Journal of Statistical Physics, Vol. 119, Nos. 1/2, April 2005 (© 2005)
DOI: 10.1007/s10955-004-2126-6

Mesoscopic Modeling for Continuous Spin Lattice Systems: Model Problems and Micromagnetics Applications

Markos A. Katsoulakis,[1] Petr Plecháč,[2] and Dimitrios K. Tsagkarogiannis[1]

Received July 5, 2004; accepted October 19, 2004

In this paper we derive deterministic mesoscopic theories for model continuous spin lattice systems both at equilibrium and non-equilibrium in the presence of thermal fluctuations. The full magnetic Hamiltonian that includes singular integral (dipolar) interactions is also considered at equilibrium. The non-equilibrium microscopic models we consider are relaxation-type dynamics arising in kinetic Monte Carlo or Langevin-type simulations of lattice systems. In this context we also employ the derived mesoscopic models to study the relaxation of such algorithms to equilibrium.

KEY WORDS: Heisenberg spin lattice system; Kac potential; large deviations; statistical equilibrium; Monte Carlo methods; relaxation dynamics.

1. INTRODUCTION

The statistical mechanics of lattice spin systems with a continuous vector-valued order parameter provide an important modeling and computational tool capable of describing magnetic materials at equilibrium and out-of-equilibrium states[1,2]. Typically such models can include detailed, material-dependent interactions such as anisotropy and exchange energy between spins, the latter derived either experimentally or from quantum mechanics calculations, as well as dipolar interactions and external magnetic and electric fields. Furthermore, thermal fluctuations are incorporated in the lattice models. On the other hand magnetic materials exhibit

[1]Department of Mathematics and Statistics, University of Massachusetts, Amherst, MA 01003–9305, USA; e-mails: markos@math.umass.edu and tsagkaro@math.umass.edu
[2]Mathematics Institute, University of Warwick, Coventry, CV4 7AL, UK; e-mail: plechac@maths.warwick.ac.uk

various types of magnetic domains, walls and magnetic vortices, at scales much larger than the length/time scales accessible in simulations of microscopic lattice spin systems. The Landau–Lifschitz model for micromagnetics[3] captures such large scale features without, however, incorporating explicitly detailed interactions and thermal fluctuations.

In relatively small scale devices and thin films, thermal fluctuations are expected to affect the nature of the magnetic domain walls and need to be incorporated in the mesoscopic modeling directly from the microscopic statistical mechanics. Therefore a systematic or even rigorous derivation of mesoscopic models that describe the system in terms of coarse variables (e.g., local average magnetization or local one-spin probability density function) would help to properly incorporate thermal fluctuations into numerical simulations performed on much larger than microscopic scales. In this paper we derive such a mesoscopic description for model spin lattice systems both at equilibrium and non-equilibrium. The full micromagnetics problem with singular (dipolar) interactions is also considered in the equilibrium case. The non-equilibrium microscopic models we consider are relaxation-type dynamics arising in Monte Carlo or Langevin-type simulations of lattice systems[1]. In this context we employ the mesoscopic models derived here to study the relaxation of such algorithms to equilibrium.

In Section 2.2 we study the asymptotics of the equilibrium state for a model Hamiltonian with a Kac interaction potential by recasting the thermodynamic limit of the canonical Gibbs measure as a Large Deviation problem. These results follow without significant modifications from recent work in[4] and[5]. We briefly state them and discuss their implications for completeness and comparison to the micromagnetics problem studied in Section 2.3, where the main new difficulty is the presence of a singular integral operator (SIO) associated with the dipolar interaction. In this case we derive, from the canonical Gibbs measure as a large deviation limit, a functional over the set of local single spin probability density functions (PDF) that includes energy interactions and entropy due to thermal fluctuations. By a single spin PDF we essentially mean the probability density of having a spin $v \in \mathbb{S}^2$ at the spatial location x. Furthermore, we obtain in Section 2.4 a finite temperature free energy for the average magnetization, i.e., the first moment of the single spin PDF. In Section 2.5 we discuss the formation of domain walls as heteroclinic orbits of the newly derived free energy.

In Section 3 we focus on relaxational dynamics of kinetic Monte Carlo (KMC). We first derive an evolution equation for the local single spin PDF and present a closed equation for the first moment of the PDF, i.e., the average magnetization. The resulting equations are reminiscent of kinetic Fokker–Planck–Vlasov equations (without the convection term) in the sense that they have a Lyapunov functional which is the functional derived for the PDFs in the equilibrium theory. The presented analysis is a continuation of work on mesoscopic equations derived for the dynamic Ising model with Kac interactions (see, e.g., Refs. 6–9). In the case studied here additional difficulty arises from the fact that spins take continuous values, instead of discrete ones as in the Ising and Potts models. Consequently we need to work mostly with an equation for the single spin PDF rather than with the local average magnetization, which is the case in Ising systems. In Section 3.2 we study analogous questions for the Langevin dynamics. Furthermore, using spectral and relative entropy estimates we obtain relaxation and relaxation rate results to the minimizing PDF of the equilibrium functional obtained in Section 2. These relaxation results in conjunction with the connection of finite-temperature free energies with the Landau-Lifschitz model established in Section 2.5, provide a rigorous link between two disparate types of modeling and simulation encountered in the literature.

2. EQUILIBRIUM THEORY

2.1. Statistical Mechanics Modeling

First we summarize the statistical mechanics background of continuous spin lattice systems in micromagnetics and subsequently we discuss a model problem that retains some of the fundamental features of the physical system. We consider a d-dimensional ($d \leqslant 3$) periodic lattice $\mathcal{L}_n = \left(\frac{1}{n}\mathbb{Z}/\mathbb{Z}\right)^d = \{\frac{i}{n}, i = 0, 1, \ldots, n-1\}^d$ with spacing $\lambda = \frac{1}{n}$ consisting of $N = n^d$ sites. At each site $x \in \mathcal{L}_n$ we have an order parameter σ which represents the spin, with values on the unit sphere $\mathbb{S}^2$, i.e. $\sigma(x) \in \mathbb{S}^2$. The configuration space is

$$\Sigma_n = \{\sigma : \sigma(x) \in \mathbb{S}^2, x \in \mathcal{L}_n\} = (\mathbb{S}^2)^{\mathcal{L}_n},$$

endowed with the product topology. The physics on the regular lattice $\mathcal{L}_n \subset \mathbb{R}^d$ is defined by means of an interaction potential between spins $\sigma(x), \sigma(x')$ at two different sites $x, x' \in \mathcal{L}_n$ with values on the unit sphere $\mathbb{S}^2$. The lattice $\mathcal{L}_n$ approximates (as $n \to \infty$) a bounded Lipschitz domain $\Omega \subset \mathbb{R}^d$, where, for simplicity, we assume to be the unit d-dimensional torus. Points on the sphere $\mathbb{S}^2$ are represented by unit vectors $\sigma \in \mathbb{R}^3$ with the components $\sigma^i, i = 1, 2, 3$. With a slight abuse of notation we write $\sigma(x)\sigma(x') \equiv \sum_i \sigma^i(x)\sigma^i(x')$ for the scalar product of spins at sites $x, x' \in \mathcal{L}_n$.

The energy of the system (for a particular configuration σ) is given by the 2-body interaction Hamiltonian:

$$H_n(\sigma) = \sum_{x,x' \in \mathcal{L}_n} U\left(x, x', \sigma(x), \sigma(x')\right) . \tag{1}$$

In the sequel the Hamiltonian remains finite as $n \to \infty$. We do not pursue an ab initio derivation and we assume that the interaction potential $U(x - x', \sigma(x), \sigma(x'))$ consists of the following terms: $U = U_e + U_d + U_a + U_h$ with particular contributions reflecting different types of interactions between magnetic moments.

2.1.1. Exchange Energy

The exchange energy is given by

$$U_e\left(\frac{k}{n}, \frac{l}{n}, \sigma\left(\frac{k}{n}\right), \sigma\left(\frac{l}{n}\right)\right) = -\frac{1}{2} \sum_{i,j=1}^{3} \mathbb{J}_n^{ij}(k-l)\sigma^i\left(\frac{k}{n}\right)\sigma^j\left(\frac{l}{n}\right), \tag{2}$$

where $k, l \in \{0, 1, \ldots, n-1\}^d$ and with the local mean-field interaction described by $\mathbb{J}_n(k-l) = \frac{1}{n^d}\mathbb{J}\left(\frac{k-l}{n}\right)$, where $\mathbb{J}$ is a positive matrix function whose entries are real smooth functions which depend on the distance, i.e., $\mathbb{J}(z) = \mathbb{J}(|z|)$. The exchange energy term is often approximated by a nearest-neighbor interaction only, but in many materials the pair-wise interaction potentials are of long-range nature, see for instance[(2)]. However, in some cases it is also physically justified (see Ref. 7) to consider the Kac interaction potentials: the scaling of the exchange energy involving the size n amounts to assuming a long-range interaction on the entire macroscopic specimen, when we focus on the $n \to \infty$ behavior. Thus if we want to express (2) on the periodic lattice $\mathcal{L}_n$(instead of $\mathbb{Z}^d$), by letting $x = \frac{k}{n}$ and $x' = \frac{l}{n}$ so that $x, x' \in \mathcal{L}_n$, we get:

$$\begin{aligned} U_e(x, x', \sigma(x), \sigma(x')) &= -\frac{1}{2N} \sum_{i,j=1}^{3} \mathbb{J}^{ij}(x-x')\,\sigma^i(x)\sigma^j(x') \\ &\equiv -\frac{1}{2N}\mathbb{J}(x-x')\sigma(x)\sigma(x') , \end{aligned} \tag{3}$$

where we recall that $N = n^d$.

2.1.2. Dipolar Energy

The dipolar energy (or more generally a long-range singular interaction) describes interactions between different classical magnetic moments and is given by

$$U_d(x - x', \sigma(x), \sigma(x')) = \frac{1}{2N} \sum_{i,j=1}^{3} \mathbb{K}_{ij}(x - x')\sigma^i(x)\sigma^j(x') \,. \tag{4}$$

The kernel $\mathbb{K}$ defines a singular integral operator (SIO) on the space $L^2(\mathbb{R}^3; \mathbb{R}^3)$, which in the context of the large deviation techniques needs to be treated in a special topology as we see in Section 2.3.

2.1.3. Anisotropy Energy

The anisotropy energy related to the crystalline structure of the material is defined by the energy density $\Psi : \mathbb{S}^2 \to \mathbb{R}$, so for $x \in \mathcal{L}_n$ we have:

$$U_a(x, \sigma(x)) = \kappa \Psi(\sigma(x)) \,. \tag{5}$$

In the sequel we often absorb the coefficient κ into the definition of Ψ. Note also that the underlying crystallographic lattice structure of the system will be incorporated through this interaction energy density Ψ rather than by using different lattice geometries.

2.1.4. External Field Energy

In the derivation of the statistical model we omit the external field to keep the notation simple and focus on the averaging procedure only. However, in the presence of an external magnetic field h_e the Hamiltonian also involves the interaction energy

$$U_h(x, \sigma(x)) = -\frac{1}{N} \sum_{i=1}^{3} h_e^i(x)\sigma^i(x) \,, \tag{6}$$

where $x \in \mathcal{L}_n$.

In addition to the micromagnetics problem described above, we also consider a model problem where the analysis is substantially easier, while some of the essential elements of the original system are preserved. In the model problem we retain only the term (2), i.e. the Kac potential, and omit the dipolar interaction term (4). For the sake of simplicity in the notation we also omit the external field, although this term can easily be added back into our calculations. Thus, the simplified Hamiltonian that we will be using is:

$$H_n(\sigma) = -\frac{1}{2N} \sum_{x,y \in \mathcal{L}_n} \mathbb{J}(x-y)\sigma(x)\sigma(y) + \sum_{x \in \mathcal{L}_n} \Psi(\sigma(x)) \,. \tag{7}$$

In spite of the simplifications mentioned above, the model Hamiltonian retains important features from the physically realistic Hamiltonian (1), such as exchange energy, anisotropy and a continuous spin variable. Moreover, for the sake of simplicity we have used the new notation for the vector product, as was introduced in formula (3).

We conclude our discussion on equilibrium models by introducing temperature as a parameter in the system. The *canonical* Gibbs measure, defined at the inverse temperature $\beta = \frac{1}{kT}$ (k is the Boltzmann constant), on the configuration space Σ_n is:

$$P_{n,\beta}(d\sigma) = \frac{1}{Z_{n,\beta}} e^{-\beta H_n(\sigma)} \prod_{x \in \mathcal{L}_n} d\sigma(x) \,, \tag{8}$$

where $Z_{n,\beta} = \int_{\Sigma_n} e^{-\beta H_n(\sigma)} \prod_{x \in \mathcal{L}_n} d\sigma(x)$ denotes the partition function. The prior distribution $\prod_{x \in \mathcal{L}_n} d\sigma(x)$ models the small scale fluctuations of the spins. We assume that spins at different lattice sites are independent, uniformly distributed random variables on $\mathbb{S}^2$. Note that we can alternatively incorporate the anisotropy energy in the prior distribution. In this paper we do not address the microcanonical ensemble and issues of ensemble equivalence,[10–12], but instead focus on the canonical ensemble. Understanding the limiting behavior for the canonical Gibbs distribution is also useful for analysis of Monte Carlo or Langevin-type dynamics that are used for sampling from this Gibbs measure. We discuss the evolution to equlibrium in more detail in Section 3.

2.2. Mesoscopic Limits at Equilibrium

We employ the theory of large deviations (see Ref. 13) to obtain an energy functional E_β in the continuum limit from the model spin Hamiltonian (7). The aim is to apply abstract large deviation type results to find the most probable configuration of the Gibbs measure given by (8). The large deviation techniques have been successfully applied to equilibrium derivations of continuum limits, for instance in[14] for magnetic systems with Kac potentials[15–17] and in[4] for the case of 2D turbulence. In this section, we follow the abstract framework presented in[4]. The derivation of the Large Deviation Principle (LDP) for the model Hamiltonian (7) follows without significant modifications from the recent work in[4] and[5]. We omit unnecessary generality and focus on the case of spatial dimension $d=3$ only. The cases $d=1,2$ can be analyzed in a similar fashion once a physically relevant Hamiltonian is formulated on the d-dimensional lattice.

We consider the space $\mathcal{P}_1(\Omega\times\mathbb{S}^2)$ of probability measures on $\Omega\times\mathbb{S}^2$ with first marginal the Lebesgue measure on Ω. The space $\mathcal{P}_1(\Omega\times\mathbb{S}^2)$ is a closed subset of $\mathcal{P}(\Omega\times\mathbb{S}^2)$ which, when endowed with an appropriate metric, is a complete, separable, metric space. Convergence in such metric is equivalent to the weak convergence of measures. According to [[13] Theorem A.5.4] we can decompose any measure $\mu\in\mathcal{P}_1(\Omega\times\mathbb{S}^2)$ as $\mu(dx\,dv)=dx\otimes\tau_\mu(x,dv)$. In particular, we define the empirical measure $\mu^N:\Sigma_N\to\mathcal{P}_1(\Omega\times\mathbb{S}^2)$ by:

$$\mu^N(\sigma;dx\,dv)=dx\otimes\tau_{\mu^N}(\sigma;x,dv)\,,\;\;\tau_{\mu^N}(\sigma;x,dv)=\sum_{s\in\mathcal{L}_N}1_{M(s)}(x)\delta_{\sigma(s)}(dv)\,. \tag{9}$$

Later in Section 3.1.1 we will use the "fully discretized" version given by:

$$\mu^N(\sigma;dx\,dv)=\frac{1}{N}\sum_{s\in\mathcal{L}_N}\delta_{\sigma(s)}(dv)\delta_s(dx)\,. \tag{10}$$

By dx we denote the Lebesgue measure on Ω and by dv the invariant uniform measure on $\mathbb{S}^2$. Moreover, for each $x\in\mathcal{L}_N$ we denote $M(x)\subset[0,1)^3$ a cubic micro-cell with the side of length $\frac{1}{N^{1/d}}$ containing the site x as its lower, left vertex. The empirical measure (10) is the microscopic analogue of the PDF of having a spin value $\sigma(s)$ at the position s. Furthermore, we introduce a coarse-grained version of (9) by:

$$W_{r,q}(dx\,dv)=dx\otimes\sum_{k=1}^{2^r}1_{D_{r,k}}(x)\frac{1}{q}\sum_{s\in D_{r,k}}\delta_{\sigma(s)}(dv)\,, \tag{11}$$

where we have defined a coarse lattice for an integer r (such that $q2^r=N$ for some $q\in\mathbb{N}$) as a new covering of Ω by 2^r macro-cells consisting of $q=\frac{N}{2^r}$ micro-cells each. For subsequent use we denote the coarse lattice by $\mathcal{L}_{r,q}$ and the corresponding space of configurations of the spins by $\Sigma_{r,q}$.

Again, (11) is an analogue of a PDF but at the larger (than the microscopic) scale r.

We derive the LDP for the sequence $\{\mu^N(dx\,dv)\}_N$ with respect to the Gibbs measure $P_{N,\beta}$ on $\mathcal{P}_1(\Omega\times\mathbb{S}^2)$. To derive the LDP we need to approximate the Hamiltonian as

$$H_N(\sigma)=N\tilde{H}(\mu^N)+o_N(1)\,,\ \text{uniformly in } \sigma\,, \tag{12}$$

where $\tilde{H}:\mathcal{P}_1(\Omega\times\mathbb{S}^2)\to\mathbb{R}$ is a bounded continuous function. This follows easily when $\mathbb{J}$ is a Kac potential. After a simple calculation, keeping the higher order terms of the expansion of $H_N(\sigma)$, we see that $\tilde{H}$ is given by:

$$\begin{aligned}\tilde{H}(\mu^N)=&-\frac{1}{2}\int_{\Omega\times\mathbb{S}^2}\int_{\Omega\times\mathbb{S}^2}\mathbb{J}(x-x')\,v\,v'\,\mu^N(dx\,dv)\,\mu^N(dx'\,dv')\\&+\int_{\Omega\times\mathbb{S}^2}\Psi(v)\,\mu^N(dx\,dv)\,.\end{aligned}$$

We also denote the prior distribution $\prod_{x\in\mathcal{L}_N}d\sigma(x)$ by $d\Pi_N$.

The relative entropy $\mathcal{R}(\mu|\,\nu)$ of μ with respect to ν, where μ and ν are measures in $\mathcal{P}(\Omega\times\mathbb{S}^2)$ is defined by:

$$\mathcal{R}(\mu\,|\,\nu)=\begin{cases}\int_{\Omega\times\mathbb{S}^2}\left(\log\frac{d\mu}{d\nu}\right)d\mu, & \text{if } \mu<<\nu\,,\\ \infty, & \text{otherwise}\,.\end{cases}$$

By Sanov's Theorem (see Ref. 13) and using the auxiliary coarse-grained measure (11) (see Refs. 18,5) we derive the LDP for μ^N with respect to Π_N on $\mathcal{P}_1(\Omega\times\mathbb{S}^2)$ (i.e. in the absence of interactions) with rate function $\mathcal{R}(\cdot\,|\lambda\times\rho)$ (where λ is the Lebesgue measure on Ω and ρ is an invariant measure on $\mathbb{S}^2$). This can be stated rigorously: for every Borel subset $F\subset\mathcal{P}_1(\Omega\times\mathbb{S}^2)$,

$$-\mathcal{R}(F^\circ)\leqslant\liminf_{N\to\infty}\frac{1}{N}\log\Pi_N\{\mu^N\in F^\circ\}\quad\text{and}\quad\limsup_{N\to\infty}\frac{1}{N}\log\Pi_N\{\mu^N\in\bar{F}\}\leqslant-\mathcal{R}(\bar{F}),$$

where $\mathcal{R}(F)=\inf_{\mu\in F}\mathcal{R}(\mu|\lambda\times\rho)$ and F°, $\bar{F}$ denote the interior and closure of F respectively. Note that if μ is absolutely continuous with respect to $\lambda(dx)\times\rho(dv)\equiv dx\,dv$, then the relative entropy is given by: $\mathcal{R}(\mu\,|\,dx\,dv)=\int_{\Omega\times\mathbb{S}^2}f(x,v)\log f(x,v)\,dxdv$, for $d\mu=f\,dxdv$. Now we are ready to present the main result which follows from the theorems, proved in[(4)]:

Theorem 2.1. The sequence $\{\mu^N(dx\,dv)\}_N$ with respect to the Gibbs measures $\{P_{N,\beta}\}_N$ satisfies an LDP on $\mathcal{P}_1(\Omega\times\mathbb{S}^2)$ with the rate functional

$$I_\beta[\mu]=E_\beta[\mu]-\inf_{\nu\in\mathcal{P}_1(\Omega\times\mathbb{S}^2)} E_\beta[\nu], \tag{13}$$

where $E_\beta[\mu]=\beta\tilde{H}(\mu)+\mathcal{R}(\mu\,|\lambda\times\rho)$.

In particular,

(i) the asymptotic behavior of the partition function $Z_{N,\beta}$ is given by

$$\lim_{N\to\infty}\frac{1}{N}\log Z_{N,\beta}=-\inf_{\mu\in\mathcal{P}_1(\Omega\times\mathbb{S}^2)} E_\beta[\mu].$$

(ii) For every continuous, bounded function $\Phi:\mathcal{P}_1(\Omega\times\mathbb{S}^2)\to\mathbb{R}$

$$\lim_{N\to\infty}\frac{1}{N}\log\int_{\Sigma_N} e^{-\Phi(\mu^N)}\,P_{N,\beta}(d\sigma)=-\inf_{\mu\in\mathcal{P}_1(\Omega\times\mathbb{S}^2)}\left\{\Phi(\mu)+I_\beta[\mu]\right\}.$$

(iii) Let $\mathcal{E}_\beta:=\{\mu\in\mathcal{P}_1(\Omega\times\mathbb{S}^2): I_\beta(\mu)=0\}$. Then $\mathcal{E}_\beta$ is a non-empty compact subset of $\mathcal{P}_1(\Omega\times\mathbb{S}^2)$, and if B is a Borel subset of $\mathcal{P}_1(\Omega\times\mathbb{S}^2)$ whose closure $\bar{B}$ has empty intersection with $\mathcal{E}_\beta$ then $I_\beta(\bar{B}):=\inf_{\mu\in\bar{B}} I_\beta(\mu)>0$ and thus for some $C<\infty$:

$$P_{N,\beta}(\mu^N\in B)\leqslant C\,e^{-I_\beta(\bar{B})}\to 0,\quad \text{as } N\to\infty.$$

Proof. We present the main calculation for the sake of clarity. For (i) we have:

$$\begin{aligned}\lim_{N\to\infty}\frac{1}{N}\log Z_{N,\beta} &= \lim_{N\to\infty}\frac{1}{N}\log\int_{\Sigma_N} e^{-\beta N\tilde{H}(\mu^N)}\prod_{x\in\mathcal{L}_N} d\sigma(x)\\ &=\lim_{N\to\infty}\frac{1}{N}\log\int_{\mathcal{P}_1(\Omega\times\mathbb{S}^2)} e^{-\beta N\tilde{H}(\mu)}\Pi_N\left(\{\sigma:\mu^N\in d\mu\}\right)\\ &=\sup_{\mu\in\mathcal{P}_1(\Omega\times\mathbb{S}^2)}\{-\beta\tilde{H}(\mu)-\mathcal{R}(\mu\,|\lambda\times\rho)\},\end{aligned}\tag{14}$$

where the first equality is due to (12) and the second is a change of the variable of integration. The last equality is due to the fact that $\mathbb{J}$ is smooth and thus $\tilde{H}$ is a bounded continuous function (see Theorem 1.2.1 in Ref. 13) and the LDP for μ^N with respect to the prior distribution Π_N.

The proof of (ii) is similar since Φ is a bounded continuous function. For (iii) first note that by Prokhorov's theorem $\mathcal{P}(\Omega \times \mathbb{S}^2)$ is compact since both Ω and $\mathbb{S}^2$ are compact and thus $\mathcal{P}_1(\Omega \times \mathbb{S}^2)$ is also a compact space as a closed subset of a compact space. Hence, since I_β is a rate function as a lower semicontinuous mapping of the compact space $\mathcal{P}_1(\Omega \times \mathbb{S}^2)$ into $[0, 1)$, it assumes its minimum of 0 on $\mathcal{P}_1(\Omega \times \mathbb{S}^2)$ and thus $\mathcal{E}_\beta$ is non-empty. Moreover, if $\bar{B}$ has empty intersection with $\mathcal{E}_\beta$, then $I_\beta(\bar{B}) > 0$ since I_β is a rate function and $I_\beta(\mu) > 0$ for every $\mu \in \bar{B}$. ∎

Thus, the LDP for the Gibbs measure, as $N \to \infty$, can be formally written as:

$$P_{N,\beta}(\mu^N = \mu) \approx e^{-N[\beta \tilde{H}(\mu) + \mathcal{R}(\mu|\lambda \times \rho) - \inf_\nu \{\beta \tilde{H}(\nu) + \mathcal{R}(\nu|\lambda \times \rho)\}]}. \tag{15}$$

Remark 2.1. The expression (15) is interpreted as follows: the most probable configuration μ of the Gibbs measure is the *minimizer* of $\beta \tilde{H}(\mu) + \mathcal{R}(\mu\,|\lambda \times \rho)$, yielding the large scale structure (at the scale of the domain $\Omega \times \mathbb{S}^2$) at equilibrium. This leads to the identification of $\mathcal{E}_\beta$ with the set of equilibrium macrostates. As we show below, this minimizer turns out to be the probability density (28) derived for the case where we have included the micromagnetic interactions. Note that $\beta \tilde{H}(\mu) + \mathcal{R}(\mu\,|\lambda \times \rho)$ is finite if and only if the measure μ has a density f, in which case we define the energy functional by

$$\begin{aligned} E_\beta[f] = &-\frac{1}{2} \int_{\Omega \times \mathbb{S}^2} \int_{\Omega \times \mathbb{S}^2} \mathbb{J}(x - x')\, v\, v'\, f(x, v) f(x', v')\, dx\, dv\, dx'\, dv' \\ &+ \frac{1}{\beta} \int_{\Omega \times \mathbb{S}^2} f(x, v) \log f(x, v)\, dx\, dv + \int_{\Omega \times \mathbb{S}^2} \Psi(v)\, f(x, v)\, dx\, dv\,. \end{aligned} \tag{16}$$

Finally, the formal expression (15) captures the microscopic *spatial* random fluctuations of the empirical measure μ^N as a function of N (i.e., the total number of spins in the specimen), around the most probable macrostate given by the minimizer of (16).

2.3. Dipolar Interactions in Micromagnetic Hamiltonians

To study the properties of the micromagnetic model we have to include the additional term U_d in the interaction potential $U(x - x', \sigma(x), \sigma(x'))$, $x, x' \in \mathcal{L}_N$, given by (4) which describes a non-local, long-range interaction between different classical magnetic moments. Then the corresponding Hamiltonian, neglecting all other contributions (which are easy to handle as in the previous section), is given by:

$$H_N(\sigma) = \frac{1}{2N} \sum_{x,y \in \mathcal{L}_N} \mathbb{K}(x-y)\sigma(x)\sigma(y) .$$

The kernel $\mathbb{K}(z)$ is smooth except at the origin and is given by:

$$\mathbb{K}_{ij}(z) = -\frac{1}{4\pi |z|^3}\left(\delta_{ij} - 3\frac{z_i z_j}{|z|^2}\right), \quad z \in \mathbb{R}^3, \quad \text{or equivalently,}$$
$$\mathbb{K}_{ij}(z) = \nabla_{z_i}\nabla_{z_j}\left(\frac{1}{|z|}\right) . \tag{17}$$

Moreover, $\mathbb{K}$ defines a singular integral operator $T{:}L^2(\mathbb{R}^3, \mathbb{R}^3) \to L^2(\mathbb{R}^3, \mathbb{R}^3)$ given by:

$$[Tu(x)]_i = \int_{\mathbb{R}^3} \sum_{j=1}^{3} \mathbb{K}_{ij}(x-y)u_j(y)dy, \quad i = 1, 2, 3 .$$

While the derivation of the LDP in the previous section has used standard results, a rigorous treatment of the Hamiltonian based on the SIO is more involved. In this paper we present only partial results derived for a specific *regularization* of the lattice Hamiltonian. We consider a new regularized interaction kernel at the level r of the coarse-graining:

$$\bar{\mathbb{K}}_{r,q}(x, y) = \sum_{k,k'=1}^{2^r} 1_{D_{r,k}}(x) 1_{D_{r,k'}}(y) \frac{1}{|D_{r,k}|^2} \int_{D_{r,k}} \int_{D_{r,k'}} \mathbb{K}(|x'-y'|)\, dy'\, dx' . \tag{18}$$

We denote the corresponding operator by $\bar{T}^{r,q}$. Using the new kernel we redefine the original Hamiltonian by:

$$\bar{H}_{r,q}(\sigma) := \langle \bar{T}^{r,q} m^N, m^N \rangle = \sum_{k,k'=1}^{2^r} \frac{1}{q^2} \sum_{s \in D_{r,k}} \sigma(s) \sum_{s' \in D_{r,k'}} \sigma(s') \int_{D_{r,k}} \int_{D_{r,k'}} \mathbb{K}(|x'-y'|)\, dy'\, dx' , \tag{19}$$

where $\sigma \in \Sigma_{r,q}$ and $m^N(x) = \sum_{s \in \mathcal{L}_N} \sigma(s) 1_{M(s)}(x)$. Before presenting the corresponding LDP Theorem, we observe that as in the proof of Theorem 2.1, in order to make the third equality in (14) valid, we need that the limiting operator is a bounded continuous functional on $\mathcal{P}_1(\Omega \times \mathbb{S}^2)$.

This is not true for the weak topology used in[4] since T is an SIO. This problem can be resolved by considering a stronger topology introduced in[5], where the LDP for the case where there are no interactions was first proved. More precisely, we consider the set

$$\mathcal{P}_{1,p}(\Omega\times\mathbb{S}^2)=\left\{\nu\in\mathcal{P}_1(\Omega\times\mathbb{S}^2): x\mapsto\int_{\mathbb{S}^2} v\,\tau_\nu(x,\,dv)\in L^p(\Omega,dx)\right\}.$$

Then let $\mathcal{T}_{1,p}$ be the strongest topology on $\mathcal{P}_{1,p}(\Omega\times\mathbb{S}^2)$ that makes continuous all maps $\nu\mapsto\left(x\mapsto\int_{\mathbb{S}^2} v\,\tau_\nu(x,\,dv)\right)\in L^p(\Omega,dx)$. The open sets of the topology under consideration are:

$$O_{\rho,\epsilon}=\left\{\nu\in\mathcal{P}_{1,p}(\Omega\times\mathbb{S}^2): \int_\Omega\left|\int_{\mathbb{S}^2} v\,\tau_\nu(x,\,dv)-\int_{\mathbb{S}^2} v\,\tau_\rho(x,\,dv)\right|^p dx<\epsilon\right\},$$

for every $\rho\in\mathcal{P}_{1,p}(\Omega\times\mathbb{S}^2)$ and $\epsilon>0$. Then for $p=2$, we have that: if $W_{r,q}\overset{\mathcal{T}_{1,2}}{\to}\omega$ then $\bar{W}_{r,q}\overset{L^2}{\to}\bar{\omega}$ and consequently $T\bar{W}_{r,q}\overset{L^2}{\to}T\bar{\omega}$, since T is continuous in the L^2-norm (see Ref. 19). Recall that $\bar{W}_{r,q}(x)=\sum_{k=1}^{2^r}1_{D_{r,k}}(x)\ \frac{1}{q}\sum_{s\in D_{r,k}}\sigma(s)$ and $\bar{\omega}=\int_{\mathbb{S}^2}v\tau_\omega(x,dv)$. The Gibbs measure $\bar{P}_{r,q,\beta}$ that corresponds to the new Hamiltonian (19) is given by:

$$\bar{P}_{r,q,\beta}(d\sigma)=\frac{1}{\bar{Z}_{r,q,\beta}}e^{-\beta\bar{H}_{r,q}(\sigma)}\prod_{x\in\mathcal{L}_{r,q}}d\sigma(x),\tag{20}$$

where $\bar{Z}_{r,q,\beta}=\int_{\Sigma_{r,q}}e^{-\beta\bar{H}_{r,q}(\sigma)}\prod_{x\in\mathcal{L}_{r,q}}d\sigma(x)$ is the new partition function. Using the stronger topology introduced above we can derive the LDP stated in the following theorem:

Theorem 2.2. The sequence $\{W_{r,q}(dx\,dv)\}_{r,q}$ with respect to the scaled Gibbs measures $\{\bar{P}_{r,q,q2^r\beta}\}_{r,q}$ satisfies a LDP in the double limit ($q\to\infty$ and $r\to\infty$) on $\mathcal{P}_{1,2}(\Omega\times\mathbb{S}^2)$ with the rate functional

$$\bar{I}_\beta[\omega]=\bar{E}_\beta[\omega]-\inf_{\nu\in\mathcal{P}_{1,2}(\Omega\times\mathbb{S}^2)}\bar{E}_\beta[\nu],\tag{21}$$

where $\bar{E}_\beta[\omega]=\beta\langle T\omega,\omega\rangle+\mathcal{R}(\omega\,|\,\lambda\times\rho)$, and we define $T\omega:=T\bar{\omega}$ for the corresponding average $\bar{\omega}$ of $\omega\in\mathcal{P}_{1,2}(\Omega\times\mathbb{S}^2)$. In particular,

(i) the asymptotic behavior of the partition function $\bar{Z}_{r,q,q2^r\beta}$ is given by

$$\lim_{r\to\infty}\lim_{q\to\infty}\frac{1}{q2^r}\log\bar{Z}_{r,q,q2^r\beta}=-\inf_{\omega\in\mathcal{P}_{1,2}(\Omega\times\mathbb{S}^2)}\bar{E}_\beta[\omega].$$

(ii) For every continuous, bounded (with respect to the new topology $\mathcal{T}_{1,2}$) functional $\Phi:\mathcal{P}_{1,2}(\Omega\times\mathbb{S}^2)\to\mathbb{R}$

$$\lim_{r\to\infty}\lim_{q\to\infty}\frac{1}{q2^r}\log\int_{\Sigma_{r,q}}e^{-q2^r\beta\Phi(W_{r,q})}\,d\bar{P}_{r,q,q2^r\beta}$$
$$=-\inf_{\omega\in\mathcal{P}_{1,2}(\Omega\times\mathbb{S}^2)}\left\{\Phi(\omega)+\bar{I}_\beta[\omega]\right\}.$$

Proof. We first observe that

$$\bar{H}_{r,q}(\sigma):=\langle\bar{T}^{r,q}m^N,m^N\rangle=\langle T\bar{W}_{r,q},\bar{W}_{r,q}\rangle. \tag{22}$$

From the above calculation and the definition of $\bar{P}_{r,q}$ we observe that for (18) we have:

$$\begin{aligned}\lim_{r\to\infty}\lim_{q\to\infty}\frac{1}{q2^r}\log\bar{Z}_{r,q,q2^r\beta}&\stackrel{(22)}{=}\lim_{r\to\infty}\lim_{q\to\infty}\frac{1}{q2^r}\log\int_{\Sigma_{r,q}}e^{-q2^r\beta\langle TW_{r,q},W_{r,q}\rangle}\prod_{x\in\mathcal{L}_{r,q}}d\sigma(x)\\&=\lim_{r\to\infty}\lim_{q\to\infty}\frac{1}{q2^r}\log\int_{\mathcal{P}_{1,2}(\Omega\times\mathbb{S}^2)}e^{-q2^r\beta\langle T\omega,\omega\rangle}\Pi_{r,q}\left(\{\sigma:W_{r,q}\in d\omega\}\right)\\&=\sup_{\omega\in\mathcal{P}_{1,2}(\Omega\times\mathbb{S}^2)}\{-\beta\langle T\omega,\omega\rangle-\mathcal{R}(\omega\,|\,\lambda\times\rho)\}.\end{aligned}\tag{23}$$

The last equality is true if we assume the stronger topology introduced above. Part (ii) can be proved similarly. Such double limit LDP's have been first introduced in[18] for the weak topology of $\mathcal{P}_1(\Omega\times\mathbb{S}^2)$. ∎

In order to treat the SIO without regularization one needs to establish, in addition to the new topology introduced above, convergence properties of the discretized SIO in the presence of fluctuations. Rigorous treatment in this case is related to the work in[20] and is presented in[21]. The main difficulty is due to the fact that for the microscopic sequence m^N we cannot use the stronger topology introduced above. This means that in order to study the non-regularized SIO appearing in the micromagnetic applications we have to account for all energy correction terms that come as a result of the different scales of oscillations on the lattice.

2.4. Free Energy and Variational Principle in Micromagnetics

In Section 2.2 we showed that the functional (16) is the rate functional for the measure-valued process $\mu^N(dx\,dv)$. In this section we derive the corresponding free energy F_β of the system for the average magnetization which is the rate functional of the measure valued process $\sum_{k=1}^{2^r} 1_{D_{r,k}}(x)\frac{1}{q}\sum_{s\in D_{r,k}}\sigma(s)$. We also see that the infimum of the free energy functional constrained on the magnetization m coincides with the infimum of the energy functional E_β. Moreover, we prove that this infimum is attained at the density $\mathrm{M}(x,v)$ which represents the equilibrium macrostate of the microscopic system. Note also that in this section we include both the smooth potential and the singular integral interaction term as obtained in Theorem 2.2.

Theorem 2.3. Let

$$F_\beta[m] = -\frac{1}{2}\int_\Omega \mathbb{J} * m\cdot m\,dx + \frac{1}{2}\int_\Omega |\nabla u|^2\,dx + \int_\Omega a_\beta^*(m)\,dx - \frac{1}{\beta}\log Z_\Psi\,, \tag{24}$$

where the function u solves (in the weak sense)

$$\Delta u = \mathrm{div}(1_\Omega m)\,, \quad \text{in } \mathbb{R}^3\,. \tag{25}$$

Moreover,

$$a_\beta^*(m) = \sup_{p\in\mathbb{R}^d}\{m\cdot p - a_\beta(p)\} \tag{26}$$

is the Legendre–Fenchel transform of the function

$$a_\beta(p) = \frac{1}{\beta}\log\int_{\mathbb{S}^2} e^{\beta v\cdot p}\tilde{\rho}(dv)\,, \tag{27}$$

where:

$$\tilde{\rho}(dv) = \frac{1}{Z_\Psi}e^{-\beta\Psi(v)}dv, \quad Z_\Psi = \int_{\mathbb{S}^2} e^{-\beta\Psi(v)}\,dv\,.$$

Then

(i)

$$F_\beta[m] = \inf_{\mu\in\mathcal{P}_{1,2}(\Omega\times\mathbb{S}^2)} \left\{ E_\beta[\mu] \,|\, m : m(x) = \int_{\mathbb{S}^2} v\,\tau_\mu(x, dv) \right\}$$

and

$$\inf\left\{ E_\beta[f] \,|\, f(x,v)\,dxdv \in \mathcal{P}_{1,2}(\Omega\times\mathbb{S}^2)\right\} = \inf\left\{ F_\beta[m] \,|\, m\in L^2(\Omega,\mathbb{R}^3)\right\}.$$

(ii) The measure-valued process $\sum_{k=1}^{2^r} 1_{D_{r,k}}(x)\frac{1}{q}\sum_{s\in D_{r,k}}\sigma(s)$ satisfies a double limit LDP (see Ref. 5, Corollary 2.5) with the rate functional F_β in $L^2(\Omega;\mathbb{R}^3)$.

(iii) The infimum is attained when $f(x,v)$ equals to:

$$\mathrm{M}(x,v) = \frac{1}{Z_\beta(x)} e^{-\beta[\Psi(v)+v(\nabla u(x)-\mathbb{J}*m(x))]}, \quad \text{and} \quad m(x) = \int_{\mathbb{S}^2} v\, f(x,v) dv\,, \tag{28}$$

where u, m are related by (25).

Proof. (i) We first observe that

$$\begin{aligned} &\inf\left\{ E_\beta[f] \,|\, f\,dxdv \in \mathcal{P}_{1,2}(\Omega\times\mathbb{S}^2)\right\} \\ &\quad = \inf_{m\in L^2(\Omega,\mathbb{R}^d)} \inf\left\{ E_\beta[f] \,|\, f\,dxdv\in\mathcal{P}_{1,2}(\Omega\times\mathbb{S}^2),\, \int_{\mathbb{S}^2} vf(x,v)\,dv = m(x)\right\}. \end{aligned}$$

Next we have to show that the inner infimum on the right hand side is equal to $F_\beta[m]$. The solution operator for Poisson's Eq. (25) and the gradient of u yield the SIO $\nabla\Delta^{-1}\mathrm{div}$ which has the kernel $\mathbb{K}(x-y)$ and we can rewrite the energy E_β in the following form

$$\begin{aligned} E_\beta[f] &= \frac{1}{2}\int_\Omega (\mathbb{K}-\mathbb{J})*m\cdot m\,dx + \int_{\Omega\times\mathbb{S}^2} f(x,v)\Psi(v)\,dvdx \\ &\quad + \frac{1}{\beta}\int_{\Omega\times\mathbb{S}^2} f(x,v)\log f(x,v)\,dv\,dx \\ &= \frac{1}{2}\int_\Omega (\mathbb{K}-\mathbb{J})*m\cdot m\,dx + \frac{1}{\beta}\int_{\Omega\times\mathbb{S}^2} f(x,v)\log\frac{f(x,v)}{N(v)}\,dvdx - \frac{1}{\beta}\log Z_\Psi\,, \end{aligned}$$

where we denote

$$N(v) = \frac{1}{Z_\Psi} e^{-\beta\Psi(v)}\,, \quad Z_\Psi = \int_{\mathbb{S}^2} e^{-\beta\Psi(v)}\,dv\,. \tag{29}$$

We consider the parametrized probability measure $f(x,v)\,dv$ and write the relative entropy

$$\mathcal{R}(f\,dv\,|\,N\,dv) \equiv \int_{\mathbb{S}^2} f(x,v)\log\frac{f(x,v)}{N(v)}\,dv\,.$$

Using the relative entropy we can rewrite the energy E_β as

$$E_\beta[f] = \frac{1}{2}\int_\Omega (\mathbb{K}-\mathbb{J})*m\cdot m\,dx + \frac{1}{\beta}\int_\Omega \mathcal{R}(f\,|\,N)\,dx - \frac{1}{\beta}\log Z_\Psi\,.$$

Thus

$$\inf\left\{E_\beta[f]\,|\,f\,dxdv\in\mathcal{P}_{1,2}(\Omega\times\mathbb{S}^2),\ \int_{\mathbb{S}^2} vf(x,v)\,dv = m(x)\right\}$$
$$= \frac{1}{2}\int_\Omega (\mathbb{K}-\mathbb{J})*m\cdot m\,dx + \inf_f \frac{1}{\beta}\int_\Omega \mathcal{R}(f\,|\,N)dx - \frac{1}{\beta}\log Z_\Psi\,.$$

Since the mapping $f\mapsto\mathcal{R}(f\,|\,N)$ is convex and non-negative we have:

$$\inf_f \frac{1}{\beta}\int_\Omega \mathcal{R}(f\,|\,N)\,dx = \frac{1}{\beta}\int_\Omega \inf_f \mathcal{R}(f\,|\,N)\,dx\,.$$

Next, using Lemma 3.3.3 from [13], we obtain

$$\inf\left\{\mathcal{R}(f|N)\,|\,f\,dv\in\mathcal{P}(\mathbb{S}^2),\ \int_{\mathbb{S}^2} vf(x,v)\,dv = m(x)\right\} = a_\beta^*(m(x))\,,$$

where $\mathcal{P}(\mathbb{S}^2)$ denotes the probability measures on $\mathbb{S}^2$. Assertion (i) of the proposition then follows.

(ii) We apply the contraction principle (see Ref. 13, p. 18) by setting $G:\mathcal{P}_{1,2}(\Omega\times\mathbb{S}^2)\to L^2(\Omega)$ to be $G(\mu)=\int_{\mathbb{S}^2} v\,\tau_\mu(x,\,dv)$. Since the functional G is continuous we obtain, from the contraction principle, that the rate function of

$$\sum_{k=1}^{2^r} 1_{D_{r,k}}(x)\frac{1}{q}\sum_{s\in D_{r,k}}\sigma(s) = \int_{\mathbb{S}^2} v\,\tau_{W_{r,q}}(x,\,dv)$$

is given in terms of the rate function $E_\beta[f]$ of $W_{r,q}(dx\,dv)$, i.e., $F_\beta[m] = \inf_{\mu\in\mathcal{P}_{1,2}(\Omega\times\mathbb{S}^2)}\{E_\beta[\mu]\,|\,G(\mu)=m\}$, which together with (i) yields (24).

(iii) To check that the infimum is attained we first observe that $E_\beta[f]$ is a lower-semicontinuous function of f. This is true since due to the stronger topology (for $\mathcal{P}_{1,2}(\Omega\times\mathbb{S}^2)$) introduced in the previous section, the mapping $f\mapsto\int_\Omega(\mathbb{K}-\mathbb{J})*m\cdot m\,dx$, with $m(x)=\int_{\mathbb{S}^2} v\,f(x,v)dv$ is continuous and a^*_β is lower semicontinuous being a convex function of m. Hence there exists a minimum. By calculations similar to the ones in Ref. 22 we derive the Euler–Lagrange equations that correspond to this minimization problem and we find that the infimum is attained when $f(x,v)$ equals to the Maxwellian density (28). ▮

2.5. Connections with the Landau–Lifschitz Theory

A natural question is how the presented derivation relates to the Landau–Lifschitz theory[23–25] widely used in micromagnetic modeling. The equilibrium states in the Landau–Lifschitz theory are described by minimizers of the energy functional

$$F_{\rm LL}[m]=\int_\Omega\frac{A}{2}|\nabla m|^2\,dx+\int_{\mathbb{R}^d}\frac{1}{2}|\nabla u|^2\,dx+\int_\Omega\tilde\Psi(m)\,dx-\int_\Omega h_e m\,dx\,, \quad (30)$$

where A is a material-dependent constant and $\tilde\Psi$ a function on $\mathbb{S}^2$ describing the crystalline anisotropy. The magnetization field m in $F_{\rm LL}$ is subject to the non-convex constraint $|m|=1$. The micromagnetics theory assumes that the macroscopic magnetization field $m(x)$ is slowly varying in space and has constant modulus $|m(x)|=m_s(T)$ throughout the specimen. Consequently at a constant temperature T the field can be normalized to the unit vector and the energy functional (30) is an accepted starting point of the static micromagnetic modeling.

The equilibrium theory discussed in the previous sections presents a rigorous derivation of a *local mean-field* approximation that explicitly connects microscopic interactions with parameters in the continuum model. The difference when compared to the Landau–Lifschitz theory is caused by the presence of the thermal fluctuations and the averaged nature of the magnetization field $m(x)=\int_{\mathbb{S}^2} v\,f(x,v)\,dv$ manifested in the relaxed constraint $|m(x)|\leqslant 1$. However, it is not difficult, at least on a formal level, to show a connection with the Landau–Lifschitz theory. Assuming that $m(x)$ is sufficiently regular and expanding the convolution $\mathbb{J}*m$ in the free energy (24), $\mathbb{J}*m=J_0m+J_2\,\Delta m+\ldots$, where $J_0=\int_\Omega\mathbb{J}(r)\,dr$, $J_2=1/2\int_\Omega|r|^2\mathbb{J}(r)\,dr$, and substituting into (24) we obtain

$$\tilde F_\beta[m]=\int_\Omega\frac{J_2}{2}|\nabla m|^2\,dx+\int_{\mathbb{R}^d}\frac{1}{2}|\nabla u|^2\,dx+\int_\Omega\left(a^*_\beta(m)-\frac{J_0}{2}|m|^2\right)dx-\int_\Omega h_e m\,dx, \quad (31)$$

in the case when we also include the external field term. In this way the interaction defined on the microscopic level is related explicitly to the constants in the macroscopic model. Through the definitions (26) and (27) we also obtain the dependence on the local anisotropy energy $\psi_\beta(m) \equiv a^*_\beta(m) - \frac{J_0}{2}|m|^2$. Although the function $a_\beta(p)$ is difficult to be found in a closed form some properties of the function follow directly from the definition of the Legendre–Fenchel transform. Direct calculation implies that since $|\partial_p a_\beta(p)| \leqslant 1$ we have $a^*_\beta(m) = \infty$ for $|m| > 1$ and consequently the energy $\tilde{F}_\beta$ (or F_β) is minimized subject to the constraint $|m| \leqslant 1$.

The following example demonstrates some features of the model for a specific choice of the lattice anisotropy $\Psi(\sigma)$.

Example (*Uniaxial anisotropy*). We define (θ, ϕ) to be the standard spherical coordinates on $\mathbb{S}^2$. The preferred direction (the easy axis) in the uniaxial anisotropy is identified with the coordinate z-axis. The simplest form of the lattice anisotropy is then given by

$$\Psi(v) \equiv \Psi(\theta, \phi) = -\kappa \cos^2 \theta \,,$$

where κ is a material dependent constant. From (27) we have

$$a_\beta(p) = \frac{1}{\beta} \log \int_0^{2\pi} \int_0^{\pi} e^{\beta(\kappa \cos^2 \theta + p_\parallel \cos\theta + p_\perp \sin\theta \cos\phi)} \sin\theta \frac{d\theta\, d\phi}{2\pi} \,, \qquad (32)$$

where $p_\parallel = |p| \cos\theta_p$ (projection of p on the easy axis) and $p_\perp = |p| \sin\theta_p$, with $(\theta_p, \phi_p, |p|)$ being spherical coordinates of the vector p. In general, the integral over $\mathbb{S}^2$ needs to be evaluated numerically. However, the form of the integral allows us to obtain closed formulae at certain asymptotic regimes. The calculations of asymptotics follow standard procedures. One would expect that in the asymptotic regime $\kappa \to \infty$ the system exhibits behavior of the Ising model as the spins tend to take only two values. Indeed, as $\kappa \to \infty$ we obtain asymptotic expansion

$$a_\beta(p) = \frac{1}{\beta} \log\left(\frac{e^{\beta\kappa}}{\kappa\beta} \cosh \beta p_\parallel \right) + \frac{1}{4\beta^2 \kappa} \left[(2 + \beta^2 p_\perp^2) - 2\beta p_\parallel \tanh \beta p_\parallel \right] + O(\kappa^{-2}) \,,$$

and the leading term gives $a_\beta(p) = \frac{1}{\beta} \cosh \beta p_\parallel$. From (26) the direct calculation yields

$$a^*_\beta(m) = \begin{cases} +\infty, & \text{if } m_\perp \neq 0 \\ \frac{1}{\beta} \left(\frac{1+m_\parallel}{2} \log\left(\frac{1+m_\parallel}{2}\right) + \frac{1-m_\parallel}{2} \log\left(\frac{1-m_\parallel}{2}\right) \right), & \text{if } m_\perp = 0. \end{cases} \qquad (33)$$

Hence the minimizing magnetization m has a non-zero component only in the direction of the easy axis, i.e., $m_\parallel$. Furthermore, the local free energy in this limit is defined as

$$\psi_\beta(m_\parallel) = \frac{1}{\beta}\left(\frac{1+m_\parallel}{2}\log\left(\frac{1+m_\parallel}{2}\right) + \frac{1-m_\parallel}{2}\log\left(\frac{1-m_\parallel}{2}\right)\right) - \frac{1}{2}J_0 m_\parallel^2 .$$

which is identical to the Ising case with Kac interactions (see Ref. 9). In general, the functions $a^*_\beta(m)$ and $\psi_\beta(m)$ will depend on both components $m_\parallel$ and $m_\perp$ and they have to be evaluated numerically. The function $a_\beta(p)$ is computed by numerical quadrature and the Legendre–Fenchel transform is performed on the discrete approximation of $a_\beta(p)$. This is done by using fast Legendre–Fenchel algorithm, see Ref. 26. The fast numerical evaluation of ψ_β allows us to use the implicit definition of the energy in standard minimization algorithms or for numerical solution of non-linear equations. We present only two applications of such numerical approach: (a) investigation of the free energy surface, (b) approximation of a one-dimensional domain wall in an infinite domain.

Figure 1 depicts the behavior of the local free energy $\psi_\beta(m)$ and its dependence on the inverse temperature β. The free energy can be conveniently parametrized by the modulus $|m|$ and the angle θ (the angle between the easy axis and the vector m). The numerical approximation of ψ_β allows us to find the value of the critical temperature. The critical inverse temperature β_c represents the Curie point of the phase transition between paramagnetic and ferromagnetic phase. In other words the free

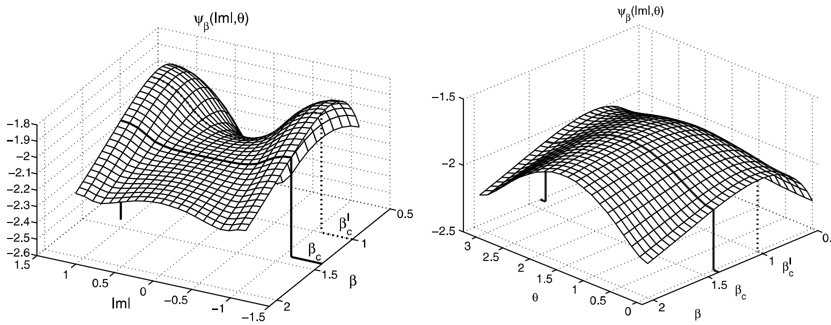

Fig. 1. The free energy surface $\psi_\beta(|m|, \theta)$ and its dependence on the inverse temperature β. Model parameters: $\kappa = 3.0$, $J_0 = 1.0$. The solid line depicts the phase transition between ferromagnetic and paramagnetic behavior, the dotted line is the critical case for the mean-field approximation of the Ising model with the same Kac potential ($\beta_c^I J_0 = 1$).

energy is convex for $\beta < \beta_c$, or equivalently $I - J_0\partial^2_{pp}a_\beta(0)$ is positive definite. The solid line indicates the position of a phase transition point at the temperature β_c where the function changes from non-convex to convex.

Existence of magnetic domains and different types of magnetic walls in the Landau–Lifschitz theory is attributed to the competition of different contributions in the free energy and to the non-convex constraint $|m| = 1$. Due to the thermal agitation incorporated in the finite-temperature free energy F_β (or $\tilde{F}_\beta$) the norm $|m|$ also fluctuates. Although for low temperatures $|m|$ can be close to the unit sphere, in general, we have the convex constraint $|m| \leqslant 1$. The brief discussion above of the free energy ψ_β suggests that the finite-temperature model allows for domain formation whenever the exchange energy is strong enough compared to the temperature. A first test in this direction is the existence of non-zero constant states as minimizers of F_β in the spatial dimension $d = 1$.

In the absence of an external field h_e, the condition that $I - J_0\partial^2_{pp}a_\beta(0)$ is positive definite guarantees that $m \equiv 0$ and $u \equiv 0$ is the minimizer of F_β (as well as $\tilde{F}_\beta$) implying a uniform state, i.e., no domain formation. In this case the probability distribution of spins is x-independent, yielding the equilibrium Gibbs distribution $N(v)$.

When $I - J_0\partial^2_{pp}a_\beta(0)$ is not positive definite, then for a suitable domain Ω and anisotropy, there exist non-trivial constant state solutions $\{m^k_\beta \neq 0,\ k = \pm\}$ and $u \equiv 0$ to the Lagrange–Euler equation for the energy $\tilde{F}_\beta$. For instance, in the special case of a hard ($\kappa \to \infty$) uniaxial material we have the usual Ising-type model where the condition for multiple steady states reduces to $\beta^I J_0 > 1$,[27]. This condition follows from the explicit form of a_β in the asymptotic limit $\kappa \to \infty$. Notice in Fig. 1 that there can be a significant difference between critical temperatures for Ising and Heisenberg models.

In the general case, the vectors m^k_β solve the algebraic equation

$$J_0 m = \partial_p a^*_\beta(m)\,, \tag{34}$$

and have lower energy than the state $m \equiv 0, u \equiv 0$. Since the magnetization m in our model is computed by averaging over the thermal fluctuations of spins the magnetic domains are identified with regions where $|m(x)| \approx |m^k_\beta| \leqslant 1$. The structure of separating domain walls can be explored numerically with relative ease since the constraint $|m| = 1$ is absent in the finite temperature case. Indeed, a one-dimensional standing wave of $\tilde{F}_\beta$ along the easy axis in a one-dimensional uniaxial material of infinite length is the minimizer of the one-dimensional functional (when $h_e = 0$),

$$\bar{\tilde{F}}_\beta[m] = \int_{-\infty}^{\infty} \left[\frac{J_2}{2}|m_x|^2 + a^*_\beta(m) - \frac{J_0}{2}|m|^2 + \frac{1}{2}m_1^2 \right] dx\,, \tag{35}$$

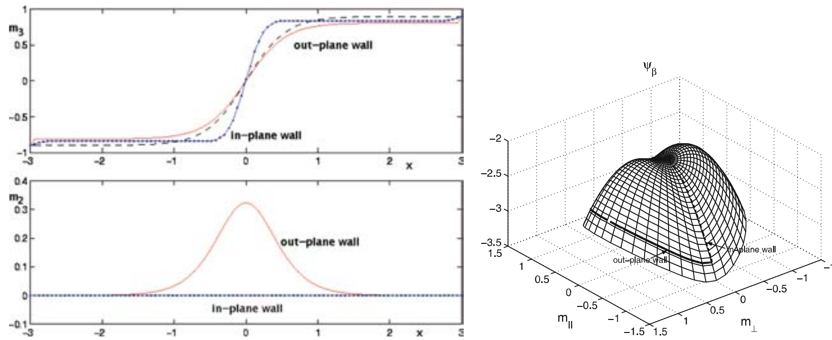

Fig. 2. Magnetic domains and domain walls in one-dimensional infinitely long chain (with $\kappa = 3.0$). The solid line represents an out-plane wall (both components m_2, m_3 change inside the wall), while the dotted line depicts an in-plane wall (only the component m_3, parallel to the easy axis, changes), that represents a metastable state. The dashed line shows the wall for mean-field approximation of the Ising model with the same Kac potential and at the same β, i.e., the case $\kappa \to \infty$. The Ising wall is given explicitly as $m_3(x) = m_\beta \tanh(\beta m_\beta x)$. The two solutions are also plotted on the energy surface of the free energy ψ_β.

where $m_\parallel(\pm\infty) = \pm m_\beta^k$, $m_\perp(\pm\infty) = m_1(\pm\infty) = 0$. The existence of such heteroclinic connections follows directly from the results on Hamiltonian systems[28], since we only impose $|m| \leqslant 1$. Using numerically computed free energies ψ_β we computed the minimizers of (35) by direct minimization (using the truncated Newton method in the minimization algorithm), see Fig. 2. We observe two types of domain walls even in the one-dimensional case. The in-plane wall is observed when the energy ψ_β has a local minimizer at $|m| = 0$. This wall represents a metastable state since in the case of one-dimensional domain Ω the global minimisers of ψ_β below the critical β_c are at states with $|m| \neq 0$. Note that in higher dimensions the energy balance will involve the shape of the domain through the magnetic potential u. Consequently, the phase diagram and existence of different phases will also depend on the domain Ω and both in-plane and out-plane phases may appear as stable.

3. NON-EQUILIBRIUM

Monte Carlo simulations for lattice spin systems are used primarily as a means for sampling from the canonical Gibbs measure[1], as well as in some cases as a caricature of real non-equilibrium dynamics,[29,30]. These algorithms are set up as ergodic Markov jump processes, which have the canonical Gibbs measure (8) as their invariant measure, see for instance[31]. Using the same considerations as in the Monte Carlo jump dynamics, we can also formulate a system of stochastic differential equations which are of Langevin-type and define relaxation dynamics converging, as $t \to \infty$, to the canonical Gibbs measure (8). In the sequel we restrict our attention to the model Hamiltonian (7). We first derive an evolution equation for the local single spin PDF and present a closed equation for the first spin moment of the PDF, i.e., the average magnetization. The resulting equations are reminiscent of kinetic Fokker–Planck–Vlasov equations (without the convection term) in the sense that they have a Lyapunov functional which is the functional derived for the PDFs in the equilibrium theory. We also study analogous questions for the Langevin dynamics. Furthermore, using spectral and relative entropy estimates we obtain relaxation, and relaxation rate results to the minimizing PDF of the equilibrium functional obtained in Section 2.

3.1. Kinetic Monte Carlo Dynamics

The jump process σ can be constructed as follows: suppose that at time t the configuration is σ_t, then the probability that in the time interval $[t, t+\Delta t]$ the spin at the site $x \in \mathcal{L}_n$ will spontaneously change from $\sigma(x)$ to an arbitrary value $v \in \mathbb{S}^2$ is $c^n(x,\sigma;v)\Delta t + O((\Delta t)^2)$, where $c^n(x,\sigma;v)$ is the jump rate of the process. We also denote by $\sigma^{x,v}$ the resulting new configuration. The dynamics need to be selected so that they guarantee convergence of an arbitrary initial measure to the canonical Gibbs measure. A sufficient condition is known as Detailed Balance (DB):

$$c^n(x,\sigma;v)e^{-\beta H_n(\sigma)} = c^n(x,\sigma^{x,v};\sigma(x))e^{-\beta H_n(\sigma^{x,v})}. \tag{36}$$

Note that $c^n(x,\sigma;v)$ is the rate of converting $\sigma(x)$ to a prescribed spin v and $c^n(x,\sigma^{x,v};\sigma(x))$ the rate of converting v back to $\sigma(x)$. Metropolis-type dynamics satisfying (36) are of the type

$$c^n(x,\sigma;v) = G(\beta\Delta_{x,v}H_n(\sigma)), \quad \text{with} \quad \Delta_{x,v}H_n(\sigma) = H_n(\sigma^{x,v}) - H_n(\sigma). \tag{37}$$

where G is a continuous function which satisfies the condition: $G(r) = G(-r)e^{-r}, \forall r \in \mathbb{R}$. Typical choices of G are $G(r) = 1/1+e^r, G(r) = e^{-r/2}$ and $G(r) = e^{-r+}$; the first case is known under the name of Glauber dynamics and the third are the Metropolis dynamics. The generator of the process is given by

$$L_n f(\sigma) = \int_{\mathbb{S}^2} dv \sum_{x \in \mathcal{L}_n} c^n(x,\sigma;v)\left(f(\sigma^{x,v}) - f(\sigma)\right),$$

where f is any real, bounded test function on Σ_n and

$$\sigma^{x,v}(y) = \begin{cases} v, & \text{if } y = x \\ \sigma(y), & \text{if } y \neq x\,. \end{cases}$$

Another interesting case, that will not be studied in the present paper, is to consider the conservative exchange-spin dynamics (see Ref. 32 for the Ising case).

3.1.1. Derivation of a Kinetic Equation

We consider as our observable the empirical measure (probability measure on $\Omega \times \mathbb{S}^2$) given by (10). Note that there is a one-to-one correspondence between configurations σ and the empirical measures $\mu_t^n(\sigma; dx\,dv)$. In particular, μ_t^n inherits the Markov property from σ. Then considering a test function $g : \Omega \times \mathbb{S}^2 \to \mathbb{R}$ in some class of functions (to be specified later) we have:

$$f(\sigma) = \langle \mu_t^n, g \rangle = \frac{1}{N} \sum_{x \in \mathcal{L}_n} g(x, \sigma(x))\,,$$

where $\langle \cdot, \cdot \rangle$ represents the usual dual pairing: $\langle \mu, f \rangle = \int f(x)\, d\mu(x)$ for f appropriately chosen and again $N = n^d$. For later use we calculate:

$$\begin{aligned} f(\sigma^{x,v}) - f(\sigma) &= \frac{1}{N} \sum_{y \in \mathcal{L}_n} \big(g(y, \sigma^{x,v}(y)) - g(y, \sigma(y)) \big) \\ &= \frac{1}{N} \left(g(x, v) - g(x, \sigma(x)) \right)\,. \end{aligned} \tag{38}$$

We introduce the path space in which we analyze the dynamics: note that $\{t \mapsto \mu_t^n\} \in D([0,T], \mathcal{P})$, where $\mathcal{P} = \mathcal{P}(\Omega \times \mathbb{S}^2)$ is the space of probability measures on $\Omega \times \mathbb{S}^2$ endowed with the weak topology and D is the space of right continuous functions with left limits (this space follows from the fact that μ exhibits jumps). For a more detailed description of these spaces we refer to Ref. 33. We define for every $n \geqslant 1$ a probability measure on $D([0,T], \mathcal{P})$ denoted by Q^n corresponding to the Markov process μ_t^n. Our goal is to prove that, for each fixed time t, the empirical measure $\mu_t^n(dx\,du)$ converges (as $n \to \infty$) in probability to $f(x,u,t)\,dx\,du$, where $f(x,u,t)$ is a solution of an equation to be yet found, with some initial condition $f_0(x,u)$. It suffices to show convergence in distribution of the process μ_t^n to a probability measure which is concentrated on a deterministic path (namely $f(x,u,t)dxdu$), since by standard results (see Ref. 34) convergence in distribution to a deterministic weakly continuous trajectory implies convergence in probability at any fixed time $0 \leqslant t \leqslant T$. Moreover, realizing that a deterministic trajectory can be interpreted as the support of a Dirac probability measure on $D([0,T],\mathcal{P})$, the problem reduces to showing the convergence of the probability measure Q^n to the Dirac measure concentrated on the solution of the equation that we are looking for. We have the theorem:

Theorem 3.1. Let $f_0 : \Omega \times \mathbb{S}^2 \to [0,1]$ be an initial density profile, where $\Omega \equiv [0,1)^d$, and let μ^n be a sequence of probability measures on the configuration space Σ_n, associated to the initial profile f_0, in the following sense:

$$\limsup_{n\to\infty} \mu^n \left\{ \left| \frac{1}{N} \sum_{x\in\mathcal{L}_n} g(x,\sigma^n(x)) - \int g(x,u)\, f_0(x,u)\, dx\, du \right| > \delta \right\} = 0, \quad (39)$$

for every continuous function $g : \Omega \times \mathbb{S}^2 \to \mathbb{R}$ and every $\delta > 0$. Then, for every $t > 0$, the sequence of random measures $\mu_t^n(dy\,du)$ given in (10) converges in probability to the absolutely continuous (with respect to the Lebesgue measure) measure $f(y,u,t)\,dy\,du$, whose density is a solution of the following equation:

$$\frac{d}{dt} f(y,u,t) = \int_{\mathbb{S}^2} S(v,u;y)\, f(y,v,t)\, dv - f(y,u,t) \int_{\mathbb{S}^2} S(u,v;y)\, dv \quad (40)$$

$$f(y,u,0) = f_0(y,u), \quad (41)$$

where $S(u,v;y) = G[\beta((u-v)\cdot \mathbb{J} * m(y,t) + \Psi(v) - \Psi(u))]$, $m(y,t) = \int_{\mathbb{S}^2} u f(u,y,t)\, du$ and $\mathbb{J}$ is a smooth potential.

Remark 3.1. Recalling (36) we get the relation:

$$S(u,v;y) M(y,u,t) = S(v,u;y) M(y,v,t), \quad (42)$$

where M is the Maxwellian defined via (28) without the dipolar term and for time-dependent magnetization m as in the statement of Theorem 3.1.

Using this relation we can rewrite Eq. (40) in the BGK form (see Ref. 35 for BGK-type equations in kinetic theory):

$$\frac{d}{dt} f(y, v, t) = \int_{\mathbb{S}^2} du\, S(v, u; y) M(y, v, t) \left[\frac{f(y, u, t)}{M(y, u, t)} - \frac{f(y, v, t)}{M(y, v, t)} \right]. \tag{43}$$

Proof. To start we need to establish the martingale problem and derive some estimates on the martingale term: under Q^n, for every test function g on $\Omega \times \mathbb{S}^2$, $\langle \mu_t^n, g \rangle$ verify the identity (martingale problem):

$$\langle \mu_t^n, g \rangle = \langle \mu_0^n, g \rangle + \int_0^t L_n \langle \mu_s^n, g \rangle\, ds + M_t^{g,n},$$

where $M_t^{g,n}$ are martingales with respect to the natural filtration $\mathcal{F}_t = \sigma(\sigma_s^n, s \leqslant t)$. From (38) we have:

$$L_n \langle \mu_s^n, g \rangle = \int_{\mathbb{S}^2} dv \sum_{x \in \mathcal{L}_n} c^n(x, \sigma; v) \frac{1}{N} (g(x, v) - g(x, \sigma(x))) =: I - II,$$

where by (37) we obtain:

$$c^n(x, \sigma; v) = G\left(\beta(\sigma(x) - v) \frac{1}{N} \sum_y \mathbb{J}(x - y)\sigma(y) + \beta\Psi(v) - \beta\Psi(\sigma(x)) \right).$$

In particular,

$$\begin{aligned} I &= \int_{\mathbb{S}^2} dv \int_{\Omega \times \mathbb{S}^2} g(y, v)\, G_n[u, v; y]\, \mu_s^n(dy\, du) \\ &= \frac{1}{N} \int_{\mathbb{S}^2} dv \sum_{x \in \mathcal{L}_n} g(x, v)\, G_n[\sigma(x), v; x], \\ II &= \int_{\mathbb{S}^2} dv \int_{\Omega \times \mathbb{S}^2} g(y, u)\, G_n[u, v; y]\, \mu_s^n(dy\, du) \\ &= \frac{1}{N} \int_{\mathbb{S}^2} dv \sum_{x \in \mathcal{L}_n} g(x, \sigma(x))\, G_n[\sigma(x), v; x]. \end{aligned}$$

Thus the martingale problem becomes:

$$
\begin{aligned}
\langle \mu_t^n, g\rangle = \langle \mu_0^n, g\rangle + \int_0^t ds \Bigg\{ \int_{\mathbb{S}^2} dv \int_{\Omega\times\mathbb{S}^2} g(y,v) G_n[u,v;y]\, \mu_s^n(dy\, du) \\
- \int_{\mathbb{S}^2} dv \int_{\Omega\times\mathbb{S}^2} g(y,u) G_n[u,v;y]\, \mu_s^n(dy\, du) \Bigg\} + M_t^{g,n} . \qquad (44)
\end{aligned}
$$

Next we derive some estimates on the martingale term that we will use in the sequel of the proof. Consider the processes:

$$
B_t^{g,n} = L_n \langle \mu_t^n, g\rangle^2 - 2\langle \mu_t^n, g\rangle L_n \langle \mu_t^n, g\rangle , \quad N_t^{g,n} = (M_t^{g,n})^2 - \int_0^t B_s^{g,n} ds .
$$

It is easy to show that $N_t^{g,n}$ is a new martingale (see, for example, Ref. 33 Lemma 5.1, Appendix 1). Then the quadratic variation of $M_t^{g,n}$ is given by:

$$
\langle M_t^{g,n}\rangle = \int_0^t \Big[L_n \langle \mu_s^n, g\rangle^2 - 2\langle \mu_s^n, g\rangle L_n \langle \mu_s^n, g\rangle \Big] ds ,
$$

where $L_n \langle \mu_s^n, g\rangle^2 =$

$$
\begin{aligned}
\int_{\mathbb{S}^2} dv \sum_{x\in\mathcal{L}_n} c^n(x,\sigma;v) \frac{1}{N^2} \Bigg[g(x,v)^2 - g(x,\sigma(x))^2 + 2g(x,v) \sum_{y\in\mathcal{L}_n, y\neq x} g(y,\sigma(y)) \\
-2g(x,\sigma(x)) \sum_{y\in\mathcal{L}_n, y\neq x} g(y,\sigma(y)) \Bigg] .
\end{aligned}
$$

Moreover,

$$
\begin{aligned}
2\langle \mu_s^n, g\rangle L_n \langle \mu_s^n, g\rangle = 2\frac{1}{N} \sum_{x\in\mathcal{L}_n} g(x,\sigma(x)) \\
\times \int_{\mathbb{S}^2} dv \sum_{x\in\mathcal{L}_n} c^n(x,\sigma;u) \frac{1}{N} \left(g(x,v) - g(x,\sigma(x))\right) .
\end{aligned}
$$

Thus,

$$
\langle M_t^{g,n}\rangle = \int_0^t ds \int_{\mathbb{S}^2} dv \sum_{x\in\mathcal{L}_n} c^n(x,\sigma;v) \frac{1}{N^2} \left(g(x,v) - g(x,\sigma(x))\right)^2 \qquad (45)
$$

and since $(M_T^{g,n})^2 = N_T^{g,n} + \langle M_T^{g,n} \rangle$, for g bounded we get that:

$$\mathbb{E}_{Q^n}(|M_T|^2) = \mathbb{E}_{Q^n}\langle M_t^{g,n} \rangle = O\left(\frac{1}{N}\right).$$

By Kolmogorov's generalized inequality and Doob's maximal inequality we have that for every $\epsilon > 0$ and for any time horizon T:

$$Q^n\left[\sup_{0\leqslant t\leqslant T} |M_t^{g,n}| > \epsilon\right] \leqslant \frac{1}{\epsilon^2}\mathbb{E}_{Q^n}\left(\sup_{0\leqslant t\leqslant T} |M_t^{g,n}|^2\right) \leqslant \frac{4}{\epsilon^2}\mathbb{E}_{Q^n}(|M_T|^2) = \frac{1}{\epsilon^2} O\left(\frac{1}{N}\right). \tag{46}$$

Now we are ready to proceed with the steps presented in [Ref. 33 Chapter 4] (all theorems, propositions etc. cited in the following proof refer there):

Step 1. (Relative compactness) We show that the sequence $\{Q^n\}$ is relatively compact. By Proposition 1.7 it suffices to check that $\{Q^n g^{-1}\}$ is relatively compact for all $g \in C^2(\Omega \times \mathbb{S}^2)$ (since $C^2(\Omega \times \mathbb{S}^2)$ is dense in $C(\Omega \times \mathbb{S}^2)$). To check it we need to apply Theorem 1.3 and Proposition 1.6 (for the second condition of the theorem) with $\mathcal{E} = \mathbb{R}$ and δ the usual distance in $\mathbb{R}$. Hence we need to check the following:

(i) $\forall t \in [0, T]$ and every $\epsilon > 0$, there is a compact $K(t, \epsilon) \subset \mathbb{R}$ s.t.

$$\sup_n Q^n g^{-1}[\mathbb{R} \backslash K(t,\epsilon)] \equiv \sup_n Q^n\left[\mu^n : \langle \mu_t^n, g\rangle \notin K(t,\epsilon)\right] \leqslant \epsilon$$

(ii) $\lim_{\gamma\to 0} \limsup_{n\to\infty} \sup_{\tau\in\mathcal{T}_T, \theta\leqslant\gamma} Q^n\left[\mu^n : \left|\langle \mu_{\tau+\theta}^n, g\rangle - \langle \mu_\tau^n, g\rangle\right| > \epsilon\right] = 0, \forall \epsilon > 0$, where $\mathcal{T}_T$ is the family of all stopping times bounded by T.

Condition (i) is trivially verified, since $\langle \mu_t^n, 1\rangle \leqslant 1$. For Condition (ii) we have:

$$\left|\langle \mu_{\tau+\theta}^n, g\rangle - \langle \mu_\tau^n, g\rangle\right| \leqslant \left|\int_\tau^{\tau+\theta} L_n\langle \mu_s^n, g\rangle\, ds\right| + \left|M_{\tau+\theta}^{g,n} - M_\tau^{g,n}\right| =: I + II,$$

where

$$I \leqslant \int_\tau^{\tau+\theta} |L_n\langle \mu_s^n, g\rangle|\, ds \leqslant C(g)\theta$$

and $C(g)$ is a finite constant depending only on g (or $\|g\|_\infty$). For II by (45) we have the estimate:

$$\mathbb{E}_{Q^n}\left[\left(M^{g,n}_{\tau+\theta}-M^{g,n}_{\tau}\right)^2\right]=\mathbb{E}_{Q^n}\left[\int_\tau^{\tau+\theta} B^{g,n}_s ds\right]\leqslant \frac{1}{N}C^2(g)\theta$$

From these two estimates and Kolmogorov's generalized inequality we conclude the proof of (ii) and of Step 1.

Step 2. (Uniqueness of the limit points) We characterize all limit points of Q^n in order to show that all limit points Q^* are concentrated on absolutely continuous (with respect to the Lebesgue measure) measures whose density is a weak solution of the Eq. (40). Moreover, we will need to prove a uniqueness theorem for the weak solutions of the Eq. (40). We split the proof of Step 2 into the following:

(i) All limit points Q^* of the sequence $\{Q^n\}_n$ are concentrated on weak solutions of Eq. (40).

(ii) All Q^*'s are concentrated on absolutely continuous measures with respect to the Lebesgue.

We begin with the proof of (i): Let Q^* be a limit point and let $\{Q^{n_k}\}$ be the subsequence of $\{Q^n\}$ converging to it. We want to show that for every $\epsilon > 0$:

$$Q^*\left\{\mu : \sup_{t\leqslant T}\left|\langle \mu_t, g\rangle - \langle \mu_0, g\rangle - \int_0^t A(s)\, ds\right| > \epsilon\right\} = 0,$$

where $A(s)$ represents the terms in the curly brackets in the weak form of Eq. (40), which is given by:

$$\begin{aligned}\langle \mu_t, g\rangle = \langle \mu_0, g\rangle + \int_0^t ds \left\{\int_{\mathbb{S}^2} du \int_{\Omega\times\mathbb{S}^2} g(y,u)\, S(v,u;y)\, \mu_s(dy\, dv)\right.\\ \left. - \int_{\mathbb{S}^2} dv \int_{\Omega\times\mathbb{S}^2} g(y,u)\, S(u,v;y)\, \mu_s(dy\, du)\right\}, \qquad (47)\end{aligned}$$

where $S(u,v;y) = G\left[\beta\left((u-v)\mathbb{J}*\alpha_t(\sigma;y) + \Psi(v) - \Psi(u)\right)\right]$ and $\alpha_t(\sigma;dy) := \int_{\mathbb{S}^2} v\, \mu_t(dv\, dy)$ is the limit value (as $n\to\infty$) of α^n_t and corresponds to the limit value μ_t of μ^n_t. The convolution $\mathbb{J}*\alpha_t$ is defined by: $\mathbb{J}*\alpha_t(x) := \int_\Omega \mathbb{J}(x-y)\alpha_t(dy)$. The mapping from $D([0,T],\mathcal{P})$ to $\mathbb{R}$ given by:

$$\{\mu_t, 0\leqslant t\leqslant T\} \mapsto \sup_{t\leqslant T}\left|\langle \mu_t, g\rangle - \langle \mu_0, g\rangle - \int_0^t A(s)\, ds\right|$$

is continuous whenever $g \in C^2(\Omega \times \mathbb{S}^2)$. Therefore $\forall \epsilon > 0$ the set $C_{T,\epsilon} := \{\mu : \sup_{t \leqslant T} |\langle \mu_t, g\rangle - \langle \mu_0, g\rangle - \int_0^t A(s)\,ds| > \epsilon\}$ is open and by the "Portmanteau" Theorem (see e.g. Ref. 34) since $Q^{n_k} \overset{*}{\rightharpoonup} Q^*$, we have that: $\liminf_{k\to\infty} Q^{n_k}(C_{T,\epsilon}) \geqslant Q^*(C_{T,\epsilon})$. Hence it suffices to show that the left hand side vanishes. We have:

$$\begin{aligned}\left|\langle \mu_t, g\rangle - \langle \mu_0, g\rangle - \int_0^t A(s)\,ds\right| &\leqslant \left(\left|\langle \mu_t^n, g\rangle - \langle \mu_t, g\rangle\right| + \left|\langle \mu_0^n, g\rangle - \langle \mu_0, g\rangle\right|\right) \\ &\quad + |M_t^{g,n}| + \left|\int_0^t A^n(s)\,ds - \int_0^t A(s)\,ds\right| \\ &=: I_1 + I_2 + I_3,\end{aligned}$$

where $A^n(s)$ includes the terms in the curly bracket in the martingale problem (44). We calculate each term separately. For I_1:

$$\liminf_{k\to\infty} Q^{n_k}\left\{\mu : \sup_{t\leqslant T} I_1 > \frac{\epsilon}{3}\right\} \leqslant \lim_{n\to\infty} Q^n\left\{\mu : \sup_{t\leqslant T} I_1 > \frac{\epsilon}{3}\right\} = 0.$$

For the martingale term I_2 we use the estimate from before, i.e.: $Q^n\left\{\sup_{0\leqslant t\leqslant T} I_2 \geqslant \frac{\epsilon}{3}\right\} \leqslant \frac{1}{\epsilon^2}\frac{1}{N}C(g)T$. So the last term to be controlled is the difference $A^n(s) - A(s)$, whose essential part is: $G_n[u, v; y]\,\mu_s^n(dy\,du) - G[u, v; y]\,\mu_s(dy\,du)$. By adding and subtracting $G[u, v; y]\,\mu_s^n(dy\,du)$ it suffices to estimate $G_n\mu_s^n - G\mu_s^n$. But G and G_n are continuous functions of their arguments and the total mass of μ_s^n is bounded by 1, so we get that:

$$\int_{\Omega\times\mathbb{S}^2} g(y, u)\left(G_n[u, v; y]\,\mu_s^n(dy\,du) - G[u, v; y]\,\mu_s(dy\,du)\right) = o_n(1).$$

This completes the proof of (i).

Now for (ii) we have:

$$|\langle \mu_t, g\rangle| = \lim_{n\to\infty}\left|\langle \mu_t^n, g\rangle\right| \leqslant \limsup_{n\to\infty}\frac{1}{N}\sum_{y\in\mathcal{L}_n} g(y, \sigma_t^n(y)) \leqslant \int g(y, u)\,dy\,du.$$

Thus, if we pick a set whose $dydu$ measure is zero then the same will be true for its measure with respect to μ_t. Moreover, we need to check that all limit points of $\{Q^n\}$ are concentrated on trajectories that at $t = 0$ are equal to $f_0(y, u)dydu$, i.e., that for every $\epsilon > 0$:

$$Q^*\left\{\left|\langle \mu_0, g\rangle - \int f_0(y, u)g(y, u)\,dy\,du\right| > \epsilon\right\} = 0,$$

which follows immediately by the weak convergence of Q^{n_k} to Q^* and by the hypothesis (39) of the theorem. This concludes the proof of (ii). Before going to the next step we also need to show that for any $T > 0$ the weak form of Eq. (40) is also true for time-dependent test functions g, which, given the above calculations, is straightforward by following the corresponding strategy in Ref. 33.

Step 3. (Uniqueness of weak solutions of the Eq. (40)) Uniqueness of solutions of Eq. (40) can be proved by the standard technique of the fixed point theorem, since the right-hand side of (40) is Lipschitz continuous.

So far (Steps 1–3) we have shown that all limit points Q^* of the sequence Q^n are concentrated on absolutely continuous measures with respect to the Lebesgue measure. Hence we have shown that $\mu_t^n(dx\,dv) \overset{*}{\rightharpoonup} f(x,v,t)\,dx\,dv$. Considering now the case where $\mathbb{J}$ is smooth, symmetric and with fast decay at infinity we have that $\alpha_t^n(dx) \overset{*}{\rightharpoonup} m(x,t)\,dx$ which implies that $\mathbb{J} * \alpha_t^n(x) \to \mathbb{J} * m(x,t)$ strongly. This is a result of long-range interactions and smoothness of $\mathbb{J}$. Basically the former is a crucial assumption that allows us to close the equation. Moreover, G is continuous so by letting $n \to \infty$ we get:

$$G_n[u,v;x] \to G\left[\beta\left((u-v)\cdot \mathbb{J} * m(x) + \Psi(v) - \Psi(u)\right)\right] =: S(u,v;x)\,.$$

Hence for $g \in C^2(\Omega \times \mathbb{S}^2)$ since $\mu_t^n(dx\,dv) \overset{*}{\rightharpoonup} f(x,v,t)\,dx\,dv$ we get the convergence (as $n \to \infty$): $\langle \mu_t^n, g\,G_n\rangle \to \langle f, gS\rangle$. By passing to the limit $n \to \infty$ in (44) we get Eq. (40) in weak form. ∎

We observe that due to the generality of equation (40) we cannot get a closure in the first moment to obtain a closed equation for the magnetization $m(x)$.

Example (*Evolution of the magnetization*). Under specific assumptions on the rate function (37) it is possible to obtain a closure for the first moment of the PDF, i.e., $m(x,t)$. In Ref. 36 the authors consider a continuous scalar spin ($\sigma(x) \in [-b,b]$, $x \in \mathcal{L}_n$) lattice system and derive a kinetic equation for the PDF together with a closed equation for the magnetization (the first moment of f). The main difference is the special choice of the rate function which is given by:

$$c^n(x,v;\sigma) = e^{\beta v \mathbb{J} * \alpha^n(x) + o_n(1)}\,. \tag{48}$$

We see that this is not a choice of Glauber or Metropolis type as in Section 3.1. The lower order terms in (48) are necessary in order the DB condition to be satisfied. The leading part of the exponent in (48) allows us to close the derived kinetic Eq. (40) in the first moment and obtain an equation for the magnetization $m(x,t)$ as $n\to\infty$. Recalling Eq. (40), we first rephrase it to be consistent with the setup of Ref. 36. To start with, we first include the anisotropy into the prior distribution measure so that instead of the uniform measure on $\mathbb{S}^2$ we now have the anisotropic $\tilde{\rho}(dv)=N(v)dv$ where $N(v)$ is given by (29). Then, using the same notation f, we now consider as f the Radon–Nikodym derivative of the initial PDF with respect to the measure $\tilde{\rho}(dv)$. Hence the new Maxwellian solution will be:

$$M(x,v,t)=\frac{1}{Z_\beta(x,t)}e^{\beta v\cdot\mathbb{J}*m(x,t)},\quad Z_\beta(x,t)=\int_{\mathbb{S}^2}e^{\beta v\cdot\mathbb{J}*m(x,t)}\tilde{\rho}(dv)\,.$$

Then applying the special choice of the rate (48) in our problem we get the following equation (in a BGK form) for the new probability density function $f(x,v,t)$:

$$\frac{d}{dt}f^v=\exp\left\{\beta a_\beta(\mathbb{J}*m)\right\}(M^v-f^v)\,, \tag{49}$$

where the function a_β is given by (27). Note that for the sake of simplicity we have used the notation: $f^u:=f(x,u,t)$ and $M^u:=M(x,u,t)$. Moreover, because of this choice we can also obtain a closed equation for the first moment of f, i.e., for the magnetization $m(x,t)$:

$$\frac{d}{dt}m=\exp\{\beta a_\beta(\mathbb{J}*m)\}\left[\partial\, a_\beta(\mathbb{J}*m)-m\right]. \tag{50}$$

Observe that (50) can be viewed as generalization of the usual equation that appears in the Ising case (see for instance[7]). We can also see that, if we neglect the micromagnetic contribution, the free energy (24) is in fact a Liapunov functional.

3.1.2. Relaxation to the Equilibrium

In Section 2.2 we have derived from statistical mechanics considerations the energy of the system given by (16). Moreover, we have seen that the energetically most favorable configuration describing the large-scale features on $\mathcal{L}_n\approx\Omega$ is given by the Maxwellian distribution (28) without the dipolar term, yielding the probability of having a spin $v\in\mathbb{S}^2$ at the location $x\in\Omega$. In this section we study the long-time behavior of the solution $f(x,v,t)$ of the Eq. (40) and we want to show that f relaxes to the local Maxwellian. Note that M depends on t through the average magnetization m. But m, in turn, seems to relaxe to $\bar{m}(x)$ which is the minimizer of the free energy functional (24) derived in Theorem 2.3. Our ultimate goal is to show the convergence $f \to \bar{M}$ in some appropriate norm, where by $\bar{M}$ we denote the Maxwellian that corresponds to the time-independent magnetization $\bar{m}$. In this context, we derive the corresponding H-Theorem (where we use the notation $S(u, v; m)$ instead of $S(u, v; x)$ in order to emphasize the dependence on $m(x)$):

$$\frac{d}{dt}E[f] = \frac{1}{\beta}\int_{\Omega} dx \int_{\mathbb{S}^2}\int_{\mathbb{S}^2} du\, dv\, S(v, u; m) M^v \left[\frac{f^u}{M^u} - \frac{f^v}{M^v}\right] \log \frac{f^v}{M^v}.$$

Then, applying a change of variable together with the DB condition (42) we have that:

$$\frac{d}{dt}E[f] = -I(f|M), \tag{51}$$

where:

$$I(f|M) = \frac{1}{2\beta}\int_{\Omega} dx \int_{\mathbb{S}^2\times\mathbb{S}^2} du\, dv\, S(v, u; m) M^v \times \left[\frac{f^u}{M^u} - \frac{f^v}{M^v}\right]\left[\log \frac{f^u}{M^u} - \log \frac{f^v}{M^v}\right].$$

Since $S(v, u; m)M^v$ is positive and the function $g(x) = (x-1)\log x$ nonnegative with $g(x) = 0$ iff $x = 1$, we have that $I(f|M) \geqslant 0$ with "=" iff $f \equiv M$.

Moreover, recalling from Theorem 2.3 that $E[f]$ is bounded from below by $E[\bar{M}]$ and given the above fact that it decreases unless f is a Maxwellian, it is tempting to conclude that f tends to $\bar{M}$ when $t \to \infty$. Furthermore, we have the following formula:

$$E[f] - E[\bar{M}] = \frac{1}{\beta}e(f|\bar{M}) - \frac{1}{2}\int_{\Omega} \mathbb{J} * (m - \bar{m})(m - \bar{m})\, dx,$$

where:

$$e(f|\bar{M}) = \int_{\Omega\times\mathbb{S}^2} f \log \frac{f}{\bar{M}}\, dx\, dv \tag{52}$$

is the physical relative entropy. The above formula suggests a connection between the convergence in the relative entropy, the convergence of $E[f]$

to $E[\bar{M}]$ and of m to $\bar{m}$ (or equivalently of f to $\bar{M}$). In the following example we show this convergence in a very simple case, which is the equation without interactions (i.e. $\mathbb{J} \equiv 0$). Although this seems to be a trivial example from the physics point of view, it highlights the difficulty introduced to the problem by the continuous vectorial spins.

Example (*Special case* $\mathbb{J} \equiv 0$). We have the equation:

$$\frac{d}{dt} f(v) = \int_{\mathbb{S}^2} S(v,u) M(v) \left(\frac{f(u)}{M(u)} - \frac{f(v)}{M(v)} \right) du\,. \tag{53}$$

We notice that since $\mathbb{J} = 0$ the expression (53) is x-independent. In this case $M(v) = \exp\{-\beta a_\beta(0) - \beta \Psi(v)\}$, where $a_\beta(0) = \frac{1}{\beta} \log \int_{\mathbb{S}^2} e^{-\beta \Psi(v)}\, dv$ and $S(v,u) = G\,(\beta \Psi(u) - \beta \Psi(v))$.

We want to show that $f \to M$ as $t \to \infty$ in the L^1-norm. We will prove first something stronger, $\|f - M\|_{L^2(M^{-1}(v)dv)} \to 0$, which is equivalent to proving $\left|\left| \frac{f}{M} - 1 \right|\right|_{L^2(M(v)dv)} \to 0$. We let $h := \frac{f}{M}$ and we study the following linear integro-differential equation in the space $L^2(M(v)dv)$:

$$\frac{d}{dt} h(v) = \int_{\mathbb{S}^2} S(v,u)\,(h(u) - h(v))\; du \;=: (Lh)(v)\,. \tag{54}$$

We have the following:

Proposition 3.1. (i) The spectrum $\Sigma(L) = \Sigma_d \cup \Sigma_c$ of the operator L consists of a discrete $\Sigma_d \subset (-\nu_0, 0]$ and an essential part $\Sigma_c = (-\infty, -\nu_0]$, where ν_0 is a positive constant. (ii) The solution h of Eq. (54) decays exponentially fast in time to the constant solution 1.

Proof. We can easily check that the operator L on $L^2(M(v)dv)$ is linear, bounded, symmetric and negative semidefinite. For the later we have:

$$\langle h, Lh \rangle_{L^2(M(v)dv)} = -\frac{1}{2} \int_{\mathbb{S}^2 \times \mathbb{S}^2} S(v,u) M(v) (h(u) - h(v))^2\, du\, dv \leqslant 0\,, \tag{55}$$

being zero iff $h(u) = h(v) =$ const., which implies that $\ker(L) = \mathrm{span}\{1\}$. Moreover, L has the decomposition $L = K - \nu(v)\, Id$, where $(Kh)(v) = \int_{\mathbb{S}^2} S(v,u) h(u)\, du$ is a compact operator (Hilbert–Schmidt operator in $L^2(M(v)dv, \mathbb{S}^2)$). The function $\nu(v) = \int_{\mathbb{S}^2} S(v,u)\, du$ satisfies the bounds $0 < \nu_0 \leqslant \nu(v) \leqslant \nu_1(v)$ for some ν_0 and $\nu_1(v)$ which depend on the specific choice of the function $G(r)$ and can be calculated explicitly in each case. Given these properties (i) follows by applying Theorem IV.5.35 in Ref. 37.

(ii) By (i) we have the following spectral gap property:

$$\langle g, Lg\rangle_{L^2(M(v)dv)} \leqslant -\alpha \|g\|^2_{L^2(M(v)dv)}, \tag{56}$$

for some α such that $0<\alpha<\nu_0$ and for every g with $g \perp_{L^2(M(v)dv)}$const. Note that $h-1$ has this property since: $\int_{\mathbb{S}^2} f(v)\,dv = \int_{\mathbb{S}^2} M(v)\,dv = 1$ if and only if $\int_{\mathbb{S}^2}\left(\frac{f(v)}{M(v)}-1\right)M(v)\,dv=0$. Thus, from (56) we get:

$$\frac{1}{2}\frac{d}{dt}\|h-1\|^2_{L^2(M(v)dv)} \leqslant -\alpha\|h-1\|^2_{L^2(M(v)dv)} \Rightarrow$$
$$\|h-1\|_{L^2(M(v)dv)} \leqslant e^{-\alpha t}\|h_0-1\|_{L^2(M(v)dv)}. \quad \blacksquare$$

Remark 3.2. From [Ref. 38 Lemma 2.6] we see that the convergence in $L^2(M(v)dv)$ for $h=\frac{f}{M}$ to the constant 1 implies convergence in the relative entropy (52) which in turn implies convergence in the L^1-norm, by applying the Csiszár–Kullback inequality.

3.2. Relaxation Langevin Dynamics

We analyze a microscopic stochastic model where the change of spin is not given by a spontaneous jump from one value to another (leading to a stochastic jump process), but it is dictated by the Langevin dynamics which are described by a Stochastic Differential Equation (SDE) of the form:

$$d\theta_i = -\frac{\delta H_N}{\delta\theta_i}dt + \tau\, dW^i, \quad i=0,\dots,N-1 \tag{57}$$

where $\theta_i \in [0,2\pi)$ is the angle of the i-th spin, which now takes values on $\mathbb{S}^1$, on the periodic lattice $\mathcal{L}_N$ that for simplicity we consider to be one-dimensional (i.e. $N=n$ and $\mathcal{L}_N=\{\frac{i}{N}; i=0,\dots,N-1\}$).$H_N$ is the interaction Hamiltonian (7) which in the set-up of this section will be given by:

$$H_N(\sigma) = -\frac{1}{2N}\sum_{j,k=0}^{N-1}\mathbb{J}(j-k)\cos(\theta_j-\theta_k) + \sum_{j=0}^{N-1}\Psi(\theta_j),$$

where Ψ is the anisotropy function $\Psi:[0,2\pi)\to\mathbb{R}$ and τ is a scalar depending on the inverse temperature β. We also denote by $\mathbf{W}=(W^0,\dots,W^{N-1})$ the standard N-dimensional Wiener process. This SDE is defined by a generator through the usual martingale problem (see

Ref. 39), so the fact that $\theta_i \in [0, 2\pi)$ is ensured. Note also that in this section we are assuming the spins to be on the unit circle rather than the unit sphere. We calculate:

$$\frac{\delta H_N}{\delta \theta_k} = \frac{1}{N} \sum_{j=0,\, j\neq k}^{N-1} \mathbb{J}(j-k) \sin(\theta_k - \theta_j) + \frac{\partial}{\partial \theta_k} \Psi(\theta_k), \quad k = 0, \ldots, N-1\,,$$

where $\mathbb{J}(r) = \mathbb{J}(|r|)$, $r \in \mathbb{R}$. We consider as our observable the usual empirical measure: $\mu_t^N(dx\, d\theta) = \frac{1}{N} \sum_{i=0}^{N-1} \delta_{\frac{i}{N}}(dx) \delta_{\theta_i(t)}(d\theta)$, where $d\theta$ is the Lebesgue measure on $[0, 2\pi)$. If ϕ is an appropriate test function, we have:

$$\langle \mu_t^N, \phi \rangle = \frac{1}{N} \sum_{i=0}^{N-1} \phi\left(\frac{i}{N}, \theta_i(t)\right).$$

We want to derive a closed equation for the quantity $\langle \mu_t^N, \phi \rangle$. Using Itô's formula from (57) we obtain:

$$\begin{aligned}
\frac{d}{dt} \langle \mu_t^N, \phi \rangle &= -\frac{1}{N} \sum_{i=0}^{N-1} \frac{\partial \phi(\frac{i}{N}, \theta_i)}{\partial \theta_i} \frac{1}{N} \sum_{j \neq i} \mathbb{J}\left(\frac{i}{N} - \frac{j}{N}\right) \sin(\theta_i - \theta_j) \\
&\quad + \frac{1}{2N} \sum_{i=0}^{N-1} \tau^2 \frac{\partial^2 \phi(\frac{i}{N}, \theta_i)}{\partial \theta_i^2} \\
&\quad - \frac{1}{N} \sum_{i=0}^{N-1} \frac{\partial \phi(\frac{i}{N}, \theta_i)}{\partial \theta_i} \frac{\partial \Psi(\theta_i)}{\partial \theta_i} \\
&\quad + \frac{1}{N} \sum_{i=0}^{N-1} \frac{\partial \phi(\frac{i}{N}, \theta_i)}{\partial \theta_i} \frac{dW^i}{dt} =: I_1 + I_2 + I_3 + I_4\,. \qquad (58)
\end{aligned}$$

Now we derive an equation for the limit of the empirical measure by calculating the mesoscopic limit of (58). We have the theorem:

Theorem 3.2. Let $f_0 : \Omega \times [0, 2\pi) \to [0, 1]$ be an initial density profile, with $\Omega \equiv [0, 1)$, and let μ_0^N be a sequence of probability measures on the configuration space $[0, 2\pi)^{\mathcal{L}_N}$, associated to the initial profile f_0, in the following sense:

$$\limsup_{N\to\infty} \mu_0^N \left\{ \left| \frac{1}{N} \sum_{i=0}^{N-1} \phi\left(\frac{i}{N}, \theta_i\right) - \int \phi(x, \theta)\, f_0(x, \theta)\, dx\, d\theta \right| > \delta \right\} = 0\,, \quad (59)$$

for every continuous function $\phi : \Omega \times [0, 2\pi) \to \mathbb{R}$ and every $\delta > 0$. Then the sequence of measure valued processes $\{\mu_t^N\}_N$ defined above is tight and any limit μ_t is a continuous probability measure-valued process absolutely continuous with respect to the Lebesgue measure $dx\,d\theta$, whose density $f(x, \theta, t)$ satisfies the following equation:

$$\begin{aligned}\frac{\partial}{\partial t} f &= \partial_\theta (f (\partial_\theta K) * f) + \partial_\theta (f \partial_\theta \Psi(\theta)) + \frac{1}{2}\tau^2 \partial^2_{\theta\theta} f \\ &= \partial_\theta \left(f \partial_\theta \left(K *_{x,\theta} f + \frac{1}{\beta} \log f + \Psi\right)\right), \end{aligned} \tag{60}$$

with the initial condition $f(x, \theta, 0) = f_0(x, \theta)$. Moreover, the kernel K is given by

$$K(x - x', \theta - \theta') = \mathbb{J}(x - x') \cos(\theta - \theta') \tag{61}$$

and therefore, $(\partial_\theta K) * \mu(x, \theta) := -\iint \mathbb{J}(x - x') \sin(\theta - \theta')\mu(dx', d\theta')$.

Proof. Given the fact that the spins are on the lattice, the proof follows the lines of the proof of Theorem 3.1. We need to prove the convergence in law of the corresponding probability measure Q^N on the path space $D([0, T], \mathcal{P})$ (see the proof of Theorem 3.1) by showing tightness, identification of the limit values, uniqueness of the limit value. In order not to repeat the proof we just present the main calculations. We write the right-hand side of (58) in terms of $\langle \mu_t^N, \phi \rangle$:

$$\begin{aligned} I_1 &= -\frac{1}{N} \sum_{i=0}^{N-1} \frac{\partial \phi(\frac{i}{N}, \theta_i)}{\partial \theta_i} \iint \mathbb{J}\left(y - \frac{i}{N}\right) \sin(\theta_i - \theta') \frac{1}{N} \sum_{j=0}^{N-1} \delta_{\frac{j}{N}}(dy) \delta_{\theta_j(t)}(d\theta') \\ &= -\frac{1}{N} \sum_{i=0}^{N-1} \frac{\partial \phi(\frac{i}{N}, \theta_i)}{\partial \theta_i} (\partial_\theta K) * \mu^N \left(\frac{i}{N}, \theta_i, t\right) \\ &= -\iint \frac{\partial \phi(x, \theta)}{\partial \theta} K * \mu^N(x, \theta, t) \frac{1}{N} \sum_{i=0}^{N-1} \delta_{\frac{i}{N}}(dx) \delta_{\theta_i}(d\theta) = -\left\langle \frac{\partial \phi}{\partial \theta} K * \mu^N, \mu^N \right\rangle. \end{aligned}$$

Similarly, $I_2 = \langle \mu^N, \tau^2 tr(\nabla^2 \phi) \rangle$ and $I_3 = -\langle \frac{\partial \phi}{\partial \theta} \partial_\theta \Psi, \mu^N \rangle$. Hence (58) will give the following formulation of the martingale problem:

$$\langle \mu_t^N, \phi \rangle - \langle \mu_0^N, \phi \rangle = \int_0^t A_N \langle \mu_s^N, \phi \rangle \, ds + M_t^{N,\phi}, \tag{62}$$

where A_N is the generator of the Markov process $\{\mu_t^N\}$ suggested by (58), and $M_t^{N,\phi} = \frac{1}{N} \sum_{i=0}^{N-1} \tau \int_0^t \frac{\partial \phi(\frac{i}{N}, \theta_i)}{\partial \theta_i} \frac{dW^i}{ds} \, ds$ is the martingale term. Since

$\{W^i\}_{i=0}^{N-1}$ are independent Wiener processes it is easy to show that the quadratic variation of $M_t^{N,\phi}$ is given by:

$$\langle M_t^{N,\phi}\rangle = \frac{1}{N^2}\sum_{i=0}^{N-1}\tau^2\left(\frac{\partial\phi(\frac{i}{N},\theta_i)}{\partial\theta_i}\right)^2$$

and therefore, $\mathbb{E}_{Q^N}\langle M_t^{N,\phi}\rangle = O(\frac{1}{N})$, since ϕ is bounded. From this estimate we can show tightness and therefore convergence up to a subsequence. To complete the proof we need to show uniqueness of the limit point μ of the solution of the weak version of Eq. (60). This follows from a similar result in Ref. 40. Moreover, we can show that all limit points of the sequence of the measures Q^N on the path space are supported on measures μ_t which are absolutely continuous with respect to the Lebesgue measure $f(x,\theta,t)\,dx\,d\theta$. Then the probability density function f satisfies the nonlinear Fokker–Planck Eq. (60). ∎

The Eq. (60) can be interpreted as a gradient flow: "$\frac{d}{dt}f = -\nabla\mathcal{F}(f)$". By ∇ we represent the gradient which is a tangent vector field on the space of probability density functions f and $\mathcal{F}(f)$ is the corresponding free energy of the system, given by:

$$\begin{aligned}\mathcal{F}[f] = \frac{1}{\beta}\int f\log f\,dx\,d\theta + \int\Psi(\theta)f\,dx\,d\theta\\ +\frac{1}{2}\int K(x-x',\theta-\theta')f(x,\theta)f(x',\theta')dx\,dx'\,d\theta\,d\theta'.\end{aligned} \tag{63}$$

Equation (60) can be written as:

$$\frac{d}{dt}f = -\partial_\theta\left(f\,\partial_\theta\left(-\frac{\delta\mathcal{F}}{\delta f}\right)\right).$$

Then we see that $\mathcal{F}$ decays along $f(t)$ (H-Theorem for Eq. (60)):

$$\begin{aligned}\frac{d}{dt}\mathcal{F}(f(t)) = \int\frac{\delta\mathcal{F}}{\delta f}\left(\frac{df}{dt}\right)dx\,d\theta = \frac{1}{\beta}\int\frac{\delta\mathcal{F}}{\delta f}\,\partial_\theta\left(f\,\partial_\theta\left(\frac{\delta\mathcal{F}}{\delta f}\right)\right)dx\,d\theta\\ = -\frac{1}{\beta}\int f\left|\partial_\theta\left(\frac{\delta\mathcal{F}}{\delta f}\right)\right|^2 dx\,d\theta \leqslant 0,\end{aligned} \tag{64}$$

with "=" iff $\frac{\delta\mathcal{F}}{\delta f} = \text{const}$ in θ. Hence for the equilibrium solutions we have: $K *_{x,\theta} f + \frac{1}{\beta}\log f + \Psi = c(x)$, where

$$K *_{x,\theta} f(x,\theta) = -\begin{pmatrix}\cos\theta\\ \sin\theta\end{pmatrix}\cdot \mathbb{J} *_x m(x) = -\begin{pmatrix}\cos\theta\\ \sin\theta\end{pmatrix}\cdot \int \mathbb{J}(x-y)m(y)dy\,. \tag{65}$$

Thus,

$$\bar{M}(x,\theta) = \frac{1}{Z(x)}\exp\left\{-\beta\Psi(\theta) + \beta\begin{pmatrix}\cos\theta\\ \sin\theta\end{pmatrix}\cdot \mathbb{J} * \bar{m}(x)\right\} \tag{66}$$

$$\bar{m}(x) = \int_0^{2\pi}\begin{pmatrix}\cos\theta\\ \sin\theta\end{pmatrix}\bar{M}(x,\theta)\,d\theta \tag{67}$$

which is the steady state solution of Eq. (60).

Next we study the asymptotic behavior of solutions to the Eq. (60). We observe that neither Ψ nor K are convex functions of the variable θ. Furthermore, the x,θ-convolution $K * f$ is merely a convolution with respect to the x-variable only, as can be seen in (65). Hence, in our case, the terms involving Ψ and $\mathbb{J}$ will be treated as L^∞ perturbations of a strictly convex potential and subsequently we derive an exponential decay for an appropriate relative entropy functional.

The strategy to show relaxation to equilibrium is similar to the one presented in Ref. 38 for the drift-diffusion-Poisson model. Note also that we get this decay for the case where the mapping $u \mapsto \langle \mathbb{J} * u,\, u\rangle$ is negative definite. We return to this assumption in Remark 3.3 below. We start by defining a t-local state:

$$M(x,\theta,t) = \frac{1}{Z(x,t)}\exp\left\{-\beta\Psi(\theta) + \beta\begin{pmatrix}\cos\theta\\ \sin\theta\end{pmatrix}\cdot \mathbb{J} *_x m(x,t)\right\}, \tag{68}$$

where

$$Z(x,t) = \int_0^{2\pi}\exp\{-\beta\Psi(\theta) + \beta\begin{pmatrix}\cos\theta\\ \sin\theta\end{pmatrix}\cdot \mathbb{J} *_x m(x,t)\}\,d\theta\,.$$

Then the equation has the BGK form:

$$\frac{\partial}{\partial t} f = \partial_\theta\left(f\,\partial_\theta\left(\frac{1}{\beta}\log\frac{f}{M}\right)\right) = \frac{1}{\beta}\partial_\theta\left(M\partial_\theta\left(\frac{f}{M}\right)\right).$$

Moreover, by the H-Theorem we have (at least formally):

$$\frac{d}{dt}\mathcal{F}(f(t)) = -\frac{1}{\beta}\iint f\left|\partial_\theta\left(\log\frac{f}{M}\right)\right|^2 dx\,d\theta\,.$$

For any fixed $x \in \Omega$, if M were of the form $M(\theta) = e^{-\beta V(\theta)}$, with V strictly convex (say $\partial_\theta^2 V \geqslant \lambda_1$, for some positive constant λ_1), then we have the logarithmic-Sobolev inequality:

$$\int f \log\frac{f}{M}\,d\theta \leqslant \frac{1}{2\lambda_1}\int f\left|\partial_\theta\left(\log\frac{f}{M}\right)\right|^2 d\theta\,. \tag{69}$$

Nevertheless this can still be done for M as in (68) since by the perturbation Theorem 3.2 in Ref. 38 we can express the exponent of M as a bounded perturbation of a strictly convex function V (which is the sricly convex version of the convex envelope of Ψ), chosen in the process of the proof. Then, we need to control the difference between the exponent $V(\theta)$ and the exponent of M. Note that (68) can be written in the form:

$$M(x,\theta,t) = \exp\left\{-\beta\Psi(\theta) + \beta\begin{pmatrix}\cos\theta\\ \sin\theta\end{pmatrix}\cdot \mathbb{J} *_x m(x,t) - \beta a_\beta(\mathbb{J} *_x m(x,t))\right\},$$

where

$$a_\beta(p) = \frac{1}{\beta}\log\int_0^{2\pi}\exp\left\{\beta\begin{pmatrix}\cos\theta\\ \sin\theta\end{pmatrix}\cdot p - \beta\Psi(\theta)\right\}d\theta\,,$$

for $p \in \mathbb{R}^2$, as in (27). Then we want to estimate the difference between the two exponents. We define:

$$v(x,\theta,t) := V(\theta) - \left(\Psi(\theta) - \begin{pmatrix}\cos\theta\\ \sin\theta\end{pmatrix}\cdot \mathbb{J} *_x m(x,t) + a_\beta(\mathbb{J} *_x m(x,t))\right).$$

It is easy to check that there are positive constants k_1, k_2 and k_3 such that:

$$\left|\begin{pmatrix}\cos\theta\\ \sin\theta\end{pmatrix}\cdot \mathbb{J} *_x m(x,t)\right| \leqslant |\mathbb{J} * m(x,t)| \leqslant k_1, \quad \forall x,\theta,t$$

(since $\|m(t,\cdot)\|_\infty$ is uniformly bounded in t and $\mathbb{J}$ is smooth) which in turn implies that $|a_\beta(\mathbb{J} *_x m(x,t))| \leqslant k_2$. Moreover, $0 \leqslant \Psi(\theta) - V(\theta) \leqslant k_3$. Thus,

$$0 < e^{-k_1-k_2} \leqslant e^{-v(x,\theta,t)} \leqslant e^{k_1+k_2+k_3} < \infty, \quad \forall x,\theta,t\,.$$

Hence, by Theorem 3.2 in Ref. 38 we have the same logarithmic-Sobolev inequality (69) for M given by (68), with a new coefficient $c>0$ depending on λ_1, k_1, k_2, k_3 and the fixed x. But, since Ω is compact, we can find another constant c (independent of x) for which:

$$\iint f \log \frac{f}{M}\, dx\, d\theta \leqslant c \iint f \left| \partial_\theta \left(\log \frac{f}{M} \right) \right|^2 dx\, d\theta\, .$$

We define now the appropriate relative entropy functional for which we will show the exponential decay. We have:

$$e_\beta(t) := \mathcal{F}(f(t)) - \mathcal{F}(\bar{M})\,, \tag{70}$$

from which, after a simple calculation, we obtain:

$$e_\beta(t) = \frac{1}{\beta} \iint f \log \frac{f}{\bar{M}}\, dx\, d\theta - \frac{1}{2} \int \mathbb{J} * (m - \bar{m}) \cdot (m - \bar{m})\, dx\, .$$

Our goal is to show exponential decay for the quantity $e_\beta(t)$. We have:

$$\begin{aligned} \frac{d}{dt} e_\beta(t) &= \frac{d}{dt} \mathcal{F}(f(t)) = -\frac{1}{\beta} \iint f \left| \partial_\theta \left(\log \frac{f}{M} \right) \right|^2 dx\, d\theta \\ &\leqslant -c \frac{1}{\beta} \iint f \log \frac{f}{M}\, dx\, d\theta \\ &= -c \frac{1}{\beta} \iint f \log \frac{f}{\bar{M}}\, dx\, d\theta - c \frac{1}{\beta} \iint f \log \frac{\bar{M}}{M}\, dx\, d\theta\, . \end{aligned} \tag{71}$$

Thus, after using Jensen's inequality for the second term of the right hand side of (71), we have:

$$\frac{d}{dt} e_\beta(t) \leqslant -c \frac{1}{\beta} \iint f \log \frac{f}{\bar{M}}\, dx\, d\theta + c \int \mathbb{J} * (m - \bar{m}) \cdot (m - \bar{m})\, dx\, .$$

Note that if the second term is negative, i.e. the quadratic form $u \mapsto \langle \mathbb{J} * u, u \rangle$ is negative definite in $L^2(\Omega)$, then we can bound it and get the desired right-hand side. Thus we conclude the exponential decay of the appropriate relative entropy functional $e_\beta(t)$:

$$\frac{d}{dt} e_\beta(t) \leqslant e^{-ct} e_\beta(0)\,, \tag{72}$$

where $e_\beta(0) = \frac{1}{\beta}\iint f_0 \log\frac{f_0}{\bar{M}}\, dx\, d\theta - \frac{1}{2}\int \mathbb{J} * (m_0 - \bar{m}) \cdot (m_0 - \bar{m})\, dx$ with $f_0 \in L^1_+(\Omega \times [0, 2\pi))$ the initial condition and m_0 the corresponding to f_0 mean magnetization. In fact we have proved the theorem:

Theorem 3.3. Let $f_0 \in L^1_+(\Omega \times [0, 2\pi))$. Suppose also that the quadratic form $u \mapsto \langle \mathbb{J} * u, u\rangle$ is negative definite in $L^2(\Omega)$. Then there exists a positive constant c such that the exponential convergence (72) of the relative entropy functional (70) holds for a solution of the Eq. (60).

Remark 3.3. The relative entropy type estimates prove a global relaxation to equilibrium. In Section 2.5 we have seen that for the ferromagnetic case we have multiple steady state solutions as well as corresponding multiple standing waves (out-plane waves and their symmetric counterpart in Fig. 2) for the energy functional. Consequently, we cannot expect such an estimate since this would mean that with an initial condition close to one minimizer of the energy functional the solution could relaxe to another minimizer, which is impossible. Therefore one should expect that only when a unique equilibrium measure exists such global estimates can be derived.

Remark 3.4. The same method and for the same relative entropy functional e_β can be applied to get exponential convergence for the solutions of the Eq. (49). But again, for the same reason as above we have to assume that the quadratic form $u \mapsto \langle \mathbb{J} * u, u\rangle$ is negative definite.

4. CONCLUSIONS

In this paper we have derived a deterministic mesoscopic theory for model continuous spin lattice systems both at equilibrium and non-equilibrium in the presence of thermal fluctuations. The full magnetic Hamiltonian that includes singular integral (dipolar) interactions has also been considered in the analysis of equilibrium. The non-equilibrium microscopic models we considered are relaxation-type dynamics arising in kinetic Monte Carlo or Langevin-type simulations of lattice systems. In this context we employed the mesoscopic models derived here to study the relaxation of such algorithms to equilibrium. Furthermore, such models provide a first step towards the construction of coarse-grained Monte Carlo algorithms for continuum spin systems, in the spirit of earlier work[41,30] for Ising type systems.

ACKNOWLEDGMENTS

The research of M.K. is partially supported by NSF-DMS-0100872 and NSF-ITR-0219211. The visits of P.P. to Amherst were supported by

NSF-DMS-0303565 and the University of Massachusetts. The research of D.T. is partially supported by NSF-ITR-0219211. M.K. and P.P. would like to thank the Isaac Newton Institute for Mathematical Sciences where part of this work was initiated during a visit in May 2003. We would also like to thank José Trashorras for many valuable discussions.

REFERENCES

1. D. Landau and K. Binder. *A Guide to Monte Carlo Simulations in Statistical Physics* (Cambridge University Press, 2000).
2. R. C. O'Handley, *Modern Magnetic Materials*: *Principles and Applications* (Wiley, 2000).
3. A. Hubert and R. Schafer, *Magnetic domains* (Springer, 1998).
4. C. Boucher, R. S. Ellis, and B. Turkington, Derivation of maximum entropy principles in two-dimensional turbulence via large deviations. *J. Stat. Phy.* **98**(5/6):1235–1278 (2000).
5. J. Trashorras, Principes de grandes déviations pour des processus de "coarse graining". C. R. Acad. *Sci. Paris – Mathematique* Ser. I **336**:69–74 (2003).
6. J. L. Lebowitz, E. Orlandi, and E. Presutti, A particle model for spinodal decomposition, *J. Stat. Phys*. **63**(5/6):933–974 (1991).
7. A. DeMasi, E. Orlandi, E. Presutti, and L. Triolo, Glauber evolution with the Kac potentials. I. Mesoscopic and macroscopic limits, interface dynamics. *Nonlinearity* **7**(3):633–696 (1994).
8. M. A. Katsoulakis and P. E. Souganidis, Stochastic Ising models and anisotropic front propagation. *J. Stat. Phys*. **87**(1–2):63–89 (1997).
9. G. Giacomin and J. L. Lebowitz and E. Presutti, Deterministic and stochastic hydrodynamic equations arising from simple microscopic model systems in *Stochastic Partial Differential Equations*: *Six Perspectives*, Math. Surveys Monogr, 64, (Am Math. Soc., Providence, RI, 1999). pp. 107–152.
10. O. Lanford, Entropy and equilibrium states in classical statistical mechanics, *Statistical Mechanics and Mathematical problems*, volume 20 *Lecture Notes in Physics,* in A. Lenard ed. (Springer-Verlag, 1971), pp. 1–113.
11. J. T. Lewis, C. E. Pfister, and W. G. Sullivan. The equivalence of ensembles for lattice systems: some examples and a counterexample. *J. Stat. Phys*. **77**(1/2):397–419 (1994).
12. R. S. Ellis, K. Heaven, and Turkington, Large deviation principles and complete equivalence and nonequivalence results for pure and mixed ensembles. *J. Stat. Phys*. **101**(5/6):9999–1064 (2000).
13. P. Dupuis and R. S. Ellis, *A weak convergence approach to the theory of large deviations* (John Wiley & Sons Inc., New York, 1997). A Wiley-Interscience Publication.
14. T. Eisele and R. S. Ellis, Symmetry breaking and random waves for magnetic systems on a circle *Z. Wahrsch. Verw. Gebiete* **63**(3):297–348 (1983).
15. M. Cassandro, E. Orlandi, and E. Presutti, Interfaces and typical Gibbs configurations for one-dimensional Kac potentials. *Probab. Theory Relat. Fields* **96**(1):57–96 (1993).
16. A. Bovier and M. Zahradník, The low-temperature phase of Kac-Ising models. *J. Stat. Phys*. **87**(1–2):311–332 (1997).
17. M. Cassandro, E. Orlandi and P. Picco, Typical configurations for one-dimensional random field Kac model *Ann. Probab*. **27**(3):1414–1467 (1999).

18. C. Boucher, R. S. Ellis, and B. Turkington, Spatializing random measures: Doubly indexed processes and the large deviation principle. *Ann. Probab.* **27**(1):297–324 (1999).
19. E. M. Stein, *Singular Integrals and Differentiability Properties of Functions* (Princeton, 1970).
20. R. D. James and S. Müller, Internal variables and fine-scale oscillations in micromagnetics, *Contin. Mech. Thermodyn.* **6**(4):291–336 (1994).
21. M. Katsoulakis, P. Plecháč, J. Trashorras, and D. Tsagkarogiannis, Large deviations for lattice systems with singular interactions, in preparation.
22. R. Jordan, D. Kinderlehrer, and F. Otto. The variational formulation of the Fokker–Planck equation, *SIAM J. Math. Anal.* **29**(1):1–17 (electronic) (1998).
23. L. D. Landau and E. M. Lifschitz, On the theory of dispersion of magnetic permeability in ferromagnetic bodies. *Phys. Z. Sowjetunion* 8 (1935).
24. C. Kittel. Physical theory of ferromagnetic domains *Rev. Mod. Phys.* **21**. 541–583 (1949).
25. W. F. Brown, *Micromagnetics* (Wiley, New York, 1963).
26. Y. Brenier, Un algorithme rapide pour le calcul de transformees de Legendre–Fenchel discretes *C. R. Acad. Sci. Paris Serie I,* **308**:587–589 (1989).
27. A. De Masi and E. Presutti, *Mathematical Methods for Hydrodynamic Limits,* volume 1501 of *Lecture Notes in Mathematics* (Springer-Verlag, 1991).
28. P. H. Rabinowitz, Periodic and heteroclinic orbits for a periodic Hamiltonian system. *Ann. Inst. H. Poincaré Anal. Non Linéaire* **6**(5):331–346 (1989).
29. S. M. Auerbach, Theory and simulation of jump dynamics diffusion and phase equilibrium in nanopores. *Int. Rev. Phys. Chem.* **19**:155 (2000).
30. M. A. Katsoulakis, A. J. Majda and D. Vlachos, Coarse-grained stochastic processes and Monte Carlo simulations in lattice systems *J. Comp. Phys.* **186**:250–278 (2003).
31. B. Gidas, Metropolis-type Monte Carlo simulation algorithms and simulated annealing, in *Topics in Contemporary Probability and its Applications*, J. Laurie Snell, ed. (CRC Press, 1995).
32. G. Giacomin and J. L. Lebowitz, Phase segregation dynamics in particle systems with long range interactions. I. Macroscopic limits. *J. Stat. Phys.* **87**(1/2):37–61 (1997).
33. C. Kipnis and C. Landim, *Scaling Limits of Interacting Particle Systems* (Springer-Verlag, 1999).
34. A. N. Shiryaev, *Probability* (Springer, Graduate Textes in Mathematics, No 95, 1996).
35. C. Cercignani. *The Boltzmann Equation and Its Applications*, volume 67 *Applied Mathematical Sciences* (Springer-Verlag, 1988).
36. F. Comets and Th. Eisele, Asymptotic dynamics, non-critical and critical fluctuations for a geometric long-range interacting model. *Commun. Math. Phys.* **118**:531–567 (1988).
37. T. Kato *Perturbation Theory for Linear Operators* (Springer-Verlag 1984).
38. A. Arnold, P. Markowich, G. Toscani, and A. Unterreiter, On convex Sobolev inequalities and the rate of convergence to equilibrium for Fokker–Planck type equations. *Commun. Partial Diff. Equations* **26**(1–2):43–100 (2001).
39. D. W. Stroock and S. R. S. Varadhan, *Multidimensional Diffusion Processes* (Springer-Verlag, New York, 1979).
40. E. Cépa and D. Lépingle, Diffusing particles with electrostatic repulsion. *Prob. Theory Relat. Fields* **107**:429–449 (1997).
41. M. Katsoulakis, A. Majda, and D. Vlachos, Coarse-grained stochastic processes for microscopic lattice systems, *Proc Natl. Acad. Sci,* **100**(3):782 (2003).