^E \rho(\eta_t^1(0), \eta_t^2(0)) \, dt.$$

As the proof is a variation of the proof of Theorem 5.3.1, it is delayed to the end of Section 5.5.

5.3.4 Speed of convergence to equilibrium in the environment process

We already know that under Assumption 1a the environment process has a unique ergodic distribution. However, we do not know at what speed this process converges to its unique stationary measure. Given the speed of convergence for the environment it is natural to believe that the environment process inherits that speed with some form of slowdown due to the additional self-interaction which is induced from the random walk. For example, if the original speed of convergence were exponential, then the environment process would also converge exponentially fast. This is indeed the case.

Theorem 5.3.4. *Let $\varphi : [0, \infty[\to \mathbb{R}$ be a monotone increasing and continuous function satisfying $\varphi(0) = 1$ and $\varphi(s+t) \leq \varphi(s)\varphi(t)$.*

a) Suppose the coupling $\widehat{\mathbb{P}}^E$ satisfies

$$\int_0^\infty \varphi(t) t^d \sup_{\eta,\xi\in\Omega} \widehat{\mathbb{E}}_{\eta,\xi}^E \rho(\eta_t^1(0), \eta_t^2(0)) \, dt < \infty.$$

Then there exists a constant $K_0 > 0$ *and a decreasing function* $C_a :]K_0, \infty[\to [0, \infty[$ *so that for any* $K > K_0$ *and any* $f : \Omega \to \mathbb{R}$ *with* $|||\, f \,||| < \infty$,

$$\sup_{\eta,\xi\in\Omega} \int_0^\infty \varphi\left(\frac{t}{K}\right) \left| S_t^{EP} f(\eta) - S_t^{EP} f(\xi) \right| dt \leq C_a(K) \,|||\, f \,|||\,.$$

b) Suppose the coupling $\widehat{\mathbb{P}}^E$ *satisfies*

$$\int_0^\infty \varphi(t) t^d \sum_{x\in\mathbb{Z}^d} \sup_{(\eta,\xi)\in(\Omega\times\Omega)_0} \widehat{\mathbb{E}}^E_{\eta,\xi} \rho(\eta_t^1(x), \eta_t^2(x))\, dt < \infty.$$

Then there exists a constant $K_0 > 0$ *and a decreasing function* $C_b :]K_0, \infty[\to [0, \infty[$ *so that for any* $K > K_0$ *and any* $f : \Omega \to \mathbb{R}$ *with* $|||\, f \,||| < \infty$,

$$\sum_{x\in\mathbb{Z}^d} \sup_{(\eta,\xi)\in(\Omega\times\Omega)_x} \int_0^\infty \varphi\left(\frac{t}{K}\right) \left| S_t^{EP} f(\eta) - S_t^{EP} f(\xi) \right| dt \leq C_b(K) |||\, f \,|||.$$

Canonical choices for φ are $\varphi(t) = \exp(\beta t^\alpha), 0 < \alpha \leq 1$, or $\varphi(t) = (1 + t)^\beta$, $\beta > 0$. This leads to the following transfer of convergence speed to equilibrium from the environment to the environment process:

- exponential decay: $e^{-\lambda t} \longrightarrow e^{-\frac{\lambda}{K_0+\epsilon} t}$,
- stretched exponential decay: $e^{-\lambda t^\alpha} \longrightarrow e^{-\frac{\lambda}{(K_0+\epsilon)^\alpha} t^\alpha}$,
- polynomial decay: $t^{-\lambda} \longrightarrow t^{-(\lambda-d-\epsilon)}$,

with $\epsilon > 0$ arbitrary, and, in the case of polynomial decay, $\lambda > d + 1$.

5.3.5 Consequences for the walker

The strong convergence of the environment process to its stationary measure obtained in Theorem 5.3.1 implies various facts for the random walker. The most basic fact is that the random walker has a limiting speed, i.e., $\mathbb{P}_\eta$-a.s.

$$v := \lim_{t\to\infty} \frac{X_t}{t} = \int \sum_{z\in\mathbb{Z}^d} z\alpha(\eta, z)\, \mu^{EP}(d\eta), \quad \text{uniformly in } \eta.$$

The convergence is a direct consequence of ergodicity.

In the following theorem we prove the functional central limit theorem for the position of the walker. The convergence to Brownian motion via martingales is a rather straightforward consequence of the ergodicity given by Theorem 5.3.1. The issue of non-degeneracy of the variance is less standard and hence we give a proof.

Theorem 5.3.5. *Assume Assumption 1a), $\| \alpha \|_2 < \infty$, $||| \alpha |||_1 < \infty$. Then the scaling limit of the random walk is a Brownian motion with drift v, i.e.,*

$$\frac{X_{tT} - vtT}{\sqrt{T}} \underset{T \to \infty}{\longrightarrow} W_D(t),$$

where W_D is a Brownian motion with covariance matrix D.

Let $e \in \mathbb{R}^d$ be a unit vector. Assume that either

a) there exists a $z \in \mathbb{Z}^d, \langle e, z \rangle \neq 0$, so that for all $t > 0$ and $\eta \in \Omega$ the probability $\mathbb{P}_\eta(\alpha(\eta_t, z) > 0)$ is positive;

b) $\mu^{EP}(\alpha(\cdot, z)) > 0$ for $z \in \mathbb{Z}^d$ with $\langle e, z \rangle$ arbitrary large.

Then $\lim_{T\to\infty} \frac{1}{T} \operatorname{Var}(\langle X_T, e \rangle) > 0$. In particular, if i) or ii) is satisfied for all e then the covariance matrix D is non-degenerate.

Proof. Notice that $\sum_{z\in\mathbb{Z}^d} z\alpha(\cdot, z) - v$ is in the domain of $(L^{EP})^{-1}$ because of Theorem 5.3.1. Decompose

$$\begin{aligned} X_t - vt = &\left(X_t - \int_0^t \sum_{z\in\mathbb{Z}^d} z\alpha(\theta_{-X_s}\eta_s, z)\, ds \right) \\ &+ \left(\int_0^t \sum_{z\in\mathbb{Z}^d} z \left[\alpha(\theta_{-X_s}\eta_s, z) - \mu^{EP}(\alpha(\cdot, z)) \right] \, ds \right). \end{aligned}$$

The first term on the right hand side is a martingale and the second one is one as well up to a uniformly bounded error. Both converge to Brownian motion with finite variance by standard arguments when $\| \alpha \|_2 < \infty$. However, as the two terms are not independent, an argument is needed to prove that they do not annihilate. To prove that we show that $\frac{1}{T} \operatorname{Var}(\langle X_T, e \rangle)$ is bounded away from 0 under the assumed conditions. Assume $T > 0$ integer and let $(\mathcal{F}_t)_{t\geq 0}$

be the canonical filtration. Introduce the discrete-time martingale

$$\begin{aligned} M_n^T &:= \mathbb{E}[\langle X_T, e\rangle|\mathcal{F}_n] - \mathbb{E}[\langle X_T, e\rangle|\mathcal{F}_0] = X_n + \Psi_{T-n}(\theta_{-X_n}\eta_n) - \Psi_t(\eta_0), \\ \Psi_S(\eta) &:= \mathbb{E}_{0,\eta}\int_0^S \sum_{z\in\mathbb{Z}^d} \langle z, e\rangle \alpha(\theta_{-X_t}\eta_t, z)\,dt = \int_0^S S_t^{EP}\varphi(\eta)\,dt; \\ \varphi(\eta) &:= \sum_{z\in\mathbb{Z}^d} \langle z, e\rangle \alpha(\eta, z). \end{aligned}$$

With this, by stationarity of the environment process started from μ^{EP},

$$\begin{aligned} &\mathrm{Var}_{\mu^{EP}}(\langle X_T, e\rangle) \geq \mathbb{E}_{\mu^{EP}}\left(\langle X_T, e\rangle - \mathbb{E}[\langle X_T, e\rangle|\mathcal{F}_0]\right)^2 \\ &= \sum_{n=1}^T \mathbb{E}_{\mu^{EP}}\left(M_n - M_{n-1}\right)^2 \\ &= \sum_{n=1}^T \mathbb{E}_{\mu^{EP}}\big(\langle X_n, e\rangle - \langle X_{n-1}, e\rangle + \Psi_{T-n}(\theta_{-X_n}\eta_n) \\ &\qquad\qquad - \Psi_{T-(n-1)}(\theta_{-X_{n-1}}\eta_{n-1})\big)^2 \\ &= \sum_{n=1}^T \mathbb{E}_{\mu^{EP}}\left(\langle X_1, e\rangle + \Psi_{T-n}(\theta_{-X_1}\eta_1) - \Psi_{T-(n-1)}(\eta_0)\right)^2. \end{aligned}$$

What has to be shown is that the above term is not 0. By Theorem 5.3.1 and $\|\,\varphi\,\|_\infty \leq |||\,\alpha\,|||_1 < \infty$,

$$\sup_{\eta,\xi\in\Omega}\sup_{T\geq 0} |\,\Psi_T(\xi) - \Psi_{T+1}(\eta)\,| =: C < \infty.$$

Therefore, using $|\,a+b\,| \geq |\,|a| - |b|\,|$,

$$Var_{\mu^{EP}}(\langle X_T, e\rangle) \geq T\mathbb{E}_{\mu^{EP}}\mathbb{1}_{|\,\langle X_1, e\rangle\,|>C}\left(|\,\langle X_1, e\rangle\,| - C\right)^2.$$

What remains to show is that $\mathbb{P}_{\mu^{EP}}(|\,\langle X_1, e\rangle\,| > C) > 0$. If ii) is satisfied this is immediate. If i) is satisfied then there is a positive probability that X_t performs sufficiently many jumps of size z (and no other jumps) up to time 1. □

Remark The convergence to Brownian motion with a non-degenerate variance also provides information about the recurrence behaviour of the walker. If $v = 0$, supposing $d = 1$ (in higher dimensions, project onto a line), the limiting Brownian motion is centred. Hence there exists an infinite sequence $t_1 < t_2 < \dots$ of times with $X_{t_{2n}} < 0$ and $X_{t_{2n+1}} > 0$, $n \in \mathbb{N}$. Supposing the walker has only jumps of size 1, it will traverse the origin between t_n, t_{n+1} for any $n \in \mathbb{N}$, i.e., it is recurrent. (If the walker also has larger jumps, then one needs an argument to actually hit the origin with some positive probability in $[t_n, t_{n+1}]$.) Particularly, the recurrence implies that there exists no regime where the random walk is transient but with 0 speed.

5.4 Examples: Layered Environments

There are many examples of environments which satisfy both Assumption 1a and Assumption 1b. Naturally, exponential convergence to the ergodic measure is sufficient independent of the dimension d. Therefore interacting particle systems in the so-called $M < \epsilon$-regime or weakly interacting diffusions on a compact manifold belong to the environments to which this method is applicable.

To exploit the fact that only sufficient polynomial decay of correlations is required we will construct a class of environments which we call *layered environments*. One can think of layered environments as a weighted superposition of a sequence of independent environments. Those kind of environments are fairly natural objects to study. Suppose the random walk moves only on a single layer but switches layers at a very high rate. Then in the limit of the speed of layer jumping going to infinity one obtains a random walk which sees a superposition of the layers weighted according to the frequency of visits. A second scenario where layered environments are relevant is when there is a natural reduction of influence, say by increasing distance of the layers to the walker.

Here we focus on layers which still have exponential decay of correlations, but each layer does converge to its stationary measure at a layer specific rate α_n, with n being the index of the layer. When α_n tends to 0 as $n \to \infty$ this introduces some form of arbitrary slow decay of correlations. We counterbalance this by weighting the superposition in such a way that the individual influence of a layer goes to 0 as well. Note that such a counterbalancing is only possible because of the Lipschitz nature of the assumptions. A uniform decay estimate does not hold because of the arbitrary slow decay in deep layers.

More formally, for each $n \in \mathbb{N}$ let $(\eta^n)_{t\geq 0}$ be a Markov process on $\Omega_0 := \{0,1\}^{\mathbb{Z}^d}$, the environment on layer n. This process should have a coupling $\widehat{\mathbb{P}}^n_{\eta,\xi}$ with

$$\sup_{\eta,\xi\in\Omega_0} \widehat{\mathbb{E}}^n_{\eta,\xi} \left| \eta^{n,1}_t(0) - \eta^{n,2}_t(0) \right| \leq 2e^{-\alpha_n t}, \quad \alpha_n > 0. \tag{5.7}$$

The *layered environment* $(\eta_t)_{t\geq 0}$ then consists of the stack of independent layers $(\eta^n_t)_{t\geq 0}$. The single site state space is $E = \{0,1\}^{\mathbb{N}}$ and space of all configurations $\Omega = E^{\mathbb{Z}^d}$.

The superposition of the environments is weighted by the distance ρ on E, which we choose in the following way. Fix a sequence $\gamma_1 > \gamma_2 > ... > 0$ with $\sum_{n\in\mathbb{N}} \gamma_n = 1$. For $(a_n)_{n\in\mathbb{N}}, (b_n)_{n\in\mathbb{N}} \in E$ the distance is

$$\rho((a_n),(b_n)) := \sum_{n\in\mathbb{N}} \gamma_n \left| a_n - b_n \right|. \tag{5.8}$$

The coupling $\widehat{\mathbb{P}}^E$ of the layered environments is simply the independent coupling of the individual layer couplings $\widehat{\mathbb{P}}^n$. The layer decay (5.7) and the choice of distance (5.8) then provide the following decay of coupling distance for the layered environment:

$$\sup_{\eta,\xi\in\Omega} \widehat{\mathbb{E}}^E_{\eta,\xi} \rho(\eta^1_t(0), \eta^2_t(0)) \leq 2 \sum_{n\in\mathbb{N}} \gamma_n e^{-\alpha_n t}. \tag{5.9}$$

The sum on the right hand side of (5.9) can have arbitrary slow decay depending on α_n, γ_n. For example, if one fixes $\alpha_n = n^{-1}$, then $\gamma_n = n^{-\gamma-1}$ leads to decay of order $t^{-\gamma}$.

We did not specify the exact nature of the individual layers, as those did not matter for the construction. A natural example is when individual layers consist of Ising model Glauber dynamics at inverse temperature $\beta_n < \beta_c$, and $\beta_n \to \beta_c$ as $n \to \infty$.

5.5 Proofs

In this section we always assume that Assumption 1a holds.

We start with an outline of the idea of the proofs. We have a coupling of the environments (η_t^1, η_t^2), which we extend to include two random walkers (X_t^1, X_t^2), driven by their corresponding environment. We maximize the probability of both walkers performing the same jumps. Then Assumption 1a is sufficient to obtain a positive probability of both walkers staying together forever. If the walkers stay together, one just has to account for the difference in environments, but not the walkers as well. When the walkers split, the translation invariance allows for everything to shifted that both walkers are back at the origin, and one can try again. After a geometric number of trials it is then guaranteed that the walkers stay together.

Proposition 5.5.1 (Coupling construction)**.** *Given the coupling $\widehat{\mathbb{P}}^E_{\eta,\xi}$ of the environments, we extend it to a coupling $\widehat{\mathbb{P}}_{\eta,x;\xi,y}$. This coupling has the following properties:*

a) (Marginals)The coupling supports two environments and corresponding random walkers:

a) $\widehat{\mathbb{P}}_{\eta,x;\xi,y}((\eta_t^1, X_t^1) \in \cdot) = \mathbb{P}_{\eta,x}((\eta_t, X_t) \in \cdot)$;

b) $\widehat{\mathbb{P}}_{\eta,x;\xi,y}((\eta_t^2, X_t^2) \in \cdot) = \mathbb{P}_{\xi,y}((\eta_t, X_t) \in \cdot)$;

b) (Extension of $\widehat{\mathbb{P}}^E_{\eta,\xi}$) The environments behave as under $\widehat{\mathbb{P}}^E$:

$$\widehat{\mathbb{P}}_{\eta,x;\xi,y}((\eta_t^1, \eta_t^2) \in \cdot) = \widehat{\mathbb{P}}^E_{\eta,\xi}((\eta_t^1, \eta_t^2) \in \cdot);$$

c) (Coupling of the walkers) X_t^1 and X_t^2 perform identical jumps as much as possible, the rate of performing a different jump is

$$\sum_{z\in\mathbb{Z}^d} \left| \alpha(\theta_{-X_t^1}\eta_t^1, z) - \alpha(\theta_{-X_t^2}\eta_t^2, z) \right|;$$

d) (Minimal and maximal walkers) In addition to the environments η_t^1 and η_t^2 and random walkers X_t^1 and X_t^2, the coupling supports minimal and maximal walkers Y_t^+, Y_t^- as well. These two walkers have the following properties:

a) $Y_t^- \leq X_t^1 - x, X_t^2 - y \leq Y_t^+$ $\widehat{\mathbb{P}}_{\eta,x;\xi,y} - a.s.$ (in dimension $d > 1$, this is to be interpreted coordinate-wise);

b) Y_t^+, Y_t^- are independent of η_t^1, η_t^2;

c) $\widehat{\mathbb{E}}_{\eta,x;\xi,y} Y_t^+ = t\gamma^+$ *for some* $\gamma^+ \in \mathbb{R}^d$*;*

d) $\widehat{\mathbb{E}}_{\eta,x;\xi,y} Y_t^- = t\gamma^-$ *for some* $\gamma^- \in \mathbb{R}^d$*.*

Proof. The construction of this coupling $\widehat{\mathbb{P}}_{\eta,x;\xi,y}$ is done in the following way: We extend the original coupling $\widehat{\mathbb{P}}^E_{\eta,\xi}$ to contain an independent sequence of Poisson processes $N^z, z \in \mathbb{Z}^d$, with rates $\lambda_z := \sup_\eta \alpha(\eta, z)$, as well as sufficient supply of independent uniform $[0,1]$ variables. The walkers X^1, X^2 then start from x resp. y and exclusively (but not necessarily) jump when one of the Poisson clocks N^z rings. When the clock N^z rings the walkers jumps from X_t^i to $X_t^i + z$ only if a uniform $[0,1]$ variable U satisfies $U < \alpha(\theta_{-X_t^1}\eta_t^1, z)/\lambda_z$, $i = 1, 2$. Note that both walkers share the same U, but Us for different rings of the Poisson clocks are independent.

The upper and lower walkers Y_t^+, Y_t^- are constructed from the same Poisson clocks N^z. They always jump on these clocks, however they jump by $\max(z, 0)$ or $\min(z, 0)$ respectively.

The properties of the coupling arise directly from the construction plus the fact that $\| \alpha \|_1 < \infty$. $\square$

To ease notation we will call $\widehat{\mathbb{P}}_{\eta,0;\xi,0}$ simply $\widehat{\mathbb{P}}_{\eta,\xi}$ and the law of Y_t^+, Y_t^- $\widehat{\mathbb{P}}$ whenever there is no fear of confusion.

Now we show how suitable estimates on the coupling speed of the environment translate to properties of the extended coupling.

Lemma 5.5.2.

$$\widehat{\mathbb{E}}_{\eta,x;\xi,y} \rho(\eta_t^1(X_t^1), \eta_t^2(X_t^1)) < (\| \gamma^+ - \gamma^- \|_\infty t + 1)^d \sup_{\eta,\xi \in \Omega} \widehat{\mathbb{E}}^E_{\eta,\xi} \rho(\eta_t^1(0), \eta_t^2(0)).$$

Proof. Denote with $R_t \subset \mathbb{Z}^d$ the set of sites $z \in \mathbb{Z}^d$ with $Y_t^- \leq z \leq Y_t^+$ (coordinate-wise). Then

$$\begin{aligned}
&\sup_{\eta,\xi,x,y} \widehat{\mathbb{E}}_{\eta,x;\xi,y} \rho(\eta_t^1(X_t^1), \eta_t^2(X_t^1)) \\
&\quad \leq \sup_{\eta,\xi,x,y} \widehat{\mathbb{E}}_{\eta,x;\xi,y} \sum_{z \in R_t} \rho(\eta_t^1(x+z), \eta_t^2(x+z)) \\
&\quad \leq \widehat{\mathbb{E}} \left[\sum_{z \in R_t} 1 \right] \sup_{\eta,\xi,z} \widehat{\mathbb{E}}^E_{\eta,\xi} \rho(\eta_t^1(z), \eta_t^2(z)) \\
&\quad \leq (\| \gamma^+ - \gamma^- \|_\infty t + 1)^d \sup_{\eta,\xi \in \Omega} \widehat{\mathbb{E}}^E_{\eta,\xi} \rho(\eta_t^1(0), \eta_t^2(0)).
\end{aligned}$$

$\square$

Lemma 5.5.3. *Denote by* $\tau := \inf\{t \geq 0 : X_t^1 \neq X_t^2\}$ *the first time the two walkers are not at the same position. Under Assumption 1a,*

$$\inf_{\eta,\xi\in\Omega} \widehat{\mathbb{P}}_{\eta;\xi}(\tau = \infty) > 0,$$

i.e., the walkers X^1 *and* X^2 *never decouple with strictly positive probability.*

Proof. Both walkers start in the origin, therefore $\tau > 0$. The probability that a Poisson clock with time dependent rate λ_t is has not yet rung by time T is $\exp(-\int_0^T \lambda_t \, dt)$. As the rate of decoupling is given by Proposition 5.5.1,c), we obtain

$$\widehat{\mathbb{P}}_{\eta,\xi}(\tau > T) = \widehat{\mathbb{E}}_{\eta,\xi} \exp\left(- \int_0^T \sum_{z\in\mathbb{Z}^d} \left| \alpha(\theta_{-X_t^1}\eta_t^1, z) - \alpha(\theta_{-X_t^1}\eta_t^2, z) \right| dt \right) \tag{5.10}$$
$$\geq \exp\left(-\widehat{\mathbb{E}}_{\eta,\xi} \int_0^T \sum_{z\in\mathbb{Z}^d} \left| \alpha(\theta_{-X_t^1}\eta_t^1, z) - \alpha(\theta_{-X_t^1}\eta_t^2, z) \right| dt \right).$$

By telescoping over single site changes,

$$\begin{aligned}
&\widehat{\mathbb{E}}_{\eta,\xi} \sum_{z\in\mathbb{Z}^d} \left| \alpha(\theta_{-X_t^1}\eta_t^1, z) - \alpha(\theta_{-X_t^1}\eta_t^2, z) \right| \\
&\leq \widehat{\mathbb{E}}_{\eta,\xi} \sum_{z\in\mathbb{Z}^d} \sum_{x\in\mathbb{Z}^d} \rho(\eta_t^1(X_t^1 + x), \eta_t^2(X_t^1 + x)) \delta_{\alpha(\cdot,z)}(x) \\
&\leq \sup_{x\in\mathbb{Z}^d} \widehat{\mathbb{E}}_{\eta,\xi} \rho(\eta_t^1(X_t^1 + x), \eta_t^2(X_t^1 + x)) \, ||| \, \alpha \, ||| \\
&\leq ||| \, \alpha \, ||| \, (\left\| \gamma^+ - \gamma^- \right\|_\infty t + 1)^d \sup_{\eta,\xi\in\Omega} \widehat{\mathbb{E}}^E_{\eta,\xi} \rho(\eta_t^1(0), \eta_t^2(0)),
\end{aligned}$$

where the last line follows from Lemma 5.5.2. With this estimate and Assumption 1a we obtain

$$\begin{aligned}
&\widehat{\mathbb{P}}_{\eta,\xi}(\tau = \infty) \\
&\geq \exp\left(- ||| \, \alpha \, ||| \int_0^\infty (\left\| \gamma^+ - \gamma^- \right\|_\infty t + 1)^d \sup_{\eta,\xi\in\Omega} \widehat{\mathbb{E}}^E_{\eta,\xi} \rho(\eta_t^1(0), \eta_t^2(0)) \, dt \right) \\
&> 0 \quad \text{uniformly in } \eta, \xi.
\end{aligned}$$

□

Proof of Theorem 5.3.1, part a). The idea of the proof is to use the coupling of Proposition 5.5.1: We wait until the walkers X_t^1 and X_t^2, which are initially at the same position, decouple, and then restart everything and try again. By Lemma 5.5.3 there is a positive probability of never decoupling, so this scheme is successful. Using the time of decoupling τ (as in Lemma 5.5.3) and the strong marginal Markov property (5.6),

$$
\begin{aligned}
&\int_0^T \left| \widehat{\mathbb{E}}_{\eta,0;\xi,0} \mathbb{1}_{t\geq\tau} \left(f(\theta_{-X_t^1}\eta_t^1) - f(\theta_{-X_t^2}\eta_t^2) \right) \right| \, dt \\
&\quad = \int_0^T \left| \widehat{\mathbb{E}}_{\eta 0;\xi,0} \mathbb{1}_{t\geq\tau} \mathbb{E} \left[f(\theta_{-X_t^1}\eta_t^1) - f(\theta_{-X_t^2}\eta_t^2) \,\middle|\, \mathfrak{F}_\tau \right] \right| \, dt \\
&\quad \leq \int_0^T \widehat{\mathbb{E}}_{\eta,0;\xi,0} \mathbb{1}_{t\geq\tau} \left| S_{t-\tau}^{EP} f(\theta_{-X_\tau^1}\eta_\tau^1) - S_{t-\tau}^{EP} f(\theta_{-X_\tau^2}\eta_\tau^2) \right| \, dt \\
&\quad = \widehat{\mathbb{E}}_{\eta,0;\xi,0} \int_0^{(T-\tau)\vee 0} \left| S_t^{EP} f(\theta_{-X_\tau^1}\eta_\tau^1) - S_t^{EP} f(\theta_{-X_\tau^2}\eta_\tau^2) \right| \, dt && (5.11) \\
&\quad \leq \widehat{\mathbb{P}}_{\eta,0;\xi,0}(\tau<\infty) \sup_{\eta,\xi\in\Omega} \int_0^T \left| S_t^{EP} f(\eta) - S_t^{EP} f(\xi) \right| \, dt. && (5.12)
\end{aligned}
$$

And therefore

$$
\begin{aligned}
&\int_0^T \left| S_t^{EP} f(\eta) - S_t^{EP} f(\xi) \right| \, dt \\
&\quad = \int_0^T \left| \widehat{\mathbb{E}}_{\eta,\xi} f(\theta_{-X_t^1}\eta_t^1) - f(\theta_{-X_t^2}\eta_t^2) \right| \, dt \\
&\quad \leq \int_0^T \widehat{\mathbb{E}}_{\eta,\xi} \mathbb{1}_{t<\tau} \left| f(\theta_{-X_t^1}\eta_t^1) - f(\theta_{-X_t^1}\eta_t^2) \right| \, dt \\
&\qquad + \widehat{\mathbb{P}}_{\eta,\xi}(\tau<\infty) \sup_{\eta,\xi\in\Omega} \int_0^T \left| S_t^{EP} f(\eta) - S_t^{EP} f(\xi) \right| \, dt && (5.13) \\
&\quad \leq \int_0^\infty \widehat{\mathbb{E}}_{\eta,\xi} \left| f(\theta_{-X_t^1}\eta_t^1) - f(\theta_{-X_t^1}\eta_t^2) \right| \, dt \\
&\qquad + \widehat{\mathbb{P}}_{\eta,\xi}(\tau<\infty) \sup_{\eta,\xi\in\Omega} \int_0^T \left| S_t^{EP} f(\eta) - S_t^{EP} f(\xi) \right| \, dt, && (5.14)
\end{aligned}
$$

which gives us the upper bound

$$\sup_{\eta,\xi\in\Omega}\int_0^\infty \left| S_t^{EP} f(\eta) - S_t^{EP} f(\xi) \right| dt$$
$$\leq \left(\inf_{\eta,\xi\in\Omega} \widehat{\mathbb{P}}_{\eta,\xi}(\tau=\infty)\right)^{-1} \sup_{\eta,\xi\in\Omega}\int_0^\infty \widehat{\mathbb{E}}_{\eta,\xi}\left| f(\theta_{-X_t^1}\eta_t^1) - f(\theta_{-X_t^1}\eta_t^2)\right| dt. \tag{5.15}$$

To show that the last integral is finite, we telescope over single site changes, and get

$$\int_0^\infty \widehat{\mathbb{E}}_{\eta,\xi}\left| f(\theta_{-X_t^1}\eta_t^1) - f(\theta_{-X_t^1}\eta_t^2)\right| dt$$
$$\leq \int_0^\infty \widehat{\mathbb{E}}_{\eta,\xi}\sum_{x\in\mathbb{Z}^d}\rho(\eta_t^1(x+X_t^1),\eta_t^2(x+X_t^1))\delta_f(x)\,dt$$
$$\leq \||\, f \,\|| \sup_{\eta,\xi,x}\int_0^\infty \widehat{\mathbb{E}}_{\eta,\xi}\rho(\eta_t^1(x+X_t^1),\eta_t^2(x+X_t^1))\,dt,$$

which is finite by Lemma 5.5.2 and Assumption 1a). Choosing

$$C_a = \left(\inf_{\eta,\xi\in\Omega} \widehat{\mathbb{P}}_{\eta,\xi}(\tau=\infty)\right)^{-1} \sup_{\eta,\xi,x}\int_0^\infty \widehat{\mathbb{E}}_{\eta,\xi}\rho(\eta_t^1(x+X_t^1),\eta_t^2(x+X_t^1))\,dt \tag{5.16}$$

completes the proof. □

To prove part b) of the theorem, we need the following analogue to Lemma 5.5.3 using Assumption 1b.

Lemma 5.5.4. *Under Assumption 1b, for every site-weight function* $w:\mathbb{Z}^d\to[0,\infty[$ *with* $\| w \|_1 := \sum_x w(x) < \infty$, *we have*

$$\sum_{x\in\mathbb{Z}^d}\sup_{(\eta,\xi)\in(\Omega\times\Omega)_x}\int_0^\infty\sum_{y\in\mathbb{Z}^d} w(y)\widehat{\mathbb{E}}_{\eta,\xi}\rho(\eta_t^1(y+X_t^1),\eta_t^2(y+X_t^1))\,dt \leq const\cdot \| w \|_1 .$$

Proof. Denote with $R_t\subset\mathbb{Z}^d$ the set of sites whose jth coordinate lies between

$Y_t^{j,-}$ and $Y_t^{j,+}$. Then,

$$\begin{aligned}
&\sum_{y\in\mathbb{Z}^d} w(y)\widehat{\mathbb{E}}_{\eta,\xi}\rho(\eta_t^1(y+X_t^1),\eta_t^2(y+X_t^1)) \\
&\quad = \sum_{y\in\mathbb{Z}^d} \widehat{\mathbb{E}}_{\eta,\xi} w(y-X_t^1)\rho(\eta_t^1(y),\eta_t^2(y)) \\
&\quad \le \sum_{y\in\mathbb{Z}^d} \widehat{\mathbb{E}}_{\eta,\xi} \sum_{z\in R_t} w(y-z)\rho(\eta_t^1(y),\eta_t^2(y)) \\
&\quad = \sum_{y\in\mathbb{Z}^d} \widehat{\mathbb{E}}\left[\sum_{z\in R_t} w(y-z)\right]\widehat{\mathbb{E}}^E_{\eta,\xi}\rho(\eta_t^1(y),\eta_t^2(y))
\end{aligned}$$

by independence of R_t and (η_t^1,η_t^2). Therewith,

$$\begin{aligned}
&\sum_{x\in\mathbb{Z}^d}\sup_{(\eta,\xi)\in(\Omega\times\Omega)_x}\int_0^\infty \sum_{y\in\mathbb{Z}^d} w(y)\widehat{\mathbb{E}}_{\eta,\xi}\rho(\eta_t^1(y+X_t^1),\eta_t^2(y+X_t^1))\,dt \\
&\quad \le \int_0^\infty \sum_{y\in\mathbb{Z}^d}\widehat{\mathbb{E}}\left[\sum_{z\in R_t} w(y-z)\right]\sum_{x\in\mathbb{Z}^d}\sup_{(\eta,\xi)\in(\Omega\times\Omega)_x}\widehat{\mathbb{E}}^E_{\eta,\xi}\rho(\eta_t^1(y),\eta_t^2(y))\,dt.
\end{aligned}$$

Note that by translation invariance the right part is equal to

$$\sum_{x\in\mathbb{Z}^d}\sup_{(\eta,\xi)\in(\Omega\times\Omega)_0}\widehat{\mathbb{E}}^E_{\eta,\xi}\rho(\eta_t^1(x),\eta_t^2(x)).$$

By construction of R_t and Proposition 5.5.1.d ,

$$\begin{aligned}
\sum_{y\in\mathbb{Z}^d}\widehat{\mathbb{E}}\left[\sum_{z\in R_t} w(y-z)\right] &= \widehat{\mathbb{E}}\left[\sum_{z\in R_t} 1\right]\|\, w\,\|_1 = \prod_{j=1}^d(\gamma^{j,+}t-\gamma^{j,-}t+1)\,\|\, w\,\|_1 \\
&\le c(t^d+1)\,\|\, w\,\|_1
\end{aligned}$$

for some suitable $c>0$. Therefore Assumption 1b completes the proof. $\square$

Proof of Theorem 5.3.1, part b). Let $\tau := \inf\{t\ge 0 : X_t^1 \ne X_t^2\}$. Then we split the integration at τ:

$$\sum_{x\in\mathbb{Z}^d}\sup_{(\eta,\xi)\in(\Omega\times\Omega)_x}\int_0^\infty \left|\, S_t^{EP}f(\eta)-S_t^{EP}f(\xi)\,\right|\,dt$$

$$\leq \sum_{x\in\mathbb{Z}^d} \sup_{(\eta,\xi)\in(\Omega\times\Omega)_x} \int_0^\infty \left| \widehat{\mathbb{E}}_{\eta,\xi} \mathbb{1}_{\tau>t} \left(f(\theta_{-X_t^1}\eta_t^1) - f(\theta_{-X_t^1}\eta_t^2) \right) \right| dt$$
$$+ \sum_{x\in\mathbb{Z}^d} \sup_{(\eta,\xi)\in(\Omega\times\Omega)_x} \int_0^\infty \left| \widehat{\mathbb{E}}_{\eta,\xi} \mathbb{1}_{\tau\leq t} \left(f(\theta_{-X_t^1}\eta_t^1) - f(\theta_{-X_t^2}\eta_t^2) \right) \right| dt$$

We estimate the first term by moving the expectation out of the absolute value and forgetting the restriction to $\tau > t$:

$$\sum_{x\in\mathbb{Z}^d} \sup_{(\eta,\xi)\in(\Omega\times\Omega)_x} \int_0^\infty \sum_{y\in\mathbb{Z}^d} \delta_f(y)\widehat{\mathbb{E}}_{\eta,\xi}\rho(\eta_t^1(y+X_t^1), \eta_t^2(y+X_t^1))\, dt.$$

By Lemma 5.5.4 with $w = \delta_f$, this is bounded by some constant times $|||\, f\,|||$. For the second term we start by using the strong marginal Markov property (5.6):

$$\begin{aligned}
&\int_0^\infty \left| \widehat{\mathbb{E}}_{\eta,\xi} \mathbb{1}_{\tau\leq t} \left(f(\theta_{-X_t^1}\eta_t^1) - f(\theta_{-X_t^2}\eta_t^2) \right) \right| dt \\
&= \int_0^\infty \left| \widehat{\mathbb{E}}_{\eta,\xi} \mathbb{1}_{\tau\leq t} \left(S_{t-\tau}^{EP} f(\theta_{-X_\tau^1}\eta_\tau^1) - S_{t-\tau}^{EP} f(\theta_{-X_\tau^2}\eta_\tau^2) \right) \right| dt \\
&\leq \widehat{\mathbb{E}}_{\eta,\xi} \mathbb{1}_{\tau<\infty} \int_\tau^\infty \left| \left(S_{t-\tau}^{EP} f(\theta_{-X_\tau^1}\eta_\tau^1) - S_{t-\tau}^{EP} f(\theta_{-X_\tau^2}\eta_\tau^2) \right) \right| dt && (5.17) \\
&\leq \widehat{\mathbb{P}}_{\eta,\xi}(\tau<\infty) \sup_{\eta,\xi\in\Omega} \int_0^\infty \left| S_t^{EP} f(\eta) - S_t^{EP} f(\xi) \right| dt. && (5.18)
\end{aligned}$$

By part a) of Theorem 5.3.1 the integral part is uniformly bounded by $C_a\, |||\, f\,|||$. So what remains to complete the proof is to show that

$$\sum_{x\in\mathbb{Z}^d} \sup_{(\eta,\xi)\in(\Omega\times\Omega)_x} \widehat{\mathbb{P}}_{\eta,\xi}(\tau<\infty) < \infty. \tag{5.19}$$

To do so we first use the same idea as in the proof of Lemma 5.5.3 to obtain

$$\begin{aligned}
&\widehat{\mathbb{P}}_{\eta,\xi}(\tau<\infty) \\
&\quad = 1 - \exp\left(- \int_0^\infty \widehat{\mathbb{E}}_{\eta,\xi} \sum_{z\in\mathbb{Z}^d} \left| \alpha(\theta_{-X_t^1}\eta_t^1, z) - \alpha(\theta_{-X_t^1}\eta_t^2, z) \right| dt \right)
\end{aligned}$$

$$\leq \int_0^\infty \widehat{\mathbb{E}}_{\eta,\xi} \sum_{z\in\mathbb{Z}^d} \Big| \alpha(\theta_{-X_t^1}\eta_t^1, z) - \alpha(\theta_{-X_t^1}\eta_t^2, z) \Big| \, dt$$
$$\leq \int_0^\infty \sum_{y\in\mathbb{Z}^d} w_\alpha(y) \widehat{\mathbb{E}}_{\eta,\xi} \rho(\eta_t^1(y+X_t^1), \eta_t^2(y+X_t^1)) \, dt$$

with

$$w_\alpha(x) := \sup_{(\eta,\xi)\in(\Omega\times\Omega)_x} \sum_{z\in\mathbb{Z}^d} |\, \alpha(\eta,z) - \alpha(\xi,z) \,|$$

and $\sum_{x\in\mathbb{Z}^d} w_\alpha(x) < \infty$. So we get

$$\sum_{x\in\mathbb{Z}^d} \sup_{(\eta,\xi)\in(\Omega\times\Omega)_x} \widehat{\mathbb{P}}_{\eta,\xi}(\tau<\infty)$$
$$\leq \sum_{x\in\mathbb{Z}^d} \sup_{(\eta,\xi)\in(\Omega\times\Omega)_x} \int_0^\infty \sum_{y\in\mathbb{Z}^d} w_\alpha(y) \widehat{\mathbb{E}}_{\eta,\xi} \rho(\eta_t^1(y+X_t^1), \eta_t^2(y+X_t^1)) \, dt,$$

and Lemma 5.5.4 completes the proof, where C_b is the combination of the various factors in front of $|||\, f \,|||$. □

Proof of Theorem 5.3.3. Let α, α' be two different transition rates. The goal is to show that

$$\big|\, \mu_\alpha^{EP}(f) - \mu_{\alpha'}^{EP}(f) \,\big| \leq C \; |||\, f \,|||$$

for all $f : \Omega \to \mathbb{R}$ with $|||\, f \,||| < \infty$.

The idea is now to use a coupling $\widehat{\mathbb{P}}$ similar to the one in Proposition 5.5.1. The coupling contains as objects two copies of the environment, η^1 and η^2, and three random walks, X^1, X^{12} and X^2. The random walk X^1 moves on the environment η^1 with rates α, and correspondingly the random walk X^2 moves on η^2 with rates α'. The mixed walker X^{12} moves on the environment η^2 as well, but according to the rates α. The walkers X^1, X^2 will perform the same jumps as X^{12} with maximal probability. This can be achieved with the same construction as in Proposition 5.5.1, but with Poisson clocks N^z which have rates $\lambda_z = \sup_{\eta\in\Omega} \alpha(\eta,z) \vee \alpha'(\eta,z)$.

We only consider the case where all three walkers start at the origin. We denote by $S_t^{EP,1}$, $S_t^{EP,2}$ the semigroups of the environment process which correspond to the rates α and α'. Let $\tau := \inf\{t \geq 0 : X_t^1 \neq X_t^{12} \text{ or } X_t^{12} \neq X_t^2\}$.

$$\begin{aligned}
&S_t^{EP,1} f(\eta) - S_t^{EP,2} f(\xi) \\
&\quad = \widehat{\mathbb{E}}_{\eta,\xi} \left(f(\theta_{-X_t^1}\eta_t^1) - f(\theta_{-X_t^2}\eta_t^2) \right) \\
&\quad = \widehat{\mathbb{E}}_{\eta,\xi} \mathbb{1}_{\tau > t} \left(f(\theta_{-X_t^1}\eta_t^1) - f(\theta_{-X_t^1}\eta_t^2) \right) \\
&\qquad + \widehat{\mathbb{E}}_{\eta,\xi} \mathbb{1}_{\tau \leq t} \left(f(\theta_{-X_t^1}\eta_t^1) - f(\theta_{-X_t^2}\eta_t^2) \right) \\
&\quad = \widehat{\mathbb{E}}_{\eta,\xi} \mathbb{1}_{\tau > t} \left(f(\theta_{-X_t^1}\eta_t^1) - f(\theta_{-X_t^1}\eta_t^2) \right) \\
&\qquad + \widehat{\mathbb{E}}_{\eta,\xi} \mathbb{1}_{\tau \leq t} \left(S_{t-\tau}^{EP,1} f(\theta_{-X_\tau^1}\eta_\tau^1) - S_{t-\tau}^{EP,2} f(\theta_{-X_\tau^2}\eta_\tau^2) \right).
\end{aligned}$$

Therefore,

$$\begin{aligned}
&\Psi(T) := \sup_{0 \leq T' \leq T} \sup_{\eta,\xi \in \Omega} \int_0^{T'} S_t^{EP,1} f(\eta) - S_t^{EP,2} f(\xi)\, dt \\
&\leq \sup_{0 \leq T' \leq T} \sup_{\eta,\xi \in \Omega} \int_0^{T'} \widehat{\mathbb{E}}_{\eta,\xi} \mathbb{1}_{\tau > t} \left(f(\theta_{-X_t^1}\eta_t^1) - f(\theta_{-X_t^1}\eta_t^2) \right) \\
&\qquad + \widehat{\mathbb{E}}_{\eta,\xi} \mathbb{1}_{\tau \leq t} \sup_{\eta,\xi \in \Omega} \left(S_{t-\tau}^{EP,1} f(\eta) - S_{t-\tau}^{EP,2} f(\xi) \right) dt \\
&\leq \sup_{0 \leq T' \leq T} \sup_{\eta,\xi \in \Omega} \left(\widehat{\mathbb{E}}_{\eta,\xi} \int_0^{\tau} f(\theta_{-X_t^1}\eta_t^1) - f(\theta_{-X_t^1}\eta_t^2)\, dt + \mathbb{1}_{\tau \leq T'} \Psi(T' - \tau) \right) \\
&\leq \sup_{\eta,\xi \in \Omega} \widehat{\mathbb{E}}_{\eta,\xi} \left(\int_0^{\infty} f(\theta_{-X_t^1}\eta_t^1) - f(\theta_{-X_t^1}\eta_t^2)\, dt + \mathbb{1}_{\tau \leq T} \Psi(T - \tau) \right).
\end{aligned} \tag{5.20}$$

We will now exploit this recursive bound on Ψ.

Lemma 5.5.5. *Let $\tau_1 := \inf\{t \geq 0 : X_t^1 \neq X_t^{12}\}$ and $\tau_2 := \inf\{t \geq 0 : X_t^{12} \neq X_t^2\}$. Set*

$$\beta := \sum_{z \in \mathbb{Z}^d} \sup_{\eta \in \mathbb{Z}^d} |\, \alpha(\eta, z) - \alpha'(\eta, z)\, |,$$

$$p(\alpha) := \inf_{\eta,\xi\in\Omega} \widehat{\mathbb{P}}_{\eta,\xi}(\tau_1 = \infty),$$
$$C(\alpha) := \int_0^\infty \left(\|\,\gamma^+(\alpha) - \gamma^-(\alpha)\,\|_\infty \, t + 1\right)^d \sup_{\eta,\xi\in\Omega} \widehat{\mathbb{E}}^E_{\eta,\xi}\rho(\eta^1_t(0), \eta^2_t(0))\,dt,$$

where $\gamma^+(\alpha), \gamma^-(\alpha)$ are as in Proposition 5.5.1 for the rates α.

Let $Y \in \{0,1\}$ be Bernoulli with parameter $p(\alpha)$ and Y' exponentially distributed with parameter β. Let $Y_1, Y_2, \dots$ be iid. copies of $Y \cdot Y'$ and $N(T) := \inf\{N \geq 0 : \sum_{n=1}^N Y_n > T\}$. Then

$$\Psi(T) \leq C(\alpha)\,|||\; f\;|||\,\mathbb{E}N(T)$$

Proof. By construction of the coupling, τ_2 stochastically dominates Y'. As we have $\tau = \tau_1 \wedge \tau_2$ it follows that $\tau \succeq Y_1$. Using this fact together with the monotonicity of Ψ in (5.20),

$$\begin{aligned}\Psi(T) &\leq \sup_{\eta,\xi\in\Omega}\left(\widehat{\mathbb{E}}_{\eta,\xi}\int_0^\infty f(\theta_{-X^1_t}\eta^1_t) - f(\theta_{-X^1_t}\eta^2_t)\,dt + \mathbb{1}_{\tau\leq T}\Psi(T-\tau)\right)\\ &\leq \sup_{\eta,\xi\in\Omega}\widehat{\mathbb{E}}_{\eta,\xi}\int_0^\infty f(\theta_{-X^1_t}\eta^1_t) - f(\theta_{-X^1_t}\eta^2_t)\,dt + \mathbb{E}\;\mathbb{1}_{Y_1\leq T}\Psi(T-Y_1).\end{aligned}$$

As $p(\alpha) > 0$ by Lemma 5.5.3 we can iterate this estimate until it terminates after $N(T)$ steps. Therefore we obtain

$$\Psi(T) \leq \mathbb{E}N(T)\sup_{\eta,\xi\in\Omega}\widehat{\mathbb{E}}_{\eta,\xi}\int_0^\infty f(\theta_{-X^1_t}\eta^1_t) - f(\theta_{-X^1_t}\eta^2_t)\,dt.$$

The integral is estimated by telescoping over single site changes and Lemma 5.5.2 in the usual way, yielding

$$\Psi(T) \leq C(\alpha)\,|||\; f\;|||\,\mathbb{E}N(T). \qquad \square$$

To finally come back to the original question of continuity,

$$\begin{aligned}&\left|\,\mu^{EP}_\alpha(f) - \mu^{EP}_{\alpha'}(f)\,\right|\\ &\quad= \frac{1}{T}\left|\int\int\int_0^T S^{EP,1}_t f(\eta) - S^{EP,2}_t f(\xi)\,dt\,\mu^{EP}_\alpha(d\eta)\,\mu^{EP}_{\alpha'}(d\xi)\right|\end{aligned}$$

$$\leq \frac{1}{T}\Psi(T) \leq \frac{1}{T}\mathbb{E}N(T)C(\alpha) \mid\mid\mid f \mid\mid\mid$$
$$\underset{T\to\infty}{\longrightarrow} \frac{1}{\mathbb{E}YY'}C(\alpha) \mid\mid\mid f \mid\mid\mid$$
$$= \frac{C(\alpha)}{p(\alpha)} \sum_{z\in\mathbb{Z}^d} \sup_{\eta\in\mathbb{Z}^d} \mid \alpha(\eta, z) - \alpha'(\eta, z) \mid \mid\mid\mid f \mid\mid\mid .$$

By sending α' to α, the right hand side tends to 0 so that the ergodic measure of the environment process is indeed continuous in the rates α. It is also interesting to note that both $p(\alpha)$ and $C(\alpha)$ are rather explicit given the original coupling of the environment. Notably when $\alpha(\eta, z) = \alpha(z)$, i.e., the rates do not depend on the environment, $p(\alpha) = 1$ and $C(\alpha) = \int_0^\infty \sup_{\eta,\xi\in\Omega} \widehat{\mathbb{E}}^E_{\eta,\xi}\rho(\eta^1_t(0), \eta^2_t(0))\, dt$.

□

Proof of Theorem 5.3.4. The proof of this theorem is mostly identical to the proof of Theorem 5.3.1. Hence instead of copying the proof, we just state where details differ.

A first fact is that the conditions for a) and b) imply Assumptions 1a) and b). In the adaptation of the proof for part a), in most lines it suffices to add a $\varphi\left(\frac{t}{K}\right)$ to the integrals. However, in line (5.11), we use

$$\varphi\left(\frac{t}{K}\right) \leq \varphi\left(\frac{t-\tau}{K}\right)\varphi\left(\frac{\tau}{K}\right) \tag{5.21}$$

to obtain the estimate

$$\widehat{\mathbb{E}}_{\eta,\xi}\varphi\left(\frac{\tau}{K}\right)\int_0^{(T-\tau)\vee 0} \varphi\left(\frac{t}{K}\right) \left| S^{EP}_t f(\theta_{-X^1_\tau}\eta^1_\tau) - S^{EP}_t f(\theta_{-X^2_\tau}\eta^2_\tau) \right| dt$$

instead. Thereby in lines (5.12), (5.13) and (5.14) we have to change $\widehat{\mathbb{P}}_{\eta,\xi}(\tau < \infty)$ to $\widehat{\mathbb{E}}_{\eta,\xi}\varphi\left(\frac{\tau}{K}\right)\mathbb{1}_{\tau<\infty}$. This change then leads to the replacement of

$$\inf_{\eta,\xi\in\Omega} \widehat{\mathbb{P}}_{\eta,\xi}(\tau = \infty)$$

by the term

$$1 - \sup_{\eta,\xi\in\Omega} \widehat{\mathbb{E}}_{\eta,\xi}\varphi\left(\frac{\tau}{K}\right)\mathbb{1}_{\tau<\infty}$$

in the lines (5.15) and (5.16) (where naturally C_a becomes $C_a(K)$). So all we have to prove that for sufficiently big K

$$\sup_{\eta,\xi\in\Omega} \widehat{\mathbb{E}}_{\eta,\xi}\varphi\left(\frac{\tau}{K}\right)\mathbb{1}_{\tau<\infty} < 1.$$

In a first step, we show that

$$\sup_{\eta,\xi\in\Omega} \widehat{\mathbb{E}}_{\eta,\xi}\varphi(\tau)\mathbb{1}_{\tau<\infty} < \infty.$$

As we already saw in the proof of Lemma 5.5.3, we can view the event of decoupling as the first jump of a Poisson process with time-dependent and random rates (equation (5.10)). Hence we have

$$\begin{aligned}
\widehat{\mathbb{E}}_{\eta,\xi}\varphi(\tau)\mathbb{1}_{\tau<\infty} &= \int_0^\infty \varphi(t)\, d\widehat{\mathbb{P}}_{\eta,\xi}(\tau > t) \\
&= \int_0^\infty \varphi(t)\widehat{\mathbb{E}}_{\eta,\xi} \sum_{z\in\mathbb{Z}^d} \Big| \alpha(\theta_{-X_t^1}\eta_t^1, z) - \alpha(\theta_{-X_t^1}\eta_t^2, z) \Big| \\
&\cdot \exp\left(- \int_0^t \sum_{z\in\mathbb{Z}^d} \big| \alpha(\theta_{-X_s^1}\eta_s^1, z) - \alpha(\theta_{-X_s^1}\eta_s^2, z) \big| \, ds \right) dt \\
&\le \int_0^\infty \varphi(t)\widehat{\mathbb{E}}_{\eta,\xi} \sum_{z\in\mathbb{Z}^d} \Big| \alpha(\theta_{-X_t^1}\eta_t^1, z) - \alpha(\theta_{-X_t^1}\eta_t^2, z) \Big| \, dt.
\end{aligned}$$

By telescoping over single site discrepancies and using Lemma 5.5.2, this is less than

$$\int_0^\infty \varphi(t)(\| \gamma^+ - \gamma^- \|_\infty + 1)^d t^d \sup_{\eta,\xi\in\Omega} \widehat{\mathbb{E}}^E_{\eta,\xi}\rho(\eta_t^1(0), \eta_t^2(0))\, dt < \infty$$

by assumption. Since $\varphi(t/K)$ decreases to 1 as $K \to \infty$, monotone convergence implies

$$\lim_{K\to\infty} \widehat{\mathbb{E}}_{\eta,\xi}\varphi\left(\frac{\tau}{K}\right)\mathbb{1}_{\tau<\infty} = \widehat{\mathbb{E}}_{\eta,\xi}\mathbb{1}_{\tau<\infty} < 1$$

by Lemma 5.5.3. Consequently, there exists a $K_0 \ge 0$ such that for all $K > K_0$

$$\widehat{\mathbb{E}}_{\eta,\xi}\varphi\left(\frac{\tau}{K}\right)\mathbb{1}_{\tau<\infty} < 1.$$

This completes the adaptation of part a).

The adaptation of the proof of part b) follows the same scheme, where we add the term $\varphi\left(\frac{t}{K}\right)$ to all integrals. Note that this gives a version of Lemma 5.5.4 as well. Then, in line (5.17) we use (5.21) again and then have to replace $\widehat{\mathbb{P}}_{\eta,\xi}(\tau<\infty)$ by $\widehat{\mathbb{E}}_{\eta,\xi}\varphi\left(\frac{\tau}{K}\right)\mathbb{1}_{\tau<\infty}$ in lines (5.18) and (5.19). To estimate (5.19), we use

$$\begin{aligned}
&\widehat{\mathbb{E}}_{\eta,\xi}\varphi\left(\frac{\tau}{K}\right)\mathbb{1}_{\tau<\infty}\\
&\quad\le\int_0^\infty\varphi\left(\frac{t}{K}\right)\widehat{\mathbb{E}}_{\eta,\xi}\sum_{z\in\mathbb{Z}^d}\Big|\alpha(\theta_{-X_t^1}\eta_t^1,z)-\alpha(\theta_{-X_t^1}\eta_t^2,z)\Big|\,dt\\
&\quad\le\int_0^\infty\varphi\left(\frac{t}{K}\right)\sum_{y\in\mathbb{Z}^d}w_\alpha\widehat{\mathbb{E}}_{\eta,\xi}\rho(\eta_t^1(y+X_t^1),\eta_t^2(y+X_t^1))\,dt
\end{aligned}$$

with w_α as in the original proof. Therefore

$$\begin{aligned}
&\sum_{x\in\mathbb{Z}^d}\sup_{(\eta,\xi)\in(\Omega\times\Omega)_x}\widehat{\mathbb{E}}_{\eta,\xi}\varphi\left(\frac{\tau}{K}\right)\mathbb{1}_{\tau<\infty}\\
&\quad\le\sum_{x\in\mathbb{Z}^d}\sup_{(\eta,\xi)\in(\Omega\times\Omega)_x}\int_0^\infty\varphi\left(\frac{t}{K}\right)\\
&\qquad\qquad\cdot\sum_{y\in\mathbb{Z}^d}w_\alpha(y)\widehat{\mathbb{E}}_{\eta,\xi}\rho(\eta_t^1(y+X_t^1),\eta_t^2(y+X_t^1))\,dt,
\end{aligned}$$

which is finite by Lemma 5.5.4. $\square$

Acknowledgement

We thank NWO for financial support, under the project "vrije competitie – coupling, concentration and stochastic dynamics" number 600.065.100.07N14.

Bibliography

[1] M. Aizenman and R. Holley. Rapid convergence to equilibrium of stochastic Ising models in the Dobrushin Shlosman regime. Percolation theory and ergodic theory of infinite particle systems, Proc. Workshop IMA, Minneapolis/Minn. 1984/85, IMA Vol. Math. Appl. 8, 1-11 (1987)., 1987.

[2] Oskari Ajanki. Transport and diffusion in disordered media. 2012.

[3] Sebastian Andres. Invariance principle for the random conductance model with dynamic bounded conductances. *arXiv:1202.0803v1*, preprint.

[4] R. Arratia. Symmetric exclusion processes: a comparison inequality and a large deviation result. *Annals of Probability*, 13(1):53–61, 1985.

[5] L. Avena, F. den Hollander, and F. Redig. Large deviation principle for one-dimensional random walk in dynamic random environment: attractive spin-flips and simple symmetric exclusion. *Markov Processes and Related Fields*, 16:139–168, 2010.

[6] L. Avena, F. den Hollander, and F. Redig. Law of large numbers for a class of random walks in dynamic random environments. *Electronic Journal of Probability*, 16:587–617, 2011.

[7] L. Avena and P. Thomann. Continuity and anomalous fluctuations in random walks in dynamic random environments: numerics, phase diagrams and conjectures. *arXiv:1201.2890v2*, preprint.

[8] Luca Avena, Renato dos Santos, and Florian Vï£¡llering. Transient random walk in symmetric exclusion: limit theorems and an einstein relation. *arXiv:1102.1075v2*, preprint.

[9] A. Bandyopadhyay and O. Zeitouni. Random walk in dynamic Markovian random environment. *ALEA Lat. Amer. J. Probab. Math. Stat.*, 1:205–224, 2006.

[10] Matthias Birkner, Jiri Cerny, Andrej Depperschmidt, and Nina Gantert. Directed random walk on an oriented percolation cluster. *arXiv:1204.2951v1*, preprint.

[11] C. Boldrighini, I.A. Ignatyuk, V. Malyshev, and A. Pellegrinotti. Random walk in dynamic environment with mutual influence. *Stoch. Proc. their Appl.*, 41:157–177, 1992.

[12] C. Boldrighini, R.A. Minlos, and A. Pellegrinotti. Random walk in random (fluctuating) environment. *Russian Math. Reviews*, 62:663–712, 2007.

[13] Jean Bricmont and Antti Kupiainen. Random walks in space time mixing environments. *Journal of Statistical Physics*, 134:979–1004, 2009.

[14] P. Cattiaux and A. Guillin. Deviation bounds for additive functionals of markov processes. *ESAIM. Probability and Statistics*, 12:12–29, 2008.

[15] S. Chatterjee. Concentration inequalities with exchangeable pairs (phd. thesis). 2005.

[16] J. Chazottes, P. Collet, C. Külske, and F. Redig. Concentration inequalities for random fields via coupling. *Probability Theory and Related Fields*, 137(1-2):201–225, 2007.

[17] J-R Chazottes, P Collet, and B Schmitt. Statistical consequences of the devroye inequality for processes. applications to a class of non-uniformly hyperbolic dynamical systems. *Nonlinearity*, 18(5):23–41, 2005.

[18] Jean-René Chazottes, Pierre Collet, , and Frank Redig. Coupling, concentration inequalities, and stochastic dynamics. *Journal of Mathematical Physics*, 49(12), 2008.

[19] Jean-René Chazottes, Frank Redig, and Florian Völlering. The poincaré inequality for markov random fields proved via disagreement percolation. *Indagationes Mathematicae*, 22(3âĂŞ4):149 – 164, 2011.

[20] J.R. Chazottes and F. Redig. Concentration inequalities for markov processes via coupling. *Electronic Journal of Probability*, 14:1162–1180, 2009.

[21] D. Dolgopyat and C. Liverani. Random walk in determistically changing environment. *ALEA*, 4:89–116, 2008.

[22] D. Dolgopyat and C. Liverani. Non-perturbative approach to random walk in Markovian environment. *Electronic Communications in Probability*, 14:245–251, 2009.

[23] Dmitry Dolgopyat, Gerhard Keller, and Carlangelo Liverani. Random walk in Markovian environment. *Ann. Probab.*, 36(5):1676–1710, 2008.

[24] Renato dos Santos and Frank den Hollander. Random walk on a supercritical contact process. paper in preparation.

[25] Devdatt Dubhashi and Alessandro Panconesi. *Concentration of Measure for the Analysis of Randomized Algorithms*. Cambridge University Press, New York, NY, USA, 1st edition, 2009.

[26] P.A. Ferrari, A. Galves, and C. Landim. Rate of converge to equilibrium of symmetric simple exclusion processes. *Markov processes and related fields*, 6:73–88, 2000.

[27] Hans-Otto Georgii, Olle Hï£¡ggstrï£¡m, and Christian Maes. The random geometry of equilibrium phases. volume 18 of *Phase Transitions and Critical Phenomena*, pages 1 – 142. Academic Press, 2001.

[28] N. Gozlan and C. Lï£¡onard. Transport inequalities: A survey. *Markov Processes and Related Fields*, 16:635ï£¡736, 2010.

[29] G. Grimmett. *Percolation*. Springer-Verlag, Berlin, 1999.

[30] A. Guionnet and B. Zegarlinski. Lectures on logarithmic sobolev inequalities. In *Séminaire de probabilités de Strasbourg*, volume 1801 of *Lecture Notes in Math.*, pages 1–134. Springer, Berlin, 2003.

[31] Richard Holley and Daniel Stroock. Logarithmic sobolev inequalities and stochastic ising models. *Journal of Statistical Physics*, 46:1159–1194, 1987. 10.1007/BF01011161.

[32] Mathew Joseph and Firas Rassoul-Agha. Almost sure invariance principle for continuous-space random walk in dynamic random environment. *ALEA Lat. Am. J. Probab. Math. Stat.*, 8:43–57, 2011.

[33] Harry Kesten. Aspects of first passage percolation. In *École d'Été de Probabilités de Saint Flour XIV - 1984*, volume 1180 of *Lecture Notes in Mathematics*, pages 125–264. Springer Berlin / Heidelberg, 1986.

[34] Christof Kï£¡lske. Concentration inequalities for functions of gibbs fields with application to diffraction and random gibbs measures. *Communications in Mathematical Physics*, 239:29–51, 2003. 10.1007/s00220-003-0841-5.

[35] C. Landim. Occupation time large deviations for the symmetric simple exclusion process. *The Annals of Probability*, 20(1):206–231, 1992.

[36] M. Ledoux. *The Concentration of Measure Phenomenon*. American Mathematical Society, 2001.

[37] T.M. Liggett. *Interacting Particle Systems*. Springer, 2005.

[38] T. Lindvall. *Lectures on the coupling method*. Wiley, 1992.

[39] Eyal Lubetzky, Fabio Martinelli, Allan Sly, and Fabio Lucio Toninelli. Quasi-polynomial mixing of the 2d stochastic ising model with "plus" boundary up to criticality. *preprint*, 2010.

[40] F. Martinelli and E. Olivieri. Approach to equilibrium of Glauber dynamics in the one phase region. II. The general case. *Comm. Math. Phys.*, 161(3):487–514, 1994.

[41] Fabio Martinelli. Lectures on glauber dynamics for discrete spin models. In Pierre Bernard, editor, *Lectures on Probability Theory and Statistics*, volume 1717 of *Lecture Notes in Mathematics*, pages 93–191. Springer Berlin / Heidelberg, 2004.

[42] Yann Ollivier. Ricci curvature of markov chains on metric spaces. *Journal of Functional Analysis*, 256(3):810 – 864, 2009.

[43] Yao-Feng Ran and Fan-Ji Tian. On the Rosenthal's inequality for locally square integrable martingales. *Stochastic processes and their applications*, 104:107–116, 2003.

[44] F Rassoul-Agha. The point of view of the particle on the law of large numbers for random walks in a mixing random environment. *Annals of Probability*, 31:1441–1463, 2003.

[45] Firas Rassoul-Agha, Timo Seppalainen, and Atilla Yilmaz. Quenched free energy and large deviations for random walks in random potentials. *http://arxiv.org/abs/1104.3110*, preprint.

[46] Michael Röckner and Feng-Yu Wang. Weak Poincaré Inequalities and L2-Convergence Rates of Markov Semigroups. *Journal of Functional Analysis*, 185(2):564 – 603, 2001.

[47] Ya. G. Sinaĭ. The limit behavior of a one-dimensional random walk in a random environment. *Teor. Veroyatnost. i Primenen.*, 27(2):247–258, 1982.

[48] Fred Solomon. Random walks in a random environment. *Annals of Probability*, 3:1–31, 1975.

[49] Daniel W. Stroock and Boguslaw Zegarlinski. The logarithmic sobolev inequality for discrete spin systems on a lattice. *Commun. Math. Phys*, 149:175–194, 1992.

[50] Michel Talagrand. Concentration of measure and isoperimetric inequalities in product spaces. *Publications Mathï£¡matiques de L'IHï£¡S*, 81:73–205, 1995. 10.1007/BF02699376.

[51] Michel Talagrand. A new look at independence. *Annals of Probability*, 24(1):1–34, 1996.

[52] J. van den Berg and C. Maes. Disagreement percolation in the study of Markov fields. *Ann. Probab.*, 22(2):749–763, 1994.

[53] C. Villani. *Topics in optimal transportation.* American Mathematical Society, 2003.

[54] C. Villani. *Optimal transport. Old and new.* Springer-Verlag, Berlin, 2009.

[55] L. Wu and N. Yao. Large deviation principles for markov processes via φ-sobolev inequalities. *Electronic Communications in Probability*, (13):10–23, 2008.

[56] Liming Wu. Poincar
'{e} and transportation inequalities for gibbs measures under the dobrushin uniqueness condition. *ANNALS OF PROBABILITY*, 34:1960, 2006.

[57] Boguslaw Zegarlinski. Dobrushin uniqueness theorem and logarithmic sobolev inequalities. *Journal of Functional Analysis*, 105(1):77 – 111, 1992.

Curriculum Vitae

Florian Völlering was born in Berlin, Germany, on 11 August 1982. He completed his high school studies at the Herder Gymnasium Berlin in 2002, obtaining the Abitur with a score of 1.3. After that he began his master at the Technische Universität Berlin, studying mathematics with a minor in computer science. During his studies he did an intership at the Zuse Institut Berlin and was a student assistant for two years. In 2008 he graduated cum laude. His Diplomarbeit (master thesis) was written under Prof. Dr. Jürgen Gärtner, with the topic "Quenched asymptotics of the parabolic Anderson model with a single catalyst". Later 2008 he moved to Leiden to start his PhD at the mathematics institute of Leiden University under the supervision of Prof. Dr. Frank Redig. His PhD research was focused on concentration inequalities and random walks in dynamic random environments. After his promoter moved to Delft Delft University of Technology it was decided to defend the thesis there. In October 2012 he will start post-doctoral research in Göttingen with Prof. Dr. Anja Sturm.