Shaken dynamics: an easy way to parallel MCMC

Valentina Apollonio[1] Roberto D'Autilia[1] Benedetto Scoppola[2]
Elisabetta Scoppola[1] Alessio Troiani[3]

May 15, 2021

[1] Dipartimento di Matematica e Fisica, Università Roma Tre
Largo San Murialdo, 1 - 00146 Roma, Italy

[2] Dipartimento di Matematica, Università di Roma "Tor Vergata"
Via della Ricerca Scientifica, 1 - 00133 Roma, Italy

[3] Dipartimento di Matematica "Tullio Levi–Civita", Università di Padova
Via Trieste, 63 - 35121 Padova, Italy

Abstract

We define a class of Markovian parallel dynamics for spin systems on general interaction graphs. In this dynamics, beside the usual set of parameters J_{xy}, the strength of the interaction between the spins σ_x and σ_y, and λ_x, the external field at site x, there is an inertial parameter q measuring the tendency of the system to remain locally in the same state. This dynamics is reversible with an explicitly defined stationary measure. For suitable choices of parameter this invariant measure concentrates on the ground states of the Hamiltonian. This implies that this dynamics can be used to solve, heuristically, difficult problems in the context of combinatorial optimization. We also study the dynamics on $\mathbb{Z}^2$ with homogeneous interaction and external field and with arbitrary boundary conditions. We prove that for certain values of the parameters the stationary measure is close to the related Gibbs measure. Hence our dynamics may be a good tool to sample from Gibbs measure by means of a parallel algorithm. Moreover we show how the parameters allow to interpolate between spin systems defined on different regular lattices.

1 Introduction

We introduce a class of parallel dynamics to study spin systems on arbitrary graphs $G = (V, E)$ with general interaction given by

$$H(\sigma) = -\sum_{e=\{x,y\}\in E} J_{xy}\sigma_x\sigma_y - 2\sum_{x\in V}\lambda_x\sigma_x$$

with J_{xy} and λ_x in $\mathbb{R}$, and $\sigma \in \{-1,+1\}^V$ configuration on G. We called them *shaken dynamics*. The main feature of these dynamics is that the interaction is divided into two parts acting alternatively. They are reversible and we control explicitly the invariant measure $\pi(\sigma)$ - see (1) below. However shaken dynamics share with irreversible dynamics the asymmetric interaction, retaining some possible advantages. This class of shaken dynamics includes and generalizes parallel dynamics (PCA) defined in previous papers. We present here consequences and application of this generalized definition also related to discrete optimization problems. Even if the convergence to equilibrium of the shaken dynamics is not analyzed in this paper, a preliminary comparison is discussed both from a theoretical and numerical point of view.

Parallel Markovian dynamics on spin systems, often introduced in terms of Probabilistic Cellular Automata (PCA), have been the object of deep investigations in the past decades. The first attempts at studying PCA in the context of Equilibrium Statistical Mechanics date back to [10], where various features of the infinite-volume limit have been investigated, in particular its space-time Gibbsian nature. Also in more recent times the interest on PCA remained on. We just mention here some references, especially in order to recall that the determination of the stationary measure can be non trivial. Examples of infinite volume PCA whose invariant measures are not Gibbsian are given in [8]. In [14] explicit conditions for the existence of reversible and ergodic PCA are provided. In [5] it has been proven that, in general, the stationary measure defined by a local PCA may have nothing in common with the Gibbs measure, giving rise to stable checkerboard configurations.

One of the main reasons of interest on PCA is related to their numerical applications. Indeed, parallelization could, at least in principle, speed up MCMC dynamics. Even though until a few years ago parallel computing was expensive and tricky, we have now powerful and cheap parallel architectures, for instance based on GPU or even FPGA.

In the past, several approaches have been proposed to exploit the capabilities of parallel architecture to simulate statistical mechanics lattice models in an effective way and to apply Monte Carlo methods to solve discrete Optimization Problems. However, these methods are strongly tied to the particular architecture used for the simulation and make use of the similarities between the structure of the graph and the structure of the hardware in use. The literature on the topic is quite vast, especially in the computer engineering and applied physics communities. Some attention, though, should be paid to the theoretical foundation of these methods. Indeed, the control of the stochastic dynamics, or at least of its stationary measure, seems to be a minimal requirement in order to develop random algorithms in combinatorial optimization. For the Markov Chain that we describe here we are able to control the stationary measure. Moreover the algorithm is natively parallel and is not bound to any particular architecture or graph structure. Therefore its performances are likely to benefit from the development of parallel computing often driven by applications not necessarily linked to academic research.

It is reasonable to think that updating all the spins at every single step, parallel dynamics exhibit a faster convergence to equilibrium compared to single spin flip dynamics. Actually, parallelization can drastically change the relaxation time. In some particular cases of dynamics updating all the spins at every step, rigorous results are available. One of the main examples is given by the Swendsen-Wang (SW) dynamics where a polynomial relaxation time in the size of the problem is proven [11]. However, SW dynamics is a cluster dynamics which is not really parallel in the sense that its probability transition kernel can not be factorized as $P(\sigma, \tau) = \prod_{x \in \Lambda} P(\tau_x|\sigma)$ for $\sigma, \tau \in \{-1, +1\}^\Lambda$. There are other examples of fast mixing parallel dynamics, see for instance [7] where a parallel irreversible dynamics is defined for the 2d Ising model and fast mixing is proven in the low temperature regime and [9] for results on the mixing time in the high temperature regime.

In some cases the acceleration of convergence to equilibrium is related to irreversibility (see for instance [13]). However general and ultimate results on the effects of parallelization and/or irreversibility on the convergence to equilibrium are still an open problem.

The *shaken dynamics* introduced in this paper contains a class of PCA on spin systems, introduced in previous papers [6, 7, 15, 18], characterized by the presence of an inertial term preventing the simultaneous update of a too large set of spins. The spin at each site x is updated with an heat-bath rule according to the local field $h_x(\sigma)$ given by the configuration at its nearest neighbors plus a self-interaction term tuned by an inertial parameter $q > 0$. More precisely this means that the transition probability can be written in terms of a doubled Hamiltonian

$$P(\sigma, \tau) = \frac{e^{-H(\sigma,\tau)}}{\sum_{\tau'} e^{-H(\sigma,\tau')}}$$

with $H(\sigma, \tau) = \sum_x h_x(\sigma)\tau_x$.

When the local field is computed considering the contribution of all the nearest neighbors it is possible to verify that the resulting dynamics is reversible with stationary measure given by

$$\pi(\sigma) = \frac{\sum_\tau e^{-H(\sigma,\tau)}}{\sum_{\sigma',\tau'} e^{-H(\sigma',\tau')}}. \tag{1}$$

If only a subset of the nearest neighbors of the site is considered, for instance the spins in a predefined direction, the dynamics becomes irreversible. The results on its mixing properties are quite interesting: irreversibility can drastically change the mobility of the dynamics, supporting mixing. In general, however, the explicit computation of the stationary measure is highly nontrivial also on regular graphs.

The idea bearing to the dynamics introduced in this paper is the following: we define a two-step update alternating the direction of the nearest neighbors defining the local field $h_x(\sigma)$. The dynamics defined in this way is reversible but has many features in common with the irreversible dynamics introduced in [15, 7]. In particular the result cited above on the fast mixing of the 2d Ising model at low temperature, was obtained in [7] under the hypothesis of periodic boundary conditions. This was a crucial ingredient to prove that $\pi(\sigma)$ was the invariant measure of the irreversible dynamics. For the shaken dynamics discussed here reversibility holds and it allows to identify $\pi(\sigma)$ given in (1) as stationary measure also for non periodic boundary conditions. The results obtained in [7] could be extended to the shaken dynamics in a straightforward manner. We believe that this general class of shaken dynamics provides a set of tools useful in numerical simulations.

In [6] the convergence, in the thermodynamical limit, of $\pi(\sigma)$ to the Gibbs measure with Hamiltonian $H(\sigma)$ has been proven in the case $\lambda = 0$ in the high temperature regime (under Dobrushin conditions and q sufficiently large). Similar results, in the low temperature regime, are harder to obtain. In [18] the case $\mathbb{Z}^2$ is discussed, again in the case of no external field. As shown in [1] this measure $\pi(\sigma)$ enables us to play with the geometry of the system and, considering different limits for the parameters appearing in the Hamiltonian, to compare the properties of spin systems defined on different lattices. Embedding the state space into a doubled space of pairs of configurations, with the doubled Hamiltonian $H(\sigma, \tau)$, the interaction graph becomes a doubled bipartite graph and when the self-interaction parameter q tends to zero or tends to infinity, the interaction graph changes, cutting the corresponding edge or collapsing the two extremal vertices respectively. Starting from the Ising model on $\mathbb{Z}^2$ we identified in [1] the critical line of this class of systems parameterized by q. In the limit $q \to \infty$ the measure π converges in total variation to the Gibbs measure corresponding to the Hamiltonian $H(\sigma) := H(\sigma, \sigma)$. In this paper we analyze these systems from a dynamical point of view defining an easy and effective way to sample them with the shaken dynamics. We extend the result of the convergence of π to the Gibbs measure and the geometrical interplay among different regular lattices is applied not only starting from $\mathbb{Z}^2$.

The paper is organized as follows. In Section 2 we define the shaken dynamics on a general graph and state the main result. In Section 3 we consider the particular case of $\mathbb{Z}^2$ with homogeneous interaction. In Section 4 we generalize the definition of shaken dynamics and in Section 5 we present an application of the shaken dynamics to optimization problems constructing a class of parallel MCMC. In Section 6 we give the proofs of the results while Section 7 is devoted to final remarks and possible applications of the shaken dynamics to tidal dissipative effects in planetary systems.

2 The shaken dynamics on a general graph

Let $G = (V, E)$ be a finite weighted graph and $\mathcal{X}_V = \{-1, 1\}^V$ be the set of spin configurations on V. We consider the nearest neighbor interaction between spins given by the Ising Hamiltonian in the general form:

$$\begin{aligned} H(\sigma) &= -\sum_{e=\{x,y\}\in E} J_{xy}\sigma_x\sigma_y - 2\sum_{x\in V}\lambda_x\sigma_x \\ &= \sum_x\sum_y \frac{1}{2}J_{xy}\mathbb{1}_{\{x,y\}\in E}\sigma_x\sigma_y - 2\sum_{x\in V}\lambda_x\sigma_x = -\langle\frac{1}{2}\mathcal{J}\sigma + 2\lambda, \sigma\rangle \end{aligned} \tag{2}$$

where the weight $J_{xy} \in \mathbb{R}$ associated to the edge $\{x, y\}$, represents the interaction, and can be written in compact form as a symmetric matrix $\mathcal{J}$ and we denote by $\langle\cdot,\cdot\rangle$ the scalar product. The vector $\lambda = \{\lambda_x\}_{x\in V}$ is an external field, possibly non constant.

We introduce a class of bipartite weighted graphs $G^b = (V^b, E^b)$ *doubling the interaction graph* G. The idea is to duplicate the vertex set into two identical copies, $V^{(1)}$ and $V^{(2)}$, representing the two parts of the vertex set of the bipartite graph. For each $x \in V$ we denote by $x^{(1)}, x^{(2)}$ the vertices corresponding to $x \in V$ in $V^{(1)}$ and in $V^{(2)}$ respectively. The edges between $x^{(1)}$ and $x^{(2)}$ are all present, for any $x \in V$. On the other hand the edges between $x^{(1)}$ and $y^{(2)}$, with $x \neq y$, or between $y^{(1)}$ and $x^{(2)}$, can be present only if $\{x, y\} \in E$. Exactly one edge among the two possibilities $(x^{(1)}, y^{(2)})$ and $(y^{(1)}, x^{(2)})$ is in E^b if $\{x, y\} \in E$. This means that for any graph G there are many *doubling graphs* G^b. Note that similar doubling graphs have already been introduced in literature for different purposes (see [14]). More precisely:

Definition 2.1 *A bipartite weighed graph $G^b = (V^b, E^b)$ is the* **doubling graph** *of $G = (V, E)$ if*

- *the vertex set $V^b = V^{(1)} \cup V^{(2)}$ where the two parts $V^{(1)}$ and $V^{(2)}$ are two identical copies of V;*
- *for any $x \in V$ the edge $(x^{(1)}, x^{(2)}) \in E^b$ with weight q and we call it a* **self-**

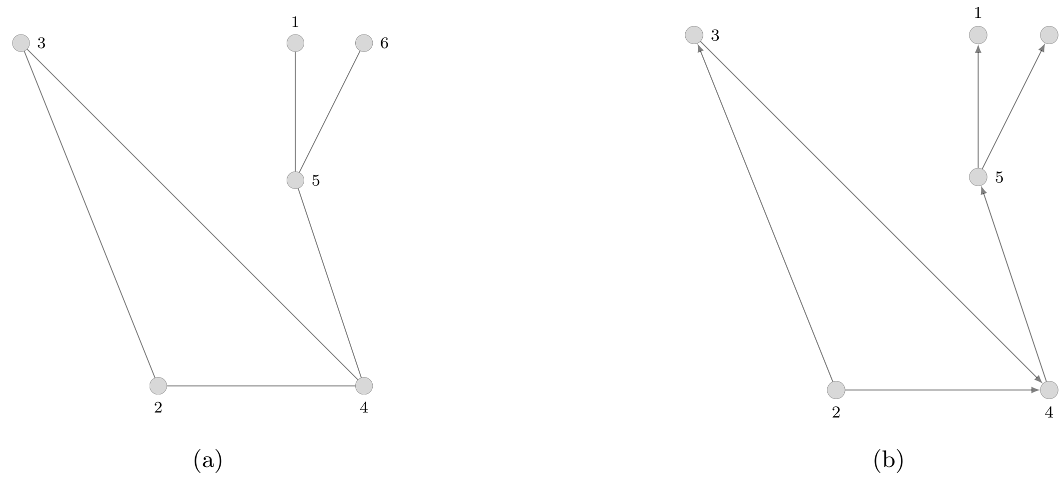


Figure 1: An undirected graph (a) and a possible choice for the related directed graph (b)

interaction edge;

- *if* $\{x,y\} \in E$ *then one, and only one, between the two edges* $\{x^{(1)}, y^{(2)}\}$ *and* $\{y^{(1)}, x^{(2)}\}$ *is in* E^b. *We call this kind of edge an* **interaction edge**.

To construct a doubling graph starting from the interaction graph $G = (V, E)$, define a new oriented graph $G^o = (V, E^o)$ simply orienting the edges in an arbitrary way. Using the oriented edges the set E^b is constructed as follows. For any $x \in V$ we have the self-interaction edge $(x^{(1)}, x^{(2)}) \in E^b$ with weight $w(x^{(1)}, x^{(2)}) = q$ and for $x \neq y \in V$ we have $(x^{(1)}, y^{(2)}) \in E^b$ if and only if $(x, y) \in E^o$ with weight $w(x^{(1)}, y^{(2)}) = J_{xy}$.

Note that the edges in E^b are not oriented. However, by construction, the graph is bipartite, so that for any $e = \{x, y\} \in E^b$ we have $x \in V^{(1)}, y \in V^{(2)}$ or viceversa and so we consider in the definition the natural order in the edges in E^b by setting $e = (e^{(1)}, e^{(2)})$ with $e^{(1)} \in V^{(1)}, e^{(2)} \in V^{(2)}$. For this reason we can use the oriented edges in E^o in order to define E^b.

We will sometimes omit the superscripts $^{(1)}$ and $^{(2)}$ and we will always consider (x, y) the ordered pair with $x \in V^{(1)}$, $y \in V^{(2)}$, and $\{x, y\}$ the unordered pair with $x, y \in V^b$.

Definition 2.2 *The pair Hamiltonian* $H(\sigma^{(1)}, \sigma^{(2)})$ *is the* **doubling** *of the Hamiltonian (2) with interaction graph* G *if there exists a doubling graph* $G^b = (V^b, E^b)$ *of* G *such that* $H(\boldsymbol{\sigma})$*, defined on the spin configurations* $\boldsymbol{\sigma} \equiv (\sigma^{(1)}, \sigma^{(2)}) \in \mathcal{X}_{V^b} = \{-1, 1\}^{V^b}$*, can be written as*

$$H(\boldsymbol{\sigma}) = -\sum_{\{x,y\} \in E^b} w(x,y) \boldsymbol{\sigma}_x \boldsymbol{\sigma}_y - \sum_{x \in V^b} \lambda_x \boldsymbol{\sigma}_x \tag{3}$$

with $w(x, y) = q$ *if* $\{x, y\}$ *is a self interaction edge and* $w(x, y) = J_{xy}$ *otherwise and with* $\lambda_{x^{(1)}} = \lambda_{x^{(2)}} = \lambda_x$.

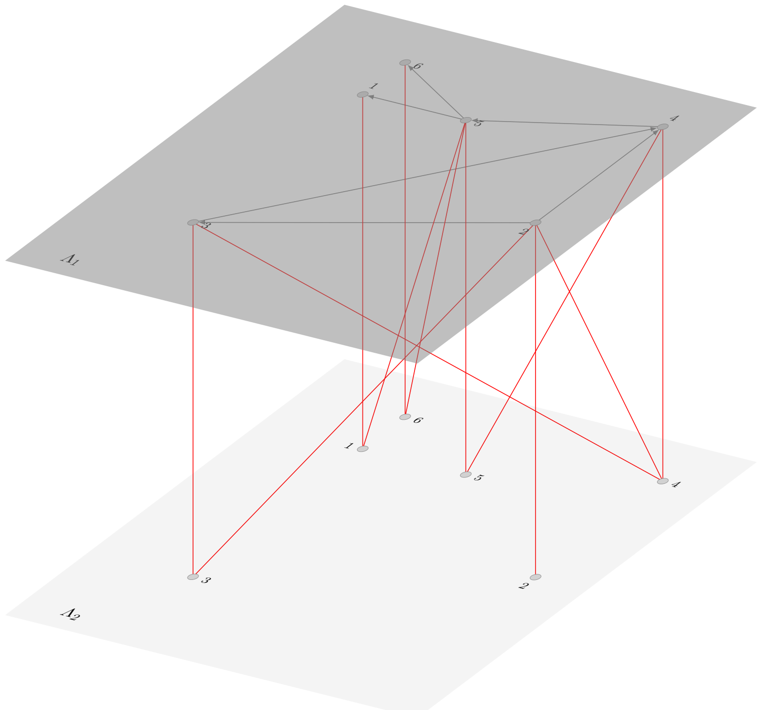


Figure 2: The doubling of the graph of Fig. 1(a) obtained from the directed graph of Fig. 1(b)

In a more explicit way we can write

$$\begin{aligned} H(\boldsymbol{\sigma}) &\equiv H(\sigma^{(1)}, \sigma^{(2)}) \\ &= - \sum_{\{x^{(1)}, y^{(2)}\} \in E^b} J_{xy} \sigma_x^{(1)} \sigma_y^{(2)} - \sum_{x \in V} \left(q \sigma_x^{(1)} \cdot \sigma_x^{(2)} + \lambda_x (\sigma_x^{(1)} + \sigma_x^{(2)}) \right) \\ &= - \sum_{x \in V} \left(\sigma_x^{(1)} h_x^{2\to1}(\sigma^{(2)}) + \lambda_x \sigma_x^{(2)} \right) \\ &= - \sum_{x \in V} \left(\sigma_x^{(2)} h_x^{1\to2}(\sigma^{(1)}) + \lambda_x \sigma_x^{(1)} \right) \end{aligned} \tag{4}$$

with

$$h_x^{2\to1}(\sigma^{(2)}) = \sum_{y \in V : \{x^{(1)}, y^{(2)}\} \in E^b} \left(J_{xy} \sigma_y^{(2)} \right) + q \sigma_x^{(2)} + \lambda_x$$

and

$$h_x^{1\to2}(\sigma^{(1)}) = \sum_{y \in V : \{y^{(1)}, x^{(2)}\} \in E^b} \left(J_{xy} \sigma_y^{(1)} \right) + q \sigma_x^{(1)} + \lambda_x$$

By defining $\mathcal{J}^o$ the matrix of oriented interaction, i.e., ${\mathcal{J}^o}_{xy} = J_{xy} \mathbb{1}_{(x,y) \in E^o}$, and its transposed $\mathcal{J}^{oT}$ corresponding to the opposite orientation, we can write

$$h_x^{2\to1}(\sigma^{(2)}) = (\mathcal{J}^o \sigma^{(2)})_x + q \sigma_x^{(2)} + \lambda_x$$

$$h_x^{1\to2}(\sigma^{(1)}) = (\mathcal{J}^{oT} \sigma^{(1)})_x + q \sigma_x^{(1)} + \lambda_x$$

and

$$
\begin{aligned}
H(\sigma^{(1)},\sigma^{(2)}) &= -\langle \sigma^{(1)}, \mathcal{J}^o \sigma^{(2)}\rangle + q\langle \sigma^{(1)},\sigma^{(2)}\rangle + \langle \lambda, \sigma^{(1)}\rangle + \langle \lambda, \sigma^{(2)}\rangle \\
&= -\langle \mathcal{J}^{oT}\sigma^{(1)}, \sigma^{(2)}\rangle + q\langle \sigma^{(1)},\sigma^{(2)}\rangle + \langle \lambda, \sigma^{(1)}\rangle + \langle \lambda, \sigma^{(2)}\rangle
\end{aligned}
$$

If we consider the case $\sigma^{(1)} = \sigma^{(2)} = \sigma$, i.e., $\sigma_x^{(1)} = \sigma_x^{(2)}$ for any $x \in V$, then we have $H(\boldsymbol{\sigma}) \equiv H(\sigma,\sigma) = H(\sigma) - q|V|$. Indeed we have immediately $\mathcal{J} = \mathcal{J}^o + \mathcal{J}^{oT}$.

We construct now the *shaken dynamics* on the state space $\mathcal{X}_V$ by considering two subsequent updating defined as follows:

$$
P^{1\to 2}(\sigma,\sigma') := \frac{e^{-H(\sigma,\sigma')}}{\overrightarrow{Z}_\sigma} = \prod_{x\in V} \frac{e^{h_x^{1\to 2}(\sigma)\sigma'_x}}{2\cosh h_x^{1\to 2}(\sigma)} \quad \text{with} \quad \overrightarrow{Z}_\sigma = \sum_{\zeta\in\mathcal{X}_V} e^{-H(\sigma,\zeta)} \tag{5}
$$

$$
P^{2\to 1}(\sigma',\tau) := \frac{e^{-H(\tau,\sigma')}}{\overleftarrow{Z}_{\sigma'}} = \prod_{x\in V} \frac{e^{h_x^{2\to 1}(\tau)\sigma'_x}}{2\cosh h_x^{2\to 1}(\tau)} \quad \text{with} \quad \overleftarrow{Z}_{\sigma'} = \sum_{\zeta\in\mathcal{X}_V} e^{-H(\zeta,\sigma')} \tag{6}
$$

$$
P^{sh}(\sigma,\tau) = \sum_{\sigma'\in\mathcal{X}_V} P^{1\to 2}(\sigma,\sigma')P^{2\to 1}(\sigma',\tau) = \sum_{\sigma'\in\mathcal{X}_V} \frac{e^{-H(\sigma,\sigma')}}{\overrightarrow{Z}_\sigma}\frac{e^{-H(\tau,\sigma')}}{\overleftarrow{Z}_{\sigma'}} \tag{7}
$$

We state the result on the shaken dynamics in this general context.

Theorem 2.3 *The stationary measure of the shaken dynamics is*

$$
\pi(\sigma) = \frac{\overrightarrow{Z}_\sigma}{Z} \quad \textit{with} \quad \overrightarrow{Z}_\sigma := \sum_\tau e^{-H(\sigma,\tau)} \quad \textit{and} \quad Z := \sum_{\sigma,\tau} e^{-H(\sigma,\tau)} \tag{8}
$$

and reversibility holds. This stationary measure is the marginal of the Gibbs measure on the space $\mathcal{X}_{V^b}$ of pairs of configurations $\boldsymbol{\sigma} := (\sigma^{(1)},\sigma^{(2)})$ defined by:

$$
\pi^b(\boldsymbol{\sigma}) := \frac{1}{Z} e^{-H(\boldsymbol{\sigma})}. \tag{9}
$$

The shaken dynamics on $\mathcal{X}_V$ corresponds to an alternate dynamics on G^b in the following sense

$$
P^{sh}(\sigma^{(1)},\tau^{(1)}) = \sum_{\tau^{(2)}\in\{-1,+1\}^{V^{(2)}}} P^{alt}(\boldsymbol{\sigma},\boldsymbol{\tau}) \tag{10}
$$

with

$$
P^{alt}(\boldsymbol{\sigma},\boldsymbol{\tau}) = \frac{e^{-H(\sigma^{(1)},\tau^{(2)})}}{\overrightarrow{Z}_{\sigma^{(1)}}}\frac{e^{-H(\tau^{(1)},\tau^{(2)})}}{\overleftarrow{Z}_{\tau^{(2)}}} \tag{11}
$$

the stationary measure of P^{alt} is $\pi^b(\boldsymbol{\sigma})$. This dynamics is in general non reversible.

3 Example: the shaken dynamics on $\mathbb{Z}^2$

Let Λ be a two-dimensional $L\times L$ square lattice in $\mathbb{Z}^2$ and let $\mathcal{B}_\Lambda$ denote the set of all nearest neighbors in Λ with periodic boundary conditions.

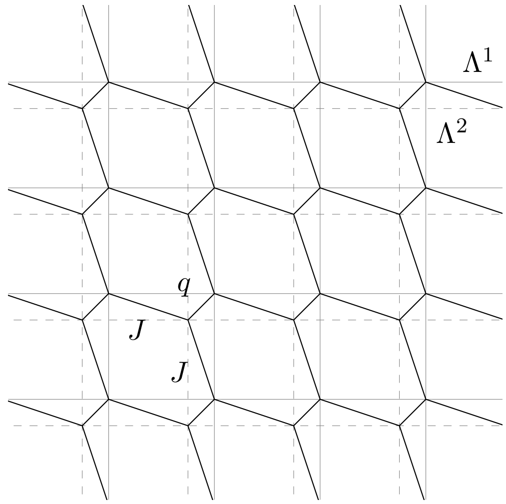


Figure 3: The doubling graph of $\mathbb{Z}^2$ represented in the figure turns out to be a hexagonal lattice.

In Λ we identify a set B where the value of the spins is frozen throughout the evolution and that plays the role of boundary conditions. This means that we will consider the state space $\mathcal{X}_{\Lambda,B} = \{\sigma \in \mathcal{X}_\Lambda : \ \sigma_x = +1 \quad \forall x \in B\}$.

Following the construction of the shaken dynamics of the previous section we can define

$$\begin{aligned} H(\sigma,\tau) &= -\sum_{x\in\Lambda} [J\sigma_x(\tau_{x^\uparrow} + \tau_{x^\rightarrow}) + q\sigma_x\tau_x + \lambda(\sigma_x + \tau_x)] \\ &= -\sum_{x\in\Lambda} [J\tau_x(\sigma_{x^\downarrow} + \sigma_{x^\leftarrow}) + q\tau_x\sigma_x + \lambda(\sigma_x + \tau_x)] \end{aligned} \tag{12}$$

where $x^\uparrow, x^\rightarrow, x^\downarrow, x^\leftarrow$ are, respectively, the up, right, down, left neighbors of the site x on the torus $(\Lambda, \mathcal{B}_\Lambda)$, $J > 0$ is the ferromagnetic interaction, $q > 0$ is the inertial constant and λ represents the external field. We can write

$$H(\sigma,\tau) = -\sum_{x\in\Lambda} \sigma_x h^{ur}(\tau) - \lambda \sum_{x\in\Lambda} \tau_x = -\sum_{x\in\Lambda} \tau_x h^{dl}(\sigma) - \lambda \sum_{x\in\Lambda} \sigma_x \tag{13}$$

where the local up-right field $h_x^{ur}(\tau)$ due to the configuration τ is given by

$$h_x^{ur}(\tau) = \Big[J(\tau_{x^\uparrow} + \tau_{x^\rightarrow}) + q\tau_x + \lambda\Big] \tag{14}$$

and the local down-left field $h_x^{dl}(\sigma)$ due to the configuration σ is given by

$$h_x^{dl}(\sigma) = \Big[J(\sigma_{x^\downarrow} + \sigma_{x^\leftarrow}) + q\sigma_x + \lambda\Big] \tag{15}$$

Define the asymmetric updating rule

$$P^{dl}(\sigma,\tau) := \frac{e^{-H(\sigma,\tau)}}{\overrightarrow{Z}_\sigma} \quad \text{with} \quad \overrightarrow{Z}_\sigma = \sum_{\sigma'\in\mathcal{X}_{\Lambda,B}} e^{-H(\sigma,\sigma')} \tag{16}$$

Due to the definition of the pair Hamiltonian, the updating performed by the transition probability $P^{dl}(\sigma,\tau)$ is parallel: given a configuration σ, at each site $x\in\Lambda$ the spin τ_x of the new configuration τ is chosen with a probability proportional to $e^{h_x^{dl}(\sigma)\tau_x}$ so that

$$P^{dl}(\sigma,\tau):=\frac{e^{-H(\sigma,\tau)}}{\overrightarrow{Z}_\sigma}=\prod_{x\in\Lambda}\frac{e^{h_x^{dl}(\sigma)\tau_x}}{2\cosh h_x^{dl}(\sigma)}$$

We have $H(\sigma,\tau)\neq H(\tau,\sigma)$ and actually, by (13), $H(\tau,\sigma)$ corresponds to the opposite direction of the interaction for the transition from σ to τ. We define

$$P^{ur}(\sigma,\tau):=\frac{e^{-H(\tau,\sigma)}}{\overleftarrow{Z}_\sigma}\quad\text{with}\quad\overleftarrow{Z}_\sigma=\sum_{\sigma'\in\mathcal{X}_{\Lambda,B}}e^{-H(\sigma',\sigma)}\tag{17}$$

Similarly for $P^{ur}(\sigma,\tau)$ with the up-right field $h_x^{ur}(\sigma)$ we get

$$P^{ur}(\sigma,\tau):=\frac{e^{-H(\tau,\sigma)}}{\overleftarrow{Z}_\sigma}=\prod_{x\in\Lambda}\frac{e^{h_x^{ur}(\sigma)\tau_x}}{2\cosh h_x^{ur}(\sigma)}$$

Note that in the definition of $H(\sigma,\tau)=-\sum_{x\in\Lambda}\tau_x h^{dl}(\sigma)-\lambda\sum_{x\in\Lambda}\sigma_x$ the last term could be canceled obtaining the same value for the transition probability $P^{dl}(\sigma,\tau)$. However we added it in the pair Hamiltonian for symmetry reasons: in particular the fact that $H(\tau,\sigma)$ is the correct pair Hamiltonian to define $P^{ur}(\sigma,\tau)$ is due to this symmetry. Note also that

$$H(\sigma,\sigma)=H(\sigma)-q|\Lambda|$$

where we define $H(\sigma)$ to be the usual Ising Hamiltonian with magnetic field 2λ

$$H(\sigma)=-\sum_{\{x,y\}\in\mathcal{B}_\Lambda}J\sigma_x\sigma_y-2\lambda\sum_{x\in\Lambda}\sigma_x\tag{18}$$

We define

$$P^{sh}(\sigma,\tau)=\sum_{\sigma'\in\mathcal{X}_{\Lambda,B}}P^{dl}(\sigma,\sigma')P^{ur}(\sigma',\tau)=\sum_{\sigma'\in\mathcal{X}_{\Lambda,B}}\frac{e^{-H(\sigma,\sigma')}}{\overrightarrow{Z}_\sigma}\frac{e^{-H(\tau,\sigma')}}{\overleftarrow{Z}_{\sigma'}}\tag{19}$$

Reversing the order of the "down–left" and the "up–right" updating rules one would obtain the chain with transition probabilities

$$P^{sh'}(\sigma,\tau)=\sum_{\sigma'\in\mathcal{X}_{\Lambda,B}}P^{ur}(\sigma,\sigma')P^{dl}(\sigma',\tau).$$

Clearly, by choosing a different orientation instead of down-left and up-right in $\mathbb{Z}^2$, a different pair Hamiltonian can be obtained with a resulting different graph for the interaction.

In this square case we could have directly used the alternate dynamics, since $\mathbb{Z}^2$ is already a bipartite graph. Indeed we can consider the chessboard splitting of the sites in $\Lambda=V^{(1)}\cup V^{(2)}$, in black and white sites, with $|V^{(1)}|=|V^{(2)}|=|V|=|\Lambda|/2$. Black sites interact only with white sites and viceversa with the usual Ising Hamiltonian

$$
\begin{aligned}
H(\boldsymbol{\sigma}) &\equiv H(\sigma^{(1)}, \sigma^{(2)}) \\
&= -\sum_{x \in V^{(1)}} \left(\sigma_x^{(1)} h_x^{2 \to 1}(\sigma^{(2)}) + \lambda_x \sigma_x^{(2)} \right) \\
&= -\sum_{x \in V^{(2)}} \left(\sigma_x^{(2)} h_x^{1 \to 2}(\sigma^{(1)}) + \lambda_x \sigma_x^{(1)} \right).
\end{aligned}
$$

By Theorem 2.3 we immediately obtain that the invariant measure of the alternate dynamics is the Gibbs measure $\pi^G(\boldsymbol{\sigma}) = e^{-H(\boldsymbol{\sigma})}/Z$. The idea of alternate dynamics on even and odd sites is already present in the literature (see [4]).

3.1 Relation with the Gibbs measure

Remaining in $\Lambda \in \mathbb{Z}^2$ with $J > 0$ and $B = \emptyset$, i.e. with the standard periodic boundary conditions, and denoting by $\pi_\Lambda = \frac{\overrightarrow{Z}_\sigma}{Z}$ the invariant measure of the shaken dynamics we have $\overrightarrow{Z}_\sigma = \overleftarrow{Z}_\sigma = Z_\sigma$ (see [7, 15]).

We denote by π_Λ^G the Gibbs measure

$$
\pi_\Lambda^G(\sigma) = \frac{e^{-H(\sigma)}}{Z^G} \qquad \text{with} \qquad Z^G = \sum_{\sigma \in \mathcal{X}_\Lambda} e^{-H(\sigma)} \tag{20}
$$

with $H(\sigma)$ defined in (18) and we define the total variation distance, or L_1 distance, between two arbitrary probability measures μ and ν on $\mathcal{X}_{\Lambda,B}$ as

$$
\|\mu - \nu\|_{TV} = \frac{1}{2} \sum_{\sigma \in \mathcal{X}_{\Lambda,B}} |\mu(\sigma) - \nu(\sigma)| \tag{21}
$$

In the following Theorem 3.1 we control the distance between the invariant measure of the shaken dynamics and the Gibbs measure at low temperature and for q positive and large. We notice that this theorem is an extension of Theorem 1.2 in [18] to the case of Hamiltonian with the non zero external field. This result could be extended to the case $B \neq \emptyset$.

Theorem 3.1 *Set $\delta = e^{-2q}$, and let δ be such that*

$$
\lim_{|\Lambda| \to \infty} \delta^2 |\Lambda| = 0 \tag{22}
$$

Under the assumption (22), there exists $\bar{J}$ such that for any $J > \bar{J}$

$$
\lim_{|\Lambda| \to \infty} \|\pi_\Lambda - \pi_\Lambda^G\|_{TV} = 0. \tag{23}
$$

3.2 Convergence to equilibrium: a comparison

The efficiency of the shaken dynamics and its convergence to equilibrium will be analyzed in future papers, both from the theoretical and from the numerical point of view. Actually by preliminary simulations it seems that comparing the single spin flip (SSF) dynamics, the PCA dynamics and the shaken dynamics (Sh) their efficiency depends on the particular considered regime. We present here, as a remark, a comparison among these three dynamics in a simple case of "metastable regime": finite volume Λ with periodic boundary conditions and low temperature. More precisely let $0 < q < \lambda < J$ and consider, for each $x \in \lambda$, the local fields

$$h_x^{SSF}(\sigma) = \Big[\frac{J}{2}(\sigma_{x^\uparrow} + \sigma_{x^\rightarrow} + \sigma_{x^\downarrow} + \sigma_{x^\leftarrow}) + 2\lambda\Big]$$

$$h_x^{PCA}(\sigma) = \Big[\frac{J}{2}(\sigma_{x^\uparrow} + \sigma_{x^\rightarrow} + \sigma_{x^\downarrow} + \sigma_{x^\leftarrow}) + q\sigma_x + \lambda\Big]$$

$h_x^{dl}(\sigma)$ and $h_x^{ur}(\sigma')$ defined in (15) and (14), and the local transition probabilities

$$p_x^*(\sigma, \tau) := \frac{e^{\beta h_x^*(\sigma)\tau_x}}{2\cosh \beta h_x^*(\sigma)}, \qquad * = SSF, PCA, dl, ur$$

we have

$$\begin{aligned}
P^{SSF}(\sigma,\tau) &= \frac{1}{|\Lambda|} p_x^{SSF}(\sigma,\tau), \qquad \text{with } \tau_y = \sigma_y \quad \forall y \neq x \\
P^{PCA}(\sigma,\tau) &= \prod_{x\in\Lambda} p_x^{PCA}(\sigma,\tau) \\
P^{Sh}(\sigma,\tau) &= \sum_{\sigma'\in\mathcal{X}_\Lambda} \prod_{x\in\Lambda} p_x^{dl}(\sigma,\tau) \prod_{x\in\Lambda} p_x^{ur}(\sigma,\tau).
\end{aligned}$$

For large inverse temperature β we have $p_x^*(\sigma,\tau) \sim 1$ if τ_x is parallel to the local field $h_x^*(\sigma)$. We call such a local move "along the drift". On the other hand $p_x^*(\sigma,\tau) \sim e^{-2\beta|h_x^*(\sigma)|}$ if τ_x is anti-parallel to the local field $h_x^*(\sigma)$. We call such a local move "against the drift".

The SSF dynamics is reversible with Gibbs invariant measure

$$\pi_\Lambda^G(\sigma) = \frac{e^{\beta \sum_x h_x^{SSF}(\sigma)\sigma_x}}{Z}$$

and the invariant measure of PCA and shaken dynamics is $\pi(\sigma)$ given in (8). All these measures, in the regime of large β, concentrate on the configuration with all positive spins $+\mathbb{1}$ representing the stable state. The configuration $-\mathbb{1}$ with all spins -1 represents, in this regime of low temperature, a metastable state. Indeed by considering the first hitting time $\tau_{+\mathbb{1}}$ to $+\mathbb{1}$ starting from $-\mathbb{1}$, for the SSF dynamics we have for any $\delta > 0$ (see for instance [17], [2]):

$$\lim_{\beta\to\infty} P_{-\mathbb{1}}^{SSF}(\tau_{+\mathbb{1}} > e^{\beta(\Gamma-\delta)}) = 1 \tag{24}$$

with

$$\Gamma = 4J\ell_c - 2\lambda\ell_c^2 + 2\lambda(\ell_c - 1)$$

and critical size $\ell_c = \left[\frac{J}{2\lambda}\right] + 1$, where $\left[\,\cdot\,\right]$ denotes the integer part. The typical exit paths from the metastable state $-\mathbb{1}$ follow a sequence of growing squares and rectangles (quasi squares) of plus spins up to the critical size ℓ_c. Starting from a rectangular droplet of plus spins a move against the drift is necessary to create a new line, and the line is completed with subsequent moves along the drift. A similar result holds for the PCA dynamics following the same arguments since again moves along the drift lead to rectangular droplets of plus spins and parallel updating against the drift has small probability for β large.

A different growth takes place in the case of shaken dynamics. Indeed using a similar argument as in [7] it is simple to prove that configurations with complete diagonals of plus spins can be used to construct a competitive way to go from the metastable to the stable state. Starting from the metastable state we have $h_x^{dl}(-\mathbb{1}) = -2J - q + \lambda$ for any x and so with a probability of order $e^{-2\beta(2J+q-\lambda)}$ a spin is flipped in a site x_0. In the subsequent semi-step (up–right interaction) of the dynamics with probability of order one we have plus spins in the sites $x_0^{\leftarrow}$ and $x_0^{\downarrow}$. The diagonal containing these sites grows with probability of order one in the subsequent moves of the shaken dynamics and it is complete after $L/2$ steps, with L the side of the volume Λ. To destroy a complete diagonal of plus spins a first move of probability

$$e^{-2\beta(2J-q+\lambda)}$$

is necessary and every successive erosion has a probability $e^{-2\beta(\lambda-q)}$. On the other hand the probability to construct a new plus diagonal near the first one has a probability

$$e^{-2\beta(2J-q-\lambda)}.$$

Comparing these probabilities, we can obtain the estimate

$$P_\sigma^{Sh}(\tau_{+\mathbb{1}} < T_0) > a$$

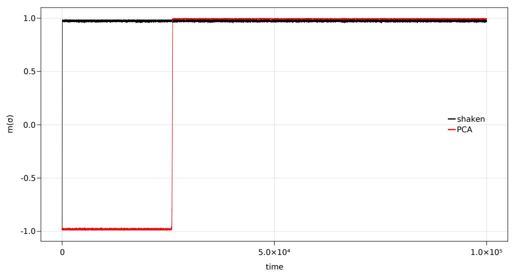
with $T_0 = e^{2\beta[(2J+q-\lambda)]}$ and a not exponentially small in β, for any starting configuration σ. Indeed it is sufficient to require that no move of probability asymptotically smaller than $e^{-2\beta(2J+q-\lambda)}$ takes place within T_0, a first complete diagonal is formed in the time interval $[0, T_0/2]$ and $L-1$ other complete diagonals are formed in the remaining time. We can conclude that for any $\delta > 0$

$$\lim_{\beta\to\infty} P_{-\mathbb{1}}^{Sh}(\tau_{+\mathbb{1}} < e^{2\beta[(2J+q-\lambda)+\delta]}) = 1.$$

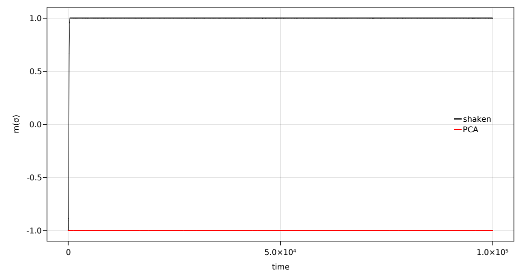
This means that the crossover takes place for the shaken dynamics, typically, within a time corresponding to the time it takes, for the SSF and PCA, to flip the first spin to $+1$. In other words the metastable behavior is no more present in the shaken dynamics (see Figure 4 for a numerical simulation). The asymmetric nature of the interaction gives the shaken dynamics a higher mobility with respect to its symmetric counterpart ("standard" PCA). This is the reason of shorter tunneling times. Note also that this higher mobility causes a slightly smaller magnetization at equilibrium.

This fact has been highlighted in [15, 16] where a comparison between the symmetric PCA and an irreversible PCA with totally asymmetric interaction has been performed in the case $\lambda = 0$. In Figure 5 a comparison between the evolution of the magnetization for a system subject to a symmetric PCA evolution and to shaken dynamics is shown for the same values of the parameters of [16]. It is clear that, with respect to the tunneling behavior, the shaken dynamics retains the same features of the irreversible PCA. However, in the case of the shaken dynamics the control of the invariant measure is more manageable thanks to reversibility.

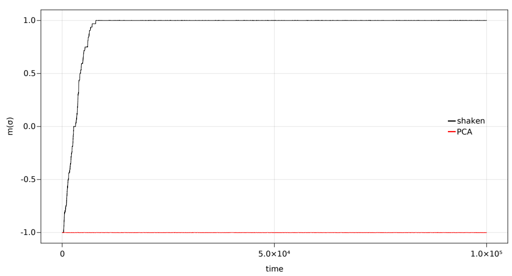
On the other hand, with the same choice of parameters, if we compare the time necessary to reach the stable state $+\mathbb{1}$ starting from a configuration given by a supercritical square of



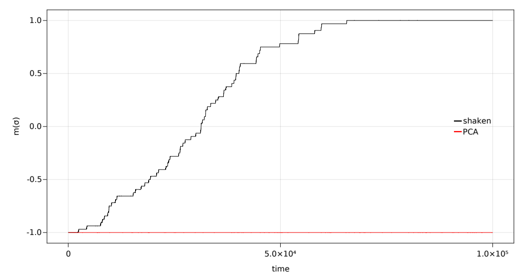
(a) $\beta = 1.3, q = 0.1, \lambda = 0.15$



(b) $\beta = 2.3, q = 0.1, \lambda = 0.15$



(c) $\beta = 3.3, q = 0.1, \lambda = 0.15$



(d) $\beta = 4.0, q = 0.1, \lambda = 0.15$

Figure 4: Comparison of the magnetization over time for PCA and shaken dynamics for several values of the inverse tempeature β.

plus spins in a see of minuses, the PCA dynamics is more rapid than the shaken dynamics if $J > 2q$.

This means that it is not possible to establish, a priori, whether the PCA or the shaken dynamics is faster, but it is necessary to take into account both the application and the starting configuration. This will be the subject of further investigations.

3.3 Geometrical discussion

In the shaken dynamics the idea of alternate dynamics is combined with that of the doubling Hamiltonian. Indeed considering only part of the interaction (for instance down-left first and then up-right in the case of $\Lambda \in \mathbb{Z}^2$ presented at the beginning of the Section) and introducing the inertial parameter of self interaction q it is possible to interpolate between different lattice geometries induced by the doubling graph as already discussed in [1].

Indeed the alternate dynamics on the hexagonal lattice makes possible to interpolate between the square ($q \to \infty$) and the 1-dimensional lattice ($q \to 0$). The interpolation between lattices induced by the shaken dynamics may be applied in general, and in the case of planar graphs the results concerning the critical behavior contained in [1] can be extended, using [3].

Consider for instance the Ising model on the triangular lattice. On this lattice we divide the 6 nearest neighbors of each vertex x into two sets, e.g. $\ell(x)$ left and $r(x)$ right nearest neighbors of x, and define a shaken dynamics with self interaction q. Hence the doubled

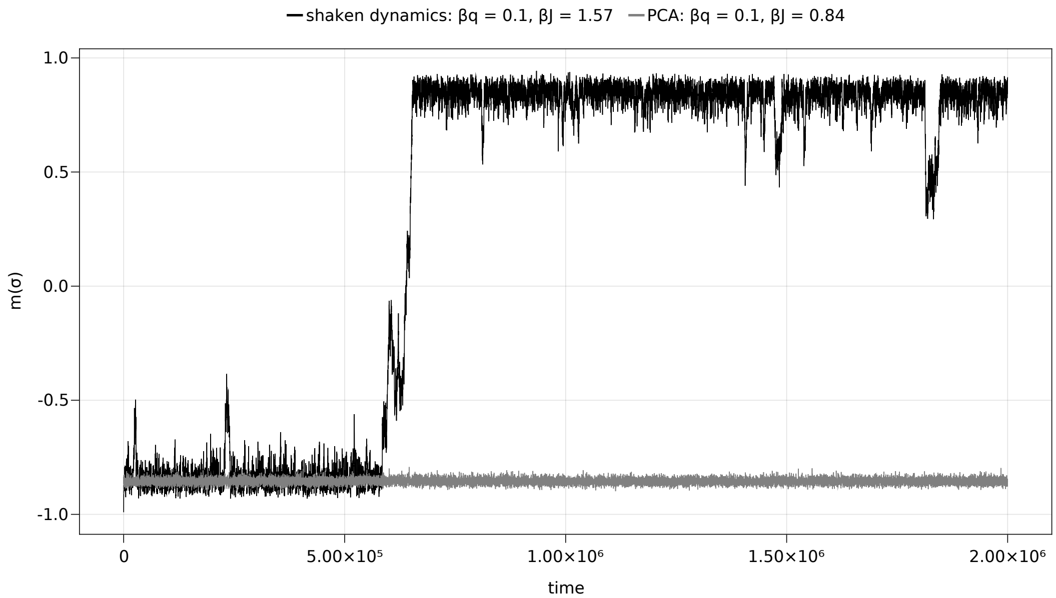


Figure 5: Comparison of the evolution of the magnetization for a spin system evolving according to a shaken dynamics (black) and according to a symmetric PCA (gray). The values of the parameters are such that both dynamics exhibit the same spontaneous magnetization and are consistent with those of [16].

Hamiltonian is

$$H^{\triangle}(\sigma,\tau) = -\sum_x \Big[\sum_{y\in\ell(x)} \big(J\sigma_y\tau_x\big) + q\sigma_x\tau_x \Big] = -\sum_x \Big[\sum_{y\in r(x)} \big(J\tau_y\sigma_x\big) + q\sigma_x\tau_x \Big]$$

The corresponding alternate dynamics turns out to be defined on the square lattice (see Fig. 6) with invariant measure the Gibbs one. In particular the square lattice is regular when we set $J = q$. In this case the parameter q can be used to move through different geometries. The triangular lattice ($q \to \infty$) and the hexagonal lattice ($q = 0$) can be derived from the original square lattice just tuning the value of q. A more precise statement of this interpolation is given by the following

Theorem 3.2 *For the shaken dynamics on the triangular lattice the critical equation relating the parameters J and q is given by*

$$1 + \tanh^3(J)\tanh(q) = 3\tanh(J)\tanh(q) + 3\tanh^2(J) \qquad (25)$$

In the case $q = J$ we obtain the Onsager critical temperature for the square lattice, for $q = 0$ we obtain the critical temperature for the hexagonal lattice and in the limit $q \to \infty$ we obtain the critical temperature for the triangular lattice.

4 The generalized shaken dynamics

We can generalize the construction of the shaken dynamics. Starting from a symmetric interaction $\mathcal{J}$ defining the Hamiltonian $H(\sigma)$, as in (2), we can define an arbitrary

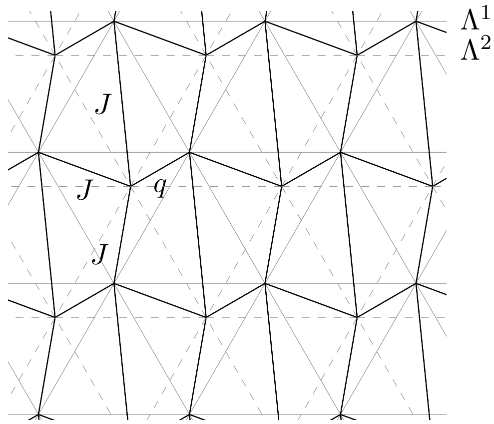


Figure 6: Interaction in the pair Hamiltonian for the shaken dynamics on the triangular lattice. Each spin of configuration σ (living on the solid lattice) interacts with the spin at the same location and the three spins on its left in τ (living on the dashed lattice). The red lines show that the pair interaction lives on a square lattice. For $q = J$ this lattice is homogeneous. As $q \to \infty$ the square lattice collapses onto the triangular lattice. If $q = 0$ the interaction graph becomes the homogeneous hexagonal lattice.

decomposition of the interaction matrix $\mathcal{J}$ in a sum of two matrices with non negative entries

$$\mathcal{J} = \mathcal{J}^o + \mathcal{J}^{oT}. \tag{26}$$

This means that every non oriented edge $\{x, y\}$ with weight J_{xy} is decomposed in a pair of oriented edges (x, y) and (y, x) with weight respectively $\mathcal{J}^o_{xy}$ and $\mathcal{J}^o_{yx}$. Call E^o the set of all these oriented edges and apply the construction presented in Section 2 to construct the doubling graph by using this set E^o of oriented edges.

We proceed as before defining the doubling Hamiltonian

$$\begin{aligned} H(\sigma^{(1)}, \sigma^{(2)}) &= -\langle \sigma^{(1)}, \mathcal{J}^o \sigma^{(2)} \rangle + q\langle \sigma^{(1)}, \sigma^{(2)} \rangle + \langle \lambda, \sigma^{(1)} \rangle + \langle \lambda, \sigma^{(2)} \rangle \\ &= -\langle \mathcal{J}^{oT} \sigma^{(1)}, \sigma^{(2)} \rangle + q\langle \sigma^{(1)}, \sigma^{(2)} \rangle + \langle \lambda, \sigma^{(1)} \rangle + \langle \lambda, \sigma^{(2)} \rangle. \end{aligned}$$

In the case $\sigma^{(1)} = \sigma^{(2)} = \sigma$ by equation (26) we have again $H(\sigma, \sigma) = H(\sigma) - q|V|$.

The corresponding alternate dynamics on the state space $\mathcal{X}_V$ is defined with two subsequent updating as follows:

$$P^{1\to 2}(\sigma, \sigma') := \frac{e^{-H(\sigma,\sigma')}}{\overrightarrow{Z}_\sigma}, \qquad P^{2\to 1}(\sigma', \tau) := \frac{e^{-H(\tau,\sigma')}}{\overleftarrow{Z}_{\sigma'}} \tag{27}$$

and

$$P^{sh}(\sigma, \tau) = \sum_{\sigma' \in \mathcal{X}_V} P^{1\to 2}(\sigma, \sigma') P^{2\to 1}(\sigma', \tau) = \sum_{\sigma' \in \mathcal{X}_V} \frac{e^{-H(\sigma,\sigma')}}{\overrightarrow{Z}_\sigma} \frac{e^{-H(\tau,\sigma')}}{\overleftarrow{Z}_{\sigma'}} \tag{28}$$

The results obtained in Theorem 2.3 can be immediately extended to this more general case.

The choice of the shaken dynamics discussed in Section 2 is a particular case of generalized shaken dynamics in which $\mathcal{J}^o_{xy}\mathcal{J}^o_{yx} = 0$ for any pair x, y. In the general case the geometrical discussion of the doubling graph of interaction is much more complicated. Also the interpolation between different geometries obtained for different values of the parameter q, as discussed in Section 3, is more involved in this generalized case.

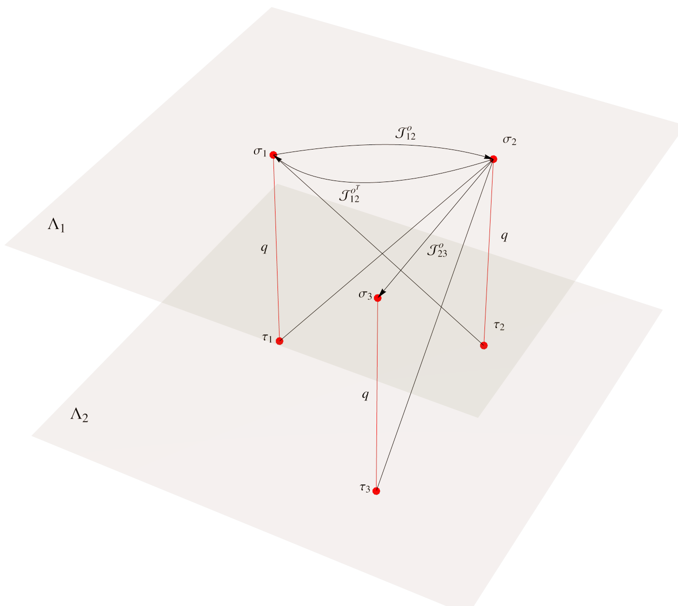


Figure 7: The construction of the doubling graph in the case of the generalized shaken dynamics.

Another particular choice in this class of generalized shaken dynamics is $\mathcal{J}^o = \frac{1}{2}\mathcal{J}$ corresponding to the PCA discussed in [6].

5 Application to optimization problems

The shaken dynamics on a general graph, and its generalization, gets the possibility to look for the minimum of a general Hamiltonian $H(\sigma)$ defined on $\{-1, +1\}^V$ by means of a parallel dynamics, by using the following result that could be considered a corollary of Theorem 2.3. In combinatorial optimization this can be used as a parallel approach to the Quadratic Unconstrained Binary Optimization (QUBO) i.e., the problem of minimizing a quadratic polynomial of binary variables (see [**?**] for a survey).

Theorem 5.1 *Given a Hamiltonian $H(\sigma)$ of the form given in (2) on $\{-1, +1\}^V$, for any Hamiltonian $H(\sigma, \tau)$ which is the doubling of $H(\sigma)$, corresponding to a bipartite graph*

$G^b = (V^b, E^b)$, *if*

$$q > \max_{x \in V} \Big\{ \sum_{y:\{x,y\} \in E^b} |J_{xy}| + |\lambda_x| \Big\} \tag{29}$$

then the alternate dynamics defined with $H(\sigma, \tau)$ is a parallel algorithm to find configurations σ minimizing $H(\sigma)$.

To assess the effectiveness of the strategy presented in Theorem 5.1, we put forward some preliminary tests on a simplified version of the Edwards-Anderson model where the weight of the edges connecting neighboring sites is set to $J = +1$ with probability $\frac{1}{2}$ and $J = -1$ with probability $\frac{1}{2}$ and where the external field is zero. In this case, setting $q > 2$ is sufficient to satisfy the hypotheses of the corollary. We compared the results with those obtained with a single spin flip heat bath dynamics and considered "grids" with side length 128 and 256. With this setting, the heuristic minima that we obtained with the shaken dynamics are essentially equivalent to those obtained with the single spin flip dynamics. However the speed up with respect to the single spin flip dynamics was significant. To be as fair as possible in this comparison, we renormalized the time of the single spin flip dynamics with the number of vertices in the graph so to have the same number of "attempted spin flips". We observed a speed-up of about 10 times when considering, for both algorithms, a CPU implementation and up to 200 times when comparing the CPU implementation of the single spin flip dynamics with a GPU implementation of the shaken dynamics. We believe these preliminary results to be rather encouraging and we plan to perform a more thorough investigation of the performances of the shaken dynamics to find the minimizers of $H(\sigma)$ in a future work taking into account the considerations concerning the mobility of the dynamics.

6 Proofs of the results

6.1 Proof of Theorem 2.3

We have immediately the detailed balance condition w.r.t. the measure $\pi(\sigma)$ indeed

$$\sum_{\sigma' \in \mathcal{X}_V} \frac{e^{-(H(\sigma,\sigma')+H(\tau,\sigma'))}}{\overleftarrow{Z}_{\sigma'}} = \overrightarrow{Z}_\sigma P^{sh}(\sigma, \tau) = \overrightarrow{Z}_\tau P^{sh}(\tau, \sigma) = \sum_{\sigma' \in \mathcal{X}_V} \frac{e^{-(H(\tau,\sigma')+H(\sigma,\sigma'))}}{\overleftarrow{Z}_{\sigma'}} \tag{30}$$

It is straightforward to prove that $\pi^b(\sigma^{(1)}, \sigma^{(2)})$ is the stationary measure of P^{alt}

$$\begin{aligned}\sum_{\sigma^{(1)},\sigma^{(2)}} \pi^b(\sigma^{(1)}, \sigma^{(2)}) P^{alt}(\boldsymbol{\sigma},\boldsymbol{\tau}) &= \sum_{\sigma^{(1)},\sigma^{(2)}} \frac{e^{-H(\sigma^{(1)},\sigma^{(2)})}}{Z} \frac{e^{-H(\sigma^{(1)},\tau^{2})}}{\overrightarrow{Z}_{\sigma^{(1)}}} \frac{e^{-H(\tau^{(1)},\tau^{(2)})}}{\overleftarrow{Z}_{\tau^{(2)}}} \\ &= \frac{e^{-H(\tau^{(1)},\tau^{(2)})}}{Z} = \pi^b(\tau^{(1)}, \tau^{(2)})\end{aligned} \tag{31}$$

□

Note that, in general

$$\pi^b(\sigma^{(1)}, \sigma^{(2)}) P^{alt}(\boldsymbol{\sigma},\boldsymbol{\tau}) \neq \pi^b(\tau^{(1)}, \tau^{(2)}) P^{alt}(\boldsymbol{\tau},\boldsymbol{\sigma}).$$

For instance consider the bipartite graphs $K_{n,n}$ with equal weights on all edges and where, for all i, $(\sigma_i^{(1)}, \sigma_i^{(2)}) = (+1, +1)$ and $(\tau_i^{(1)}, \tau_i^{(2)}) = (+1, -1)$

6.2 Proof of Theorem 3.1

To prove Theorem 3.1 it is possible to argue as in the proof of Theorem 1.2 in [18].

In our notation π_Λ and π_Λ^G have the role, respectively, of π_{PCA} and π_G used in [18]. Further let $g_x(\sigma) := J(\sigma_{x^\downarrow} + \sigma_{x^\leftarrow})$ be the analogue of $h_i(\sigma)$ in [18]. Here we assume $\lambda < 0$. The case $\lambda > 0$ can be treated likewise.

Recalling that $\delta = e^{-2q}$, it is possible to write Z_σ in the following way:

$$\begin{aligned} Z_\sigma &= \sum_\tau e^{-H(\sigma,\tau)} = \sum_\tau e^{-H(\sigma,\sigma)} e^{-[H(\sigma,\tau)-H(\sigma,\sigma)]} \\ &= e^{q|\Lambda|} e^{-H(\sigma)} \sum_\tau e^{\sum_{x:\sigma_x \neq \tau_x} -2g_x(\sigma)\sigma_x - 2q - 2\lambda\sigma_x} \\ &= e^{q|\Lambda|} e^{-H(\sigma)} \sum_{I \subset \Lambda} \delta^{|I|} \prod_{x \in I} e^{-2g_x(\sigma)\sigma_x - 2\lambda\sigma_x} \\ &= e^{q|\Lambda|} e^{-H(\sigma)} \prod_{x \in \Lambda} (1 + \delta e^{-2g_x(\sigma)\sigma_x - 2\lambda\sigma_x}) \end{aligned} \tag{32}$$

where the sum over τ has been rewritten as the sum over all subsets $I \subset \Lambda$ such that $\tau_x = -\sigma_x$ if $x \in I$ and $\tau_x = \sigma_x$ otherwise. The factor $e^{q|\Lambda|}$ does not depend on σ and cancels out in the ratio $\frac{Z_\sigma}{Z}$.

Call $f(\sigma) := \prod_{x\in\Lambda}(1 + \delta e^{-2g_x(\sigma)\sigma_x - 2\lambda\sigma_x})$, $\;w(\sigma) := e^{-H(\sigma)} f(\sigma) = w^G(\sigma) f(\sigma)$. Then (32) can be rewritten as

$$\pi_\Lambda(\sigma) = \frac{w(\sigma)}{\sum_\tau w(\tau)} = \frac{w^G(\sigma) f(\sigma)}{\sum_\tau w^G(\tau) f(\tau)} = \frac{\frac{w^G(\sigma)}{Z^G} f(\sigma)}{\sum_\tau \frac{w^G(\tau)}{Z^G} f(\tau)} = \frac{\pi_\Lambda^G(\sigma) f(\sigma)}{\pi_\Lambda^G(f)}$$

with $\pi_\Lambda^G(f) = \sum_\sigma \pi_\Lambda^G(\sigma) f(\sigma)$.

As in [18], using Jensen's inequality the total variation distance between π_Λ and π_Λ^G can be bounded as

$$\|\pi_\Lambda - \pi_\Lambda^G\|_{TV} \leq \sqrt{\frac{\pi_\Lambda^G(f^2)}{(\pi_\Lambda^G(f))^2} - 1} =: \sqrt{(\Delta(\delta))}.$$

To prove the theorem, it will be shown that $\Delta(\delta) = O(\delta^2|\Lambda|)$.

By writing $\Delta(\delta) = e^{\log(\pi_\Lambda^G(f^2)) - 2\log(\pi_\Lambda^G(f))} - 1$, the claim follows by showing that the argument of the exponential divided by $|\Lambda|$ is analytic in δ and that the first order term of its expansion in δ cancels out.

In other words the claim follows thanks to the following lemma.

Lemma 6.1 *There exists J_c such that, for all $J > J_c$*

1. $\frac{\log(\pi_\Lambda^G(f^2))}{|\Lambda|}$ *and* $\frac{\log(\pi_\Lambda^G(f))}{|\Lambda|}$ *are analytic in δ for $|\delta| < \delta_J$*

2. $\frac{\log(\pi_\Lambda^G(f^2))}{|\Lambda|} - 2\frac{\log(\pi_\Lambda^G(f))}{|\Lambda|} = O(\delta^2)$

Proof: The analyticity of $\frac{\log(\pi_\Lambda^G(f^2))}{|\Lambda|}$ and $\frac{\log(\pi_\Lambda^G(f))}{|\Lambda|}$ is proven by showing that these quantities can be written as partition functions of an abstract polymer gas. The analyticity is obtained using standard cluster expansion.

To carry over this task, we will rewrite $\pi_\Lambda^G(f^k)$ in terms of standard Peierls contours. Divide the sites in Λ according to the value of the spins and number of edges of the Peierls contour left and below the site in the following way:

- $\Lambda_{-\underset{-}{-}}$: $\{x \in \Lambda : \sigma_x = -1 \wedge (\sigma_{x^\leftarrow} = -1, \sigma_{x^\downarrow} = -1)\}$;
- $\Lambda_{+\underset{-}{-}}$: $\{x \in \Lambda : \sigma_x = -1 \wedge ((\sigma_{x^\leftarrow} = +1, \sigma_{x^\downarrow} = -1) \vee (\sigma_{x^\leftarrow} = -1, \sigma_{x^\downarrow} = +1))\}$;
- $\Lambda_{+\underset{+}{-}}$: $\{x \in \Lambda : \sigma_x = -1 \wedge \sigma_{x^\leftarrow} = +1, \sigma_{x^\downarrow} = +1\}$;
- $\Lambda_{+\underset{+}{+}}$: $\{x \in \Lambda : \sigma_x = +1 \wedge (\sigma_{x^\leftarrow} = +1, \sigma_{x^\downarrow} = +1)\}$;
- $\Lambda_{-\underset{+}{+}}$: $\{x \in \Lambda : \sigma_x = +1 \wedge ((\sigma_{x^\leftarrow} = +1, \sigma_{x^\downarrow} = -1) \vee (\sigma_{x^\leftarrow} = -1, \sigma_{x^\downarrow} = +1))\}$;
- $\Lambda_{-\underset{-}{+}}$: $\{x \in \Lambda : \sigma_x = +1 \wedge (\sigma_{x^\leftarrow} = -1, \sigma_{x^\downarrow} = -1)\}$;

With this notation, $f(\sigma)$ can be written as

$$\begin{aligned} f(\sigma) = (1+\delta e^{-4J+2\lambda})^{|\Lambda|} & \prod_{x\in\Lambda_{+\underset{-}{-}}} \frac{(1+\delta e^{+2\lambda})}{(1+\delta e^{-4J+2\lambda})} \prod_{x\in\Lambda_{+\underset{+}{-}}} \frac{(1+\delta e^{+4J+2\lambda})}{(1+\delta e^{-4J+2\lambda})} \\ & \prod_{x\in\Lambda_{+\underset{+}{+}}} \frac{(1+\delta e^{-4J-2\lambda})}{(1+\delta e^{-4J+2\lambda})} \prod_{x\in\Lambda_{-\underset{+}{+}}} \frac{(1+\delta e^{-2\lambda})}{(1+\delta e^{-4J+2\lambda})} \prod_{x\in\Lambda_{-\underset{-}{+}}} \frac{(1+\delta e^{+4J-2\lambda})}{(1+\delta e^{-4J+2\lambda})} \\ & = (1+\delta e^{-4J+2\lambda})^{|\Lambda|}\tilde{\xi}(\sigma,\lambda) \end{aligned} \tag{33}$$

with

$$\begin{aligned} \tilde{\xi}(\sigma,\lambda) = & \left[\frac{(1+\delta e^{+2\lambda})}{(1+\delta e^{-4J+2\lambda})}\right]^{\left|\Lambda_{+\underset{-}{-}}\right|} \left[\frac{(1+\delta e^{+4J+2\lambda})}{(1+\delta e^{-4J+2\lambda})}\right]^{\left|\Lambda_{+\underset{+}{-}}\right|} \\ & \left[\frac{(1+\delta e^{-4J-2\lambda})}{(1+\delta e^{-4J+2\lambda})}\right]^{\left|\Lambda_{+\underset{+}{+}}\right|} \left[\frac{(1+\delta e^{-2\lambda})}{(1+\delta e^{-4J+2\lambda})}\right]^{\left|\Lambda_{-\underset{+}{+}}\right|} \left[\frac{(1+\delta e^{+4J-2\lambda})}{(1+\delta e^{-4J+2\lambda})}\right]^{\left|\Lambda_{-\underset{-}{+}}\right|} \end{aligned} \tag{34}$$

For a given a configuration $\sigma \in \mathcal{X}_\Lambda$, we denote by $\gamma(\sigma)$ its Peierls contour in the dual $\mathcal{B}_\Lambda^* = \cup_{(x,y)\in\mathcal{B}_\Lambda}(x,y)^*$

$$\gamma(\sigma) := \{(x,y)^* \in \mathcal{B}_\Lambda^* : \ \sigma_x\sigma_y = -1\} \tag{35}$$

Noting that $e^{-H(\sigma)} = e^{(2J-2\lambda)|\Lambda|}e^{-2J|\gamma(\sigma)|+4\lambda|V_+(\sigma)|}$, with $|V_+(\sigma)| = \sum_{x\in\Lambda} 1_{\{\sigma_x=+1\}}$ is the number of plus spins in Λ of configuration σ, we have

$$\pi_\Lambda^G(f^k) = \frac{1}{Z^G}e^{(2J-2\lambda)|\Lambda|}(1+\delta e^{-4J+2\lambda})^{k|\Lambda|}\sum_\sigma \left[e^{-2J|\gamma(\sigma)|+4\lambda|V_+(\sigma)|}\tilde{\xi}^k(\sigma,\lambda)\right] \tag{36}$$

Setting

$$
\xi(\sigma,\lambda) = \left[\frac{(1+\delta e^{+2\lambda})}{(1+\delta e^{-4J+2\lambda})}\right]^{\left|\Lambda_{+\substack{-\\-}}\right|} \left[\frac{(1+\delta e^{+4J+2\lambda})}{(1+\delta e^{-4J+2\lambda})}\right]^{\left|\Lambda_{+\substack{-\\+}}\right|}
\left[\frac{e^{+2\lambda}(1+\delta e^{-4J-2\lambda})}{(1+\delta e^{-4J+2\lambda})}\right]^{\left|\Lambda_{+\substack{+\\+}}\right|} \left[\frac{e^{+2\lambda}(1+\delta e^{-2\lambda})}{(1+\delta e^{-4J+2\lambda})}\right]^{\left|\Lambda_{-\substack{+\\+}}\right|} \left[\frac{e^{+2\lambda}(1+\delta e^{+4J-2\lambda})}{(1+\delta e^{-4J+2\lambda})}\right]^{\left|\Lambda_{-\substack{+\\-}}\right|} \tag{37}
$$

allows us to write, for $k \in \{1,2\}$,

$$
\sum_\sigma \left[e^{-2J|\gamma(\sigma)|+4\lambda|V_+(\sigma)|}\tilde{\xi}^k(\sigma,\lambda)\right] = \sum_\sigma \left[e^{-2J|\gamma(\sigma)|}\left(e^{+2\lambda|V_+(\sigma)|}\right)^{2-k}\xi^k(\sigma,\lambda)\right] \tag{38}
$$

A straightforward computation yields $\xi^k(\sigma,\lambda) \le \xi^k(\sigma,0)$ and then

$$
\sum_\sigma \left[e^{-2J|\gamma(\sigma)|}\left(e^{+2\lambda|V_+(\sigma)|}\right)^{2-k}\xi^k(\sigma,\lambda)\right] \le \sum_\sigma e^{-2J|\gamma(\sigma)|}\xi^k(\sigma,0) = 2\sum_\gamma e^{-2J|\gamma|}\xi^k(\gamma,0)
$$

where $\xi^k(\gamma,0)$ coincides with $\xi^I_k(\Gamma)$ in the proof of Lemma 2.3 in [18], with $\left|\Lambda_{+\substack{-\\-}}\right| + \left|\Lambda_{-\substack{+\\+}}\right| = |l_1(\Gamma)|$ and $\left|\Lambda_{+\substack{-\\+}}\right| + \left|\Lambda_{-\substack{+\\-}}\right| = |l_2(\Gamma)|$.

This implies that the proof can be concluded following the same steps as in [18]. □

6.3 Proof of Theorem 3.2

This is an application of Theorem 1.1 in [3] holding for a finite planar, non degenerate and doubly periodic weighted graph $G = (V, E)$. Denote by $\mathcal{E}(G)$ the set of all even subgraphs of G, that is, those subgraphs where the degree of each vertex is even. Further call $\mathcal{E}_0(G)$ the set of even subgraphs of the lattice winding an even number of times around each direction of the torus and $\mathcal{E}_1(G) = \mathcal{E}(G) \setminus \mathcal{E}_0(G)$. Then the critical curve relating the parameters J and q of the Hamiltonian is the solution of the equation

$$
\sum_{\gamma\in\mathcal{E}_0(G)}\prod_{e\in\gamma}\tanh J_e = \sum_{\gamma\in\mathcal{E}_1(G)}\prod_{e\in\gamma}\tanh J_e. \tag{39}
$$

The square lattice induced by the shaken dynamics on the triangular lattice, with $J_e = q$ for the self–interaction edges and $J_e = J$ for the other edges, satisfies the hypotheses of this theorem and can be obtained by periodically repeating the elementary cell of Figure 8.

A direct application of (39) yields the claim. □

6.4 Proof of Theorem 5.1

Let $H(\sigma)$ be a Hamiltonian of the form given in (2) and let $H(\sigma,\tau)$ be its doubling.

The invariant measure π^b of the alternate dynamics P^{alt} defined on the bipartite graph G^b

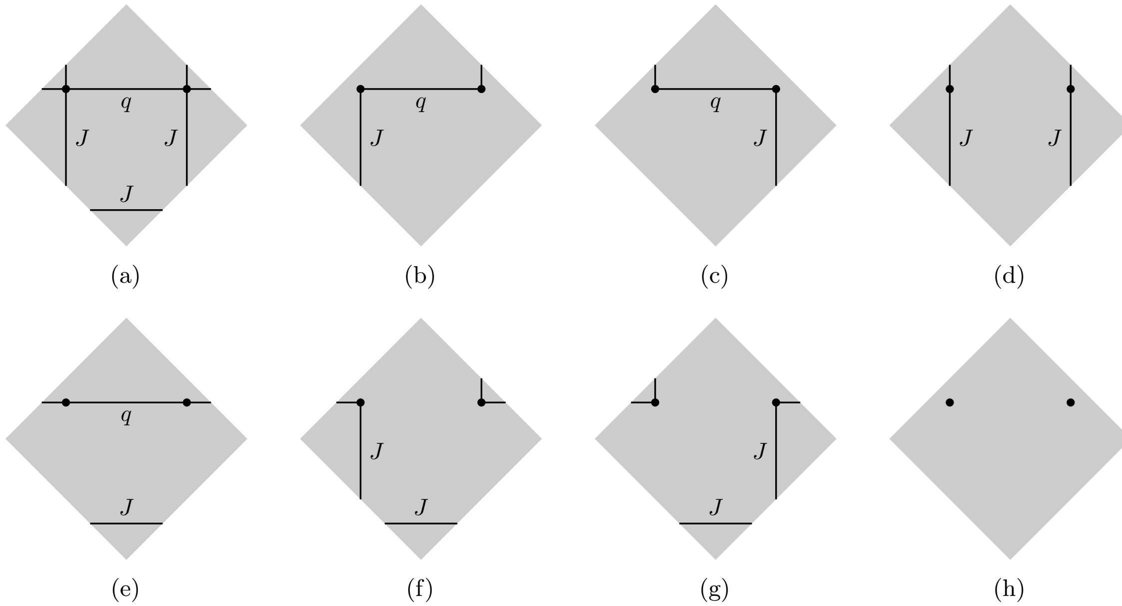


Figure 8: The elementary cell (a) for the shaken interaction on the triangular lattice and the corresponding even subgraphs. Subgraphs (a) and (h) wind around the torus an even number of times and are, therefore, in $\mathcal{E}_0$ whereas the remaining subgraphs are in $\mathcal{E}_1$.

with Hamiltonian $H(\sigma, \tau)$ is identified in Theorem 2.3:

$$\pi^b(\boldsymbol{\sigma}) := \frac{1}{Z} e^{-H(\boldsymbol{\sigma})}.$$

At very low temperature, this measure concentrates on the set of configurations minimizing the Hamiltonian $H(\sigma, \tau)$. If the following condition is satisfied

$$\min_{\sigma,\tau} H(\sigma, \tau) = \min_{\sigma} H(\sigma, \sigma) \tag{40}$$

the parallel algorithm provided by the alternate dynamics may be used to find configurations minimizing $H(\sigma)$, since $\min_\sigma H(\sigma, \sigma) = \min_\sigma H(\sigma) - q|V|$.

Equation (40) can be verified by contradiction. Assume that there exists a pair configuration $(\bar{\sigma}, \bar{\tau})$ such that

$$\min_{\sigma,\tau} H(\sigma, \tau) = H(\bar{\sigma}, \bar{\tau})$$

and $\bar{\sigma} \neq \bar{\tau}$ at least in a vertex $x \in V$. If the parameter q satisfies condition (29), a spin flip at vertex x leads to a lower value for the doubling Hamiltonian, contradicting the hypothesis that the pair $(\bar{\sigma}, \bar{\tau})$ is the minimizer of $H(\sigma, \tau)$. $\square$

7 Conclusions and open problems

We briefly conclude our paper with some general comments and open problems.

With the shaken dynamics we have constructed a reversible parallel dynamics and we control its invariant measure with arbitrary boundary conditions. The advantages of the shaken dynamics can be summarized as follows:

- The shaken prescription can be applied to general interaction graphs. This allows to construct parallel algorithms to tackle a large class of optimization problems.
- The shaken prescription, modifying suitably the parameters appearing in the doubled Hamiltonian, allows to compare the spin systems defined on different geometries. In particular, when $q \to 0$, the system tends to a geometry with a different dimension.
- The dynamics can be interpreted as a model for systems in which some kind of interaction alternates its direction on short timescale. See below for an example referring to the tidal dissipation.

An open problem concerns the estimate of the mixing time for the shaken dynamics. In this regard, it is interesting to investigate the role of the parameter q when studying decay of metastable states in the low temperature regime. This could lead to detailed estimates of the mixing time and it will be the subject of further investigation.

As noted by an anonymous referee, the shaken dynamics could be extended in order to consider spin systems with more general summable interactions, not necessarily limited to two or one body terms. This could have important applications in combinatorial optimization problems such as set covering problem.

The construction of the shaken dynamics and, in particular, of its generalization, is not a unique prescription. This freedom in the definition of the oriented graph defining the dynamics and in the choice of the parameters involved could be usefully exploited in applications to speed up the dynamics.

Finally we want to outline that the presence of an alternate interaction suggests that the shaken dynamics, with $B \neq \emptyset$ and q large, could be a good model to take into account the effects of Earth's tides in geodynamics and other tidal dissipative phenomena in Solar System. We assume that the inner structure of the Earth and of the satellites of the major planets may be described in terms of constraints that can be randomly broken, with a probability depending on the state of the nearest neighbors of each constraint. Tidal effects could give a dependence of this breaking probability on an alternate direction, related to the tidal state and to the related tidal currents. This geological and astronomical application will be developed in forthcoming papers.

Acknowledgments: The authors are grateful to Emilio Cirillo, Alexandre Gaudillière and Gabriella Pinzari for useful discussions. B.S. and E.S. thank the support of the A*MIDEX project (n. ANR-11-IDEX-0001-02) funded by the “Investissements d'Avenir" French Government program, managed by the French National Research Agency (ANR). B.S. acknowledges the MIUR Excellence Department Project awarded to the Department of Mathematics, University of Rome Tor Vergata, CUP E83C18000100006. E.S. has been supported by the PRIN 20155PAWZB “Large Scale Random Structures". A.T. has been supported by Project FARE 2016 Grant R16TZYMEHN.

References

[1] V. Apollonio, R. D'Autilia, B. Scoppola, E. Scoppola, A. Troiani, *Criticality of measures on 2-d Ising configurations: from square to hexagonal graphs*, J. Statist. Phys., **177**, 1009-1021 (2019).

[2] A. Bovier, F. den Hollander, *Metastability: a potential-theoretic approach* Vol. **351**. Springer, 2016.

[3] D. Cimasoni, H. Duminil-Copin, *The critical temperature for the Ising model on planar doubly periodic graphs*, Electron. J. Probab., **18**, no. 44, 1–18 (2013).

[4] E.N.M. Cirillo, *A Note on the Metastability of the Ising Model: The Alternate Updating Case*, J. Statist. Phys., **106**, 385-390 (2002).

[5] E.N.M. Cirillo, F.R. Nardi, C. Spitoni, *Metastability for reversible probabilistic cellular automata with self-interaction*, J. Statist. Phys., **132**, 431-471 (2008).

[6] P. Dai Pra, B. Scoppola, E.Scoppola, *Sampling from a Gibbs measure with pair interaction by means of PCA*, J. Statist. Phys., **149**, 722-737 (2012).

[7] P. Dai Pra, B. Scoppola, E. Scoppola, *Fast mixing for the low-temperature 2D Ising model through irreversible parallel dynamics* J. Statist. Phys., **159**, 1-20 (2015).

[8] R. Fernández, A. Toom, *Non Gibbsiannes of the invariant measures of non-reversible cellular automata with totally asymmetric noise*, Astérisque, **287**, 71-87 (2003).

[9] B.H. Fukushima–Kimura, S. Handa, K. Kamakura, Y. Kamijima, A. Sakai, *Mixing time and simulated annealing for the stochastic cellular automata*, arXiv:2007.11287.

[10] S. Goldstein, R. Kuik, J.L. Lebowitz, C. Maes, *From PCA's to equilibrium systems and back*, Commun. Math. Phys., **125**, 71-79 (1989).

[11] H. Guo, M. Jerrum *Random Cluster Dynamics for the Ising Model is Rapidly Mixing* The Annals of Applied Probability **28**, 1292-1313 (2018).

[12] O. Häggström, *Finite Markov chains and algorithmic applications*, **52**, Cambridge University Press, (2002).

[13] M. Kaiser, R.L. Jack, J. Zimmer, *Acceleration of Convergence to Equilibrium in Markov Chains by Breaking the Detailed Balance* J. Statist. Phys., **168**, 259-287 (2017).

[14] O. Kozlov, N. Vasilyev, *Reversible Markov chains with local interaction, multi-component random systems* Adv. Probab. Related Topics, **6**, 451–469 (1980).

[15] C. Lancia, B. Scoppola, *Equilibrium and non-equilibrium Ising models by means of PCA*, J. Statist. Phys., **153**, 641-653 (2013).

[16] C. Lancia C., B. Scoppola *Ising Model on the Torus and PCA Dynamics: Reversibility, Irreversibility, and Fast Tunneling*, In: Louis PY., Nardi F. (eds) Probabilistic Cellular Automata. Emergence, Complexity and Computation, vol **27**, Springer (2018).

[17] E. Olivieri, M.E. Vares, *Large deviations and metastability*, **100**, Cambridge University Press, (2005).

[18] A. Procacci, B. Scoppola, E. Scoppola, *Probabilistic Cellular Automata for the low-temperature 2d Ising Model*, J. Statist. Phys., **165**, 991-1005 (2016).

[19] A. Procacci, B. Scoppola, E. Scoppola, *Effects of boundary conditions on irreversible dynamics* Annales Henri Poincaré, **19** (2018).

[20] B. Scoppola, A. Troiani, *Gaussian mean fields lattice gas*, J. Statist. Phys., **170** 1161–1176 (2018).