

## STOCHASTIC HYDRODYNAMICAL LIMITS OF PARTICLE SYSTEMS\*

MARKOS A. KATSOULAKIS<sup>†</sup> AND ANDERS SZEPESSY<sup>‡</sup>

**Abstract.** Even small noise can have substantial influence on the dynamics of differential equations, e.g. for nucleation/coarsening and interface dynamics in phase transformations. The aim of this work is to establish accurate models for the noise in macroscopic differential equations, related to phase transformations/reactions, derived from more fundamental microscopic master equations.

For this purpose the mathematical paradigm of the dynamic Ising model is considered in the relatively tractable case of stochastic spin flip dynamics and long range spin/spin interactions. More specifically, this paper shows that localized spatial averages, with width  $\epsilon$ , of solutions to such Ising systems with long range interaction of range  $\mathcal{O}(1)$ , are approximated with error  $\mathcal{O}(\epsilon + (\gamma/\epsilon)^{2d})$  in distribution by a solution of an Ito stochastic differential equation, with drift as in the corresponding mean field model and a small diffusion coefficient of order  $(\gamma/\epsilon)^{d/2}$ , generating noise with spatial correlation length  $\epsilon$ , where  $\gamma$  is the distance between neighboring spin sites on a uniform periodic lattice in  $\mathbb{R}^d$ . To determine the correct noise is subtle in the sense that there are expected values, i.e. observables, that require different noise: the expected values that can be accurately approximated by the Einstein-diffusion and the expected values that need an alternative diffusion related to large deviation theory are identified; for instance dendrite dynamics up to a bounded time needs Einstein diffusion while transition rates need a different diffusion model related to invariant measures.

The elementary proofs use  $\mathcal{O}((\gamma/\epsilon)^{2d})$  consistency of the Kolmogorov-backward equations for the averaged spin and the stochastic differential equation and show that the long range interaction yields smoothing, which contributes with the  $\mathcal{O}(\epsilon)$  error. A new aspect of the derivation is that the error, based on residuals and weights, is computable and suitable for adaptive refinements and modeling.

**Key words.** hydrodynamical limit, stochastic coarse-graining, mean-field approximation, master equation, Ising model, large deviation theory, kinetic Monte Carlo method, fluctuation, dendrite, invariant measure

**AMS subject classifications.** 65C30, 82C20, 60K35

### 1. Introduction to hydrodynamical limits and spin systems

What is the right noise in an Allen-Cahn/Ginzburg-Landau equation

$$\partial_t m = \Delta m - V'(m) + \text{noise?} \quad (1.1)$$

This stochastic partial differential equation is a model used for the description of dynamic phase transformations in materials science. Noise is important e.g. for nucleation/coarsening effects of clusters [19, 20] [8], reaction rates [22], and dendrite dynamics in phase transformations [24, 27]. The aim of this work is to derive macroscopic continuum models with noise, i.e. stochastic differential equations (SDEs), from more fundamental master equation models given on smaller (nano or micro) scales. A motivation for continuum models is their advantage to be more computational efficient on larger scales; an additional purpose here is to introduce tools

---

\*Received: December 6, 2005; accepted (in revised version): May 19, 2006. Communicated by Weinan E.

The research of M.K. is partially supported by the U.S. Department of Energy through the grant DE-FG02-05ER25702 and the National Science Foundation through the grants NSF-DMS-0413864 and NSF-ITR-0219211. The work of A. S. is supported by the Swedish Research Council grants 2002-6285 and 2002-4961, the Swedish Foundation for Strategic Research and the European network HYKE, funded by the EC as contract HPRN-CT-2002-00282.

<sup>†</sup>Department of Mathematics and Statistics, University of Massachusetts, Amherst MA 01003, USA (markos@math.umass.edu).

<sup>‡</sup>Matematiska Institutionen, Kungl. Tekniska Högskolan, S-100 44 Stockholm, Sweden (szepessy@math.kth.se).

for adaptive coarse-graining in computations, coupling microscopic and macroscopic models. It turns out that the SDE-noise model depends on the type of expected value one wants to approximate: an application of our derivation suggests that the SDE approximation for lattice gas dendrite dynamics, up to a bounded time, is related to the Einstein diffusion; while the SDE-model to compute long transition times, e.g. for reaction rates, instead is based on invariant measures for processes in equilibrium. Our work also finds that such noise has small positive correlation length, which implies well posedness for several stochastic partial differential equations, cf. [3], as compared to the often ad hoc suggested space-time white noise model in partial differential equations, which lacks mathematical well posedness theory for nonlinear problems in higher than one space dimension.

One of our primary motivations is the series of papers [4, 5, 6], by De Masi, Orlandi, Presutti and Triolo, on a stochastic Ising model, with long range interaction on length scales of order one and Glauber flip rates [16]; they prove that spacial averages  $\bar{\sigma}$ , of the spin variable on distance of order one, converge in distribution to the mean field solution,  $m$ , of a deterministic Allen-Cahn related equation

$$\partial_t m = -m + \tanh(\beta(J_0 * m + h)),$$

as the distance of neighboring spin sites  $\gamma$  tends to zero, in a uniform lattice in  $\mathbb{R}^d$ . Their proof is a kind of law of large numbers. Here  $J_0: \mathbb{R}^d \rightarrow \mathbb{R}$  is the interaction potential,  $\beta$  is the inverse temperature and  $*$  denotes the standard convolution in  $L^2(\mathbb{R}^d)$ . They also prove a central limit theorem: the scaled deviation  $\gamma^{-d/2}(\bar{\sigma} - m)$  tends to an Ornstein-Uhlenbeck process with space-time (Einstein) white noise, linearized around the mean field solution. The work [15] provides an inspiring overview on asymptotic limits of stochastic interacting particle models with deterministic and stochastic differential equations in different scalings; for example, the central limit result can be scaled differently to include non-linear effects. By writing the flux in the mean field ordinary differential equation as

$$-m + \tanh(\beta J_0 * m) = \left(-m + J_0 * m\right) + \left(\tanh(\beta J_0 * m) - J_0 * m\right),$$

we identify, on scales larger than the interaction length, a diffusion related first term and a second non-linear reaction term as in the partial differential equation (1.1), with  $J_0$  scaled such that  $J_0 * 1 = 1$ , see [39].

In the physical chemistry literature, [23] is a paper related this work, where the authors study stochastic lattice systems with long range interactions modeling the adsorption/ desorption and surface diffusion mechanisms on a catalytic surface. These micro mechanisms are essentially identical to the spin flip and spin exchange lattice dynamics studied in [4, 5, 6, 15], except for the choice of the spin flip/exchange rate functions, where a more appropriate (for surface processes modeling) Arrhenius rate is selected. In [23] the authors include random fluctuations in the mesoscopic mean field equations by a formal derivation of stochastic partial differential equations with space/time white noise which is of Einstein-diffusion type; their approach relies on direct derivation from the microscopic Ising systems by means of a small noise, or equivalently large system-size, expansion in the corresponding high dimensional Fokker-Planck equation. Another important background result for our work is [21], by Hanggi, Grabert, Talkner and Thomas who show in the case of small noise, by direct calculation for a 1D example with uniform interactions, that accurate SDE approximation of transition rates for a master equation requires different diffusion

than the Einstein-diffusion for jump process in [4, 5, 6]: instead of a small noise expansion in a Fokker-Planck equation, which results in the Einstein noise, the work [21] points out that correct approximation of the corresponding invariant measures is essential to obtain consistent transition rates in the small noise limit.

Our work focuses mainly on the stochastic Ising model in [4, 5, 6], with long range interaction and Glauber flip rates; we identify which expected values that can be accurately approximated by Einstein-diffusion and which expected values that need the approximation in [21]. We present in section three a variant of the combination of the law of large numbers and central limit theorem results in [4]; we show an elementary derivation of an approximating Einstein-SDE for spacial averages of the spin variable, with averaging width  $\epsilon$ , where  $\gamma \ll \epsilon \ll 1$ . The drift term in the approximating SDE coincides with the mean field drift and the diffusion coefficient of order  $\mathcal{O}((\gamma/\epsilon)^{d/2})$  generates noise with correlation length  $\epsilon$ . The main result is a simple proof that the solution to the SDE approximates these  $\epsilon$ -localized spacial averages of the spin, up to a bounded time, with error  $\mathcal{O}(\epsilon + (\gamma/\epsilon)^{2d})$ , while the mean field approximation has the larger error  $\mathcal{O}(\epsilon + (\gamma/\epsilon)^d)$ . The proof is based on three ideas:

- No law of large numbers is applied; instead the proof uses  $\mathcal{O}((\gamma/\epsilon)^{2d})$  consistency of the Kolmogorov-backward equations for the averaged spin variable and the SDE, following the work on adaptive methods for SDE in [42].
- the assumption of regular long range interaction yields smoothing and contributes with the  $\mathcal{O}(\epsilon)$  error.
- derivatives of the solution to the Kolmogorov backward equation are estimated by stochastic flows.

Similar expansions have been applied before; the problem of replacing a high rate of small jump process with a diffusion process is classical: the first derivation of an approximating diffusion equation from a master equation was by Einstein [9], continued by Kolmogorov in the mathematical foundation of diffusion processes [33], with numerous applications cf. [43]. Our approach follows along well-known ideas from real space renormalization methods [26] using localized averages, in contrast to the usual expansions in the global system size [43]; in comparison, the new aspect here is to have a mathematically precise a priori error estimate, based on a general error representation, that is also computable and suitable for adaptive refinements and modeling. Computable means that the error representation is based on residuals and weights which can be computed with similar amount of work as the solution itself, as the adaptive SDE method in [42].

A similar hierarchical approach that avoids hydrodynamic limits and law of large numbers-type asymptotics [32], was taken in [28] where the authors developed coarse-grained Monte Carlo methods by deriving an approximating stochastic process to the detailed microscopic Monte Carlo algorithm. On the other hand this type of renormalization for non equilibrium stochastic systems is related to and inspired by the work on hydrodynamic limits of interacting particle systems, where a PDE is obtained in the asymptotic limit of a locally averaged quantity, corresponding for instance to density and described probabilistically by an empirical measure [32]. A key estimate needed in hydrodynamic limits is to show that in a suitable rescaling the joint probability distribution function of the particle system is in local equilibrium, i.e. that the process is close to its invariant measure parameterized by the solution of a macroscopic PDE, in direct analogy to such ideas in kinetic theory, cf. [40], [45].

The alternative diffusion, needed for accurate simulation of transition rates, is obtained by a WKB expansion of the probability density, relevant for large deviation

of rare events and long time approximation: the diffusion is chosen so that the SDE invariant measure is consistent with the master equation invariant measure. We show in Section 5 that a WKB expansion of the expected exit time leads to consistent Hamilton–Jacobi equations for the SDE and the master equation.

Asymptotic Ito stochastic differential equations have also been derived from differential equations with stochastic parameters from a Markov process and well separated fast and slow time scales, using asymptotic expansions, see ([35]) which includes a master equation as an example. A related setting with deterministic differential equations with well separated fast and slow time scales gives asymptotic Ito stochastic differential equations when the fluxes averaged over the invariant measure for the fast dynamics vanish in the leading order term [44]; the invariant measure corresponds to the probability measure for the Markov process in [35].

Stochastic interacting particle models, such as the Ising model, are used not only to gain qualitative theoretical understanding of noise-driven physical phenomena like nucleation and coarsening of clusters, but also for quantitative prediction, e.g. of phase transformations and reactions in physical chemistry, materials science, biochemistry. The key ingredient of such Monte Carlo simulation is to determine order parameters, interaction potentials and rates, which is a demanding modeling task including quantum calculations or fitting to experimental data, cf. [46].

Kinetic Monte Carlo computations of interacting particle models are computationally expensive for large scale effects, cf. [25]. The systematic coarse-grained kinetic Monte Carlo method introduced in [28] and [29] offers more efficient approximations to larger systems by reducing the number of variables, while at the same time it gives rise to the asymptotically correct invariant measure, at least for potentials with long enough interactions, see [28]. In addition, rigorous approximation estimates on the loss of information in the transition from finer to coarser descriptions in a non-equilibrium setting were also obtained in [31]; however these estimates involve, as a measure of the loss of information, the specific relative entropy between the fine and coarse-grained PDFs, which is a computationally intractable quantity for high dimensional systems such as most interacting particle systems. In contrast, the approach presented in this paper is based on estimations of arbitrary observables (i.e. weak convergence estimates) which are computationally verifiable. The Ito stochastic differential equation developed here is a further step in model simplification: the stochastic differential equation is a small noise perturbation of the mean field equation and avoids therefore the small time scale in the coarse-grained Monte Carlo process due to its high rate of small jumps. The coarse-grained Monte Carlo method and its relation to stochastic differential equations is discussed in Section 3.3 and Section 4 on invariant measures. Further coarse-graining of the stochastic differential equation can be obtained by adaptive methods, see Remark 3.9.

Our presentation focuses on the stochastic Ising model with Glauber dynamics, however the proof uses nothing particular for this example and seems applicable to more general master equations with long range interaction potentials. Further work in this direction and on spin exchange dynamics will appear in a forthcoming publication.

The next two sections introduce the model, prove the main approximation result for projected averages and end with some remarks on dendrite dynamics, coarse-grained Monte Carlo methods, reflected diffusion and alternative approximations. Section 3.1 derives required estimates of the solution to the high dimensional Kolmogorov equation, and shows how to handle the complication that the identity matrix becomes a delta function in high dimension, so that the standard  $L^2$  analysis needs

to be modified. Section 4 presents a comparison between the two invariant measures for the coarse grained Monte Carlo method and the stochastic differential equation, for the Curie-Weiss model. Section 5 shows that computation of some exit time problems for mean transition times needs a different approximating stochastic differential equation. The Appendix partly extends the results on projected averages in Section 3 to centered averages.

**2. Stochastic interacting particle systems**

Define a periodic lattice  $\mathcal{L} := \gamma Z^d / Z^d$ , with neighboring sites on distance  $\gamma$ , and consider spin configurations  $\sigma : \mathcal{L} \times [0, T] \rightarrow \{-1, 1\}$  defined on this lattice. Introduce a stochastic spin system where the spin  $\sigma_t(x)$ , at site  $x \in \mathcal{L}$  and time  $t$ , will flip to  $-\sigma_t(x)$  with the rate  $c(x, \sigma_t(\cdot))dt$ , in the time interval  $(t, t + dt)$ , independent of possible flips at other sites, cf. [34]. Let  $\sigma^x$  denote the configuration of spins after a flip at  $x$  of state  $\sigma$ , i.e.

$$\sigma^x(y) = \begin{cases} \sigma(y) & y \neq x \\ -\sigma(x) & y = x, \end{cases}$$

the probability density  $P(\sigma, t)$  of finding the spin system in configuration  $\sigma \in \{-1, 1\}^{\mathcal{L}}$  at time  $t$  then solves the master equation

$$\frac{dP(\sigma, t)}{dt} = \sum_{x \in \mathcal{L}} (c(x, \sigma^x)P(\sigma^x, t) - c(x, \sigma)P(\sigma, t)), \tag{2.1}$$

where the gain term  $\sum_x c(x, \sigma^x)P(\sigma^x, t)$  is the probability of jumping to state  $\sigma$  at time  $t$  and the loss term  $\sum_x c(x, \sigma)P(\sigma, t)$  is the probability to leave state  $\sigma$ . Similar master equations are used for microscopic models of chemical reactions and phase transformations, cf. [43, 18], where lattice sites are occupied by different species of particles. For instance with two species the state space could be  $\{0, 1\} \times \{0, 1\}$  instead of  $\{-1, 1\}$  for the classical spin model above.

We want a spin system that has statistical mechanics relevance, which can be achieved e.g. by choosing the rate function  $c$  as follows. Consider the Hamiltonian

$$H(\sigma) = -\frac{1}{2} \sum_{x, y \in \mathcal{L}} J(x - y)\sigma(x)\sigma(y) - \sum_{x \in \mathcal{L}} h(x)\sigma(x)$$

$$J = \gamma^d J_0, \quad J_0(x) = 0 \quad \text{for } |x| \geq 1,$$

where the long range interaction potential,  $J_0 \in C^2(\mathbb{R}^d)$ , is compactly supported and the function  $h \in C^2(\mathbb{R}^d)$  is a given external field. Define the Glauber Markov process on  $\{-1, 1\}^{\mathcal{L}}$  with generator

$$\frac{d}{dt} E[f(\sigma_t) | \sigma] = Lf(\sigma) = \sum_{x \in \mathcal{L}} c(x, \sigma)(f(\sigma^x) - f(\sigma)) \tag{2.2}$$

for  $f : \{-1, 1\}^{\mathcal{L}} \rightarrow \mathbb{R}$  and the flip rate

$$\begin{aligned} c(x, \sigma) &= \frac{e^{-\beta U(x)\sigma(x)}}{e^{-\beta U(x)} + e^{\beta U(x)}} \\ &= \frac{1}{2} \left( 1 - \sigma(x) \tanh(\beta U(x)) \right), \\ U(x) &= h(x) + \sum_y J(x - y)\sigma(y) =: h(x) + J * \sigma(x), \end{aligned} \tag{2.3}$$

where  $\beta > 0$  is the inverse temperature. This flip rate has built in invariance of the Gibbs measure,  $e^{-\beta H(\sigma)} / \sum_{\sigma} e^{-\beta H(\sigma)}$ , since it satisfies the detailed balance

$$c(x, \sigma) e^{-\beta H(\sigma)} = c(x, \sigma^x) e^{-\beta H(\sigma^x)},$$

which implies that this Gibbs measure is a time independent (invariant) solution to (2.1). Having this invariant Gibbs measure implies that the model has statistical mechanics relevance, cf. [16],[34],[15]. For example in the neighborhood of  $x \in \mathcal{L}$ , where  $h$  and  $J^*(1, \dots, 1)$  are positive, the construction of the flip rate  $c$  makes the system favor phases with spins mostly equal to 1 as compared to phases with spins mostly equal to  $-1$ .

We will study two types of localized averages of  $\sigma$  on scale  $\epsilon$ : a projection average over fixed subdomains, and a centered average, respectively. The projection average is simpler to analyze, but yields a somewhat larger error estimate. In particular we will find approximations to expected values of such averages. The error analysis uses consistency with the backward equation

$$\partial_t \tilde{u} + L\tilde{u} = 0 \quad t < T, \quad \tilde{u}(\cdot, T) = g,$$

corresponding to the master equation (2.1) for expected values

$$\tilde{u}(\xi, t) := E[g(\sigma_T) \mid \sigma_t = \xi].$$

### 3. Projected averages

Define the coarse periodic lattice  $\bar{\mathcal{L}} := q\gamma Z^d / Z^d$  with neighboring sites on distance  $q\gamma =: \epsilon$ , where  $q$  is an even positive integer and  $q^d$  is the number of fine sites projected to a coarse site: the lattice points  $y \in \bar{\mathcal{L}}$  define the coarse cells

$$C_y = \{x \in \mathcal{L} : -q\gamma/2 \leq x_i - y_i < q\gamma/2\},$$

of  $q^d$  neighboring points in the fine lattice and the averaging operator

$$A_\epsilon(z, x) = \begin{cases} 1/q^d & \text{if } x \text{ and } z \text{ are in the same coarse cell } C_y \\ 0 & \text{if } x \text{ and } z \text{ are in different coarse cells.} \end{cases}$$

We will study the behavior of the localized projection averages

$$\bar{X}(\sigma; z) := \sum_{x \in \mathcal{L}} A_\epsilon(z, x) \sigma(x), \tag{3.1}$$

for  $z \in \bar{\mathcal{L}}$ . The coarse-grained average  $\bar{X}(\sigma; \cdot)$  can be interpreted as a function on the coarse lattice since the restriction of  $\bar{X}(\sigma; \cdot)$  to each coarse cell  $C_z$  is constant, i.e.  $\bar{X}(\sigma; \cdot) = \sum_{x \in C_z} \sigma(x) / q^d$ . We will sometimes omit the first argument  $\sigma$  in  $\bar{X}$ . Define the averaged interaction potential  $\bar{J} : \bar{\mathcal{L}} \rightarrow \mathbb{R}$ , of  $J$ ,

$$\bar{J}(y) := \sum_{z \in C_y} \gamma^d J_0(z) = \epsilon^d J_0(y') \quad \text{for some } y' \in C_y,$$

and the convolution on  $\bar{\mathcal{L}}$

$$\bar{J} * \bar{X}(x) = \sum_{y \in \bar{\mathcal{L}}} \bar{J}(x-y) \bar{X}(y) \quad \text{for all } x \in \bar{\mathcal{L}}.$$

We shall show that the average spin,  $\bar{X}$ , can be approximated by the solution,  $X : \bar{\mathcal{L}} \times [0, T] \times \Omega \rightarrow \mathbb{R}$ , to the Ito stochastic differential equation

$$dX_t(x) = a(X_t)(x)dt + b(X_t)(x)dW^x, \quad X_0 = \bar{X}_0, \tag{3.2}$$

with the drift,  $a : \mathbb{R}^{\bar{\mathcal{L}}} \rightarrow \mathbb{R}^{\bar{\mathcal{L}}}$ , and diffusion,  $b : \mathbb{R}^{\bar{\mathcal{L}}} \rightarrow \mathbb{R}^{\bar{\mathcal{L}}}$ , coefficients given by

$$\begin{aligned} a(X) &= -X + \tanh(\beta(\bar{J} * X + h)), \\ b(X)(x) &= \left(\frac{\gamma}{\epsilon}\right)^{d/2} \sqrt{|1 - X \tanh(\beta(\bar{J} * X + h))(x)|} \eta(X(x)), \\ \eta(r) &= \begin{cases} 1 & \text{for } x \in [-1, 1], \\ 0 & \text{for } x \in (-\infty, -\hat{r}) \cup (\hat{r}, \infty), \end{cases} \\ \hat{r} &:= \min(1 + e^{-2\beta(2|J_0|_{L^1} + \|h\|_{L^\infty})}, 3/2) \end{aligned} \tag{3.3}$$

and a Wiener process  $W : \bar{\mathcal{L}} \times [0, T] \times \Omega \rightarrow \mathbb{R}$  on a probability space  $(\Omega, P, \{\mathcal{F}_t\}_{t=0}^T)$ , with the set of outcomes  $\Omega$ , probability measure  $P$  and sigma algebra  $\mathcal{F}_t$  of events up to time  $t$ . Here  $W^x$  are independent one dimensional standard Brownian motions for  $x \in \bar{\mathcal{L}}$ , so that formally

$$\begin{aligned} E[dW_t^x] &= 0, \\ E[dW_s^x dW_t^y] &= 0 \text{ for } s \neq t, \\ E[dW_t^x dW_t^y] &= 0 \text{ for } x \neq y, \text{ and} \\ E[dW_t^x dW_t^x] &= dt. \end{aligned}$$

The  $C^\infty$  cut-off function  $\eta : \mathbb{R} \rightarrow [0, 1]$ , with compact support, is introduced to handle the complication that  $|X(x)|$  may be larger than 1, although  $|\bar{X}(x)|$  is not, so that  $1 - X \tanh(\beta(\bar{J} * X + h))(x)$  may be close to zero causing large values on derivatives of  $\sqrt{|1 - X \tanh(\beta(\bar{J} * X + h))(x)|}$ ; note that we have  $|\bar{X}(x)| \leq 1$  and consequently the cut-off  $\eta$  improves the approximation by switching off the noise before  $1 - X \tanh(\beta(\bar{J} * X + h))(x)$  becomes zero making  $b$  a  $C^\infty$  function, see Remark 3.4. An alternative with instantaneous reflection is presented in Remark 3.5.

The approximation uses the fact that the high dimensional value function  $u : \mathbb{R}^{\bar{\mathcal{L}}} \times [0, T] \rightarrow \mathbb{R}$  defined by

$$u(\xi, t) = E[g(X_T) \mid X_t = \xi] \tag{3.4}$$

solves a corresponding Kolmogorov backward equation, where the drift and diffusion coefficients in (3.3) are chosen to minimize the error  $E[g(\bar{X}_T)] - E[g(X_T)]$ , for some suitable chosen functions  $g : \mathbb{R}^{\bar{\mathcal{L}}} \rightarrow \mathbb{R}$ , where  $\bar{X}_T := \bar{X}(\sigma_T)$ . To define the Kolmogorov backward equation introduce the scalar products

$$\begin{aligned} w \cdot v &:= \sum_{y \in \bar{\mathcal{L}}} w_y v_y \quad \text{for } w, v \in \mathbb{R}^{\bar{\mathcal{L}}}, \\ w \cdot v &:= \sum_{x, y \in \bar{\mathcal{L}}} w_{xy} v_{xy} \quad \text{for } w, v \in \mathbb{R}^{\bar{\mathcal{L}}^2}, \\ w \cdot v &:= \sum_{x, y, z \in \bar{\mathcal{L}}} w_{xyz} v_{xyz} \quad \text{for } w, v \in \mathbb{R}^{\bar{\mathcal{L}}^3}. \end{aligned}$$

Then  $u$  satisfies the Kolmogorov backward equation

$$\partial_t u + a \cdot u' + D \cdot u'' = 0, \text{ for } t < T,$$

$$u(\cdot, T) = g,$$

with the diagonal diffusion matrix

$$D_{xy} = \begin{cases} (1 - X \tanh(\beta(J * X + h)))(x) \eta^2(X(x)) & y = x, \\ 0 & y \neq x, \end{cases}$$

and the first and second order derivatives  $u'(\xi, t) = \partial_\xi u(\xi, t) = (u'_1, u'_2, \dots, u'_{|\bar{\mathcal{L}}|})$  and  $u''(\xi, t) = (\partial_{x_y} u)$ . We write  $\partial g(\xi) / \partial \xi_x = \partial_x g = g'_x$  and similarly for higher order derivatives.

We consider expected values of three times differentiable functions  $g$  satisfying the bounds

$$\begin{aligned} \sup_{\xi \in \mathbb{R}^{\mathcal{L}}} |g'_i(\xi)| &= \mathcal{O}(\epsilon^d) \\ \sup_{\xi} |g''_{ii}(\xi)| &= \mathcal{O}(\epsilon^d), \quad \sup_{\xi} |g''_{ij}(\xi)| = \mathcal{O}(\epsilon^{2d}) \quad j \neq i \\ \sup_{\xi} |g'''_{iii}(\xi)| &= \mathcal{O}(\epsilon^d), \quad \sup_{\xi} |g'''_{ijj}(\xi)| = \mathcal{O}(\epsilon^{2d}) \quad j \neq i \\ \sup_{\xi} |g'''_{ijk}(\xi)| &= \mathcal{O}(\epsilon^{3d}) \quad j \neq i, k \neq i. \end{aligned} \tag{3.5}$$

This means that  $g$  measures global properties, related to thermodynamic observables. For instance, the potential energies  $\sum_{x,y \in \bar{\mathcal{L}}} \bar{J}(x-y) X(x) X(y) \epsilon^d$ ,  $\sum_{x \in \bar{\mathcal{L}}} h(x) X(x)$  and  $\sum_{x \in \bar{\mathcal{L}}} h(x) f(X(x))$  satisfy (3.5), for  $h(x) = \epsilon^d h_0(x)$  with  $h_0$  a continuous function on the periodic unit cube and  $f \in \mathcal{C}^3(\mathbb{R})$ .

Our main result is:

**THEOREM 3.1.** *Assume  $g: \mathbb{R}^{\bar{\mathcal{L}}} \rightarrow \mathbb{R}$  satisfies (3.5). At a fixed time  $T$ , the average spin,  $\bar{X}_T$ , can be approximated by the solution,  $X_T$ , to the Ito stochastic differential equation (3.2) with error*

$$E[g(\bar{X}_T)] - E[g(X_T)] = \mathcal{O}(\epsilon + (\gamma/\epsilon)^{2d}), \quad \text{as } \epsilon \text{ and } \gamma \text{ tend to zero.} \tag{3.6}$$

Note that  $a = 0$  gives  $\mathcal{O}(1)$  error, while  $b = 0$  gives  $\mathcal{O}((\gamma/\epsilon)^d)$  error so that  $b$  defined by (3.3) is justified for  $\gamma \ll \epsilon \ll \gamma^{2d/(2d+1)}$ , with  $T$  fixed. The stochastic differential equation (3.2) has  $\mathcal{C}^\infty$  coefficients, where perturbations of solutions may grow exponentially in time. In Sections 4 and 5 we study long time approximations using invariant measures.

*Proof.* The proof is in three steps: first to derive an error representation based on  $u(\xi, t) = E[g(X_T) \mid X_t = \xi]$  along  $\xi = \bar{X}_t$ , then to estimate the error using the long range interaction and finally to bound derivatives of  $u(\xi, t)$  in Lemma 3.2.



*Step 1.* The definitions of  $u$ , the generator (2.2) and the average (3.1) imply

$$\begin{aligned}
 E[g(\bar{X}_T)] - E[g(X_T)] &= E[u(\bar{X}_T, T)] - E[u(X_0, 0)] \\
 &= E\left[\int_0^T du(\bar{X}_t, t)\right] = \int_0^T E[Lu + \partial_t u] dt = \int_0^T E[E[Lu - a \cdot u' - D \cdot u'' \mid \bar{X}_t]] dt \\
 &= \int_0^T E\left[E\left[\sum_{x \in \mathcal{L}} c(x, \sigma_t) \left(u(\bar{X}(\sigma_t^x), t) - u(\bar{X}(\sigma_t), t)\right) \right. \right. \\
 &\quad \left. \left. - a \cdot u'(\bar{X}(\sigma_t), t) - D \cdot u''(\bar{X}(\sigma_t), t) \mid \bar{X}_t\right]\right] dt \\
 &= \int_0^T E\left[E\left[\sum_{x \in \mathcal{L}} c(x, \sigma_t) \left(u(\bar{X}_t - 2A_\epsilon(x, \cdot)\sigma_t(x), t) - u(\bar{X}_t, t)\right) \right. \right. \\
 &\quad \left. \left. - a \cdot u'(\bar{X}_t, t) - D \cdot u''(\bar{X}_t, t) \mid \bar{X}_t\right]\right] dt. \tag{3.7}
 \end{aligned}$$

The first step to estimate this error is to write the difference of  $u$  in terms of its derivatives by Taylor expansion, for some  $s \in [0, 1]$ ,

$$\begin{aligned}
 &u(\bar{X}(\sigma) - 2A_\epsilon(x, \cdot)\sigma(x), t) - u(\bar{X}(\sigma), t) \\
 &= -2u'(\bar{X}, t) \cdot A_\epsilon(x, \cdot)\sigma(x) \\
 &\quad + 2u''(\bar{X}, t) \cdot A_\epsilon(x, \cdot)A_\epsilon(x, \cdot)\sigma^2(x) \\
 &\quad - \frac{4}{3}u'''(\bar{X} - 2sA_\epsilon(x, \cdot)\sigma(x), t) \cdot A_\epsilon(x, \cdot)A_\epsilon(x, \cdot)A_\epsilon(x, \cdot)\sigma^3(x), \tag{3.8}
 \end{aligned}$$

so that the error representation (3.7) becomes

$$\begin{aligned}
 &E[g(\bar{X}_T)] - E[g(X_T)] \\
 &= \int_0^T E\left[E\left[u'(\bar{X}_t, t) \cdot \left(-2 \sum_{x \in \mathcal{L}} c(x, \sigma_t) A_\epsilon(x, \cdot)\sigma_t(x) - a\right) \right. \right. \\
 &\quad \left. \left. + u''(\bar{X}_t, t) \cdot \left(\sum_{x \in \mathcal{L}} 2c(x, \sigma_t) A_\epsilon(x, \cdot)A_\epsilon(x, \cdot)\sigma_t^2(x) - D\right) \right. \right. \\
 &\quad \left. \left. - \frac{4}{3} \sum_{x \in \mathcal{L}} u'''(\bar{X}_t - 2sA_\epsilon(x, \cdot)\sigma_t(x), t) \cdot c(x, \sigma_t) A_\epsilon(x, \cdot)A_\epsilon(x, \cdot)A_\epsilon(x, \cdot)\sigma_t^3(x) \mid \bar{X}_t\right]\right] dt.
 \end{aligned}$$

We note that the matrix  $F_{yz} := -2 \sum_{x \in \mathcal{L}} c(x, \sigma) A_\epsilon(x, y)A_\epsilon(x, z) - D_{yz}$  is diagonal (i.e.  $F_{yz} = 0$  for  $y \neq z$ ) and the tensor  $R_{yzv} := c(x, \sigma) A_\epsilon(x, y)A_\epsilon(x, z)A_\epsilon(x, v)\sigma^3(x)$  is also diagonal (i.e.  $R_{yzv} = 0$  if not  $y = z = v$ ). Therefore the error representation reduces to

$$\begin{aligned}
 &E[g(\bar{X}_T)] - E[g(X_T)] \\
 &= \int_0^T E\left[E\left[\sum_{y \in \bar{\mathcal{L}}} u'_y(\bar{X}_t, t) \left(-2 \sum_{x \in \mathcal{L}} c(x, \sigma_t) A_\epsilon(x, y)\sigma_t(x) - a(y)\right) \right. \right. \\
 &\quad \left. \left. + \sum_{y \in \bar{\mathcal{L}}} u''_{yy}(\bar{X}_t, t) F_{yy} - \frac{4}{3} \sum_{x \in \mathcal{L}} \sum_{y \in \bar{\mathcal{L}}} u'''_{yyy}(\bar{X}_t - 2sA_\epsilon(x, y)\sigma_t(x), t) R_{yyy}(x) \mid \bar{X}_t\right]\right] dt. \tag{3.9}
 \end{aligned}$$

*Step 2.* The next step is to determine the optimal  $a$  and  $b$  which minimize the error (3.9). For this purpose we shall in the flipping rate approximate the coupling

$J * \sigma$  by  $\bar{J} * \bar{X}$ , using the long range  $\mathcal{O}(1)$  interaction distance of  $J$ . We have

$$\begin{aligned} \sum_x A_\epsilon(x, y) &= 1, \\ \sum_x A_\epsilon(x, y) A_\epsilon(x, y) &= (\gamma/\epsilon)^d, \\ \sum_x A_\epsilon(x, y) A_\epsilon(x, y) A_\epsilon(x, y) &= (\gamma/\epsilon)^{2d}. \end{aligned} \tag{3.10}$$

The definition of the average (3.1) implies

$$J * \bar{X} = \sum_{z, y \in \mathcal{L}} J(\cdot - y) A_\epsilon(y, z) \sigma(z)$$

and consequently the coupling has the uniform error estimate

$$\begin{aligned} J * \sigma(x) - J * \bar{X}(x) &= \sum_{z \in \mathcal{L}} (J(x - z) - \sum_{y \in \mathcal{L}} J(x - y) A_\epsilon(y, z)) \sigma(z) \\ &= \sum_{y, z \in \mathcal{L}} (J(x - z) - J(x - y)) A_\epsilon(y, z) \sigma(z) = \mathcal{O}(\epsilon), \end{aligned} \tag{3.11}$$

since, for any of the  $\gamma^{-d}$  points  $z$ , either  $A_\epsilon(y, z) = 0$  or  $A_\epsilon(y, z) = (\gamma/\epsilon)^d$  and

$$|J(x - z) - J(x - y)| = \mathcal{O}(\epsilon) \gamma^d$$

for  $(\epsilon/\gamma)^d$  points  $y \in C_z$ . The definition of  $\bar{J}$  shows

$$J * \bar{X}(x) = \sum_{y \in \bar{\mathcal{L}}} \sum_{z \in C_y} \gamma^d J_0(x - z) \bar{X}(y) = \bar{J} * \bar{X}(x)$$

and consequently

$$\sup_{x \in \mathcal{L}} |J * \sigma(x) - \bar{J} * \bar{X}(x)| = \mathcal{O}(\epsilon). \tag{3.12}$$

This error estimate and the flip rate (2.3) imply

$$\begin{aligned} - \sum_{x \in \mathcal{L}} 2c(x, \sigma) A_\epsilon(x, \cdot) \sigma(x) &= -\bar{X} + A_\epsilon \cdot \tanh(\beta(J * \sigma + h)) \\ &= -\bar{X} + \tanh(\beta(\bar{J} * \bar{X} + h)) + \mathcal{O}(\epsilon), \end{aligned} \tag{3.13}$$

and

$$\begin{aligned} \sum_{x \in \mathcal{L}} 2c(x, \sigma) A_\epsilon(x, \cdot) A_\epsilon(x, \cdot) \sigma^2(x) &= \left(\frac{\gamma}{\epsilon}\right)^d \left[1 - \bar{X} \tanh(\beta(\bar{J} * \bar{X} + h))\right] \\ &\quad + \mathcal{O}((\gamma/\epsilon)^{2d} + \epsilon^2). \end{aligned} \tag{3.14}$$

We have  $\bar{X}_t \in [-1, 1]^{\bar{\mathcal{L}}}$ , therefore we need estimates of the derivatives of  $u$  in this set. Lemma 3.2 in Section 3.1 shows that

$$\sup_{\xi \in [-1, 1]^{\bar{\mathcal{L}}}} \sum_{x \in \bar{\mathcal{L}}} |u'_x(\xi, t)| + \sup_{\xi \in [-1, 1]^{\bar{\mathcal{L}}}} \sum_{x \in \bar{\mathcal{L}}} |u''_{xx}(\xi, t)| + \sup_{\xi \in [-1, 1]^{\bar{\mathcal{L}}}} \sum_{x \in \bar{\mathcal{L}}} |u'''_{xxx}(\xi, t)| = \mathcal{O}(1), \tag{3.15}$$

as  $\epsilon, \gamma \rightarrow 0+$ , which together with the expansions (3.9), (3.10), (3.13) and (3.14) proves the theorem.  $\square$

We also have

LEMMA 3.1. Assume  $g: \mathbb{R}^{\bar{\mathcal{L}}} \rightarrow \mathbb{R}$  satisfies (3.5) and that the initial spin  $\sigma_0$  has expected value  $m$ , where  $\sigma_0(x) - m_x$  are i.i.d. with bounded variance and second order difference quotients  $|d^2m/dx^2| = \mathcal{O}(1)$ . Then the deterministic mean field solution,  $\hat{X}: \mathbb{R}^{\bar{\mathcal{L}}} \times [0, T] \rightarrow \mathbb{R}$ ,

$$d\hat{X}/dt = -\hat{X} + \tanh(\beta(\bar{J} * \hat{X} + h - J(0)\hat{X})), \quad \hat{X}_0 = E[\bar{X}_0],$$

depends on  $\epsilon$  only through the initial data and satisfies

$$E[g(\bar{X}_T)] - E[g(\hat{X}_T)] = \mathcal{O}(\epsilon + (\gamma/\epsilon)^d)$$

provided the drift  $a$  is defined by (3.3).

*Proof.* Think of  $\hat{X}$  as an  $X$  with  $b=0$  and apply the corresponding expansion (3.7), (3.8) and (3.10). Then it remains to verify that the initial data satisfy

$$E[u(\bar{X}_0, 0) - u(\hat{X}_0, 0)] = \mathcal{O}((\gamma/\epsilon)^d),$$

but this is a direct consequence of the central limit theorem and the initial  $\sigma_0 - E[\sigma_0]$  being i.i.d. with bounded variance.  $\square$

The following variant of error estimates is a direct consequence of the proof of Theorem 3.1, where  $\|w\|_{\ell^1} := \sum_{x \in \mathcal{L}^n} |w(x)|$ .

COROLLARY 3.2. Suppose the first to the third order derivatives of  $u$  are uniformly bounded in  $\ell^1$ , independent of  $T$ , up to time  $T$ , then

$$\frac{1}{T} \int_0^T E[G(\bar{X}_t, t)] dt - \frac{1}{T} \int_0^T E[G(X_t, t)] dt = \mathcal{O}(\epsilon + (\gamma/\epsilon)^{2d}) \tag{3.16}$$

with the value function  $u(\xi, t) = E[\int_t^T G(X_t, t) dt \mid X_t = \xi]$  satisfying

$$\partial_t u + a \cdot u' + D \cdot u'' = -G, \quad t \leq T,$$

$$u(\cdot, T) = 0.$$

When the processes  $\bar{X}$  and  $X$  are ergodic and  $T \rightarrow \infty$ , the expected value on the left hand side of (3.16) can be removed.

**3.1. Estimates of the value function.** This section proves the bound (3.15) of the value function  $u$ , namely:

LEMMA 3.2. Assume  $g: \mathbb{R}^{\bar{\mathcal{L}}} \rightarrow \mathbb{R}$  satisfies (3.5), then the function  $u$ , defined by (3.4), satisfies

$$\sup_{\xi \in [-1, 1]^{\bar{\mathcal{L}}}} \sum_{x \in \bar{\mathcal{L}}} |u'_x(\xi, t)| + \sup_{\xi \in [-1, 1]^{\bar{\mathcal{L}}}} \sum_{x \in \bar{\mathcal{L}}} |u''_{xx}(\xi, t)| + \sup_{\xi \in [-1, 1]^{\bar{\mathcal{L}}}} \sum_{x \in \bar{\mathcal{L}}} |u'''_{xxx}(\xi, t)| = \mathcal{O}(1),$$

as  $\epsilon, \gamma \rightarrow 0+$ .

*Proof.* The proof is divided into three steps: to find representation of  $u', u'', u'''$  in terms of variations of the process  $X$ ; to estimate the first variation  $X'$  by splitting into diagonal and non diagonal contributions; and to combine the representations and the  $X'$  estimate to derive bounds on the  $\ell^1$  norms of  $u', u''$  and  $u'''$ .

*Step 1.* The proof uses variations up to order three of the process  $X$ , defined by (3.2). To define these variations consider a process  $Y$  depending on  $X$ . Define the first variation of  $Y(t)$  with respect to a perturbation in the initial location of the solution  $X$ , to (3.2) at time  $s < t$ , by the gradient

$$Y'(t; s) = \partial_\xi(Y(t); X_s = \xi).$$

Then, with  $Y = X$  and fixed  $s$ , the first variation  $X'(\cdot; s) : [s, T] \rightarrow \mathbb{R}^{\mathcal{L}^2}$  solves the Ito stochastic differential equation

$$\begin{aligned} dX'_{xy}(t) &= \sum_{z \in \mathcal{L}} X'_{zy}(t) \left( \partial_{X_z} a_x(X(t)) dt + \partial_{X_z} b_x(X(t)) dW_t^x \right), \quad t > s, \\ X'_{xy}(s) &= \delta(x - y) := \begin{cases} 0 & x \neq y, \\ 1 & x = y. \end{cases} \end{aligned} \tag{3.17}$$

We have

$$u'_x(X_t, t) = E[X'_{rx}(T; t) g'_r(X(T)) \mid X_t], \tag{3.18}$$

where we use the summation convention to sum over the index set when an index appears in different factors of a product

$$X'_{zy} g'_z := \sum_{z \in \mathcal{L}} X'_{zy} g'_z.$$

Let  $a'_{xz} = \partial_{X_z} a_x$  denote the Jacobian of  $a$  and define the matrix

$$A_{ij} := \int_s^t b'_{ij} dW^i + \int_s^t \left( a'_{ij} - \frac{1}{2} b_{ij}^2 \right) dt;$$

note that there is no sum over  $i$  in the  $dW^i$  measure here and below. Ito's formula implies that the matrix

$$X'(t, s) = e^A := I + \sum_{n=1}^{\infty} \frac{1}{n!} A^n$$

solves (3.17).

Differentiation of (3.17) shows that the second variation  $X'' := \partial_{\xi\xi}(X_T; X_s = \xi)$ , defined to be the first variation of  $Y = X'$ , is

$$X''_{ijj}(T, t) = \int_t^T X'_{kj}(s, t) \left( a''_{ikm} ds - b'_{ik} b''_{ikm} ds + b''_{ikm}(X_s) dW_s^i \right) X'_{mj}(T, t)$$

and

$$u''_{ii}(X_t, t) = E[X''_{kii}(T, t) g'_k(X(T)) + X'_{ki}(T, t) X'_{mi}(T, t) g''_{km}(X(T)) \mid X_t]. \tag{3.19}$$

Similarly  $u'''$  is determined by the third variation of  $X$ , which is

$$\begin{aligned} X'''_{ijjj}(T, t) &= \int_t^T X'_{kj}(s, t) \left( a''_{ikm} ds - b'_{ik} b''_{ikm} ds + b''_{ikm}(X_s) dW_s^i \right) X''_{mjj}(T, t) \\ &+ \int_t^T X''_{kjj}(s, t) \left( a''_{ikm} ds - b'_{ik} b''_{ikm} ds + b''_{ikm}(X_s) dW_s^i \right) X'_{mj}(T, t) \\ &+ \int_t^T X'_{kj}(s, t) X'_{rj}(s, t) \left( (a'''_{ikmr} - b''_{ikr} b''_{ikm} - b'_{ik} b'''_{ikmr}) ds + b'''_{ikmr}(X_s) dW_s^i \right) X'_{mj}(T, t) \end{aligned}$$

and

$$\begin{aligned} u'''_{iii}(X_t, t) &= E \left[ X'''_{kiii}(T, t) g'_k(X(t)) + 3X''_{kii}(T, t) X'_{mi}(T, t) g''_{km}(X(t)) \right. \\ &\quad \left. + X'_{ki}(T, t) X'_{mi}(T, t) X'_{ri}(T, t) g'''_{kmr}(X(T)) \mid X_t \right]. \end{aligned}$$

*Step 2.* The standard estimates of Ito differential equations in  $\mathbb{R}^n$ , with smooth uniformly bounded coefficients, use the Frobenius norm to estimate the matrix norm of  $X'$ . The fact that  $X'(s; s) = \delta$  has unbounded Frobenius norm

$$\sqrt{\sum_{(x,y) \in \mathcal{L}^2} (X'_{xy})^2} = \mathcal{O}(\epsilon^{-d/2})$$

as  $\epsilon \rightarrow 0+$ , shows that the standard mean square proof needs to be modified; with the aim to estimate  $X'$  we split  $A$  into its diagonal and non diagonal parts

$$D_{ij} = \begin{cases} 0 & i \neq j, \\ A_{ii} & i = j, \end{cases} \quad N_{ij} = \begin{cases} A_{ij} & i \neq j, \\ 0 & i = j, \end{cases}$$

so that

$$X' = e^N e^D = \underbrace{(e^N - I)}_{=: M} e^D + e^D \tag{3.20}$$

where  $(e^D)_{ij}$  is the diagonal matrix with  $d_i$  in the diagonal. Since the matrix elements  $|a'_{ij}|$  and  $|b'_{ij}|$  are uniformly bounded, there is for each  $p \geq 2$  a constant  $C$  such that the diagonal elements can be bounded

$$\begin{aligned} E[d_i^p] &= E \left[ \exp \left( \int_t^T p a'_{ii} - \left( \frac{p}{2} - \frac{p^2}{2} \right) b_{ii}^{\prime 2} dt \right) \exp \left( - \int_t^T \frac{p^2}{2} b_{ii}^{\prime 2} dt + \int_t^T p b'_{ii} dW^i \right) \right] \\ &\leq CE \left[ \exp \left( - \int_t^T \frac{p^2}{2} b_{ii}^{\prime 2} dt + \int_t^T p b'_{ii} dW^i \right) \right] \\ &= CE \left[ \exp \left( - \int_t^t \frac{p^2}{2} b_{ii}^{\prime 2} dt + \int_t^t p b'_{ii} dW^i \right) \right] = C, \end{aligned} \tag{3.21}$$

where the two last equalities are a consequence of  $\exp(-\int_t^T \frac{p^2}{2} b_{ii}^{\prime 2} dt + \int_t^T p b'_{ii} dW^i)$  being a martingale, as a function of  $T$ .

To estimate the non diagonal part  $M$  of  $X'$ , we will use standard  $L^p$  estimates applied to the matrix  $M$  for even positive integers  $p$ . Introduce the notation

$$\tilde{a}'_{ij} = \begin{cases} a_{ij} & i \neq j, \\ 0 & i = j, \end{cases} \quad \tilde{b}'_{ij} = \begin{cases} b'_{ij} & i \neq j, \\ 0 & i = j. \end{cases}$$

We have  $M + I = \exp(N)$ , therefore Ito's formula yields

$$dM_{ij} = M_{kj}(\tilde{a}'_{ik}dt + \tilde{b}'_{ik}dW^i) + \tilde{a}'_{ij}dt + \tilde{b}'_{ij}dW^i$$

so that

$$\begin{aligned} \sum_{ij} E[M_{ij}^p] &= \sum_{ij} E[pM_{ij}^{p-1}dM_{ij}] + \sum_{ij} E[\frac{p(p-1)}{2}M_{ij}^{p-2}dM_{ij}dM_{ij}] \\ &= p \sum_{ij} E[M_{ij}^{p-1}M_{kj}\tilde{a}'_{ik}]dt + p \sum_{ij} E[M_{ij}^{p-1}\tilde{a}'_{ij}]dt \\ &\quad + \frac{p(p-1)}{2} \sum_{ij} E[M_{ij}^{p-2}(M_{kj}\tilde{b}'_{ik} + \tilde{b}'_{ij})^2]dt. \end{aligned} \quad (3.22)$$

The next step is to use Hölder estimates for the three sums on the right hand side. Let  $q = p/(p-1)$ , then

$$|\sum_{ij} M_{ij}^{p-1}\tilde{a}'_{ij}| \leq (\sum_{ij} M_{ij}^p)^{1/q} (\sum_{ij} (\tilde{a}'_{ij})^p)^{1/p} \leq \sum_{ij} M_{ij}^p + \sum_{ij} (\tilde{a}'_{ij})^p.$$

Similarly

$$\begin{aligned} |\sum_{ijk} M_{ij}^{p-1}M_{kj}\tilde{a}'_{ik}| &\leq \sum_{ij} M_{ij}^{p-1} (\sum_k M_{kj}^p)^{1/p} (\sum_k |\tilde{a}'_{ik}|^q)^{1/q} \\ &\leq \sum_i (\sum_j M_{ij}^p)^{1/q} (\sum_{kj} M_{kj}^p)^{1/p} (\sum_k |\tilde{a}'_{ik}|^q)^{1/q} \\ &\leq (\sum_{ij} M_{ij}^p)^{1/p+1/q} \left( \sum_i (\sum_k |\tilde{a}'_{ik}|^q)^{p/q} \right)^{1/p}. \end{aligned}$$

The last sum in (3.22) we split into two

$$\frac{1}{2}(M_{kj}\tilde{b}'_{ik} + \tilde{b}'_{ij})^2 \leq (M_{kj}\tilde{b}'_{ik})^2 + (\tilde{b}'_{ij})^2$$

and let  $r = p/(p-2)$  to obtain

$$\begin{aligned} |\sum_{ij} M_{ij}^{(p-2)}(\tilde{b}'_{ij})^2| &\leq \sum_i (\sum_j M_{ij}^p)^{1/r} (\sum_j (\tilde{b}'_{ij})^p)^{2/p} \\ &\leq (\sum_{ij} M_{ij}^p)^{1/r} (\sum_{ij} (\tilde{b}'_{ij})^p)^{2/p} \\ &\leq \sum_{ij} M_{ij}^p + \sum_{ij} (\tilde{b}'_{ij})^p, \end{aligned}$$

and finally

$$\begin{aligned} \sum_{ij} M_{ij}^{p-2}(M_{kj}\tilde{b}'_{ik})^2 &\leq \sum_{ij} M_{ij}^{p-2} (\sum_k M_{kj}^p)^{2/p} (\sum_k |\tilde{b}'_{ik}|^q)^{2/q} \\ &\leq \sum_i (\sum_j M_{ij}^p)^{1/r} (\sum_{kj} M_{kj}^p)^{2/p} (\sum_k |\tilde{b}'_{ik}|^q)^{2/q} \\ &\leq (\sum_{ij} M_{ij}^p)^{1/r+2/p} \left( \sum_i (\sum_k |\tilde{b}'_{ik}|^q)^{p/q} \right)^{2/p}. \end{aligned}$$

The reason we split  $X'$  into diagonal and non-diagonal parts is that the non diagonal part has small matrix elements

$$\sup_{X,ij} |\tilde{a}'_{ij}(X)| + \sup_{X,ij} |\tilde{b}'_{ij}(X)| = \mathcal{O}(\epsilon^d)$$

due to the long range interaction, i.e.  $\partial_{X_i} \bar{J} * X = \mathcal{O}(\epsilon^d)$ , and consequently, as  $\epsilon \rightarrow 0+$ , for  $p \geq 2$

$$\begin{aligned} \sup_X \sum_{ij} (\tilde{a}'_{ij}(X))^p + \sup_X \sum_{ij} (\tilde{b}'_{ij}(X))^p &= \mathcal{O}(\epsilon^{(p-2)d}), \\ \sup_X \left( \sum_i \left( \sum_k |\tilde{a}'_{ik}(X)|^q \right)^{p/q} \right)^{1/p} + \sup_X \left( \sum_i \left( \sum_k |\tilde{b}'_{ik}(X)|^q \right)^{p/q} \right)^{2/p} &= \mathcal{O}(1). \end{aligned} \tag{3.23}$$

The combination of (3.22), (3.23),  $1/q + 1/p = 1$ ,  $1/r + 2/p = 1$  and the Hölder estimates show that there is a constant  $C$  such that

$$dE[\sum_{ij} M_{ij}^p] \leq C(E[\sum_{ij} M_{ij}^p] + \epsilon^{(p-2)d})dt$$

so that Grönwall's lemma and the initial condition  $M(s, s) = 0$  imply

$$E[\sum_{ij} M_{ij}^p] = \mathcal{O}(\epsilon^{(p-2)d}). \tag{3.24}$$

*Step 3.* The final step is to use the representations (3.18) and (3.20) combined with the estimates (3.21) and (3.24) of  $X'$ . The estimate of  $u'$  becomes

$$u'_i(x, t) = E[X'_{ki}g'_k \mid X_t = x] = E[(M + I)_{ki}d_k g'_k \mid X_t = x]$$

so that, with  $E[\cdot]$  denoting  $E[\cdot \mid X_t = x]$  and  $C$  a positive constant (not necessary the same at each occurrence) the splitting into diagonal and non diagonal terms of  $X'$  leads to the claimed  $\ell^1$  bound on  $u'$

$$\begin{aligned} \sum_i |u'_i| &\leq \sum_i \left( \sqrt{E[d_i^2]} \sqrt{E[(g'_i)^2]} + \sqrt{E[d_k^2(g'_k)^2]} \sqrt{E[M_{ki}^2]} \right) \\ &\leq C \sum_i \epsilon^d + C \epsilon^d \sum_{ik} \sqrt{E[d_k^2]} \sqrt{E[M_{ki}^2]} \\ &\leq C \left( 1 + \epsilon^d \sum_{ik} \sqrt{E[M_{ki}^2]} \right) \\ &\leq C \left( 1 + \sqrt{\sum_{ik} E[M_{ki}^2]} \right) \leq C. \end{aligned}$$

As for the Jacobian of  $a$  and  $b$  we see that each additional derivative in a non diagonal component gain a factor  $\epsilon^d$

$$\begin{aligned} b''_{iii} &= \mathcal{O}(1) \\ b''_{ijj} &= b''_{jji} = \mathcal{O}(\epsilon^d), \quad j \neq i \\ b''_{ijk} &= \mathcal{O}(\epsilon^{2d}), \quad j \neq i, k \neq i \end{aligned}$$

and similarly for  $a''$

$$\begin{aligned} a''_{iii} &= \mathcal{O}(1) \\ a''_{ijj} &= a''_{iji} = \mathcal{O}(\epsilon^d), \quad j \neq i \\ a''_{ijk} &= \mathcal{O}(\epsilon^{2d}), \quad j \neq i, k \neq i. \end{aligned}$$

The estimate of  $u''$  has several terms, let us start with the term (and no summation over  $k$  and  $j$  here)

$$\begin{aligned} &|E[\int_0^T X'_{kj} b''_{ikm} dW^i X'_{mj} g'_i]| \\ &\leq (\int_0^T E[(X'_{kj})^2 (b''_{ikm})^2] dt)^{1/2} (E[(X'_{mj})^2 (g'_i)^2])^{1/2}. \end{aligned}$$

It turns out that a main contribution to the  $\ell^1$  norm of  $u''$  comes from the diagonal term  $b''_{iii}$

$$\begin{aligned} &\sum_{ij} (\int_0^T E[(X'_{ij})^2 (b''_{iii})^2] dt)^{1/2} (E[(X'_{ij})^2 (g'_i)^2])^{1/2} \\ &\leq C \epsilon^d \sum_{ij} \sup_{s \leq T} E[(X'_{ij}(s, t))^2] \\ &\leq C \epsilon^d \sum_i E[d_i^2] + C \epsilon^d \sum_{ij} E[M_{ij}^2 d_i^2] \\ &\leq C + C \epsilon^d \sum_{ij} (E[M_{ij}^4])^{1/2} (E[d_i^4])^{1/2} \\ &\leq C + C (\sum_{ij} E[M_{ij}^4])^{1/2} \leq C. \end{aligned} \tag{3.25}$$

The non diagonal terms  $b''_{ijk}$  similarly give the same bound, where the sum over more terms is compensated by additional powers of  $\epsilon^d$  in  $b''$ . The other type of term  $X'_{kj} X'_{mj} g''_{km}$  in (3.19) also yields the contribution

$$\sum_{jk} \epsilon^d E[(X'_{kj})^2] + \sum_{jkm} \epsilon^{2d} E[|X'_{kj} X'_{mj}|]$$

which is bounded and we can conclude that the  $\ell^1$  norm of  $u''$  is bounded.

The proof of the estimate of  $E[M_{ij}^p]$  applied to an Ito integral  $\int_0^T f_t dW$ , i.e. replacing  $M_{ij}$  by  $\int_0^T f_t dW$ , shows that

$$E[(\int_0^T f_t dW)^p] \leq p(p-1) e^{Tp(p-1)/2} \int_0^T E[f_t^p] dt,$$

which we will use to estimate  $u'''$ . The estimate of  $u'''$  has several terms; we start



with  $E[\int_0^T X'_{kj} b''_{ikm} dW^i X''_{mjj} g'_i]$  which has a main term

$$\begin{aligned} & \sum_{k,m,r,n} |E[\int_0^T X'_{kj} b''_{ikm} dW^i \int_0^T X'_{rj} b''_{mrn} dW^r X'_{nj} g'_i]| \\ & \leq \sum_{k,m,r,n} C (\int_0^T E[(X'_{kj})^2 (b''_{ikm})^2] dt)^{1/2} \\ & \quad \times (\int_0^T E[(X'_{rj})^4 (b''_{mrn})^4] dt)^{1/4} (E[(X'_{nj})^4 (g'_i)^4])^{1/4}. \end{aligned}$$

As for  $u''$ , the diagonal terms yield

$$\begin{aligned} & |E[\int_0^T X'_{ij} b''_{iii} dW^i \int_0^T X'_{ij} b''_{iii} dW^i X'_{ij} g'_i]| \\ & \leq C \epsilon^d (\int_0^T E[(X'_{ij})^2] dt)^{1/2} (\int_0^T E[(X'_{ij})^4] dt)^{1/4} (E[(X'_{ij})^4])^{1/4} \end{aligned}$$

so that using (3.25)

$$\begin{aligned} & \sum_{ij} |E[\int_0^T X'_{ij} b''_{iii} dW^i \int_0^T X'_{ij} b''_{iii} dW^i X'_{ij} g'_i]| \\ & \leq C \epsilon^d \sup_{s \leq T} \sum_{ij} E[(X'_{ij})^2] + C \epsilon^d \sup_{s \leq T} \sum_{ij} E[(X'_{ij})^4] \leq C. \end{aligned}$$

The remaining terms in  $u'''$  can be bounded similarly. □

**3.2. Dendrites with Einstein diffusion.** We see by Theorem 3.1 and Lemma 3.1 that the mean field differential equation solution is also an accurate approximation to the spin dynamics; this indicates that the stochastic differential equation (3.2) then only offers a small quantitative improvement. However, if the derivatives of the value function would be large, coupled to the small scale  $\epsilon$ , the mean field solution may give a qualitatively wrong answer, with  $\mathcal{O}(1)$  error as  $\gamma/\epsilon \rightarrow 0+$ , while the stochastic differential equation still could yield an asymptotically correct limit; such an example is dendrite formation in phase transformations, cf. [27, 24, 1], [18].

Let us try to understand why the noise in Theorem 3.1 seems applicable to dendrite formation. Dendrite dynamics can be formulated by the phase field method with an Allen-Cahn/Ginzburg-Landau equation coupled to a diffusion equation for the energy, determining the temperature, and by master equations coupled to the energy equation, cf. [18]. Mean field equations related to such a phase field system have been derived from a spin system coupled to a diffusion equation, see [7]; a related model is to let the coarse-grained potential energy be defined by

$$m(\sigma, z) := \sum_x \left( \sum_{y \neq x} \frac{1}{2} J(x-y) \sigma(y) - h \right) \sigma(x) A_\epsilon(x, z),$$

where  $A$  is the average in (3.1), and replace the Glauber dynamics with Arrhenius dynamics. That is, the microscopic dynamics are given by independent spins  $\sigma(x) \in \{0, 1\}$ , for each lattice point  $x \in \mathcal{L}$ , flipping with adsorption rate

$$c_a(x) = d_0 (1 - \sigma(x)),$$

from states 0 to 1, and with desorption rate

$$c_d(x) = d_0 \sigma(x) \exp\left(-\frac{1}{k_B T} \left(\sum_{y \neq x} J(x-y) \sigma(y) - h\right)\right),$$

from states 1 to 0, where  $h$  is a surface binding energy or an external field and  $d_0$  is a given rate, cf. [28]. Arrhenius dynamics also satisfies detailed balance with the same Gibbs density

$$e^{\left(\sum_x \sum_{y \neq x} J(x-y) \sigma(x) \sigma(y) / 2 - \sum_x h \sigma(x)\right) / (k_B T)}$$

as for Glauber dynamics. The dynamics for the potential energy variable can then be coupled to the macroscopic energy equation

$$\partial_t(c_v \mathcal{T} + m) = \operatorname{div}(k \nabla \mathcal{T})$$

by letting the temperature  $\mathcal{T}$  vary on the coarse-grained scale.

The dendrite grows with a positive non vanishing speed. Without noise in the model there is no side branching, while the side branching is present with added noise to the phase field model, cf. [1], or to the mean field model derived in [18]. This noise induced side branching is explained by the high sensitivity with respect to small perturbations at the dendrite tip, cf. [27]. Therefore the derivatives of an appropriate value function are large. Here the value function,  $u$ , could for instance measure the total dendrite surface at a fixed time. The inconsistent approximation of the mean field solution could by Lemma 3.1 be explained by having

$$(\gamma/\epsilon)^d \|u''\|_{\ell^1} = \mathcal{O}(1). \tag{3.26}$$

The smallest scale in the problem is the dendrite tip radius  $\rho$ ; with a bounded value function its derivatives could then be

$$\begin{aligned} \|u'\|_{\ell^1} &= \mathcal{O}(1/\rho), \\ \|u''\|_{\ell^1} &= \mathcal{O}(1/\rho^2), \\ \|u'''\|_{\ell^1} &= \mathcal{O}(1/\rho^3). \end{aligned}$$

Consequently (3.26) yields  $(\gamma/\epsilon)^{d/2} = \rho$ , so that the noise error for the stochastic differential equation with the Einstein diffusion of Theorem 3.1 would be bounded by  $(\gamma/\epsilon)^{2d} \|u'''\|_{\ell^1} = \mathcal{O}((\gamma/\epsilon)^{d/2})$ , which tends to zero as  $\gamma/\epsilon \rightarrow 0+$ . Therefore, this adsorption/desorption kinetic Monte Carlo model with long range interaction generates an approximating stochastic differential equation, which could be applicable also to coupling with the energy equation if the derivation remains valid with slowly varying temperature. An essential and maybe more difficult question is to find accurate kinetic Monte Carlo methods for real systems with dendrite dynamics, e.g. using ideas from the molecular dynamics coarse-graining.

**3.3. Coarse-grained Monte Carlo.** Let now  $\bar{u}(\xi, t)$  be the solution to a Kolmogorov backward equation for a coarse-grained stochastic interacting particle model

$$\partial_t \bar{u} + \bar{L} \bar{u} = 0.$$

Let  $C_\eta$  be the set of all  $x \in \mathcal{L}$  in a coarse cell with  $\bar{X} = \sum_{x \in \eta} \sigma(x)/q^d$ , where  $q^d = (\epsilon/\gamma)^d$  is the number of sites in a coarse cell. Then the KMC-CGMC error between the averaged spin  $\bar{X}$  and the coarse grained Monte Carlo solution  $X$

$$\begin{aligned} E[g(\bar{X}_T) - g(X_T)] &= \int_0^T E[d\bar{u}(\bar{X}(\sigma_t), t)] = \int_0^T E[L\bar{u} + \partial_t \bar{u}] dt \\ &= \int_0^T E[L\bar{u}(\bar{X}(\sigma_t), t) - \bar{L}\bar{u}(\bar{X}_t, t)] dt \\ &= \int_0^T E \left[ \sum_{x \in \mathcal{L}} c(x, \sigma) \left( \bar{u}(\bar{X}(\sigma^x), t) - \bar{u}(\bar{X}(\sigma), t) \right) \right. \\ &\quad \left. - \sum_{\eta \in \bar{\mathcal{L}}} \bar{c}(\eta, \bar{X}) \left( \bar{u}(\bar{X}^\eta, t) - \bar{u}(\bar{X}, t) \right) \right] dt \\ &= \int_0^T E \left[ \sum_{\eta \in \bar{\mathcal{L}}} \left( \sum_{x \in \eta} c(x, \sigma) - \bar{c}(\eta, \bar{X}) \right) \left( \bar{u}(\bar{X}^\eta, t) - \bar{u}(\bar{X}, t) \right) \right] dt \end{aligned} \tag{3.27}$$

is at most of order  $\epsilon$ , cf. remark 3.6, since the coarse grained Monte Carlo has the transition rates

$$\bar{c}(\eta, \bar{X}) = \begin{cases} \frac{q^d(1-\bar{X})}{4} \left( 1 + \tanh(\beta(J * \bar{X} + h - J(0)X)) \right), & \bar{X} \text{ to } \bar{X} + 1/q^d, \text{ at } \eta \\ \frac{q^d(1+\bar{X})}{4} \left( 1 - \tanh(\beta(J * \bar{X} + h - J(0)X)) \right), & \bar{X} \text{ to } \bar{X} - 1/q^d, \text{ at } \eta. \end{cases}$$

which implies

$$\sum_{x \in \eta} c(x, \sigma) - \bar{c}(\eta, \bar{X}) = \mathcal{O}(q^d \epsilon) \tag{3.28}$$

and

$$|\bar{X}^\eta - \bar{X}| = 2/q^d.$$

Note that the error KMC -SDE splits to KMC-CGKMC + CGKMC -SDE which is  $\mathcal{O}(\epsilon) + \mathcal{O}((\gamma/\epsilon)^{2d})$ . The computational gain of the coarse-grained Monte Carlo method is to reduce the number of variables, see [28]. However, this coarse-graining does not change the number of spin jumps. Therefore the coarse-grained model has a high rate,  $\mathcal{O}(q^d)$ , of small jumps,  $2/q^d$ , which implies a fast time scale. The SDE approximation further improves the computational efficiency by avoiding this fast time scale, replacing the high rate of small jumps with small diffusion.

**3.4. Lattice systems with finite-range interactions.** Here we discuss a straightforward extension to lattice systems with finite-range interaction potentials of the form

$$J(x - y) = \frac{1}{L^d} J_0 \left( \frac{x - y}{\gamma L} \right),$$

where  $L^d$  denotes the number of interacting neighbors and  $J_0(x) = 0$ , if  $\|x\| \geq 1$ .

A closer inspection of the proof of Theorem 3.1 along with the observation that

$$\begin{aligned} - \sum_{x \in \mathcal{L}} 2c(x, \sigma) A_\epsilon(x, \cdot) A_\epsilon(x, \cdot) A_\epsilon(x, \cdot) \sigma^3(x) = \\ \left( \frac{\gamma}{\epsilon} \right)^{2d} \left[ -\bar{X} + \tanh \left( \beta(J * \bar{X} + h) \right) + \mathcal{E}_1 \right] \end{aligned}$$

yield the following, where  $\|w\|_{\ell^\infty} := \sup_x |w(x)|$ ,

**THEOREM 3.3.** *The average spin,  $\bar{X}$ , can be approximated by the solution,  $X$ , to the Ito stochastic differential equation (3.2) with error*

$$E[g(\bar{X}_T)] - E[g(X_T)] = T(\mathcal{E}_1 + \mathcal{E}_2)$$

where

$$\mathcal{E}_1 \leq C\beta E[\|J * \sigma - J * \bar{X}\|_{\ell^\infty}] = \beta \mathcal{O}(\epsilon/(L\gamma))$$

and

$$\mathcal{E}_2 = CE[\|-\bar{X} + \tanh(\beta(J * \bar{X} + h))\|_{\ell^\infty}] \mathcal{O}((\gamma/\epsilon)^{2d})$$

where  $C$  depends on the  $\ell^1(\bar{\mathcal{L}}^i)$  norms of the derivatives,  $u'(\bar{X}_t, t), u''(\bar{X}_t, t)$  and  $u'''(\bar{X}_t, t)$  on the path  $\bar{X}$ , up to time  $T$ .

The form of the error terms  $\mathcal{E}_1$  and  $\mathcal{E}_2$  suggests that, if the constant  $C$  in the Theorem remains uniformly bounded, then the SDE approximation can perform well even if we do not necessary have long range interactions. For instance, we can employ the *a priori* knowledge of the behavior of the approximated exact spin flip dynamics to ensure the errors in Theorem 3.3 are small:

- In a high temperature regime,  $\beta \ll 1$ , the error term  $\mathcal{E}_1$  is small and provide a good approximation even if the involved potentials are short ranged. Similarly the prefactor  $E[\|-\bar{X} + \tanh(\beta(J * \bar{X} + h))\|_{\ell^\infty}] \ll 1$ , since in that regime we know that  $E[\|\bar{X}\|] \ll 1$ .
- In the presence of a strong external field,  $|h| \gg 1$ , the SDE dynamics also provides a good approximation even for short ranged interactions. In this case the spins are all aligned, hence with large probability  $\sigma(\cdot) \approx \bar{X} \approx \pm 1$ ; therefore we have that  $E[\|J * \sigma - J * \bar{X}\|_{\ell^\infty}] \ll 1$ ,  $E[\|-\bar{X} + \tanh(\beta(J * \bar{X} + h))\|_{\ell^\infty}] \ll 1$ .

**3.5. Other remarks on Theorem 3.1.**

**REMARK 3.4.** (Bounds on  $X$ ). Let us verify that  $1 - X \tanh(\beta(J * X + h))$  is positive on the support of  $\eta$ , so that  $b$  becomes a  $C^\infty$  function. Let

$$r = \min(1 + 2e^{-2\beta(2|J|_{\ell^1} + |h|_{\ell^\infty})}, 2).$$

Assume that  $\max_x |X(x)| \leq r$ , then

$$\begin{aligned} 1 - X \tanh(\beta(J * X + h)) &\geq 1 - r \tanh(\beta(r|J|_{\ell^1} + |h|_{\ell^\infty})) \\ &\geq 1 - r \tanh(\beta(2|J|_{\ell^1} + |h|_{\ell^\infty})) \\ &\geq 1 - r(1 - 2e^{-2\beta(2|J|_{\ell^1} + |h|_{\ell^\infty})}) > (r - 1)^2 > 0 \end{aligned}$$

which shows that  $b$  is a  $C^\infty$  function. For  $|X(x)| > 1$  we note that the drift push  $X(x)$  towards the set  $[-1, 1]$ , in particular if  $|X(x)| = r$  we have  $b(x) = 0$ , by the construction of the cut-off function  $\eta$ , and consequently the drift prevents  $X(x)$  from leaving the set  $[-r, r]$ , so that the assumption  $\max_{x,t} |X_t(x)| \leq r$  indeed holds.

If one think of  $m = \tanh(\beta(J * \bar{X} + h - J(0)\bar{X}))$  as constant, the process for  $X$  has a drift as in an Ornstein-Uhlenbeck process with mean  $m$ , however with small

non constant diffusion depending on  $X$ ; therefore the values  $X$  will approach  $m$ , as  $\gamma \rightarrow 0+$ . A rigorous statement in this direction is  $E[g(X_T) - g(\bar{X}_T)] = T\mathcal{O}(\epsilon^2 + (\gamma/\epsilon)^d)$  obtained from Lemma 3.1 and Theorem 3.1. Note that  $|\hat{X}(y)| \leq 1$ .

REMARK 3.5. (Reflected diffusion). An alternative to the  $\eta$  cut-off in the diffusion coefficient (yielding a gradual reflection of  $X(y)$  to the set  $[-1, 1]$ ) is to let  $X$  be a diffusion process which is instantaneously reflected at the two boundaries  $X = \pm 1$

$$dX_x(t) = a_x(X(t))dt + b_x(X(t))dW^x + n(X_x(t))dZ_x(t),$$

where

$$Z_x(t) := \lim_{\delta \rightarrow 0+} \frac{1}{2\delta} \int_0^t 1_{\{X_x(s)-1 > -\delta \text{ or } X_x(s)+1 < \delta\}}(s)ds$$

is the local time  $X$  spends at the boundary and  $n(X_x(t))$  is the inward normal at  $X_x(t)$ . The work [13] includes an introduction to reflected diffusion processes. The first two steps in the proof of Theorem 3.6 are applicable also for this case of reflected diffusion, since the generator is the same. The reason that we do not only present this reflected case is that we do not know how to derive the corresponding mean square estimates of variations of reflected  $X$ , as established for the non reflected case in Section 3.1. The obstacle is the second and higher order variations, which include stopped diffusion problems that are not differentiable in the same sense as for the problem without boundary.

REMARK 3.6. (Improved KMC-CGKMC rate). Using asymptotic independence of  $\sigma(x)$ , established in [4], the rate in (3.28) improves to  $\mathcal{O}(q^d \epsilon^2)$ , which yields the KMC-CGKMC error  $\mathcal{O}(\epsilon^2)$ . This is also established directly in [30], without using the results in [4], again through a derivation in the spirit of the calculation (3.27), by reversing the role of  $u$  and  $\bar{u}$  and letting  $X$  be a jump process on the fine lattice with the coarse rates; the proof is based on an error estimate of  $u' - \bar{u}'$  and that the flips of  $X$  in the same cell are independent by construction.

REMARK 3.7. (Approximation of probability measures). A uniform error estimate (3.6), for all  $g$ , would give an estimate of the error of the probability measures of the KMC and SDE: by varying  $g$  in  $E[g(\bar{X}_T) | \bar{X}_t = \xi] = \int g(\zeta) dp_\xi(\zeta)$ . Section 4 presents an example of invariant measures.

REMARK 3.8. (Bounds on  $u$ ). Assume that the mean field equation is stable in the sense that the characteristics,  $Y$ , solving  $Y' = a(Y)$  are stable. Then the backward equation for  $u$  can to leading order be integrated by the characteristics and consequently the solution,  $u$ , to equation (3.16) will typically grow at least as  $\mathcal{O}(T)$ , so that  $u$  is not bounded uniformly in  $T$ . In the neighborhood of a local unstable equilibrium, converging backward characteristics may give large derivatives of  $u$ , see Section 4.4. In Section 5 we give an exit time example where the derivatives of  $u$  are growing exponentially as  $e^{C(\epsilon/\gamma)^d}$ .

REMARK 3.9. (Adaptivity). The CGMC and the SDE can be adaptively approximated in  $x$  and  $t$  by computing also approximations to  $u', u'', u'''$ . Such algorithms for time adaptivity of ordinary and stochastic differential equations are presented in [42, 36],[37]. Spatial adaptivity of a partial differential equation with this setting is presented in [38]. Furthermore, some initial work on spatial adaptivity for KMC based on a posteriori bounds and using variable-size CGMC simulations is demonstrated in

[2]. Ongoing work uses the separation into KMC-CGMC and CGMC-SDE error in (3.27) to construct adaptive methods for master equations and its hybrid coupling to SDE.

**4. The Curie-Weiss model and invariant measures**

**4.1. The Curie-Weiss model.** The Curie-Weiss model in one space dimension consists of the usual lattice system where the interaction potential is constant:

$$J(x-y) := \frac{J}{N}, \quad x, y \in \mathcal{L} \tag{4.1}$$

and  $N = \frac{1}{\gamma}$  is the total number of lattice sites on  $\mathcal{L}$ . Then the microscopic Hamiltonian

$$H(\sigma) = -\frac{1}{2} \sum_{x \in \mathcal{L}} \sum_{y \in \mathcal{L}} J(x-y) \sigma(x) \sigma(y) + \sum_{x \in \mathcal{L}} h \sigma(x), \tag{4.2}$$

can be written exactly as the coarse-grained Hamiltonian

$$H(\sigma) = \bar{H}(\eta) := N \left( -\frac{J}{2} \eta^2 + h \eta \right), \tag{4.3}$$

where the coarse variable  $\eta$  is defined as

$$\eta := \frac{1}{N} \sum_{x \in \mathcal{L}} \sigma(x). \tag{4.4}$$

The meaning of the previous calculation is that in the Curie-Weiss example the Hamiltonian can be coarse-grained exactly.

The microscopic dynamics is given by independent spins  $\sigma(x) \in \{0, 1\}$ , for each lattice point  $x \in \mathcal{L}$ , flipping with adsorption rate  $\tilde{c}_a(x) = 1 - \sigma(x)$ , from states 0 to 1, and with desorption rate  $\tilde{c}_d(x) = \sigma(x) \exp(-\beta(J\eta - h))$ , from states 1 to 0, where  $\beta = \frac{1}{kT}$  is the inverse temperature with the Boltzmann constant  $k$  and temperature  $T > 0$ , cf. [28].

REMARK 4.1 (Alternative Curie-Weiss model). The usual Curie-Weiss model with the Hamiltonian

$$H(\sigma) = -\frac{1}{2} \sum_{x \in \mathcal{L}} \sum_{y \neq x} J(x-y) \sigma(x) \sigma(y) + \sum_{x \in \mathcal{L}} h \sigma(x) = N \left( -\frac{J}{2} \eta(\eta - N^{-1}) + h \eta \right)$$

excluding self interaction, can be handled analogously to (4.2).

**4.2. Coarse-grained Monte Carlo dynamics and invariant measures**

In the Curie-Weiss case the microscopic spin flip dynamics coarse grain exactly: since the interaction potential is constant, the corresponding error in (3.28) becomes zero instead of  $\mathcal{O}(q^d \epsilon)$ ; this is described in more detail in [28] where the process described in ([28] pages 256-259) is exact, i.e. there is no error in ([28]-3.15) due to the lack of error in ([28]-3.10). Thus for the coarse variable  $\eta = \frac{1}{N} \sum_x \sigma(x) \in [0, 1]$  we have the Markov process defined in terms of a generator as:

$$\frac{d}{dt} E g(\eta) = E \bar{L} g(\eta) := E \left[ c_a(\eta) \left( g\left(\eta + \frac{1}{N}\right) - g(\eta) \right) + c_d(\eta) \left( g\left(\eta - \frac{1}{N}\right) - g(\eta) \right) \right], \tag{4.5}$$

where

$$c_a(\eta) = N(1 - \eta), \quad c_d(\eta) = N\eta \exp[-\beta(J\eta - h)]. \tag{4.6}$$

Let  $\mu$  denote the density of the invariant measure for the coarse-grained dynamics; its detailed balance

$$c_d(\eta + N^{-1})\mu(\eta + N^{-1}) = c_a(\eta)\mu(\eta)$$

implies that the invariant measure is

$$\begin{aligned} \mu(\eta) &= \mu(0) \prod_{i=1}^{\eta N} \frac{c_a((i-1)/N)}{c_d(i/N)} = \mu(0) e^{\sum_{i=0}^{\eta N} (\beta(Ji/N - h) - \log(i/N) + \log(1 - (i-1)/N))} \\ &= \bar{Z}^{-1} e^{-N(\beta\bar{H}(\eta) + \eta \log(\eta) + (1-\eta) \log(1-\eta) + \mathcal{O}(N^{-1}))} \\ &=: \bar{Z}^{-1} e^{-N\beta U(\eta)}, \end{aligned} \tag{4.7}$$

using Stirling's formula on the right hand side and  $\bar{Z} := N^{-1} \sum_{i=1}^N e^{-N\beta U(i/N)}$ .

A consequence of the exact coarse-graining of the Hamiltonian is that the coarse-grained invariant measure is obtained from the microscopic Gibbs measure,  $\exp(-\beta H(\sigma))/\bar{Z} = \exp(-\beta \bar{H}(\eta))/\bar{Z}$ , by

$$\mu(\eta) = \frac{1}{\bar{Z}} \exp(-\beta \bar{H}(\eta)) P_N(\eta),$$

where the product binomial distribution

$$P_N\left(\eta = \frac{k}{N}\right) = \frac{N!}{k!(N-k)!} \left(\frac{1}{2}\right)^N$$

is the prior distribution arising from the microscopic prior by including  $N$  independent sites, see [28], which by Stirling's formula yields (4.7). In the case of a general long range interaction potential, the work [28] shows that the invariant measure for the coarse-grained dynamics converges to the one from the exact microscopic invariant measure.

**4.3. SDE/MC approximating dynamics.** We now derive an SDE approximation of (4.5) in the spirit of the derivations in the previous section. We expand the smooth function  $g$  in (4.5) so that

$$g(\eta \pm \frac{1}{N}) - g(\eta) = \pm \frac{1}{N} g'(\eta) + \frac{1}{2N^2} g''(\eta) + \mathcal{O}(N^{-3}).$$

Substitution in (4.5) gives the approximation

$$\frac{d}{dt} E g(\eta) = E \left[ \frac{c_a(\eta) - c_d(\eta)}{N} g'(\eta) + \frac{c_a + c_d(\eta)}{2N^2} g''(\eta) \right] + \mathcal{O}(N^{-2}) \tag{4.8}$$

where  $c_a, c_d$  are given by (4.6). Thus we have the approximating SDE

$$d\eta = a(\eta)dt + b(\eta)dW_t, \tag{4.9}$$

where

$$a(x) = 1 - x - x \exp(-\beta(Jx - h)), \quad b(x) = \sqrt{\frac{|1 - x + x \exp(-\beta(Jx - h))|}{N}}.$$

The density  $\nu$  of the invariant measure of the 1-D stochastic differential equation (4.9) satisfies

$$(a\nu)' - \left(\frac{b^2\nu}{2}\right)'' = 0$$

and by detailed balance

$$a\nu = \left(\frac{b^2\nu}{2}\right)'$$

the solution is

$$\begin{aligned} \nu(\eta) &= \frac{1}{Z} \exp\left(2 \int_0^\eta \frac{a(x)}{b^2(x)} dx - \log b^2(\eta)\right) \\ &=: \frac{1}{Z} \exp\left(-NV(\eta)\right) \end{aligned} \tag{4.10}$$

where

$$\begin{aligned} V(\eta) &= -2 \int_0^\eta \frac{1-x-x \exp(-\beta(Jx-h))}{1-x+x \exp(-\beta(Jx-h))} dx \\ &\quad + \frac{1}{N} \log\left(1-\eta+\eta \exp(-\beta(J\eta-h))\right), \\ Z &= \int_{\mathbb{R}} \exp\left(-NV(\eta)\right) d\eta. \end{aligned}$$

**4.4. Comparison of the SDE/MC invariant measures.** We compare the asymptotics of the MC and the SDE by the corresponding different invariant measures (4.7) and (4.10). Looking carefully at the derivatives of  $U$  and  $V$  one notices that  $U$  and  $V$  have the same critical points up to order  $\mathcal{O}(N^{-1})$ . These critical points  $m_i$  are precisely the equilibrium points for the mean field equation

$$\frac{d}{dt}m(t) = a(m(t)).$$

We see that either there is one stable equilibrium point or there are two stable,  $m_{\pm}$ , and one unstable,  $m_0$ , equilibrium. At these equilibrium points there holds

$$U''(m) = V''(m) + \mathcal{O}(N^{-1}),$$

so that the stable points are local minima of  $U$  and  $V$ .

In the special case when  $U$  and  $V$  have only one local equilibrium point, which is a global minimum, the expansion

$$\begin{aligned} \int_0^1 G(x)e^{-NU(x)} dx / Z_{c,\beta} &\simeq \int_0^1 \left(G(m) + G'(m)(x-m) + \frac{1}{2}G''(m)(x-m)^2 + \dots\right) \\ &\quad \times e^{-N(U(m) + U''(m)(x-m)^2/2 + \dots)} dx / \bar{Z} \\ &= \int_0^1 \left(G(m) + G'(m)(x-m) + \frac{1}{2}G''(m)(x-m)^2 + \dots\right) \\ &\quad \times e^{-N(V(m) + V''(m)(x-m)^2/2 + \dots)} dx / Z + \mathcal{O}(N^{-2}) \\ &= \int_0^1 G(x)e^{-NV(x)} dx / Z + \mathcal{O}(N^{-2}) \end{aligned} \tag{4.11}$$



verifies that

$$\int_0^1 Gd\mu = \int_0^1 Gd\nu + \mathcal{O}(N^{-2}), \tag{4.12}$$

in accordance with Corollary 3.16 and (4.8).

The functions  $U$  and  $V$  can have two global minimum points  $m_{\pm}$  when  $h = J/2$  and  $J$  is sufficiently large; then to leading order  $U$  and  $V$  are even around  $\eta = 1/2$ . For other values of  $h$ , the functions  $U$  and  $V$  may have their global minimum at different points, making the derivation (4.11) invalid. In this case when  $U$  has a local equilibrium point  $y''$ , which is not a local minimum, the backward characteristic paths  $\{Y(t;x) \mid t < T, x \in \Omega\}$

$$Y'(t) = a(Y(t)) \quad t < T, \quad Y(T) = x,$$

may converge, in the neighborhood  $\Omega$  of  $y''$ . This can give rise to large derivatives of  $u$  so that the assumption in (3.16) does not hold for long time; the solution  $u$  is related to a linearized shock wave problem, cf. [17].

We see also that the deterministic mean field solution satisfies

$$\begin{aligned} \lim_{t \rightarrow \infty} m(t) &= m_+ && \text{if } m(0) > m_0, \\ \lim_{t \rightarrow \infty} m(t) &= m_- && \text{if } m(0) < m_0, \\ m(t) &= m_0 && \text{if } m(0) = m_0. \end{aligned}$$

Therefore the mean field behavior can be bistable, but if  $U$  has only one global minimum point, say  $m_+$  and  $m(0) > m_0$ , then

$$\int_0^1 Gd\mu = G(m_+) + \mathcal{O}(N^{-1}) = \lim_{t \rightarrow \infty} G(m(t)) + \mathcal{O}(N^{-1}), \tag{4.13}$$

in accordance with Corollary 3.16 and Lemma 5.5.

In conclusion:

- the two different invariant measures for the stochastic differential equations and the coarse-grained Monte Carlo method, respectively, have almost the same most likely states  $m_i$ , corresponding to the minima of the potentials  $U$  and  $V$ , which are the equilibrium mean states,  $a(m_i) = 0$ , with the mean field approximation (4.13);
- at the most likely states also the second derivatives of  $U$  and  $V$  are almost the same, so that the two invariant measures generate similar statistics (4.12) provided the derivatives of value function  $u$  remains bounded for all time;
- although the coarse-grained Monte Carlo method is expected to be computationally more costly than the SDE-based methods described in this paper, it generates an asymptotically correct invariant measure for potentials with long enough interactions, even in regimes that exhibit bistability, see [28] for more detailed estimates.

**5. Mean transition times with invariant measure diffusion**

Not all expected values  $E[g(\bar{X}_T)]$  can be approximated using the stochastic differential equation (3.2), due to the required bounds on the derivatives of  $u$ ; such an example is to determine the expected first exit time  $\tau(Y) = \inf\{t : Y_t \notin A\}$  from the

neighborhood  $A$  of an equilibrium point  $y' \in A$ , where  $a(y') = 0$  and  $Y_0 \in A$ . Then the expected exit time is exponentially large, i.e.

$$\lim_{\gamma/\epsilon \rightarrow 0+} (\frac{\gamma}{\epsilon})^d \log E[\tau(\bar{X})] \text{ and } \lim_{\gamma/\epsilon \rightarrow 0+} (\frac{\gamma}{\epsilon})^d \log E[\tau(X)] \tag{5.1}$$

are both strictly positive.

These expected values are related to transition rates  $k$ , called reaction rates in chemistry; and  $E[\tau] \approx 1/k$  in simple cases, see [22], [14]. Hanggi et. al. [21] have proposed a remedy by approximating the master equation by a different stochastic differential equation with the same asymptotic drift but a modified diffusion, to leading order, chosen so that the SDE invariant measure  $Z^{-1}e^{-U/(\gamma/\epsilon)^d}$  is asymptotically the same as for the master equation; we will apply this to the coarse-grained master equation. Consequently the asymptotics of the CGMC invariant measure are used to define the diffusion. Because of (5.1) the theory of large deviations for rare events is relevant, cf. [14].

Let  $\gamma_1 := \gamma/\epsilon$ . Consider an SDE

$$dX_t(x) = (a(X_t) + \gamma_1^d c(X_t))(x)dt + \gamma_1^{d/2} \tilde{b}(X_t)(x)dW_t^x,$$

with the generator

$$Lf = (a + \gamma_1^d c) \cdot f' + \gamma_1^d \tilde{D} \cdot f'', \quad D_{ij} = \tilde{b}_i \tilde{b}_j \delta_{ij};$$

the idea in [21] is to find  $c$  and  $D$  so that the corresponding SDE asymptotically has the same invariant measure  $e^{-U/\gamma_1^d}/Z$  as the master equation. Hanggi et.al. [21] determines the diagonal diffusion matrix  $\tilde{D}$  and the small contribution to the drift,  $\gamma_1^d c$ , by

$$\begin{aligned} \tilde{D}_{ii} &= -a_i/U'_i \\ c_i &= -\partial_{x_i} \tilde{D}_{ii}; \end{aligned} \tag{5.2}$$

note that since  $a$  and  $U$  have the same zeros, the constructed function  $\tilde{D}_{ii}$  is positive in general. The equation (5.2) can be obtained by the WKB expansion

$$\begin{aligned} 0 \simeq L^* e^{-U/\gamma_1^d} &= \left( \gamma_1^{-d} (a_i U'_i + \tilde{D}_{ii} U'_i U'_i) \right. \\ &+ \gamma_1^0 (\partial_i a_i + 2U'_i \partial_i \tilde{D}_{ii} + U''_{ii} \tilde{D}_{ii} + c_i U'_i) \\ &\left. + \gamma_1^d (\partial_i c - \partial_{ii} \tilde{D}_{ii}) \right) e^{-U/\gamma_1^d} \end{aligned}$$

together with the two leading order conditions that the terms of order  $\gamma_1^{-d}$  and  $\gamma_1^0$  vanish; here  $L^*$  is the Fokker-Planck operator adjoint to  $L$ . Consequently the choice (5.2) will in general generate an SDE with an invariant measure  $e^{-\tilde{U}/\gamma_1^d}/Z$ , where  $|\tilde{U} - U| = \mathcal{O}(\gamma_1^{2d})$ .

Let us investigate why good approximation of the invariant measure implies that also the expected values,  $E[\tau]$ , for exit problems related to rare events, are accurately computed: the work [14] shows that

$$\lim_{\gamma_1 \rightarrow 0+} \gamma_1^d \log E[\tau(X)] = \inf_{y \in \partial A} U(y) - U(y'), \tag{5.3}$$

for the case  $\tilde{D}_{ii} = 1$  and one stable attracting equilibrium point  $y' \in A$ ; the following section extends this to  $D$  given by (5.2) and shows that also

$$\lim_{\gamma_1 \rightarrow 0+} \gamma_1^d \log E[\tau(\bar{X})] = \inf_{y \in \partial A} U(y) - U(y'). \tag{5.4}$$

REMARK 5.1. (Einstein-diffusion versus invariant measure diffusion). Although the invariant measure diffusion (5.2) does not achieve the same high accuracy as the Einstein-diffusion (3.2) for bounded time approximation, the invariant measure diffusion solution will have the same order of accuracy as the mean field solution in Lemmas 3.1 for bounded times. On the other hand, the Einstein-diffusion (3.2) yields by (5.3) inconsistent approximation to exit times since its invariant measure (4.10) is different from the invariant measure (4.7) for the master equation. Note however that when the mean field equation is slightly unstable, such as for the dendrite dynamics in Section 3.2, the Einstein diffusion may be consistent for short times while others are not. The Einstein diffusion, valid for short time approximation, and the invariant measure diffusion, valid for very long time, leaves a gap of intermediate time where an accurate diffusion approximation remains to be found.

**5.1. Asymptotic agreement of exit times for SDE’s and master equations.** This section proves that the exit time for SDE’s with invariant measure diffusion is asymptotically the same as for the master equation for the 1D Curie-Weiss Model with the generator (4.5). More specifically assume that the master equation has the invariant measure density  $\exp(-\tilde{U}/\gamma_1)$  with  $\tilde{U} \rightarrow U$  as  $\gamma_1 \rightarrow 0+$  and consider a stochastic differential equation

$$\begin{aligned} dX_t &= \tilde{a}(X_t)dt + \sqrt{\gamma_1 \tilde{D}(X_t)}dW, \\ \tilde{D} &= D + \mathcal{O}(\gamma_1) \\ \tilde{a} &= -DU' + \mathcal{O}(\gamma_1). \end{aligned} \tag{5.5}$$

We show that

$$\lim_{\gamma_1 \rightarrow 0+} \gamma_1^d \left( \log E[\tau(X)] - \log E[\bar{\tau}(\bar{X})] \right) = 0, \tag{5.6}$$

where  $E[\bar{\tau}(\bar{X})]$  and  $E[\tau(X)]$  denote the mean exit time of the Curie-Weiss and the invariant measure diffusion respectively. Our technique to establish this asymptotic agreement is to use logarithmic (Hopf-Cole) transformations of the two mean exit times, as functions of the initial location, which transforms the corresponding two linear Kolmogorov backward equations to two nonlinear Hamilton-Jacobi equations, cf. [12]. The two processes give rise to two *different* asymptotic Hamilton-Jacobi equations (see (5.11) below); however the key observation is that they have the *same* viscosity solution since they are both convex and have the same set of zeros. Of course, the result (5.6) does not imply that the corresponding expected times will have the same pre-factor in their WKB expansion. In order to compare these pre-factors, a more detailed WKB expansion of the corresponding Hamilton-Jacobi equations would be necessary, following the asymptotic series results in [12], Section VI.7 and [11].

Next we describe in detail these results. The value function

$$u(x, t) = E[\tau(X); X_t = x] - t,$$

with exit from the domain  $(a, b) = A$ , solves

$$\partial_t u + \tilde{a}u' + \gamma_1 \tilde{D}u'' + 1 = 0 \quad \text{in } (a, b) \times (0, T), \quad u(a, \cdot) = u(b, \cdot) = 0, \quad u(\cdot, T) = 0.$$

By definition,  $u$  and  $\bar{u}$  are non-negative functions. Assume that  $\tilde{a}(y') = 0$  for precisely one point  $y' \in (a, b)$ . The time independent one dimensional first order equation for

$v = u'$  can be solved explicitly, but we will instead derive an approximating technique that can be applied to the master equation and in higher dimension, based on a WKB transformation. Assume that  $\max_{x,t} u(x,t) = M$ . Since  $u$  is non-negative, we can uniquely define the non-negative function  $w$  (depending on  $\gamma_1$ ) by  $u = M + 1 - e^{w/\gamma_1}$ . Similarly, let  $\max \bar{u} = \bar{M}$ , then  $\bar{u} = \bar{M} + 1 - e^{\bar{w}/\gamma_1}$  defines the non-negative function  $\bar{w}$ . The definitions of  $w$  and  $\bar{w}$  yield

$$\begin{aligned} \partial_t w + \tilde{a}w' + \tilde{D}w'w' + \gamma_1 \tilde{D}w'' - \gamma_1 e^{-w/\gamma_1} &= 0, \\ \partial_t \bar{w} + \gamma_1 c_a \left( e^{(\bar{w}(x+\gamma_1) - \bar{w}(x))/\gamma_1} - 1 \right) + \gamma_1 c_d \left( e^{(\bar{w}(x-\gamma_1) - \bar{w}(x))/\gamma_1} - 1 \right) - \gamma_1 e^{-\bar{w}/\gamma_1} &= 0 \end{aligned}$$

with the boundary condition that  $w = \gamma_1 \log(M + 1)$  on  $(a \times (0, T)) \cup (b \times (0, T)) \cup ((a, b) \times T)$ , the boundary values of  $w$  are constant independent of  $x$  and  $t$ , and likewise for  $\bar{w}$  we have on the boundary  $\bar{w} = \gamma_1 \log(\bar{M} + 1)$ . In some sense, it is the unknown constants  $M$  and  $\bar{M}$  we want to determine.

The advantage with these two Hamilton-Jacobi equations is that their limit as  $\gamma_1 \rightarrow 0+$  is well behaved in the viscosity solution theory. To understand these limits, observe first that  $a = -DU' + \mathcal{O}(\gamma_1)$  from condition (5.5), derived by the detailed balance in (5.2). The master equation corresponding to (4.5) has the detailed balance

$$c_d(x + \gamma_1) e^{-\tilde{U}(x+\gamma_1)/\gamma_1} = c_a(x) e^{-\tilde{U}(x)/\gamma_1}$$

so that

$$\frac{c_d(x)}{c_a(x)} = \frac{c_d(x + \gamma_1)}{c_a(x)} + \mathcal{O}(\gamma_1) = e^{(\tilde{U}(x+\gamma_1) - \tilde{U}(x))/\gamma_1} + \mathcal{O}(\gamma_1) = e^{U'(x)} + \mathcal{O}(\gamma_1).$$

Therefore, using also that  $w$  and  $\bar{w}$  are non-negative and differentiable, the two functions  $w$  and  $\bar{w}$  have limits as  $\gamma_1 \rightarrow 0+$  and these limit functions  $(w, \bar{w})$  satisfy in viscosity solutions sense, cf. [12], the equations

$$\partial_t w + \underbrace{D(-U'w' + w'w')}_{=: H(w', \cdot)} = 0, \tag{5.7}$$

and

$$\partial_t \bar{w} + \underbrace{\gamma_1 c_a (e^{\bar{w}'} - 1 + e^{U'} (e^{-\bar{w}'} - 1))}_{=: \bar{H}(\bar{w}', \cdot)} = 0, \tag{5.8}$$

respectively, with the boundary conditions (for the outward normal  $n$ )

$$\begin{aligned} w(\alpha) &= \gamma_1 \log(M_0 + 1) & \text{if } H_{w'}(w'(\alpha), \alpha) \cdot n(\alpha) &\geq 0 & \alpha \in \partial A \\ \bar{w}(\alpha) &= \gamma_1 \log(\bar{M}_0 + 1) & \text{if } \bar{H}_{\bar{w}'}(\bar{w}'(\alpha), \alpha) \cdot n(\alpha) &\geq 0 & \alpha \in \partial A \end{aligned} \tag{5.9}$$

or

$$\begin{aligned} w(\alpha) &\geq \gamma_1 \log(M_0 + 1) & \text{if } H_{w'}(w'(\alpha), \alpha) \cdot n(\alpha) &\leq 0 & \alpha \in \partial A \\ \bar{w}(\alpha) &\geq \gamma_1 \log(\bar{M}_0 + 1) & \text{if } \bar{H}_{\bar{w}'}(\bar{w}'(\alpha), \alpha) \cdot n(\alpha) &\leq 0 & \alpha \in \partial A, \end{aligned} \tag{5.10}$$

where  $M_0$  and  $\bar{M}_0$  are the limits of  $M$  and  $\bar{M}$ . We see that strict inequality in the second case (5.10) are not possible values for  $w, \bar{w}$ , therefore we need the solutions

to satisfy  $H_{w'} \cdot n|_{\partial A} \geq 0$ . These boundary conditions come from the related optimal control setting, where (5.9) means that a controlled path hits the boundary  $\partial A$ ; while if a controlled path instead exits at  $T = \infty$  the terms  $-\gamma_1 e^{-w/\gamma_1}$  and  $-\gamma_1 e^{-\bar{w}/\gamma_1}$  in the Hamilton-Jacobi equations for  $w$  and  $\bar{w}$  give an additional cost  $-\infty$  making such paths non optimal. We seek time independent solutions of the equations

$$\begin{aligned} H(w', \cdot) &= 0 \\ \bar{H}(\bar{w}', \cdot) &= 0 \end{aligned} \tag{5.11}$$

obtained by first taking the limit  $T \rightarrow \infty$  and then  $\gamma_1 \rightarrow 0+$ . The equations (5.11) determine  $w$  and  $\bar{w}$  only up to an unknown additive constant. The boundary condition (5.9) sets the constant, but since  $M_0$  and  $\bar{M}_0$  are unknown, the right constant is better found by the condition  $\lim_{a,b \rightarrow y'} w = 0$ . Since the two Hamilton-Jacobi equations (5.11) have convex Hamiltonians with the same zeros and boundary condition, the equations are equivalent and their viscosity solutions are the same. In the domain where the solution is differentiable, we have  $w'$  and  $\bar{w}'$  equal to 0 or  $U'$ .

Now let  $c > y'$  be defined by  $U(a) - U(y') = U(b) - U(c)$ . Then the unique solution to (5.11) and (5.9) is

$$\begin{aligned} w(x) = \bar{w}(x) &= U(x) - U(y') && \text{for } x < y', \\ w(x) = \bar{w}(x) &= 0 && \text{for } y' \leq x < c, \\ w(x) = \bar{w}(x) &= U(x) - U(c) && \text{for } c < x < b, \end{aligned}$$

see Figure 5.1, which solves (5.11) with the boundary condition (5.9). We have proved:

**THEOREM 5.2.** *The master equation corresponding to (4.5) and the stochastic equation (5.5) have to leading order the same mean transition times, i.e. for  $x \in A = (a, b)$*

$$\lim_{\gamma_1 \rightarrow 0+} \gamma_1 \log E[\tau(\bar{X}) \mid \bar{X}_0 = x] = \lim_{\gamma_1 \rightarrow 0+} \gamma_1 \log E[\tau(X) \mid X_0 = x] = \min_{y \in \partial A} (U(y) - U(y')).$$

**REMARK 5.3.** (Higher dimensions). The coarse-grained Ising model in Section 3.3 takes a similar form as the adsorption-desorption model. Note that by detailed balance, the derivation of (5.3) and (5.4) can be extended to multi-dimensional master equations with coarse-grained adsorption-desorption model; in higher dimension the two Hamiltonians still have the common zeros  $U'$  and 0, but there are other non-common zeros in general so that the two time independent Hamilton-Jacobi equations are not equivalent. However, in a domain  $A$  where  $U$  has only one equilibrium point the solutions  $w$  and  $\bar{w}$  will, as in the one dimensional case, have the gradient equal to  $U'$  or 0. Therefore (5.3) and (5.4) seems to hold also in several dimensions.

**Appendix A. Centered averages.** This section partly extends the results on projected averages in Section 3 to localized centered averages on scale  $\epsilon$

$$\bar{X} := A_\epsilon * \sigma = \sum_{y \in \mathcal{L}} A_\epsilon(\cdot - y) \sigma(y) \tag{A.1}$$

where  $A_\epsilon$  is the symmetric mollifier

$$A_\epsilon(x) = C_A e^{-|x|^2/(2\epsilon^2)} \tag{A.2}$$

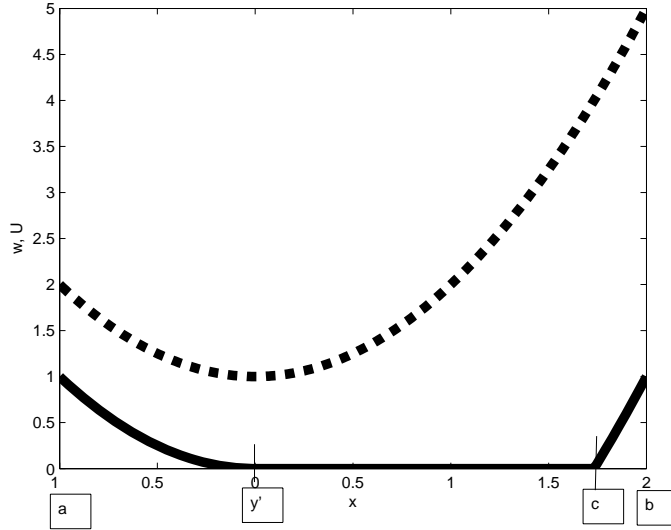


FIG. 5.1. The functions  $w = \bar{w}$  (solid) for  $U(x) = x^2 + 1$  (dashed), with  $(a, b) = (-1, 2)$ ,  $y' = 0$ , and  $c = \sqrt{3}$ .

with  $C_A$  chosen such that

$$\sum_{x \in \mathcal{L}} A_\epsilon(x) = 1, \tag{A.3}$$

and

$$\sum_{x \in \mathcal{L}} x A_\epsilon(x) = 0.$$

We assume that  $\epsilon > \gamma$ . In this section we consider a Hamiltonian which excludes self-interaction, so that  $H(\sigma) = \sum_x \sum_{y \neq x} J(x - y) X(x) X(y)$ .

We shall show that the average spin,  $\bar{X}$ , can be approximated by the solution,  $X$ , to the Ito stochastic differential equation

$$dX_t = a(X_t)dt + \sum_{z \in \mathcal{L}} b^z(X_t) dW^z, \quad X_0 = \bar{X}_0, \tag{A.4}$$

with the drift and diffusion coefficients given by

$$\begin{aligned} a(X) &= -X + \tanh(\beta(J * X + h - J(0)X)), \\ b_x^z(X) &= A_\epsilon(x - z) \sqrt{|1 - X \tanh(\beta(J * X + h - J(0)X))(x)|} \eta(X(x)), \\ \eta(r) &= \begin{cases} 1 & \text{for } x \in [-1, 1], \\ 0 & \text{for } x \in (-\infty, -\hat{r}) \cup (\hat{r}, \infty), \end{cases} \\ \hat{r} &:= \min(1 + e^{-2\beta(2|J|_{\ell^1} + \|h\|_{L^\infty})}, 3/2) \end{aligned} \tag{A.5}$$

and a Wiener process  $W : \mathcal{L} \times [0, T] \rightarrow \mathbb{R}$ , where  $W^x$  are independent one dimensional standard Brownian motions for  $x \in \mathcal{L}$ ,

The approximation uses that the value function

$$u(\xi, t) = E[g(X_T) | X_t = \xi]$$

solves a corresponding Kolmogorov backward equation, where the drift and diffusion coefficients in (A.5) are chosen to minimize the error  $E[g(\bar{X}_T)] - E[g(X_T)]$ . Introduce the weighted scalar products

$$\begin{aligned} w \cdot v &:= \sum_{y \in \mathcal{L}} w_y v_y \gamma^d && \text{for } w, v \in \mathbb{R}^{\mathcal{L}}, \\ w \cdot v &:= \sum_{x, y \in \mathcal{L}} w_{xy} v_{xy} \gamma^{2d} && \text{for } w, v \in \mathbb{R}^{\mathcal{L}^2}, \\ w \cdot v &:= \sum_{x, y, z \in \mathcal{L}} w_{xyz} v_{xyz} \gamma^{3d} && \text{for } w, v \in \mathbb{R}^{\mathcal{L}^3}. \end{aligned} \tag{A.6}$$

Then  $u$  satisfies the Kolmogorov backward equation

$$\partial_t u + a \cdot u' + D \cdot u'' = 0, \text{ for } t < T,$$

$$u(\cdot, T) = g,$$

where

$$\begin{aligned} D_{xy} &= \frac{1}{2} \sum_{z \in \mathcal{L}} b_x^z b_y^z \\ &= \sqrt{1 - X \tanh(\beta(J * X + h))(x)} \sqrt{1 - X \tanh(\beta(J * X + h))(y)} \\ &\quad \times \underbrace{\sum_{z \in \mathcal{L}} A_\epsilon(x - z) A_\epsilon(y - z)}_{= A_{\sqrt{2}\epsilon}(x - y)} \eta(X(x)) \eta(X(y)), \end{aligned} \tag{A.7}$$

and  $u'(\xi, t) = \partial_\xi u(\xi, t)$  and  $u''(\xi, t)$  are the first and second order Gâteaux derivatives of  $u$  in the weighted scalar products (A.6) on  $\mathcal{L}$  and  $\mathcal{L}^2$ , respectively. The definition of the weighted scalar product implies that the Gâteaux derivative  $u'_i(\xi, t)$  is a factor of  $\gamma^{-d}$  larger than the standard partial derivative  $\partial u(\xi, t) / \partial \xi_i$ , i.e.

$$\begin{aligned} u'(\xi, t) &= \gamma^{-d} \partial u(\xi, t) / \partial \xi \\ u''(\xi, t) &= \gamma^{-2d} \partial^2 u(\xi, t) / \partial \xi^2 \\ u'''(\xi, t) &= \gamma^{-3d} \partial^3 u(\xi, t) / \partial \xi^3. \end{aligned}$$

Note that the stochastic process  $\tilde{W}$  defined by

$$\tilde{W}^x = \sum_{z \in \mathcal{L}} A_\epsilon(x - z) W^z,$$

implies

$$\sum_{z \in \mathcal{L}} b_x^z dW^z = \sqrt{1 - X \tanh(\beta(J * X + h))(x)} \eta(X(x)) d\tilde{W}^x,$$

where  $\tilde{W}$  is a  $\ell^2(\mathcal{L})$  valued  $A_{\sqrt{2\epsilon}}$ -Wiener process with covariance  $A_{\sqrt{2\epsilon}} = \mathcal{O}((\gamma/\epsilon)^d)$ , cf. [3]. The covariance  $A_{\sqrt{2\epsilon}}$  has bounded trace and exponentially decaying eigenvalues.

Introduce the weighted  $\ell^1(\mathcal{L}^i)$  norms  $\|w\|_{\ell^1(\mathcal{L}^i)} := \sum_{x \in \mathcal{L}^i} |w_x| \gamma^{id}$ . The main result in this section is:

**THEOREM 5.4.** *The average spin,  $\bar{X}$ , can be approximated by the solution,  $X$ , to the Ito stochastic differential equation (A.4) with error*

$$E[g(\bar{X}_T)] - E[g(X_T)] = \mathcal{O}(T\epsilon^2 + (T\epsilon^2)^2 + (T+T^2)(\gamma/\epsilon)^{2d}) \tag{A.8}$$

provided that the Gâteaux derivatives,  $u'(\bar{X}_t, t), \dots, u''''(\bar{X}_t, t)$  on the path  $\bar{X}$ , are bounded in the weighted norms  $\ell^1(\mathcal{L}^i)$  up to time  $T$  and the initial spin  $\sigma_0$  has expected value  $m$ , where  $\sigma_0(x) - m_x$  are i.i.d. with bounded variance and second order difference quotients  $|d^2m/dx^2| = \mathcal{O}(1)$ .

Note that  $a=0$  gives  $\mathcal{O}(1)$  error, while  $b=0$  gives  $\mathcal{O}((\gamma/\epsilon)^d)$  error so that  $b$  defined by (A.5) is justified for  $\gamma \ll \epsilon \ll \gamma^{d/(d+1)}$ , with  $T$  fixed. The stochastic differential equation (A.4) has  $\mathcal{C}^\infty$  coefficients, where perturbations of solutions may grow exponentially in time, cf. Section 3.1. The method to estimate derivatives of  $u$ , for the projection averages in Section 3.1, could in principle be used also for the analogous stochastic differential equation for centered averages.

*Proof.* The proof proceeds as for the analogous one for the projected average in Theorem 3.1: the definitions of  $u$ , the generator (2.2) and the average (A.1) imply

$$\begin{aligned} E[g(\bar{X}_T)] - E[g(X_T)] &= E[u(\bar{X}_T, T)] - E[u(X_0, 0)] \\ &= E\left[\int_0^T du(\bar{X}_t, t)\right] = \int_0^T E[L u + \partial_t u] dt = \int_0^T E[E[L u - a \cdot u' - D \cdot u'' | \bar{X}_t]] dt \\ &= \int_0^T E\left[E\left[\sum_{x \in \mathcal{L}} c(x, \sigma_t) \left(u(\bar{X}(\sigma_t^x), t) - u(\bar{X}(\sigma_t), t)\right) - a \cdot u' - D \cdot u'' | \bar{X}_t\right]\right] dt \\ &= \int_0^T E\left[E\left[\sum_{x \in \mathcal{L}} c(x, \sigma_t) \left(u(\bar{X}(\sigma_t) - 2A_\epsilon(x - \cdot)\sigma_t(x), t) - u(\bar{X}(\sigma_t), t)\right) \right. \right. \\ &\quad \left. \left. - a \cdot u'(\bar{X}_t, t) - D \cdot u''(\bar{X}_t, t) | \bar{X}_t\right]\right] dt. \end{aligned} \tag{A.9}$$

The first step to estimate this error is to write the differences in  $u$  in terms of its derivatives by Taylor expansion for some  $s \in [0, 1]$

$$\begin{aligned} &u(\bar{X}(\sigma_t) - 2A_\epsilon(x - \cdot)\sigma_t(x), t) - u(\bar{X}(\sigma_t), t) \\ &= -2u'(\bar{X}_t, t) \cdot A_\epsilon(x - \cdot)\sigma_t(x) \\ &\quad + 2u''(\bar{X}_t, t) \cdot A_\epsilon(x - \cdot)A_\epsilon(x - \cdot)\sigma_t^2(x) \\ &\quad - \frac{4}{3}u'''(\bar{X}_t - 2sA_\epsilon(x - \cdot)\sigma_t(x), t) \cdot A_\epsilon(x - \cdot)A_\epsilon(x - \cdot)A_\epsilon(x - \cdot)\sigma_t^3(x). \end{aligned} \tag{A.10}$$

The next step is to determine the optimal  $a$  and  $b$  which minimize the error (A.9). For this purpose we shall in the flipping rate approximate the coupling  $J * \sigma$  and  $J(0)\sigma = \mathcal{O}(\gamma^d)$  by  $J * \bar{X}$  and  $J(0)\bar{X}$ , using the long range  $\mathcal{O}(1)$  interaction distance of  $J$ . The definition of the average (A.1) implies

$$J * \bar{X} = (J * A_\epsilon) * \sigma$$



and consequently the centered coupling improves (3.11) to the uniform error estimate

$$|J * \sigma - J * \bar{X}| \leq \|J_0 - J_0 * A_\epsilon\|_{\ell^1} \|\sigma\|_{\ell^\infty} = \mathcal{O}(\epsilon^2). \tag{A.11}$$

This error estimate, the flip rate (2.3) and  $J(0) = \mathcal{O}(\gamma^d)$  imply

$$\begin{aligned} - \sum_{x \in \mathcal{L}} 2c(x, \sigma) A_\epsilon(x - \cdot) \sigma(x) &= -\bar{X} + A_\epsilon * \tanh\left(\beta(J * \sigma + h - J(0)\sigma)\right) \\ &= -\bar{X} + \tanh\left(\beta(J * \bar{X} + h - J(0)\bar{X})\right) + \mathcal{O}(\epsilon^2 + \gamma^{2d}). \end{aligned} \tag{A.12}$$

The condition (A.3) shows that  $C_A = \mathcal{O}((\gamma/\epsilon)^d)$  and therefore

$$\begin{aligned} \sum_x A_\epsilon(x - \cdot) A_\epsilon(x - \cdot) &= \mathcal{O}((\gamma/\epsilon)^d), \\ \sum_x A_\epsilon(x - \cdot) A_\epsilon(x - \cdot) A_\epsilon(x - \cdot) &= \mathcal{O}((\gamma/\epsilon)^{2d}), \end{aligned} \tag{A.13}$$

which together with the expansions (A.9) and (A.10) imply:

LEMMA 5.5. *Suppose the assumptions in Theorem 5.4 hold. The deterministic mean field solution,  $\hat{X}$ ,*

$$d\hat{X}/dt = -\hat{X} + \tanh(\beta(J * \hat{X} + h - J(0)\hat{X})), \quad \hat{X}_0 = E[\bar{X}_0], \tag{A.14}$$

*depends on  $\epsilon$  only through the initial data and satisfies*

$$E[g(\bar{X}_T)] - E[g(\hat{X}_T)] = \mathcal{O}(\epsilon^2 + (\gamma/\epsilon)^d) \tag{A.15}$$

*provided the drift  $a$  is defined by (A.5).*

*Proof.* We may consider  $\hat{X}$  similarly to an  $X$  with  $b=0$  and apply the corresponding expansion (A.9), (A.10) and (A.13). Then it remains to verify that the initial data satisfy

$$E[u(\bar{X}_0, 0) - u(\hat{X}_0, 0)] = \mathcal{O}((\gamma/\epsilon)^d),$$

but this is a direct consequence of the central limit theorem and the initial  $\sigma_0 - E[\sigma_0]$  being i.i.d. with bounded variance.  $\square$

In order to derive the refined error (A.8), we shall now also use that the diffusion coefficient  $b$  is chosen precisely so that the term

$$u'' \cdot \left( \sum_{x \in \mathcal{L}} c(x, \sigma) A_\epsilon(x - y) A_\epsilon(x - z) \sigma^2(x) - D_{yz} \right)$$

in (A.9) and (A.10) becomes small. The definition (A.2) implies

$$A_\epsilon(x - y) A_\epsilon(x - z) \sigma^2(x) = A_{\sqrt{2}\epsilon}(y - z) A_{\frac{\epsilon}{\sqrt{2}}}\left(x - \frac{y+z}{2}\right) \tag{A.16}$$

and consequently by (2.3)

$$\begin{aligned} &\sum_{x \in \mathcal{L}} c(x, \sigma) A_\epsilon(x - y) A_\epsilon(x - z) \sigma^2(x) \\ &= A_{\sqrt{2}\epsilon}(y - z) \left( 1 - A_{\frac{\epsilon}{\sqrt{2}}} * \sigma \tanh(\beta(J * \sigma + h - J(0)\sigma)) \right) \left( \frac{y+z}{2} \right) \\ &= \mathcal{O}((\gamma/\epsilon)^d). \end{aligned} \tag{A.17}$$

LEMMA 5.6. *The following estimate holds*

$$E \left[ u''(\bar{X}_t, t) \cdot \left( \sum_{x \in \mathcal{L}} c(x, \sigma_t) A_\epsilon(x-y) A_\epsilon(x-z) \sigma_t^2(x) - D_{yz} \right) \right] = \mathcal{O}((\gamma/\epsilon)^{2d} + \epsilon^4). \quad (\text{A.18})$$

The bound (A.13) implies the estimate

$$\sum_{x \in \mathcal{L}} c(x, \sigma_t) u'''(\bar{X}_t - 2s A_\epsilon(x-\cdot) \sigma_t(x)) \cdot A_\epsilon(x-\cdot) A_\epsilon(x-\cdot) A_\epsilon(x-\cdot) \sigma_t^3(x) = \mathcal{O}((\gamma/\epsilon)^{2d})$$

which combined with (A.9), (A.10), (A.12) and (A.18) proves the Theorem.  $\square$

*Proof.* Let  $\bar{X}' := A_{\frac{\epsilon}{\sqrt{2}}} * \sigma =: A' * \sigma$ . The first step is to use Lemma 5.5 to relate  $\bar{X} = A' * \bar{X}'$  and  $\bar{X}'$  to the mean field solutions  $\hat{X}_\epsilon =: \hat{X}$  and  $\hat{X}_{\epsilon/\sqrt{2}} =: \hat{X}'$ , respectively, and then to use the smoothness of  $\hat{X}'$  to estimate the error  $\hat{X} - \hat{X}'$ .

Let  $f_x(Y) := \sqrt{1 - Y \tanh(\beta(J * A' * Y + h - J(0)Y))} \eta(Y)(x)$ . The estimate (A.15) applied to the error between  $\bar{X}'$  and  $\hat{X}'$ , with

$$g(\xi) = A_{\sqrt{2}\epsilon}(x-y) f_{\frac{x+y}{2}}^2(\xi) \cdot u''(A' * \xi, t),$$

shows

$$E \left[ A_{\sqrt{2}\epsilon}(x-y) f_{\frac{x+y}{2}}^2(\bar{X}') \cdot u''(A' * \bar{X}', t) - A_{\sqrt{2}\epsilon}(x-y) f_{\frac{x+y}{2}}^2(\hat{X}') \cdot u''(A' * \hat{X}', t) \right] = \mathcal{O}(\epsilon^4 + (\gamma/\epsilon)^{2d}), \quad (\text{A.19})$$

where (A.17) implies that the first term in the left hand sides of (A.18) and (A.19) are the same. Similarly choose  $g(\xi) = A_{\sqrt{2}\epsilon}(x-y) f_x(\xi) f_y(\xi) \cdot u''(\xi, t)$  to obtain

$$E \left[ A_{\sqrt{2}\epsilon}(x-y) f_x(\hat{X}) f_y(\hat{X}) \cdot u''(\hat{X}, t) - A_{\sqrt{2}\epsilon}(x-y) f_x(\bar{X}) f_y(\bar{X}) \cdot u''(\bar{X}, t) \right] = \mathcal{O}((\epsilon^4 + (\gamma/\epsilon)^{2d})), \quad (\text{A.20})$$

where definition (A.7) of  $D$ ,  $|\bar{X}| \leq 1$  and  $|\hat{X}| \leq 1$ , show that the second term in the left hand sides of (A.18) and (A.20) are the same.

Finally, it remains to estimate the difference of the second term in (A.19) and the first term in (A.20)

$$I := A_{\sqrt{2}\epsilon}(x-y) f_x(\hat{X}) f_y(\hat{X}) \cdot u''(\hat{X}, t) - A_{\sqrt{2}\epsilon}(x-y) f_{\frac{x+y}{2}}^2(\hat{X}') \cdot u''(A' * \hat{X}', t),$$

which requires bounds on  $\hat{X} - \hat{X}'$ , to evaluate  $f$  for the same function, and on second difference quotients  $d^2 \hat{X}/dx^2$  and  $d^2 \hat{X}'/dx^2$ , to evaluate  $f$  at the same point  $(x+y)/2$  and to estimate  $A' * \hat{X}' - \hat{X}'$ . The assumed smoothness of the initial data  $E[\bar{X}'_0] = A' * E[\sigma_0]$  shows that

$$\hat{X}(0) - \hat{X}'(0) = A' * \hat{X}'(0) - \hat{X}'(0) = \mathcal{O}(\epsilon^2),$$

and the smoothness of the interaction potential  $J$  and the mean field equation (A.14) imply that the mean field solution remains smooth with

$$\begin{aligned} \hat{X}(t) - \hat{X}'(t) &= \mathcal{O}(\epsilon^2), \\ \left| \frac{d^2 \hat{X}}{dx^2} \right| + \left| \frac{d^2 \hat{X}'}{dx^2} \right| &= \mathcal{O}(1), \quad \text{for the second difference quotient } d^2/dx^2. \end{aligned}$$

Consequently, we have  $I = \mathcal{O}(\epsilon^4 + (\gamma/\epsilon)^{2d})$ , which together with (A.19) and (A.20) and (A.11), to relate  $\tanh(\beta(J * A' * Y + h - J(0)Y))$  and  $\tanh(\beta(J * Y + h - J(0)Y))$  in (A.19) and (A.20), prove Lemma 5.6.  $\square$

REMARK 5.7. (Why a Gaussian average  $A_\epsilon$ ). A crucial contribution to the error from the second order term including  $u''$  in (A.10) is  $\sum_x A_\epsilon(x-z)A_\epsilon(x-y)\sigma(x)$ , which we need to relate to the average  $\bar{X} = A_\epsilon * \sigma$ . The reason we use the Gaussian average  $A_\epsilon$ , defined in (A.2), is to have this product of two translated Gaussians  $A_\epsilon(x-z)A_\epsilon(x-y)$  equal to another Gaussian centered at the mid point between the two, see (A.16). Then this centered Gaussian yields an averaged  $\sigma$ , which is related to  $\bar{X}$  by Lemma 5.6.

REMARK 5.8. (Alternative Approximations). Consider the alternative approximation  $\tilde{X}$  defined by additive noise from the deterministic mean field solution  $\bar{X}$  to (A.14)

$$d\tilde{X}_t(z) = a_z(\tilde{X}_t)dt + \tilde{b}_z(t)d\tilde{W}^z, \quad \tilde{X}_0 = \bar{X}_0, \quad (\text{A.21})$$

$$\tilde{b}_z(t) = \sqrt{1 - \hat{X}_t \tanh(\beta(J * \hat{X}_t + h - J(0)\hat{X}_t))}(z).$$

The proof of the theorem shows that  $E[g(\bar{X}_T) - g(\tilde{X}_T)] = \mathcal{O}(\epsilon^2 + (\gamma/\epsilon)^{2d})$ .

A third alternative approximation  $\tilde{X}$  is

$$d\tilde{X}_t = a(\tilde{X}_t)dt + \sum_{z \in \mathcal{L}} \tilde{b}^z(\tilde{X}_t)dW^z, \quad \tilde{X}_0 = \bar{X}_0, \quad (\text{A.22})$$

with diffusion coefficients given by

$$\tilde{b}_x^z = A_\epsilon(x-z) \sqrt{1 - \tilde{X} \tanh(\beta(J * \tilde{X} + h))}(z) \eta(\tilde{X}(z)).$$

The proof of the theorem again shows that  $E[g(\bar{X}_T) - g(\tilde{X}_T)] = \mathcal{O}(\epsilon^2 + (\gamma/\epsilon)^{2d})$ .

REMARK 5.9. (Wavelets). Finally we remark that the increments  $d\tilde{W}$  can be generated by a wavelet expansion, see [10].

#### REFERENCES

- [1] W. J. Boettinger, J. A. Warren, C. Beckermann and A. Karma, *Phasefield simulation of solidification*, Ann. Rev. Mater. Res., 32, 163-194, 2002.
- [2] A. Chatterjee, M. A. Katsoulakis and D. G. Vlachos, *Spatially adaptive grand canonical ensemble Monte Carlo simulations*, Phys. Rev. E, 71, 026702, 2005.
- [3] G. Da Prato and J. Zabczyk, *Stochastic equations in infinite dimensions*, Encyclopedia of Mathematics and its Applications, Cambridge University Press, Cambridge, 44, 1992.
- [4] A. De Masi, E. Orlandi, E. Presutti and L. Triolo, *Glauber evolution with the Kac potentials. I. Mesoscopic and macroscopic limits, interface dynamics*, Nonlinearity, 7, 633-696, 1994.
- [5] A. De Masi, E. Orlandi, E. Presutti and L. Triolo, *Glauber evolution with Kac potentials. II. Fluctuations*, Nonlinearity, 9, 27-51, 1996.
- [6] A. De Masi, E. Orlandi, E. Presutti and L. Triolo, *Glauber evolution with Kac potentials. III. Spinodal decomposition*, Nonlinearity, 9, 53-114, 1996.
- [7] N. Dirr and S. Luckhaus, *Mesoscopic limit for non-isothermal phase transition*, Markov Processes and Related Fields, 7, 355-381, 2001.
- [8] W. E, W. Ren and E. Vanden-Eijnden, *Minimum action method for the study of rare events*, Comm. Pure Appl. Math., 57, 637-656, 2004.
- [9] A. Einstein, *Über die von der molekularkinetischen theorie der wärme geforderte bewegung von in ruhenden flüssigkeiten suspendierten teilchen*, Ann. Phys., 17, 549-560, 1905.

- [10] F. W. Elliott, D. J. Horntrop and A. J. Majda, *A Fourier-wavelet Monte Carlo method for fractal random fields*, J. Comput. Phys., 384-408, 132, 1997.
- [11] W. H. Fleming and H. M. Soner, *Asymptotic expansions for Markov processes with Levy generators*, Applied Math. Opt., 19, 203-223, 1989.
- [12] W. H. Fleming and H. M. Soner, *Controlled Markov Processes and Viscosity Solutions*, Applications of Mathematics (New York), Springer-Verlag, New York, 25, 1993.
- [13] M. Freidlin, *Functional Integration and Partial Differential Equations*, Annals of Mathematics Studies, Princeton University Press, Princeton, NJ, 109, 1985.
- [14] M. I. Freidlin and A. D. Wentzell, *Random Perturbations of Dynamical Systems*, Grundlehren der Mathematischen Wissenschaften, Springer-Verlag, New York, 260, 1984.
- [15] G. Giacomin, J. L. Lebowitz and E. Presutti, *Deterministic and stochastic hydrodynamic equations arising from simple microscopic model systems*, Stochastic partial differential equations: six perspectives, Math. Surveys Monogr., Amer. Math. Soc., Providence, RI, 64, 107-152, 1999.
- [16] R. J. Glauber, *Time-dependent statistics of the Ising model*, J. Math. Phys., 4, 294-307, 1963.
- [17] J. Goodman, A. Szepessy and K. Zumbrun, *A remark on the stability of viscous shock waves*, SIAM J. Math. Anal., 25, 1463-1467, 1994.
- [18] J. F. Gouyet, M. Plapp, W. Dietrich and P. Maass, *Description of far-from-equilibrium process by mean-field lattice gas models*, Advances in Physics, 52, 523-638, 2003.
- [19] L. Granasy, T. Borzsonyi and T. Pusztai, *Nucleation and bulk crystallization in binary phase field theory*, Phys. Rev. Lett., 88, 206105, 2002.
- [20] L. Granasy, T. Pusztai, T. Borzsonyi and T. et al., *A general mechanism of polycrystalline growth*, Nature Materials 3, 645-650, 2004.
- [21] P. Hanggi, H. Grabert, P. Talkner and H. Thomas, *Bistable systems: master equation versus Fokker-Planck modeling*, Phys. Rev. A, 3, 371-378, 1984.
- [22] P. Hanggi, P. Talkner and M. Borkovec, *Reaction-rate theory - 50 years after Kramers*, Rev. Modern Phys., 62, 251-341, 1990.
- [23] M. Hildebrand and A. S. Mikhailov, *Mesoscopic modeling in the kinetic theory of adsorbates*, J. Phys. Chem., 100, 19089-19101, 1996.
- [24] J. J. Hoyt, M. Asta and A. Karma, *Atomistic and continuum modeling of dendritic solidification*, Materials Science Engineering R-Reports, 41, 121-163, 2003.
- [25] M. H. Kalos and P. A. Whitlock, *Monte Carlo Methods. Vol. I. Basics*, A Wiley-Interscience Publication. John Wiley & Sons, Inc., New York, 1986.
- [26] L. P. Kadanoff, *Statistical Physics: Statics, Dynamics and Renormalization*, World Scientific, 2000.
- [27] A. Karma and W. J. Rappel, *Phase-field model of dendritic sidebranching with thermal noise*, Phys Rev. E., 60, 3614-3625, 1999.
- [28] M. A. Katsoulakis, A. J. Majda and D. G. Vlachos, *Coarse-grained stochastic processes and Monte Carlo simulations in lattice systems*, J. Comput. Phys., 186, 250-278, 2003.
- [29] M. A. Katsoulakis, A. J. Majda and D. G. Vlachos, *Coarse-grained stochastic processes for microscopic lattice systems*, Proc. Natl. Acad. Sci. USA, 100, 782-787, 2003.
- [30] M. A. Katsoulakis, P. Plechac and A. Sopotakis, *Error analysis for coarse-graining of stochastic lattice dynamics*, SIAM J. Num. Anal., to appear.
- [31] M. A. Katsoulakis and J. Trashorras, *Information loss in coarse-graining of stochastic particle dynamics*, J. Stat. Phys., 122, 115-135, 2006.
- [32] C. Kipnis and C. Landim, *Scaling Limits of Interacting Particle Systems*, Springer-Verlag, 1999.
- [33] A. Kolmogorov, *Über die analytischen methoden in der wahrscheinlichkeitsrechnung*, Mathem. Annalen, 104, 415-458, 1931.
- [34] T. M. Liggett, *Interacting Particle Systems*, Springer-Verlag, Berlin, 2005.
- [35] G. C. Papanicolaou, *Introduction to the asymptotic analysis of stochastic equations*, Modern Modeling of Continuum Phenomena (Ninth Summer Sem. Appl. Math., Rensselaer Polytech. Inst., Troy, N. Y., 1975), Lectures in Appl. Math., Amer. Math. Soc., Providence, R. I., 16, 109-147, 1977.
- [36] K.-S. Moon, A. Szepessy, R. Tempone and G. E. Zouraris, *Convergence rates for adaptive weak approximation of stochastic differential equations*, Stoch. Anal. Appl., 23, 511-558, 2005.
- [37] K.-S. Moon, A. Szepessy, R. Tempone and G. E. Zouraris, *Convergence rates for adaptive approximation of ordinary differential equations*, Numer. Math., 96, 99-129, 2003.
- [38] K.-S. Moon, E. von Schwerin, A. Szepessy and R. Tempone, *Convergence rates of an adaptive dual weighted residual finite element algorithm*, preprint, 2005.
- [39] O. Penrose, *A mean-field equation of motion for the dynamic Ising model*, J. Statist. Phys., 63, 975-986, 1991.

- [40] B. Perthame, *Mathematical tools for kinetic equations*, Bull. Amer. Math. Soc., 41, 205-244, 2004.
- [41] J. S. Reese, S. Raimondeau and D. G. Vlachos, *Monte Carlo algorithms for complex surface reaction mechanisms: Efficiency and accuracy*, J. Comp. Phys., 173, 302-321, 2001.
- [42] A. Szepessy, R. Tempone and G. E. Zouraris, *Adaptive weak approximation of stochastic differential equations*, Comm. Pure Appl. Math., 54, 1169-1214, 2001.
- [43] N. G. van Kampen, *Stochastic Processes in Physics and Chemistry*, North-Holland, 1981.
- [44] E. Vanden-Eijnden, *Numerical techniques for multi-scale dynamical systems with stochastic effects*, Commun. Math. Sci., 1, 385-391, 2003.
- [45] C. Villani, *A review of mathematical topics in collisional kinetic theory*, Handbook of mathematical fluid dynamics, North-Holland, Amsterdam, I, 71-305, 2002.
- [46] S. Yip (Ed.), *Handbook of Materials Modeling*, Springer, 2005.