ERROR ANALYSIS OF COARSE-GRAINED KINETIC MONTE CARLO METHOD

ERROR ANALYSIS OF COARSE-GRAINED KINETIC MONTE CARLO METHOD

MARKOS A. KATSOULAKIS*, PETR PLECHÁČ†, AND ALEXANDROS SOPASAKIS‡

Abstract. The coarse-grained Monte Carlo (CGMC) algorithm was originally proposed in the series of works [20, 21, 24]. In this paper we further investigate the approximation properties of the coarse-graining procedure and provide both analytical and numerical evidence that the hierarchy of the coarse models is built in a systematic way that allows for error control in both transient and long-time simulations. We demonstrate that the numerical accuracy of the CGMC algorithm as an approximation of stochastic lattice spin flip dynamics is of order two in terms of the coarse-graining ratio and that the natural small parameter is the coarse-graining ratio over the range of particle/particle interactions. The error estimate is shown to hold in the weak convergence sense. We employ the derived analytical results to guide CGMC algorithms and we demonstrate a CPU speed-up in demanding computational regimes that involve nucleation, phase transitions and metastability.

Key words. coarse-grained stochastic processes, Monte Carlo simulations, birth-death process, detailed balance, Arrhenius dynamics, Gibbs measures, weak error estimates, kinetic Monte Carlo method

AMS subject classifications. 65C05, 65C20, 82C20, 82C26

1. Introduction.

Microscopic computational models for complex systems such as Molecular Dynamics (MD) and Monte Carlo (MC) algorithms are typically formulated in terms of simple rules describing interactions between individual particles or spin variables. The large number of variables and even larger number of interactions between them present the principal limitation for efficient simulations. Another restricting factor is illustrated by essentially sequential nature of approximating the time evolution in particle systems that yields a substantial slowdown in the resolution of dynamics, especially in metastable regimes.

In [20, 21, 24] the authors started developing systematic mathematical strategies for the coarse-graining of microscopic models, focusing on the paradigm of stochastic lattice dynamics and the corresponding MC simulators. In principle, coarse-grained models are expected to have fewer observables than the original microscopic system making them computationally more efficient than the direct numerical simulations. In these papers a hierarchy of coarse-grained stochastic models – referred to as coarse-grained MC (CGMC) – was derived from the microscopic rules through a stochastic closure argument. The CGMC hierarchy is reminiscent of Multi-Resolution Analysis approaches to the discretization of operators [4], spanning length/time scales from the microscopic to the mesoscopic. The resulting *stochastic coarse-grained processes* involve Markovian birth-death and generalised exclusion processes and their combinations, and as demonstrated in [20, 21, 24], they share the same ergodic properties with their microscopic counterparts. The full hierarchy of the coarse-grained stochastic dynamics satisfies detailed balance relations and as a result not only it yields self-consistent random fluctuation mechanism, but also consistent with the underlying microscopic fluctuations and the unresolved degrees of freedom. From the computational complexity perspective, a comparison of CGMC with conventional MC methods for the same real time shows, [20], that the CPU time can decrease approximately as $O(1/q^2)$ or faster, where q is the level of coarse-graining, as demonstrated for spin-flip lattice dynamics. Thus, while for macroscopic size systems in the millimeter length scale or larger, microscopic MC simulations are impractical on a single processor, the computational savings of CGMC make it a suitable tool

*Department of Mathematics and Statistics, University of Massachusetts, Amherst, MA 01003–9305, USA, markos@math.umass.edu

†Mathematics Institute, The University of Warwick, Coventry, CV4 7AL, United Kingdom, plechac@maths.warwick.ac.uk

‡Department of Mathematics and Statistics, University of Massachusetts, Amherst, MA 01003–9305, USA, sopas@math.umass.edu

capable of capturing large scale features, while retaining microscopic information on intermolecular forces and particle fluctuations. In the case of diffusion (spin exchange) dynamics we also observe an additional coarse-graining in time by a factor q^2, improving the hydrodynamic slowdown effect in the conservative MC, [24].

In the recent paper [23] the authors rigorously analysed CGMC models as approximations of conventional MC in *non-equilibrium*, by estimating the *information loss* between microscopic and coarse-grained adsorption/desorption lattice dynamics. In analogy to the numerical analysis for PDEs, an error analysis was carried out between the *exact microscopic process* $\{\sigma_t\}_{t\geq 0}$ and the *approximating coarse-grained process* $\{\eta_t\}_{t\geq 0}$. The key step in this direction was to use, as a quantitative measure for the loss of information in the coarse-graining from finer to coarser scales, the information-theoretic concept of the *relative entropy* between probability measures, [8]. Such relative entropy estimates give a first mathematical reasoning for the parameter regimes, i.e., the degree of coarse-graining versus the interaction range, for which CGMC is expected to give errors within a given tolerance. Using the rigorous results in [23] as a starting point, in this paper we focus on carrying out a detailed numerical analysis of the error propagation for spin flip lattice dynamics. Due to the numerical intractability of the relative entropy for a large particle system, we employ, in the numerical error calculations, suitable computable upper and lower bounds, as well as *targeted* coarse observables. The latter point of view necessitates in the use of a weak convergence framework for the study of the error between CGMC and direct numerical simulations of the stochastic lattice dynamics. We demonstrate that the numerical accuracy of the CGMC algorithm is of order two in terms of the ratio of the coarse-graining over the range of particle/particle interactions. We also refer to recent work in [22] on weak error estimates between microscopic MC algorithms and therein derived SDE approximations. Further details about a priori estimates for weak convergence of approximations to SDEs can be found in [3, 33] and [27]. Related a posteriori estimates are discussed in [32]. We further employ the derived analytical results to guide CGMC algorithms and we demonstrate a CPU speed-up in demanding computational regimes that involve nucleation, phase transitions and metastability. We demonstrate computationally that CGMC probes efficiently the energy landscape, yielding *spatial path-wise* agreement with the underlying microscopic lattice dynamics, at least for fairly long but still finite interactions.

The CGMC algorithms discussed here are related to a number of methods involving coarse-graining at various levels, for instance fast summation techniques, computational renormalization and simulation and multi-scale computational methods for stochastic systems. One of the sources of the computational complexity of molecular simulations arises in the calculation of particle/particle interactions, especially in the case where long range forces are relevant. The evaluation cost of such pairwise interactions can be significantly reduced by applying well-controlled approximation schemes and/or a hierarchical decomposition of the computation. Such ideas have been successfully applied in the development of Ewald summation techniques, multigrid (MG) , fast multipole methods (FMP) or tree-code algorithms. Typically, once the interaction terms are computed with one of these fast summation methods, they are entered in the microscopic algorithm where a simulation with a large number of individually tracked particles has still to be carried out. The point of view adopted by CGMC is related to these methods in the sense that the interaction potential or operator is approximated in terms of a truncated multi-resolution decomposition within a given tolerance. The CGMC is subsequently defined at the coarse level specified by the truncation of the decomposition. However, a notable difference is that CGMC models track much fewer coarse observables instead of simulating every individual particle. The equilibrium set-up of CGMC is essentially given by the renormalised Hamiltonian after a single iteration in the renormalisation group flow. It is not surprising that such an approach, when applied to near critical temperature simulations, has many limitations. For example, in the nearest-neighbour Ising-type models this fact is manifested in the aforementioned error estimates and the comparative simulations in [20]. On the other hand the focus of CGMC is dynamic simulations usually coupled to a macroscopic system (see for instance the hybrid systems in [34, 19]), where criticality may not be as important due to the presence of a time-varying external field. Nevertheless, further corrections to the CGMC dynamics from the renormalisation group flow given by RGMC and multigrid MC methods [5, 7, 12] can improve the order of convergence of the CGMC. We refer to [18] for higher order accurate CGMC methods based on cluster expansions, where the coarse-graining procedure described here is the model around which a cluster expansion is carried out with controlled errors. In that sense the CGMC method is of order two accurate as explained in Section 4.

In recent years there has been a growing interest in developing and analysing coarse-graining methods for the purpose of modelling and simulation across scales. Such systems arise in a broad spectrum of scientific disciplines ranging from materials science to macro-molecular dynamics, to epidemiology and to atmosphere/ocean science. Various coarse-graining approaches may yield explicitly derived stochastic coarse models using different coarse approximations, e.g., [13, 15, 16, 29, 31], or can be statistics-based [30] or may rely on on-fly simulations, e.g., the equation-free method [25], the heterogeneous multi-scale method [10], or multi-scale FE methods [14]. A systematic approach to upscaling of stochastic systems has been also developed from the multi-level perspective in [1, 2, 6], where the authors proposed algorithms for efficient multi-scale simulations using Monte Carlo methods. Other coarse-graining techniques in the polymer science literature include the bond fluctuation model and its variants [28]. Such coarse-graining methodologies often rely on parametrisation, hence at different conditions (e.g., temperature, density, composition) coarse potentials need to be re-parametrised [30].

2. Microscopic lattice models. The presented analysis applies to the class of Ising-type lattice systems. For the sake of simplicity we assume that the computational domain is defined as the discrete periodic lattice $\Lambda_N = \frac{1}{n}\mathbb{Z}^d \cap \mathbb{T}$ which represents discretion of the d-dimensional torus $\mathbb{T} = [0,1)^d$ and d denotes the spatial dimension. We restrict presentation of the results to $d = 1$, nevertheless higher dimensional cases are obtained without significant changes. However, the algorithms can also be implemented on bounded domains with usual boundary conditions. The number of lattice sites $N = n^d$ is fixed. The microscopic degrees of freedom or the microscopic order parameter is given by the spin-like variable $\sigma(x)$ defined at each site $x \in \Lambda_N$. In this paper we discuss only the case of discrete spin variables, i.e., $\sigma(x) \in \Sigma$ with $\Sigma = \{-1, 1\}$, $\Sigma = \{0, 1\}$ (Ising model) or $\Sigma = \{0, 1, \ldots s\}$ (Potts models). The case of the spin variable belonging to a compact Riemannien manifold, e.g., $\Sigma = \mathbb{S}^2$ (Heisenberg model), $\Sigma = \mathrm{SU}(2)$ (matrix model), will be studied elsewhere. We denote by $\sigma = \{\sigma(x) \,|\, x \in \Lambda_N\}$ a configuration of spins on the lattice, i.e., an element of the configuration space $\mathcal{S}_N = \Sigma^{\Lambda_N}$. The interactions between spins at a given configuration σ are defined by the microscopic Hamiltonian

$$H(\sigma) = -\frac{1}{2}\sum_{x\in\Lambda_N}\sum_{y\neq x} J(x-y)\sigma(x)\sigma(y) + \sum_{x\in\Lambda_N} h(x)\sigma(x)\,, \tag{2.1}$$

where $h(x)$ denotes the external field at the site x. The two-body inter-particle potential J accounts for interactions between individual spins. We consider the class of potentials with the following properties

$$J(x-y) = \frac{1}{L^d} V\left(\frac{n}{L}|x-y|\right)\,, \quad x,y \in \Lambda_N\,, \tag{2.2}$$
$$V:\mathbb{R}\to\mathbb{R}\,, \quad V(r)=V(-r)\,, \quad V(r)=0\,, \quad \text{if } |r|\geq 1. \tag{2.3}$$

We impose additional assumptions on V which allow us to derive explicit error estimates:

$$V \text{ is smooth on } \mathbb{R}\setminus\{0\}, \tag{2.4}$$
$$\int_{\mathbb{R}} |V(r)|\,dr < \infty\,, \text{ and } \int_{\mathbb{R}} |\partial_r V(r)|\,dr < \infty\,. \tag{2.5}$$

Note that the summability condition for V guarantees that the potential J is also summable due to the scaling factor. Hence the Hamiltonian is well defined even for $N, L \to \infty$. The canonical equilibrium state is given in terms of the Gibbs measure

$$\mu_{N,\beta}(d\sigma) = \frac{1}{Z_{N,\beta}} e^{-\beta H(\sigma)} \mathrm{P}_N(d\sigma)\,, \qquad Z_{N,\beta} = \int_{\mathcal{S}_N} e^{-\beta H(\sigma)} \mathrm{P}_N(d\sigma)\,, \tag{2.6}$$

where $\mathrm{P}_N(d\sigma) = \prod_{x\in\Lambda_N} \rho(d\sigma(x))$ is the product measure on $\mathcal{S}_N$ and the spins $\sigma(x)$ are independent identically distributed (i.i.d.) random variables with the common distribution ρ. For example, in the Ising model the prior distribution on $\Sigma = \{0,1\}$ would typically be $\rho(0) = \rho(1) = 1/2$.

The microscopic dynamics is defined as a continuous-time jump Markov process that defines a change of the spin $\sigma(x)$ with the probability $c(x,\sigma;\xi)\Delta t$ over the time interval $[t, t+\Delta t]$. The function $c : \Lambda_N \times \mathcal{S}_N \times \Sigma \to \mathbb{R}$ is called a rate of the process. The jump process $\{\sigma_t\}_{t\geq 0}$ is constructed in the following way: suppose that at the time t the configuration is σ_t, then the probability of changing the spin at the site $x\in\Lambda_N$ spontaneously from $\sigma_t(x)$ to a new value $\xi\in\Sigma$ over the time interval $[t,t+\Delta t]$ is $c(x,\sigma;\xi)\Delta t + O(\Delta t^2)$. We denote the resulting configuration by $\sigma^{x,\xi}$. In the case of the Ising-type state space and spin-flip dynamics we omit ξ in this notation. The generator $\mathcal{L} : L^\infty(\mathcal{S}_N) \to L^\infty(\mathcal{S}_N)$ of the Markov process acting on a bounded test function $\phi \in L^\infty(\mathcal{S}_N)$ defined on the space of configurations is given by

$$(\mathcal{L}\phi)(\sigma) = \sum_{x\in\Lambda_N} \int_\Sigma c(x,\sigma;\xi)\left(\phi(\sigma^{x,\xi}) - \phi(\sigma)\right)\,d\xi\,. \tag{2.7}$$

The evolution of an observable (a test function) ϕ is given by

$$\frac{d}{dt}\mathbb{E}\left[\phi(\sigma_t)\right] = \mathbb{E}\left[\mathcal{L}\phi(\sigma_t)\right]\,, \tag{2.8}$$

where the expectation operator $\mathbb{E}\,[.]$ is with respect to a measure conditioned to the initial configuration $\sigma_{t=0} = \sigma_0$. We require that the dynamics is of relaxation type such that the invariant measure of this Markov process is the Gibbs measure (2.6). The sufficient condition is known as *Detailed Balance* (DB) and it imposes condition on the form of the rate

$$c(x,\sigma;\xi)e^{-\beta H(\sigma)} = c(x,\sigma^{x,\xi};\sigma(x))e^{-\beta H(\sigma^{x,\xi})}\,. \tag{2.9}$$

This condition has a simple interpretation: $c(x,\sigma;\xi)$ is the rate of converting $\sigma(x)$ to the value ξ while $c(x,\sigma^{x,\xi};\sigma(x))$ is the rate of changing the spin with the value ξ at the site x

back to $\sigma(x)$. The widely used class of Metropolis-type dynamics satisfies (2.9) and has the rate given by

$$c(x,\sigma;\xi) = G(\beta\Delta_{x,\xi}H(\sigma))\,,\ \text{where } \Delta_{x,\xi}H(\sigma) = H(\sigma^{x,\xi}) - H(\sigma), \tag{2.10}$$

where G is a continuous function satisfying: $G(r) = G(-r)e^{-r}$ for all $r \in \mathbb{R}$. The most common choices in physics simulations are $G(r) = \frac{1}{1+e^r}$ (Glauber dynamics), $G(r) = e^{-[r]_+}$, (Metropolis dynamics), with $[r]_+ = r$ if $r \geq 0$ and $= 0$ otherwise, or $G(r) = e^{-r/2}$. Such dynamics are often used as samplers from the canonical equilibrium Gibbs measure. However, the kinetic Monte Carlo method is also used for simulations of non-equilibrium processes. The dynamics in such a case is known as *Arrhenius dynamics*, whose rates are usually derived from transition state theory or obtained from molecular dynamics simulations.

To avoid unnecessary generality we restrict the description to the Ising-type model with $\Sigma = \{0,1\}$ used for modelling adsorption/desorption processes. We also omit ξ in the notation. The Arrhenius rate is defined as follows

$$c(x,\sigma) = \begin{cases} d_0 & \text{if } \sigma(x) = 0, \\ d_0 e^{-\beta U(x,\sigma)} & \text{if } \sigma(x) = 1, \end{cases} \tag{2.11}$$

where

$$U(x,\sigma) = \sum_{y\in\Lambda_N\,, y\neq x} J(x-y)\sigma(y) - h(x)\,. \tag{2.12}$$

Furthermore, the spin-flip rule is given by

$$\sigma^x(y) = \begin{cases} 1-\sigma(x) & \text{if } y = x \\ \sigma(y) & \text{if } y \neq x. \end{cases}$$

With the introduced notation the coarse-graining algorithm can be described as an *approximation* of the microscopic dynamics, i.e., of the process $\{\sigma_t\}_{t\geq 0}$ by a coarse-grained process $\{\eta_t\}_{t\geq 0}$ where the approximation is done in a controlled way. We are interested not only in the approximation of the invariant measure $\mu_{N,\beta}(d\sigma)$ (see (2.6)) but also in the approximation of the measure on the path space.

3. Approximation of the coarse-grained process. The coarse-graining is defined in a geometric way by introducing the coarse-grained observables as block-spin variables. This approach follows the standard procedure of real-space renormalisation, see for example [17]. We remark that although we introduce block-spins our aim is not to approximate the renormalisation group flow (either on the space of Gibbs measures or on the path space) rather we want to find an approximation that is constructed with low computational cost and with controlled and computable error estimates.

In general terms we define the coarse-graining operator $\mathbf{T} : \mathcal{S}_N \to \mathcal{S}^c_{M,q}$, where the coarse configuration space $\mathcal{S}^c_{M,q}$ is defined on the coarse lattice Λ^c_M, and with the new state space Σ^c, i.e., $\mathcal{S}^c_{M,q} = (\Sigma^c)^{\Lambda^c_M}$. The coarse configuration $\eta = \mathbf{T}\sigma \in \mathcal{S}^c_{M,q}$ is defined on a smaller lattice with M lattice sites and with the coarse state space Σ^c for the new lattice spins $\eta(k)$. The parameter q defines the coarse-graining ratio. The operator $\mathbf{T}$ induces an operator $\mathbf{T}_*$ on the space of probability measures

$$\mathbf{T}_* : \mathcal{P}(\mathcal{S}_N) \to \mathcal{P}(\mathcal{S}^c_{M,q})\,, \quad \mu(\sigma) \mapsto \mu^c(\eta) := \mu\{\sigma \in \mathcal{S}_N \mid \mathbf{T}\sigma = \eta\}\,.$$

Ising-type spins. To be more specific we analyse the following case of Ising spin-flip dynamics $\mathcal{S}_N = \{0,1\}^{\Lambda_N}$. Each coarse lattice site $k \in \Lambda^c_M$ represents a cube C_k that contains

q sites of the microscopic lattice Λ_N. The projection operator defines the block spin at the coarse site k

$$(\mathbf{T}\sigma)(k) := \sum_{x\in C_k} \sigma(x)\,. \tag{3.1}$$

If the dimension d of the lattice is greater than one we understand k and x as multi-indices $k = (k_1, \ldots, k_d)$ and we index the corresponding lattice sites in the natural order. Choosing the projection operator in this way defines the coarse state space as $\Sigma^c = \{0, 1, \ldots q\}$. Given the Markov process $(\{\sigma_t\}_{t\geq 0}, \mathcal{L})$ with the generator $\mathcal{L}$ we obtain a coarse-grained process $\{\mathbf{T}\sigma_t\}_{t\geq 0}$ which is *not*, in general, a Markov process. From the computational point of view this may cause significant difficulties should sampling of such a process be implemented on the computer. Therefore we derive an *approximating* Markov process $(\{\eta_t\}_{t\geq 0}, \bar{\mathcal{L}}^c)$ which can be easily implemented once its generator is given explicitly.

For the model Ising system the projected generator of the coarse-grained process $\{\eta_t\}_{t\geq 0}$ can be evaluated explicitly by rearranging the summations on the lattice Λ_N. Given the microscopic state σ and corresponding coarse state $\eta = \mathbf{T}\sigma$

$$\begin{aligned}\mathcal{L}\psi(\mathbf{T}\sigma) = &\sum_{k\in\Lambda^c_M} \left[\sum_{x\in C_k} c(x,\sigma)(1-\sigma(x))\right] [\psi(\eta+\delta_k) - \psi(\eta)] + \\ &\sum_{k\in\Lambda^c_M} \left[\sum_{x\in C_k} c(x,\sigma)\sigma(x)\right] [\psi(\eta-\delta_k) - \psi(\eta)]\;.\end{aligned} \tag{3.2}$$

The configuration δ_k defined on the coarse state space is equal to zero at all sites except the site $k \in \Lambda^c_M$ where it is equal 1, i.e., $\delta_k(j) = 1$ for $j = k$ and $= 0$ otherwise. We see from the formula (3.2) that the exact generator for the coarse process can be written in the form

$$\mathcal{L}^c\psi(\eta) = \sum_{k\in\Lambda^c_M} c_a(k)\,[\psi(\eta+\delta_k) - \psi(\eta)] + \sum_{k\in\Lambda^c_M} c_d(k)\,[\psi(\eta-\delta_k) - \psi(\eta)]\;, \tag{3.3}$$

where the new rates

$$c_a(k) = \sum_{x\in C_k} c(x,\sigma)(1-\sigma(x))\,, \qquad c_d(k) = \sum_{x\in C_k} c(x,\sigma)\sigma(x)\,, \tag{3.4}$$

correspond to the adsorption and desorption processes. In this form the rates depend on the microscopic configuration σ and not on the coarse random variable $\mathbf{T}\sigma$. Therefore, it is reasonable to propose an approximating Markov process, which for the case of desorption/adsorption is a *birth-death* process $\{\eta_t\}_{t\geq 0}$ defined on the state space $\Sigma^c = \{0, 1, \ldots q\}$. This process is defined by the generator $\bar{\mathcal{L}}^c$ of the form (3.3) where the rates c_a and c_d are replaced by approximate rates

$$\bar{c}_a(k,\eta) = d_0(q - \eta(k))\,, \qquad \bar{c}_d(k,\eta) = d_0\eta(k)e^{-\beta\bar{U}(k,\eta)}\,. \tag{3.5}$$

For details we refer to [20]. The new rates have a simple interpretation in terms of fluctuations on each cell: $\bar{c}_a(k,\eta)$ describes the rate with which the coarse variable $\eta(k)$ is increased by one (i.e., adsorption of a single particle in the coarse cell C_k) and $\bar{c}_d(k,\eta)$ defines the rate with which it is decreased by one (desorption in C_k). The new interaction potential $\bar{U}(\eta)$ represents the approximation of the original interaction $U(\sigma)$.

DEFINITION 3.1. *We define the approximation $\bar{U}(k,\eta)$ of the potential $U(x,\sigma)$, (2.12), at the coarse level*

$$\bar{U}(k,\eta) = \sum_{\substack{l\in\Lambda_M^c \\ l\neq k}} \bar{J}(k,l)\eta(l) + \bar{J}(0,0)(\eta(k)-1) - \bar{h}(k)\,. \tag{3.6}$$

The coarse-grained interaction potential $\bar{J}$ is computed as the average of the pair-wise interactions between microscopic spins between the coarse cells C_k and C_l

$$\bar{J}(k,l) = \frac{1}{q^2}\sum_{x\in C_k}\sum_{y\in C_l} J(x-y)\,, \quad \text{for all } k,l\in\Lambda_M^c, \text{ such that } k\neq l, \text{ and} \tag{3.7}$$

$$\bar{J}(k,k) \equiv J(0,0) = \frac{1}{q(q-1)}\sum_{x\in C_k}\sum_{\substack{y\in C_k \\ y\neq x}} J(x-y)\,. \tag{3.8}$$

The error estimate for the projection follows directly from the assumptions on the regularity of J (or V) (2.4)–(2.5) and the Taylor expansion of the potential J. We state it as a separate lemma.

LEMMA 3.2. *Assume that J satisfies (2.4)–(2.5) then the coarse-grained interaction potential $\bar{J}$ at the coarse-graining level q approximates the potential J with the error*

$$|J(x-y) - \bar{J}(k,l)| \leq \frac{1}{L} c_d \sup_{\substack{x'\in C_k \\ y'\in C_l}} ||\nabla V(x'-y')|| \leq \mathrm{O}\left(\frac{q}{L^2}\right) \tag{3.9}$$

$$|J(x-y) - \bar{J}(0,0)| \leq \frac{1}{L} c_d \sup_{\substack{x',y'\in C_k \\ y'\neq x'}} ||\nabla V(x'-y')|| \leq \mathrm{O}\left(\frac{q}{L^2}\right)\,, \tag{3.10}$$

where $c_d = \max_{k\in\Lambda_M^c}\{\mathrm{diam}\,(C_k)\}$.

PROOF: Using the properties of the potential V, we expand V into the Taylor series,

$$V(z) = V(z') + (z-z').\nabla V(z') + \mathrm{O}(||z-z'||^2)\,.$$

Using the definition of J, (2.2) and setting $z = x-y$ and $z' = x'-y'$, where $x,x'\in C_k$ and $y,y'\in C_l$, we have

$$\begin{aligned} J(x-y) = {} & \frac{1}{q^2}\sum_{x'\in C_k}\sum_{y'\in C_l} J(x'-y') + \\ & + \frac{1}{Lq^2}\sum_{x'\in C_k}\sum_{y'\in C_l} ((x-y)-(x'-y')).\nabla V(x'-y') \\ & + \frac{1}{Lq^2}\sum_{x'\in C_k}\sum_{y'\in C_l} \mathrm{O}\left(||(x-y)-(x'-y')||^2\right)\,, \end{aligned}$$

and using the estimate $||(x-y)-(x'-y')|| \leq ||x-x'|| + ||y-y'|| \leq \max\{\mathrm{diam}\,(C_k)\}$ we obtain (3.9) in the case $k\neq l$ and similarly for $k = l$.

From Lemma 3.2 we derive the error bound for the approximation of the coarse-grained potential $\bar{U}$. Note that in the definition of U the principle contribution to the summation involves interactions within the interaction range L and thus we have the following estimate.

COROLLARY 3.3. *The microscopic potential $U(x,\sigma)$ is approximated by $\bar U(k,\eta)$, with the error*

$$\Delta_{q,N}(\bar U, U) \equiv |\bar U(k,\mathbf{T}\sigma) - U(x,\sigma)| = \mathrm{O}\left(\frac{q}{L}\right)\,, \quad \textit{for all } x\in C_k. \tag{3.11}$$

Note that this approximation represents the direct projection of the interaction kernel J on the coarse space and the contribution from fine scales are neglected. This procedure differs from the renormalisation group approach where fluctuations from the fine scales contribute to the transformed Hamiltonian. However, in the case of finite-range interaction kernels J, treated here, the above projection yields approximation of the order $O(q/L)^2$ as we discuss in the next section. The coarse interaction Hamiltonian is then given explicitly in terms of $\bar J$ and $\bar h$ as

$$\bar H(\eta) = -\frac{1}{2}\sum_{l\in\Lambda_M^c}\sum_{k\neq l}\bar J(k,l)\eta(k)\eta(l) - \frac{1}{2}\bar J(0,0)\sum_{l\in\Lambda_M^c}\eta(l)(\eta(l)-1) + \sum_{l\in\Lambda_M^c}\bar h(l)\eta(l)\,. \tag{3.12}$$

A direct calculation shows that the invariant measure of the Markov process $\{\eta_t\}_{t\geq 0}$ generated by $\bar{\mathcal{L}}^c$ is again a canonical Gibbs measure

$$\mu^c_{M,q,\beta}(d\eta) = \frac{1}{Z_{M,q,\beta}}e^{-\beta\bar H(\eta)}\mathrm{P}_{M,q}(d\eta)\,,$$

where the product measure $\mathrm{P}_{M,q}(d\eta)$ is the coarse-grained prior distribution. Note that the prior distribution is altered by coarse-graining procedure and different projection operators $\mathbf{T}$ may yield prior distributions that are computationally intractable.
For example, the coarse-grained prior arising from the uniform microscopic prior ($\rho(0) = \rho(1) = 1/2$) is the binomial distribution corresponding to q independent sites:

$$\mathrm{P}_{M,q}(d\eta) = \prod_{k\in\Lambda_M^c}\rho^c_q(d\eta(k))\,, \qquad \rho^c_q(\eta(k)=p) = \frac{q!}{p!(q-p)!}\left(\frac{1}{2}\right)^q\,.$$

The condition of detailed balance for $\{\eta_t\}_{t\geq 0}$ with respect to the measure $\mu_{M,q,\beta}$ is

$$\bar c_a(k,\eta)\mu_{M,q,\beta}(\eta) = \bar c_d(k,\eta+\delta_k)\mu_{M,q,\beta}(\eta+\delta_k)\,,$$
$$\bar c_d(k,\eta)\mu_{M,q,\beta}(\eta) = \bar c_a(k,\eta-\delta_k)\mu_{M,q,\beta}(\eta-\delta_k)\,.$$

We only verify the first relation, while the second identity is checked in analogous way. Using that $\bar H(\eta+\delta_k) - \bar H(\eta) = -\bar U(k)$ and the definitions of the rates (3.5), we have (assuming without loss of generality, $d_0 = 1$):

$$\begin{aligned}
&\bar c_a(k,\eta)\mu_{M,q,\beta}(\eta) - \bar c_d(k,\eta+\delta_k)\mu_{M,q,\beta}(\eta+\delta_k) = \\
&(q-\eta(k))e^{-\beta\bar H(\eta)}\mathrm{P}_{M,q}(\eta) - (\eta(k)+1)e^{-\beta\left(\bar H(\eta+\delta_k)+\bar U(k)\right)}\mathrm{P}_{M,q}(\eta+\delta_k) = \\
&e^{-\beta\bar H(\eta)}\left\{(q-\eta(k))\mathrm{P}_{M,q}(\eta) - (\eta(k)+1)\mathrm{P}_{M,q}(\eta+\delta_k)\right\} = \\
&\prod_{l=1,l\neq k}^{m}\eta(l)\left\{(q-\eta(k))\eta(k) - (\eta(k)+1)(\eta(k)+1)\right\}\,.
\end{aligned}$$

Since $(q-p)\rho_q(p) = (p+1)\rho_q(p+1)$, for all integers $0\leq p\leq q$, the last curly bracket is equal to zero, hence the detailed balance holds. This calculation shows that due to the specific form of the self-interaction term $\eta(l)(\eta(l)-1)$ the detailed balance condition is satisfied for the coarse Hamiltonian (3.12) and hence the fluctuations from microscopic dynamics are properly included into the coarse-grained process. The coarse-graining procedure described here satisfies basic criteria imposed on an approximating process:

(i) error control on a finite-time interval $[0,T]$. In particular, the derived coarse-grained stochastic process $\{\eta_t\}_{t\geq 0}$ approximates a pre-specified observable on a finite-time interval $[0,T]$, e.g., (3.1). In particular, time-dependent error estimates such as (5.2) can rigorously demonstrate that the process $\{\eta_t\}_{t\geq 0}$ keeps track of fluctuations from the microscopic level. Consequently expected values of certain path dependent (global) quantities can be properly estimated. We characterise approximation properties of $\{\mathbf{T}\sigma_t\}_{t\geq 0}$ by $\{\eta_t\}_{t\geq 0}$ using a suitable probability metric on the path space.

(ii) approximation of the invariant (equilibrium) measure. The invariant measure $\mu^c_{M,q,\beta}(d\eta)$ for the process $\{\eta_t\}_{t\geq 0}$ defined on $\mathcal{S}^c_{M,q}$ is close, in a suitable probability metric, to the projection of the microscopic measure $\mathbf{T}_*(\mu_{N,\beta}(d\sigma))$. In particular the error estimates in (5.1) below demonstrate that the coarse-grained process can preserve the ergodicity properties of the microscopic process within a prescribed tolerance. We also note that the coarse-graining modifies the microscopic prior $\mathrm{P}_N(d\sigma)$ in (2.6), yielding the coarse prior $\mathrm{P}_{M,q}(d\eta)$.

If the approximating process follows the basic principles (i) and (ii) we observe as a result of the error estimates presented here and in [23], that both the transient, as well as the long time dynamics are expected to be captured accurately by the coarse-graining. Although this is not a complete proof of a controlled error for infinite time, it constitutes a first rigorous step in this direction. The approximation properties are also supported by the numerics presented here and in the references.

4. Probability metrics and information theory tools. Since we propose the coarse-grained process $\{\eta_t\}_{t\geq 0}$ to be only an approximation of $\{\mathbf{T}\sigma_t\}_{t\geq 0}$ which can be computed in a fast and simple way it is necessary to define in what sense we evaluate the approximation properties. We view the approximation in coarse-graining procedure as information loss. Such approach is naturally connected to the actual computational implementation in the Monte Carlo algorithm. In this section we give a brief introduction to basic tools of information theory required in the error analysis. We define the basic notions on a probability space with the countable state space $\mathcal{S}$ but analogous properties and definitions hold for the relative entropy of measures on general probability spaces (see [9]). Although the exposition in this section is general we keep the notation consistent with the previous section. However, the reader may assume that the state space $\mathcal{S}$ does not necessarily refer to the space of spin configurations.

We consider two probability measures $\pi_1(\sigma)$ and $\pi_2(\sigma)$ on the countable state space $\mathcal{S}$, and we define the relative entropy

$$\mathcal{R}\left(\pi_1\,|\,\pi_2\right)=\sum_{\sigma\in\mathcal{S}}\pi_1(\sigma)\log\frac{\pi_1(\sigma)}{\pi_2(\sigma)}\,. \tag{4.1}$$

Using Jensen's inequality it is not difficult to show that

$$\begin{aligned}&\mathcal{R}\left(\pi_1\,|\,\pi_2\right)\geq 0\quad\text{and,}\\ &\mathcal{R}\left(\pi_1\,|\,\pi_2\right)=0\quad\text{if and only if } \pi_1(\sigma)=\pi_2(\sigma) \text{ for all } \sigma\in\mathcal{S}.\end{aligned}$$

Although the above properties of the relative entropy $\mathcal{R}\left(\pi_1\,|\,\pi_2\right)$ suggest that this quantity is a distance between the measures π_1 and π_2, it does not define a true metric since it is not symmetric, i.e., $\mathcal{R}\left(\pi_1\,|\,\pi_2\right)\neq\mathcal{R}\left(\pi_2\,|\,\pi_1\right)$ for all measures π_1, π_2. Nevertheless, there is an important inequality that allows us to use the relative entropy as a tool for estimating distance between two measures and hence use it for evaluating errors in the coarse-graining procedures. Using the relative entropy we can bound the total variation of the measures π_1 and π_2:

$$\mathcal{R}\left(\pi_1 \,|\, \pi_2\right) \geq \frac{1}{2}\left(\sum_{\sigma\in\mathcal{S}} |\pi_1(\sigma) - \pi_2(\sigma)|\right)^2 \equiv \frac{1}{2}||\pi_1 - \pi_2||^2_{\mathrm{TV}}\,, \tag{4.2}$$

and hence for any observable $\phi = \phi(\sigma)$ we have the bound

$$|\mathbb{E}_{\pi_1}\left[\phi(\sigma)\right] - \mathbb{E}_{\pi_2}\left[\phi(\sigma)\right]| \leq \sup_{\sigma} |\phi(\sigma)| \sqrt{2\mathcal{R}\left(\pi_1 \,|\, \pi_2\right)}\,. \tag{4.3}$$

The following *variational* characterisation of the relative entropy is useful in the error estimation. Given a bounded function (observable) $\phi \in L^\infty(\mathcal{S})$ defined on the state space $\mathcal{S}$ we have the natural dual pairing with the measures on $\mathcal{S}$

$$\langle \pi, \phi\rangle = \sum_{\sigma\in\mathcal{S}} \pi(\sigma)\phi(\sigma) \equiv \mathbb{E}_\pi\left[\phi\right].$$

The relative entropy (4.1) has the variational representation (see [26, pp. 338-339])

$$\mathcal{R}\left(\pi_1 \,|\, \pi_2\right) = \sup_{\phi\in L^\infty(\mathcal{S})} \left\{\langle \pi_1, \phi\rangle - \log\langle \pi_2, e^\phi\rangle\right\}\,. \tag{4.4}$$

The variational representation is used in the next section to obtain lower bounds on the relative entropy error of coarse-grained processes.

It is worth mentioning the relation between coarse graining, information theory and application of the relative entropy in the context of coarse graining. The information point of view also clearly explains the meaning of the relative entropy as a tool that estimates the loss of information. In information theory one is interested in encoding the random variable σ with values in the state space $\mathcal{S}$, and distributed according to the probability measure $\pi = \pi(\sigma)$, $\sigma \in \mathcal{S}$. The information should be encoded using symbols from a D-nary alphabet, for example only 0 and 1 in the case of the binary alphabet. Suppose that $C_D(\sigma)$ is a code/string corresponding to the value $\sigma \in \mathcal{S}$. We denote $\ell_D(\sigma)$ the length of the code needed for the state σ. Since the information is carried in the random variable σ we have to ask what is the *expected length* of the code required to capture the states of σ provided we know the distribution of σ. The expected length is given by

$$\mathbb{E}_\pi\left[\ell_D(S)\right] = \sum_{\sigma\in\mathcal{S}} \pi(\sigma)\ell_D(\sigma)\,. \tag{4.5}$$

It can be shown (see [8]) that the optimal (minimal) expected length is attained by choosing

$$\bar{\ell}_D(\sigma) = \log_D \frac{1}{\pi(\sigma)}\,. \tag{4.6}$$

Obviously, to set the optimal length for encoding the states of the random variable σ one needs to know the measure π. If we assume a wrong distribution $\omega = \omega(\sigma)$ to define the length of the code we obtain the expected length which would not be optimal. The relative entropy $\mathcal{R}\left(\pi \,|\, \omega\right)$ describes the increase of the length (4.6) due to using the wrong distribution for the random variable σ. In this sense $\mathcal{R}\left(\pi \,|\, \omega\right)$ is interpreted as the increase in descriptive complexity due to “wrong information”.

This information point of view is applicable to the analysis of coarse-graining procedures: the spin configurations σ are sampled by the Markov chain Monte Carlo algorithms and hence samples of a random variable σ with large-dimensional state space are generated. On the coarse level we sample an approximate process $\{\eta_t\}_{t\geq 0}$ instead of the exact projection $\{\mathbf{T}\sigma_t\}_{t\geq 0}$ and thus assuming a wrong measure/distribution for the random variable σ. Using the relative entropy for evaluating the approximation properties we estimate the loss of information arising from using samples of $\{\eta_t\}_{t\geq 0}$ instead of the exact coarse-grained process.

5. Error analysis and a priori estimates for coarse-grained processes. As described in the previous section we construct a new process which only approximates the projected process $\{\mathbf{T}\sigma_t\}_{t\geq 0}$. The approximation properties of such construction are quantified in this section.

We do not attempt to capture the effect of fine scales exactly and incorporate them into the coarse model through the renormalisation group transformation. Instead we construct an approximate process $\{\eta_t\}_{t\geq 0}$, with the invariant measure $\mu^c_{M,q,\beta}$. The first question which needs to be addressed is comparison and an error estimate for the exactly coarse-grained equilibrium measure, i.e., $\mathbf{T}_*\mu_{N,\beta}$, and its approximation $\mu^c_{M,q,\beta}$. We recall that $\mathbf{T}_*$ is the projection operator induced by the fine-to-coarse projection of spin variables.

5.1. Information theory estimates. The principal idea proposed in [24] is to control *the specific loss of information* quantified by the relative entropy $\mathcal{R}\left(\mu^c_{M,q,\beta}\,|\,\mathbf{T}_*\mu_{N,\beta}\right)$ between the coarse-grain equilibrium measure $\mu^c_{M,q,\beta}$ and the projected equilibrium measure $\mathbf{T}_*\mu_{N,\beta}$ of the microscopic process.

PROPOSITION 5.1 ([24], *A priori estimate*).

$$\begin{aligned}
&\frac{1}{N}\mathcal{R}\left(\mu^c_{M,q,\beta}\,|\,\mathbf{T}_*\mu_{N,\beta}\right) := \qquad (5.1)\\
&\frac{1}{N}\sum_{\eta\in\mathcal{S}^c_{M,q}}\log\left(\frac{\mu^c_{M,q,\beta}(\eta)}{\mu_{N,\beta}(\{\sigma\in{\mathcal{S}_N}^{\Lambda_N}\,|\,\mathbf{T}\sigma=\eta\})}\right)\mu^c_{M,q,\beta}(\eta) = O\left(\frac{q}{L}\right)\,.
\end{aligned}$$

This a priori estimate quantifies the dependence of the information distance, the specific relative entropy $\mathcal{R}\left(\mu\,|\,\nu\right)$, in terms of the coarse-graining ratio q and the interaction range L.

The procedure described in the previous section defines a hierarchy of coarse-grained algorithms parametrised by q. The fully resolved simulations correspond to the microscopic model $q = 1$ while the mean-field approximation is obtained in the case where $q \geq L$, i.e., when we coarse-grained beyond the interaction range of the potential. Each level of this hierarchy introduces an error since some fine-scale fluctuations are neglected.

For the comparison of the processes $\{\mathbf{T}\sigma_t\}_{t\geq 0}$ and $\{\eta_t\}_{t\geq 0}$ we need to carry out a similar a priori analysis on the coarse path space $\mathcal{D}(\mathcal{S}^c_{M,q})$, i.e., on the space of all right-continuous paths $\eta_t : [0,\infty)\to\mathcal{S}^c_{M,q}$. Above we have presented estimates for the exact coarse graining $\mathbf{T}_*\mu_{N,\beta}$ of the invariant measure $\mu_{N,\beta}$ and its approximation $\mu^c_{M,q,\beta}$ computed in terms of the coarse Hamiltonian. In a similar way we treat the measures on the path space: we denote $Q_{\sigma_0,[0,T]}$ the measure on $\mathcal{D}(\mathcal{S}_N)$ for the process on the interval $[0,T]$, $\{\sigma_t\}_{t\in[0,T]}$ with the initial distribution σ_0. Similarly $Q^c_{\eta_0,[0,T]}$ denotes the measure on the coarse path space $\mathcal{D}(\mathcal{S}^c_{M,q})$. With a slight abuse of notation we also use $\mathbf{T}_*Q$ to denote the projection of the measure Q on the coarse path space, i.e., the exact coarsening of the measure Q. The fully rigorous analysis on the path space is more involved and we refer to [23]. For the sake of completeness we only state the main a priori estimate.

PROPOSITION 5.2 ([23]). *Suppose the process $\{\eta_t\}_{t\in[0,T]}$, given by the coarse generator $\bar{\mathcal{L}}^c$, defines the coarse approximation of the microscopic process $\{\sigma_t\}_{t\in[0,T]}$ then for any $q<L$ and N, $Mq=N$, the information loss as $q/L\to 0$ is*

$$\frac{1}{N}\mathcal{R}\left(\mathbf{T}_* Q_{\mathbf{T}_*\sigma_0,[0,T]}\,|\,Q^c_{\eta_0,[0,T]}\right) = T\,O\left(\frac{q}{L}\right) \tag{5.2}$$

REMARK: The detailed proof of this information estimate (see [23]) reveals that no control of fluctuations of the process $\{\sigma_t\}_{t\geq 0}$ is necessary for the estimate. Consequently the estimate is very robust and, as long as q/L is small, the approximation by the coarse-graining scheme yields a small error independent of the potential V or the initial distribution σ_0. Although the estimate is for finite times $[0,T]$ only, and grows with T, in many cases the system nucleates a new phase at the initial stage of its evolution and thus the estimate ensures good approximation of the nucleation phase.

It is worth noticing that the relative entropy estimate clearly demonstrate limitations of the coarse-graining method since it gives the error of order one for short-range interactions (the nearest neighbour interaction corresponds to $L=1$). On the other hand the analysis using the relative entropy (information) distance identifies the small parameter in the asymptotic expansion of the blocking error, namely the ratio q/L.

The next estimate provides a lower bound for the loss of information in terms of coarser observables:

PROPOSITION 5.3 (*Lower bound*). *Suppose the process $(\{\eta_t\}_{t\in[0,T]},\bar{\mathcal{L}}^c)$, defined by the coarse-graining operator $\mathbf{T}$ with coarse-graining parameters $Mq=N$, is the coarse approximation of the microscopic process $\{\sigma_t\}_{t\in[0,T]}$. Let $\mathbf{T}^{M',q'}$ be another coarse-graining operator, such that $M'\leq M$, $M'q'=Mq=N$. Then the following estimate for the invariant microscopic measure $\mu_{N,\beta}$ and the coarse approximation $\mu^c_{M,q,\beta}$ holds*

$$\mathcal{R}\left(\mu^c_{M,q,\beta}\,|\,\mathbf{T}_*\mu_{N,\beta}\right) \geq \mathcal{R}\left(\mathbf{T}^{M',q'}_*\mu^c_{M,q,\beta}\,|\,\mathbf{T}^{M',q'}_*\mu_{N,\beta}\right)\,. \tag{5.3}$$

Moreover, on any finite-time interval $[0,T]$

$$\mathcal{R}\left(\mathbf{T}_* Q_{\mathbf{T}\sigma_0,[0,T]}\,|\,Q^c_{\eta_0,[0,T]}\right) \geq \mathcal{R}\left(\mathbf{T}^{M',q'}_* Q_{\mathbf{T}\sigma_0,[0,T]}\,|\,\mathbf{T}^{M',q'}_* Q^c_{\eta_0,[0,T]}\right)\,. \tag{5.4}$$

PROOF: We first recall the variational formulation for the relative entropy

$$\mathcal{R}\left(\mu\,|\,\nu\right) = \sup_f\left\{\int f\,d\mu - \log\int e^f\,d\nu\right\}\,, \tag{5.5}$$

where the supremum is over all bounded functions in the space where the measures are defined. This inequality now readily implies the result since

$$\mathcal{R}\left(\mu\,|\,\nu\right) \geq \sup_{f\circ\mathbf{T}}\left\{\int f\circ\mathbf{T}\,d\mu - \log\int e^{f\circ\mathbf{T}}\,d\nu\right\} = \mathcal{R}\left(\mathbf{T}_*\mu\,|\,\mathbf{T}_*\nu\right) \tag{5.6}$$

where $\mathbf{T}$ is the projection operator (super-scripts omitted) in the statement of the proposition.

REMARK: This estimate provides a lower bound for the loss of information in terms of coarser observables, hence the condition $M'\leq M$ where $M'q'=Mq=N$. For instance if $M'=1, q'=N$ the measures $\mathbf{T}^{M',q'}_*\mu^c_{M,q,\beta}$ and $\mathbf{T}^{M',q'}_*\mu_{N,\beta}$ are the PDFs of the total coverage with respect to the coarse-grained (essentially mean field with a noise) and the microscopic Gibbs states respectively. We characterise such an estimate as *a priori* since the bound depends on the exact microscopic process, in analogy to bounds for approximations to PDEs which depend on the Sobolev norm of the exact solution, [11]. At first glance it may appear that such an estimate is hard to implement since it depends on the exact microscopic MC. However, for relatively small systems where microscopic MC can be carried out, the bound (5.3) can provide a lower bound on the loss of information, as well as a sense on how sharp are the upper bounds given by a posteriori estimates. More specifically when M' is small , i.e., $M' = 1, 2, 3 \ldots$ etc., the PDFs can be calculated as a histogram by MC and subsequently the relative entropy in the lower bound is straightforward to compute.

5.2. Microscopic reconstruction and weak convergence estimates. In many practical MC simulations the main goal is to estimate averages (expected values) of specific observables. Therefore it is natural to analyze the weak approximation properties of the coarse-graining procedure. The weak error is defined as the quantity $e_w \equiv |\mathbb{E}_S[\psi(\mathbf{T}\sigma_t)] - \mathbb{E}_S[\psi(\eta_t)]|$, where the expectation $\mathbb{E}_S[\cdot]$ is defined for the path conditioned on the initial configuration $\eta_0 = \mathbf{T}\sigma_0 = S$. Alternatively we can compare the microscopic process $\{\sigma_t\}_{t\geq 0}$ with its synthetic process $\{\gamma_t\}_{t\geq 0}$ which is reconstructed from the coarse process $\{\eta_t\}_{t\geq 0}$. The weak error is then defined as $e_w \equiv |\mathbb{E}_S[\phi(\sigma_t)] - \mathbb{E}_S[\phi(\gamma_t)]|$, where the expectation $\mathbb{E}_S[\cdot]$ is now defined for the path conditioned on the initial configuration $\sigma_0 = S$. Here and in what follows ϕ denotes a test function (observable) on the fine level while ψ is used for a test function on the coarse level. Theorem 5.8 and Corollary 5.9 quantify the rate of convergence for the weak error on both levels as $q/L \to 0$. We refer to [22] for error estimates in the weak topology between microscopic MC algorithms and therein derived SDE approximations.

Before we formulate the proposition and proceed with the proof it is worth clarifying the difficulty of comparing the projected process $\{\mathbf{T}\sigma_t\}_{t\geq 0}$ with the approximating process $\{\eta_t\}_{t\geq 0}$. The projection $\mathbf{T}\sigma_t$ of the microscopic process on the coarse grid does not necessarily define a Markov process. On the other hand the approximating process $\{\eta_t\}_{t\geq 0}$ is constructed as a Markov process $(\{\eta_t\}_{t\geq 0}, \bar{\mathcal{L}}^c)$ with the generator $\bar{\mathcal{L}}^c$ defined by (3.5). To circumvent the technical difficulty the authors in [23] suggested to construct an auxiliary process $\{\gamma_t\}_{t\geq 0}$ as an intermediate step in the estimation of the relative entropy between the processes $\{\sigma_t\}_{t\geq 0}$ and $\{\eta_t\}_{t\geq 0}$. We adopt the same strategy in order to make comparison between observables which depend on Markovian processes $\{\sigma_t\}_{t\geq 0}$ and $\{\gamma_t\}_{t\geq 0}$. The process $\{\gamma_t\}_{t\geq 0}$ can be directly reconstructed from the coarse-grained process $\{\eta_t\}_{t\geq 0}$. Thus we are lead to the definition of the *synthetic microscopic (Markov) process* $\{\gamma_t\}_{t\geq 0}$ associated with the process $\{\sigma_t\}_{t\geq 0}$.

DEFINITION 5.4 (Synthetic microscopic process). *The auxiliary process $\{\gamma_t\}_{t\geq 0}$ is defined on the microscopic configuration space $\mathcal{S}_N$ by the generator $\mathcal{L}^\gamma : L^\infty(\mathcal{S}_N) \to \mathbb{R}$*

$$(\mathcal{L}^\gamma \phi)(\sigma) = \sum_{x\in\Lambda_N} c_\gamma(x,\sigma)(\phi(\sigma^x) - \phi(\sigma)), \tag{5.7}$$

where the rate function $c_\gamma(x,\sigma)$ is defined in terms of the coarse-grained interaction potential

$$c_\gamma(x,\sigma) = d_0(1-\sigma(x)) + d_0\sigma(x)e^{-\beta\bar{U}(k(x),\mathbf{T}\sigma)}.$$

The coarse-grained interaction potential $\bar{U}(k,\eta)$ has been defined in (3.6). The piece-wise constant interpolation is used to extend the function $\bar{U}(.,.)$ from the coarse lattice to the fine lattice. We denote $k(x)$ to be the cell index of the cell to which the site x belongs, i.e., $x \in C_{k(x)}$.

The properties of $\{\gamma_t\}_{t\geq 0}$ were studied in [23] and it was proved that:

(i) the coarse-grained projection $\{\mathbf{T}\gamma_t\}_{t\geq 0}$ of the Markov process $(\{\gamma_t\}_{t\geq 0}, \mathcal{L}^\gamma)$ is still a Markov process.

(ii) the processes $\{\mathbf{T}\gamma_t\}_{t\geq 0}$ and $\{\eta_t\}_{t\geq 0}$ have the same transition rates. Hence, whenever the processes have the same initial distribution they induce the same probability measure on the coarse-grained path space $\mathcal{D}(\mathcal{S}^c_{M,q})$. If we define $Q^c_{\eta_0}(\eta,t)$ and $Q_{\gamma_0}(\gamma,t)$ to be the probability measures of the Markov processes $\{\eta_t\}_{t\geq 0}$ and $\{\gamma_t\}_{t\geq 0}$ respectively (conditioned on the initial condition $\eta_0 = \mathbf{T}\gamma_0$), then for all $t > 0$ we have the projection

$$Q^c_{\eta_0}(\eta,t) = \mathbf{T}_* Q_{\gamma_0}(\gamma,t) \equiv \sum_{\{\gamma \mid \mathbf{T}\gamma=\eta_t\}} Q_{\gamma_0}(\gamma,t)\,,$$

provided this relation is satisfied at $t=0$. Hence this property allows us to compare the processes in a path-wise way.

(iii) the microscopic process $\{\gamma_t\}_{t\geq 0}$ can be reconstructed from the approximating coarse process $\{\eta_t\}_{t\geq 0}$. Such reconstruction is an inverse procedure to the projection from fine to coarse configuration space. In such a way we can compare the original microscopic process with the approximation on the coarse configuration space. A simple choice of a reconstruction operator is to distribute spins $\gamma_t(x)$ for $x \in C_k$ uniformly so that $\mathbf{T}\gamma_t|_{C_k} = \eta_t(k)$.

REMARK: It is conceivable that the synthetic process $\{\gamma_t\}_{t\geq 0}$ can be used not only as a technical tool but as a systematic procedure for reconstructing the microscopic process $\{\sigma_t\}_{t\geq 0}$ for the purpose of model refinement or adaptivity since, as shown in Theorem 5.8, the reconstruction is done under rigorous error estimates. In the estimates derived below we deal with a specific class of test functions $\phi \in L^\infty(\mathcal{S}_N)$ which depend only on the coarse variable $\eta = \mathbf{T}\sigma$, in other words we impose the assumption

(A1) $\phi(\sigma) = \psi(\mathbf{T}\sigma)\,,$ where $\psi \in L^\infty(\mathcal{S}^c_{M,q})$, and (5.8)

$$\sum_{x\in\Lambda_N} |\partial_x \phi(\sigma)| \leq C\,, \quad \text{where } C \text{ is a constant independent of } N. \tag{5.9}$$

REMARK: Observables, such as, for example, the total coverage, used in the numerical simulations satisfy this assumption.

The principal tool for analysing the weak error is its representation in terms of solutions to the final value problem on $\mathcal{S}_N$

$$\partial_t v(t,\sigma) + \mathcal{L}v(t,\sigma) = 0\,, \qquad v(T,.) = \phi(.)\,, \quad \text{for } t < T\,,$$

where $\mathcal{L}$ is a generator of the Markov semigroup that defines the lattice dynamics. Before we state the main estimate of the weak error and its proof we need several preliminary lemmata that characterize properties of the semigroup generated by the operator $\mathcal{L}$ defined by (2.7). The specific calculations are better presented by introducing an alternative notation for the generator $\mathcal{L}$. We define an operator of discrete differentiation for functions $f \in L^\infty(\mathcal{S}_N)$

$$\partial_x f(\sigma) \equiv f(\sigma^x) - f(\sigma)\,, \quad \text{for all } x \in \Lambda_N, \tag{5.10}$$

and we introduce two vectors indexed by the lattice sites $x \in \Lambda_N$

$$\nabla_\sigma f(\sigma) \equiv (\partial_x f(\sigma))_{x\in\Lambda_N}\,, \quad \mathbf{c}(\sigma) \equiv (c(x,\sigma))_{x\in\Lambda_N}\,.$$

The scalar product is defined in the natural way as $\mathbf{c}(\sigma)\cdot\nabla_\sigma f(\sigma) \equiv \sum_{x\in\Lambda_N} c(x,\sigma)\partial_x f(\sigma)$. Using this notation we write

$$\mathcal{L}f(\sigma) = \mathbf{c}(\sigma)\cdot\nabla_\sigma f(\sigma)\,, \quad \text{for all } \sigma \in \mathcal{S}_N. \tag{5.11}$$

The space of functions defined on the configuration space $\mathcal{S}_N$ is equipped with the strong L^∞ topology given by the norm $||f||_\infty \equiv \sup_\sigma\{f(\sigma)\}$.

To prove the estimate in Theorem 5.8 we need an estimate for the difference operator ∇_σ stated here as a separate lemma.

LEMMA 5.5. *Let $v(t,\sigma)$ be the solution of*

$$\partial_t v + \mathcal{L}v = 0\,, \quad v(T,\sigma) = \phi(\sigma)\,, \quad \textit{for } t < T, \tag{5.12}$$

on a given interval $t \leq T$, then

$$\sum_{x\in\Lambda_N} ||\partial_x v(t,.)||_\infty \leq C_T \sum_{x\in\Lambda_N} ||\partial_x \phi||_\infty\,. \tag{5.13}$$

Moreover, the constant C_T depends exponentially on the final time T.

PROOF: Using the notation introduced above and the definition of $\mathcal{L}$ we recast the evolution equation (5.12) into a familiar form of a transport equation on the configuration space

$$\partial_t v + \mathbf{c}(\sigma)\cdot\nabla_\sigma v = 0\,, \quad \sigma\in\mathcal{S}_N\,,\ t>0\,. \tag{5.14}$$

Subtracting (5.14) for $v(t,\sigma^x)$ and $v(t,\sigma)$ we have

$$\partial_t(v(t,\sigma^x) - v(t,\sigma)) + \mathbf{c}(\sigma)\cdot(\nabla_\sigma v(t,\sigma^x) - \nabla_\sigma v(t,\sigma)) + (\mathbf{c}(\sigma^x) - \mathbf{c}(\sigma))\cdot\nabla_\sigma v(t,\sigma^x) = 0\,,$$

which we write as

$$\partial_t\left(\partial_x v(t,\sigma)\right) + \mathbf{c}(\sigma)\cdot\nabla_\sigma\left(\partial_x v(t,\sigma)\right) + \partial_x\mathbf{c}(\sigma)\cdot\nabla_\sigma v(t,\sigma^x) = 0\,. \tag{5.15}$$

Next we derive L^∞-bounds for the discrete derivatives $\partial_x\mathbf{c}(\sigma)$ using the explicit definition of the rates $c(x,\sigma)$ in (2.11). For each component, indexed by $z\in\Lambda_N$, of the vector $\mathbf{c}(\sigma)$ we have

$$\partial_x c(z,\sigma) = c(z,\sigma^x) - c(z,\sigma) = (1-\sigma^x(z)) + \sigma^x(z)e^{-U(z,\sigma^x)} - (1-\sigma(z)) + \sigma(z)e^{-U(z,\sigma)}\,.$$

For the spin-flip dynamics, i.e., $\sigma^x(y) = 1-\sigma(y)$ if $x=y$ and $\sigma^x(y) = \sigma(y)$ otherwise, a straightforward calculation gives $\partial_x U(z,\sigma) \equiv U(z,\sigma^x) - U(z,\sigma) = J(z-x)(1-2\sigma(x))$ if $z\neq x$ and it is equal zero otherwise. Thus the discrete derivate $\partial_x\mathbf{c}(\sigma)$ is

$$\partial_x c(z,\sigma) = \begin{cases} (2\sigma(x)-1)(1-e^{-U(x,\sigma)})\,, & \text{for } z = x,\\ \sigma(z)e^{-U(z,\sigma)}\left(1-e^{J(x-z)(1-2\sigma(x))}\right) & \text{if } z\neq x. \end{cases}$$

Recalling the definition (2.3) of the interaction potential J we have that $J(z-x)\sim 1/L$ for $|z-x|\leq L$ and $J=0$ otherwise. Hence we derived L^∞-bounds for the discrete derivative of the rates

$$\partial_x c(z,\sigma) \sim \begin{cases} \mathrm{O}(1)\,, & \text{for } z=x,\\ \mathrm{O}(1/L)\,, & \text{for } |z-x|<L,\\ 0\,, & \text{otherwise.}\end{cases} \tag{5.16}$$

Going back to the equation (5.15) we have for all $x\in\Lambda_N$

$$\partial_t\left(\partial_x v(t,\sigma)\right) + \mathcal{L}\partial_x v(t,\sigma) + \sum_{z\in\Lambda_N}\partial_x c(z,\sigma)\partial_z v(t,\sigma^x) = 0\,. \tag{5.17}$$

The estimates in (5.16) imply

$$\partial_t\partial_x v(t,\sigma) + \mathcal{L}\partial_x v(t,\sigma) + \mathrm{O}(1)\partial_x v(t,\sigma^x) + \mathrm{O}\left(\frac{1}{L}\right)\sum_{\substack{z\in\Lambda_N\\|z-x|\leq L}}\partial_z v(t,\sigma^x) = 0\,, \quad (5.18)$$

and we have for all $\sigma\in\mathcal{S}_N$ the solution formula

$$\partial_x v(t,\sigma) = e^{t\mathcal{L}}[\partial_x v(0,\sigma)] + \int_t^T e^{(s-t)\mathcal{L}}[\mathrm{O}(1)\partial_x v(s,\sigma^x) + \mathrm{O}(1/L)\sum_{|z-x|\leq L}\partial_z v(s,\sigma^x)]\,ds\,.$$

By the contractive property of the semigroup $e^{t\mathcal{L}}$ we have the estimate

$$||\partial_x v(t,\cdot)||_\infty \leq ||\partial_x v(0,\cdot)||_\infty + \int_t^T \mathrm{O}(1)||\partial_x v(s,\cdot)||_\infty\,ds + \int_t^T \mathrm{O}(1/L)\sum_{|z-x|\leq L}||\partial_z v(s,\cdot)||_\infty\,ds\,,$$

for all $x\in\Lambda_N$. Thus summing over all $x\in\Lambda_N$ we obtain

$$\sum_{x\in\Lambda_N}||\partial_x v(t,\cdot)||_\infty \leq \sum_{x\in\Lambda_N}||\partial_x v(0,\cdot)||_\infty + \\ + \int_t^T\Big(\mathrm{O}(1)\sum_{x\in\Lambda_N}||\partial_x v(s,\cdot)||_\infty + \mathrm{O}(1/L)\sum_{x\in\Lambda_N}\sum_{|z-x|\leq L}||\partial_z v(s,\cdot)||_\infty\Big)\,ds\,,$$

where the last double sum in the integrand is bounded by $2L\sum_x||\partial_x v(s,\cdot)||_\infty$. Hence by setting $\theta(t)=\sum_x||\partial_x v(t,\cdot)||_\infty$ we have

$$\theta(t)\leq\theta(0)+\int_t^T\mathrm{O}(1)\theta(s)\,ds\,,$$

from which, by using Gronwall's inequality, we obtain the bound

$$\theta(t)\leq e^{c(T-t)}\theta(T)\,,$$

which concludes the proof of (5.5).

Next we establish an L^∞-bound for discrete derivatives of solutions generated by semigroups $e^{t\mathcal{L}}$ and $e^{t\mathcal{L}^\gamma}$.

LEMMA 5.6. *Let $u(t,\sigma)$ be the solution of*

$$\partial_t u + \mathcal{L}u = 0\,,\quad u(T,.)=\phi\,,\ \textit{for } t<T,$$

and let $v(t,\sigma)$ solves

$$\partial_t v + \mathcal{L}^\gamma v = 0\,,\quad v(T,.)=\psi\,,\ \textit{for } t<T,$$

then for any $t\leq T$ the following estimate holds

$$\sum_{x\in\Lambda_N}||\partial_x u(t,\cdot)-\partial_x v(t,\cdot)||_\infty \leq C_1(T)\sum_{x\in\Lambda_N}||\partial_x\phi-\partial_x\psi||_\infty + C_2(T)\left(\frac{q}{L}\right)\,. \quad (5.19)$$

The constants C_1 and C_2 are independent of q and L but depend exponentially on the final time T.

PROOF: We use the same approach and notation as in the proof of Lemma 5.5. Subtracting the evolution equations and defining $w_x(t,\sigma) \equiv \partial_x u(t,\sigma) - \partial_x v(t,\sigma)$, $\mathbf{w}(t,\sigma) \equiv (w_x(t,\sigma))_{x\in\Lambda_N}$ we have

$$\partial_t w_x(t,\sigma) + \mathcal{L} w_x(t,\sigma) + \tag{5.20}$$
$$(\mathbf{c}_\gamma(\sigma) - \mathbf{c}(\sigma))\cdot\nabla_\sigma v(t,\sigma^x) + \tag{5.21}$$
$$+\partial_x \mathbf{c}(\sigma)\cdot \mathbf{w}(t,\sigma^x) + \tag{5.22}$$
$$(\partial_x \mathbf{c}(\sigma) - \partial_x \mathbf{c}_\gamma(\sigma))\cdot\nabla_\sigma v(t,\sigma^x) = 0\,. \tag{5.23}$$

From Lemma 5.5 we have estimates for the terms involving $\nabla_\sigma v(t,.)$ (notice that the lemma essentially gives the estimate of $||\nabla_\sigma v(t,.)||_\infty$). Furthermore, from the definition of rates $c(x,\sigma)$ and $c_\gamma(x,\sigma)$ direct calculation (similar to that used in the proof of Lemma 5.5) yields the estimate

$$||\mathbf{c} - \mathbf{c}_\gamma||_\infty = \mathrm{O}\left(\frac{q}{L}\right)\,, \tag{5.24}$$

which allows us to control (5.21) and (5.23). Term (5.22) is treated in the same way as a similar term in the proof of Lemma 5.5. Hence, for all $x\in\Lambda_N$ we obtain

$$\partial_t w_x(t,\sigma) + \mathcal{L} w_x(t,\sigma) + \mathrm{O}(1/L)\sum_{|z-x|\le L} w_x(z,\sigma^x) \le \mathrm{O}(q/L)||\partial_x v(t,.)||_\infty\,.$$

Similarly as in the proof of Lemma 5.5 we complete the proof by summing over $x\in\Lambda_N$ and applying Gronwall's inequality.

Since we are comparing the process $\{\sigma_t\}_{t\ge0}$ with the process $\{\gamma_t\}_{t\ge0}$, which is defined only up to the equivalence given by the projection operator $\mathbf{T}$ we have to establish uniqueness of solutions for initial data satisfying the assumption $(A1)$.

LEMMA 5.7. *Let $\phi \in L^\infty(\mathcal{S}_N)$, $\psi \in L^\infty(\mathcal{S}^c_{M,q})$ be test functions satisfying (A1). Assume that $v(t,\gamma)$ is the solution of the final value problem*

$$\partial_t v + \mathcal{L}^\gamma v = 0\,,\quad v(T,\gamma) = \phi(\gamma) = \psi(\mathbf{T}\gamma)\,, \tag{5.25}$$

then for all γ, $\gamma'\in\mathcal{S}_N$ such that $\mathbf{T}\gamma = \mathbf{T}\gamma'$

$$v(t,\gamma) = v(t,\gamma')\,,\quad \textit{for all } t\le T. \tag{5.26}$$

PROOF: For convenience we write $v(t,\gamma) = v(t,\mathbf{T}\gamma)$. Given a configuration $\gamma\in\mathcal{S}_N$ we can reconstruct an arbitrary configuration $\gamma'\in\mathcal{S}_N$ such that $\mathbf{T}\gamma' = \mathbf{T}\gamma$ by considering a permutation $\pi:\Lambda_N\to\Lambda_N$, $\pi = (\pi_1,\ldots,\pi_M)$ such that

$$\pi_k : C_k \to C_k\,,\quad k = 1,\ldots,M\,.$$

The action of π on the configuration space is defined in a natural way $\gamma' = \gamma\circ\pi$, or equivalently $\gamma'(x) = \gamma(\pi x)$. Since the permutation does not change the total spin in the cell we have $\mathbf{T}\gamma\circ\pi = \mathbf{T}\gamma$. Hence we write $v(t,\gamma') = v(t,\gamma\circ\pi)$ and $v(T,\gamma\circ\pi) = v(T,\gamma) = \psi(\mathbf{T}\gamma)$. It is sufficient to show that the function $u(t,\gamma)\equiv v(t,\gamma\circ\pi)$ is a solution of (5.25). From the uniqueness of solutions to (5.25) we conclude immediately that $u(t,\gamma) = v(t,\gamma)$. From the definition of the generator $\mathcal{L}^\gamma$ we have

$$\partial_t v(t,\gamma\circ\pi) + \sum_{k\in\Lambda^c_M}\sum_{x\in C_k} c_\gamma(x,\gamma\circ\pi)(v(t,(\gamma\circ\pi)^x) - v(t,\gamma\circ\pi)) = 0\,. \tag{5.27}$$

Recall the definition of the rate c_γ

$$c_\gamma(x,\gamma) = d_0(1-\gamma(x)) + d_0\gamma(x)e^{-\beta\bar{U}(k(x),\mathbf{T}\gamma)}\,,$$

and denote $c_\gamma(x,\gamma)$ by $C_\gamma(\gamma(x),k,\mathbf{T}\gamma)$ to emphasise the dependence on $\gamma(x)$, k, and $\eta = \mathbf{T}\gamma$ only. Thus the inner summation in (5.27) becomes

$$\sum_{x\in C_k} C_\gamma(\gamma\circ\pi,k,\mathbf{T}\gamma)(v(t,(\gamma\circ\pi)^x) - v(t,\gamma\circ\pi))\,. \tag{5.28}$$

On the other hand the definition of spin-flip dynamics leads to

$$(\gamma\circ\pi)^x(z) = \begin{cases} \gamma(\pi z) & z\neq x\,, \\ 1-\gamma(\pi x) & z = x\,, \end{cases} \quad \text{while } \gamma^{(\pi x)}(\pi z) = \begin{cases} \gamma(\pi z) & z\neq x\,, \\ 1-\gamma(\pi x) & z = x\,. \end{cases} \tag{5.29}$$

Hence we obtain

$$(\gamma\circ\pi)^x(z) = \gamma^{(\pi x)}(\pi z) = \left(\gamma^{\pi x}\circ\pi\right)(z)\,, \tag{5.30}$$

and substituting to the expression (5.28) leads to

$$\begin{aligned}
&\sum_{x\in C_k} C_\gamma(\gamma(\pi x),k,\mathbf{T}\gamma)(v(t,(\gamma\circ\pi)^x) - v(t,\gamma\circ\pi)) = \\
&= \sum_{x\in C_k} C_\gamma(\gamma(\pi x),k,\mathbf{T}\gamma)(v(t,\gamma^{\pi x}\circ\pi) - v(t,\gamma\circ\pi)) = \\
&= \sum_{y\in C_k} C_\gamma(\gamma(y),k,\mathbf{T}\gamma)(v(t,\gamma^{y}\circ\pi) - v(t,\gamma\circ\pi)) = \\
&= \sum_{y\in C_k} C_\gamma(\gamma(y),k,\mathbf{T}\gamma)(u(t,\gamma^{y}) - u(t,\gamma))\,.
\end{aligned}$$

Thus we have shown that

$$\partial_t u(t,\gamma) + \sum_{k\in\Lambda_M^c}\sum_{x\in C_k} c_\gamma(x,\gamma)(u(t,\gamma^x) - u(t,\gamma)) = 0\,.$$

Recalling the definition of $u(t,\gamma)$ we obtain that $v(t,\gamma\circ\pi)$ also solves (5.25). The uniqueness of solutions to (5.25) implies that $v(t,\gamma\circ\pi) = v(t,\gamma)$ for all γ or $v(t,\gamma') = v(t,\gamma)$ for all γ' such that $\mathbf{T}\gamma' = \mathbf{T}\gamma$.

Now we can formulate and prove the weak error estimate that allows us to compare the microscopic process and its coarse-level approximation. We estimate the weak error on the microscopic level by comparing the microscopic process and its synthetic process.

THEOREM 5.8 (*Weak error*). *Let $\phi\in L^\infty(\mathcal{S}_N)$ be a test function (observable) on the microscopic space satisfying (A1) and let $(\{\gamma_t\}_{t\geq 0},\mathcal{L}^\gamma)$ be the synthetic Markov process (in the sense of Definition 5.4) of the microscopic process $(\{\sigma_t\}_{t\geq 0},\mathcal{L})$ with the initial condition $\sigma_0 = S$, then the weak error satisfies, for $0<T<\infty$,*

$$\left|\mathbb{E}_S\left[\phi(\sigma_T)\right] - \mathbb{E}_S\left[\phi(\gamma_T)\right]\right| \leq C_T\left(\frac{q}{L}\right)^2\,, \tag{5.31}$$

where the constant C_T is independent of q and L but depends on T.

PROOF: The two ingredients of the proof, the Feynman-Kac formula and the martingale property, follow from the standard properties of Markov processes (see for example [26]). If we define, for the microscopic process $\{\sigma_t\}_{t\geq 0}$ defined by the generator $\mathcal{L}$, the function

$$u(t,S)=\mathbb{E}\left[\phi(\sigma_T)\,|\,\sigma_t=S\right],$$

then from the Feynman-Kac formula with the zero potential follows that the function $u(t,S)$ solves the final value problem

$$\partial_t u+\mathcal{L}u=0\,,\qquad u(T,.)=\phi\,,\;\; t<T\,. \tag{5.32}$$

On the other hand the martingale property implies that for any smooth function $v(t,S)$ and the process $\{\gamma_t\}_{t\geq 0}$ with the generator $\mathcal{L}^\gamma$ we have

$$\mathbb{E}_S\left[v(T,\gamma_T)\right]=\mathbb{E}_S\left[v(0,\gamma_0)\right]+\int_0^T\mathbb{E}_S\left[(\partial_s+\mathcal{L}^\gamma)v(s,\gamma_s)\right]ds\,.$$

The definition of $u(t,S)$ leads to the representation of the error $|\mathbb{E}_S\left[\phi(\sigma_T)\right]-\mathbb{E}_S\left[\phi(\gamma_T)\right]|$ by $e_w=|\mathbb{E}_S\left[u(0,S)\right]-\mathbb{E}_S\left[u(T,\gamma_T)\right]|$ and hence

$$e_w=\left|\int_0^T\mathbb{E}_S\left[\left(\partial_s+\mathcal{L}^\gamma\right)u(s,\gamma_s)\right]ds\right|\,.$$

The function $u(t,S)$ solves the equation $\partial_t u=-\mathcal{L}u$ thus we obtain

$$\begin{aligned}\mathbb{E}_S\left[\phi(\sigma_T)-\phi(\gamma_T)\right]=\int_0^T\mathbb{E}_S\left[\mathcal{L}^\gamma u(t,\gamma_t)-\mathcal{L}u(t,\gamma_t)\right]dt=\\ =\int_0^T\mathbb{E}_S\left[\sum_{x\in\Lambda_N}\left(c(x,\gamma_t)-c_\gamma(x,\gamma_t)\right)\partial_x u(t,\gamma_t)\right]dt\,.\end{aligned}$$

We split the summation $\sum_{x\in\Lambda_N}$ which gives us

$$\begin{aligned}\mathbb{E}_S\left[\phi(\sigma_T)-\phi(\gamma_T)\right]=\int_0^T\mathbb{E}_S\left[\sum_{k\in\Lambda_M^c}\sum_{x\in C_k}(c(x,\gamma_t)-c_\gamma(x,\gamma_t))\partial_x u(t,\gamma_t)\right]dt=\\ =\int_0^T\mathbb{E}_S\left[\sum_{k\in\Lambda_M^c}\sum_{x\in C_k}\gamma_t(x)(e^{-\beta U(x,\gamma_t)}-e^{-\beta\bar{U}(k(x),\mathbf{T}\gamma_t)})(\partial_k v(t,\mathbf{T}\gamma_t)+R_T^{q,L}(x))\right]dt\,.\end{aligned}$$

Here we need to replace $\partial_x u$ by the $\partial_x v$, where v solves the final value problem (5.32) with $\mathcal{L}$ replaced by $\mathcal{L}^\gamma$. From Lemma 5.6 we know that the error term $R_T^{q,L}(x)=\partial_x u(t,\gamma)-\partial_x v(t,\gamma)$ is controlled by $\mathrm{O}(q/L)$ in $||\cdot||_\infty$. Furthermore, Lemma 5.7 guarantees that with the final condition ϕ which satisfies Assumption (A1) the solution depends only on $\mathbf{T}\gamma$ and hence we can replace the discrete difference $\partial_x v$ by the difference $\partial_k v(t,\eta)\equiv v(t,\eta+\delta_k)-v(t,\eta)$, where $\eta=\mathbf{T}\gamma$. Next we expand the exponentials to obtain

$$\Gamma(k,\gamma)\equiv\sum_{x\in C_k}\beta\gamma(x)e^{-\beta\bar{U}(k(x),\mathbf{T}\gamma)}\left(\Delta(\bar{U},U)+\frac{1}{2}\beta^2\Delta^2(\bar{U},U)+\mathrm{O}\left(\beta^3\Delta^3(\bar{U},U)\right)\right)\,,$$

and we recast the error representation into

$$\mathbb{E}_S\left[\phi(\sigma_T)-\phi(\gamma_T)\right]=\int_0^T \mathbb{E}_S\left[\sum_{k\in\Lambda_M^c}\Gamma(k,\gamma_t)\partial_k v(t,\mathbf{T}\gamma_t)+\sum_{x\in\Lambda_N}(c(x,\gamma_t)-c_\gamma(x,\gamma_t))R_T^{q,L}(x)\right]dt$$

$$=\int_0^T \mathbb{E}_S\left[q\sum_{k\in\Lambda_M^c}\partial_k v(t,\eta_t)\,\mathbb{E}\left[\Gamma(k,\gamma)\big|\mathbf{T}\gamma=\eta_t\right]\right]dt+ \tag{5.33}$$

$$+\int_0^T \mathbb{E}_S\left[\sum_{x\in\Lambda_N}(c(x,\gamma_t)-c_\gamma(x,\gamma_t))R_T^{q,L}(x)\right]dt\,. \tag{5.34}$$

Assumption (A1) and Lemma 5.5 imply that the term $q\sum_{k\in\Lambda_M^c}\partial_k v(t,\eta_t)$ is bounded. To estimate the conditional expectation we use the property of the reconstruction operator for the process $\{\gamma_t\}_{t\geq 0}$, in particular on each cell $\gamma_t(x)$ is reconstructed from $\eta_t(k)$ by assuming a "local" equilibrium and distributing $\gamma_t(x)$ uniformly in the cell $C_{k(x)}$. Using this property we can compute the conditional expectation explicitly and we obtain for $l\neq k$

$$\mathbb{E}\left[\sum_{x\in C_k}\gamma(x)\Delta(\bar U,U)\big|\mathbf{T}\gamma=\eta\right]=\eta_k\eta_l\sum_{\substack{x\in C_k\\ y\in C_l}}\left(J(x-y)-\bar J_{kl}\right)=0\,.$$

Similarly we handle the case $l=k$ and we conclude that, after averaging, the first-order term $\Delta(\bar U,U)$ in $\Gamma(k,\gamma)$ vanishes. We recall (see (3.3)) that

$$\Delta(\bar U,U)\equiv\bar U(k(x),\mathbf{T}\gamma)-U(x,\gamma)=\mathrm{O}\left(\frac{q}{L}\right)\,,$$

and hence we can estimate (5.33) by $O(q^2/L^2)$. For the term (5.34) we use the estimate $\sum_{x\in\Lambda_N}|R_T^{q,L}(x)|\sim\mathrm{O}(q/L)$ from Lemma 5.6 and the Hölder inequality

$$\mathbb{E}_S\left[\sum_{x\in\Lambda_N}(c(x,\gamma_t)-c_\gamma(x,\gamma_t))R_T^{q,L}(x)\right]\leq||\mathbf{c}-\mathbf{c}_\gamma||_\infty\mathbb{E}_S\left[\sum_{x\in\Lambda_N}|R_T^{q,L}(x)|\right].$$

The first term on the right-hand side is estimated from (5.24) by $C(q/L)$ and hence the left-hand side behaves as $\mathrm{O}(q^2/L^2)$. Combining the estimates of (5.33) and (5.34) we conclude the proof.

Using the estimate for the synthetic process and its reconstruction from the coarse-grained process $\{\eta_t\}_{t\geq 0}$ we can compare the projected process $\{\mathbf{T}\sigma_t\}_{t\geq 0}$ and the coarse-grained process $\{\eta_t\}_{t\geq 0}$ also on the coarse level. The weak error for observables on the coarse space is also natural in simulations where we usually project finer simulations on the coarse level and use estimators for the coarse processes.

COROLLARY 5.9. *Let $\psi\in L^\infty(\mathcal{S}_{M,q}^c)$ be a test function on the coarse level such that there exists a test function $\phi\in L^\infty(\mathcal{S}_N)$ satisfying (A1) with the property $\psi(\mathbf{T}\sigma)=\phi(\sigma)$. Given the initial configuration σ_0 we define the coarse configuration $\eta_0=\mathbf{T}\sigma_0$. Assume the microscopic process $(\{\sigma_t\}_{t\geq 0},\mathcal{L})$ with the initial condition σ_0 and the approximating coarse process $(\{\eta_t\}_{t\geq 0},\bar{\mathcal{L}}^c)$ with the initial condition $\eta_0=\mathbf{T}\sigma_0$, then the weak error satisfies, for $0<T<\infty$,*

$$|\mathbb{E}_S\left[\psi(\mathbf{T}\sigma_T)\right]-\mathbb{E}_S\left[\psi(\eta_T)\right]|\leq C_T\left(\frac{q}{L}\right)^2\,, \tag{5.35}$$

where the constant C_T is independent of q and L but depends on T.

6. Implementation of the coarse-grained Monte Carlo algorithms. The hierarchy of coarse-grained Monte Carlo processes (CGMC) parametrised by q has been designed in such a way that it is easily implemented in the unified manner. In fact, the nature of the generator $\bar{\mathcal{L}}^c$ at the level q allows us to use the same implementation as for the standard MC at the microscopic level, i.e., $q = 1$.

The stochastic system is simulated with the kinetic Monte Carlo (KMC) algorithm. Each iteration of the Monte Carlo simulation produces a variable time step Δt within which a spin flip occurs at a specific lattice node based on the transition probability,

$$[c_a(k,\eta) + c_d(k,\eta)]\Delta t + O(\Delta t^2)$$

where c_a and c_d are as in (3.5). This procedure repeats until the stopping criteria (see below) have been met. More specifically, the simulation is implementing the following *global updating* process-type kinetic Monte Carlo (KMC) algorithm for spin flip Arrhenius dynamics:

Step 1 Calculate all transition rates $c_a(k,\eta)$ (adsorption), $c_d(k,\eta)$ (desorption) from (3.5) for all nodes k in the lattice Λ_M^c

Step 2 Calculate the total $R_a = \sum_{l\in\Lambda_M^c} c_a(l,\eta)$, $R_d = \sum_{l\in\Lambda_M^c} c_d(l,\eta)$ adsorption, desorption rates respectively. Similarly obtain the total rate $R_T = R_a + R_d$.

Step 3 Obtain two random numbers ρ_1 and ρ_2.

Step 4 Use the first random number to choose between absorption or desorption based on the measure created by the rates R_a, R_d and R_T. Assume that the choice is to adsorb(desorb) and denote by $c \equiv c_a(l,\eta)$, $(c_d(l,\eta))$ and $R = R_a$, (R_d), respectively.

Step 5 Find the node at lattice position $l \in \Lambda_M^c$ such that,

$$\sum_{j=0}^{l} c(j,\eta) \geq \rho_2 R \geq \sum_{j=0}^{l-1} c(j,\eta)$$

Step 6 Update the time, $t = t + \Delta t$ where

$$\Delta t = 1/R_T\,. \tag{6.1}$$

Step 7 Repeat from Step 1 until equilibrium or dynamics of interest have been captured.

As expected a kinetic Monte Carlo algorithm produces no "null" steps and therefore every trial is accepted. A similar version of the algorithm can also be implemented with a *local* updating mechanism which can improve speed substantially at the reciprocal expense of allocating further computer memory for dynamic array allocation. In the simulation that follow we use a finite size interaction potential and lattice size $L, N < \infty$.

We produce simulations and compare observables at microscopic ($q = 1$) and coarse grained ($q > 1$) levels. For consistency purposes we use the same seed for our random number generator in order to compare simulations for different coarse grained values of q. This allows us to focus on the differences attributed only to the coarse graining variable and not on those resulting from different paths due to the initial seed. In the case of several realisations we initialise each new microscopic realisation with a different seed. Once again, for comparison purposes, we initialise each subsequent coarse grained realisation with the same seeds used in the respective microscopic simulations. All simulations are compared in the same non-dimensional time units. The corresponding non-dimensional time-step is respectively set by the Monte Carlo simulation based on the rule 6.1.

In the simulations which follow we try, when possible, to group together various parameters in the model so the results are presented with respect of variations in a smaller number

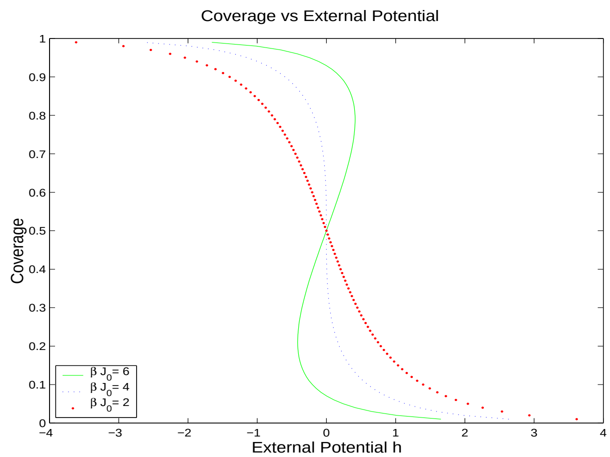

FIG. 7.1. *Equilibrium coverage $c(\sigma)$ and its dependence on the external field. The critical point for the $\{0, 1\}$ spins satisfies $\beta_c J_0 = 4$. The solid line depicts equilibria below the critical temperature. The hysteresis shape of the curve manifests existence of two equilibria in the neighbourhood of the zero external field.*

of parameters. In that respect we point out that for the fixed external field $\bar{h}$ it is possible to group together d_0 and $\bar{h}$ in (3.5) as follows,

$$\bar{c}_d(k, \eta) = c_0 \eta(k) e^{-\beta[\sum_l \bar{J}(k,l)\eta(l) + \bar{J}(0,0)(\eta(k)-1)]}$$

where $c_0 = d_0 e^{\bar{h}}$. We provide the values of all pertinent parameters as well as d_0 and c_0 in the relevant figures.

7. Numerical simulations. We use the CGMC described and analyzed in the previous sections for efficient simulations in the spin systems that undergo phase transitions. Within the context of spin-flip dynamics a typical example is nucleation of spatial regions of a new phase or a transition from one phase (all spins equal to zero) to another (all spins equal to one). In such simulations the emphasis is on the path-wise properties of the coarse-grained process so that the switching mechanism is simulated efficiently while approximation errors are controlled. We compare simulations on the microscopic level $q = 1$ with those performed on different levels of coarse-graining hierarchy parametrized by q.

The qualitative behaviour of the Ising model with a long-range potential can be understood from the mean-field approximation of the equilibrium total coverage $c(\sigma)$. Below the critical temperature the Gibbs measure is not unique (in the thermodynamic limit $N \to \infty$) and two phases can coexist. When the energy landscape is probed by changing the external field h we observe non-uniqueness of the equilibrium coverage as depicted in Figure 7.1. The fluctuations allow for transitions between the equilibrium which leads to nucleation of regions with a different phase. Changing the external field h makes the original phase unstable and a switching occurs – the system transforms into the other equilibrium configuration.

The parameters in the simulations have been chosen as follows: We use a uniform finite range potential for all examples presented. We simulate a finite lattice with a total of $N = 1000$ microscopic nodes and allow a potential interaction range of $2L + 1$ for $L = 100$. We choose the constant $d_0 = 1$ so that $c_a = 1$ and $c_d = 1$. Hence in this case the critical value β_c is given by $\beta_c J_0 = 4$. If $\beta J_0 > \beta_c J_0 = 4$ the system is in the phase transition regime and the two phases can coexist. In this region we typically observe a transition from one phase (e.g., zero (low) coverage) to the other phase (e.g., full coverage). For the phase transition

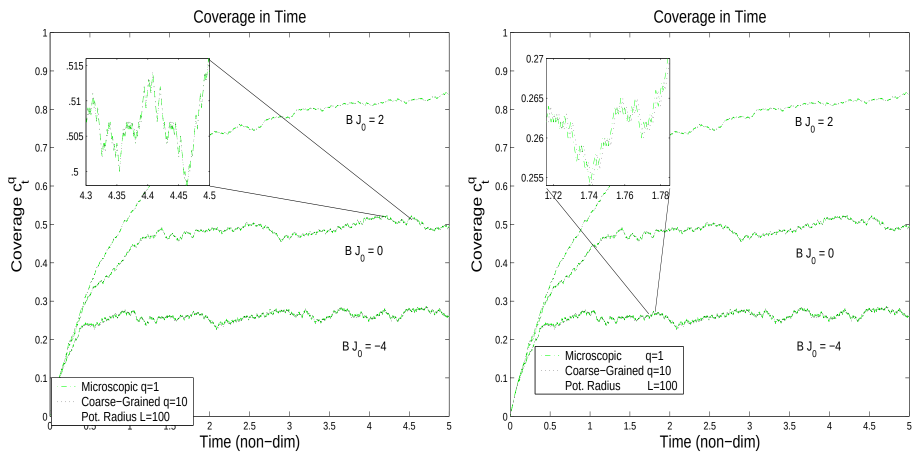

FIG. 7.2. *Relaxation dynamics. Comparison of microscopic ($q = 1$) and coarse grained ($q = 10$) simulations. The plot depicts a short time simulation in order to calibrate the code and compare to Figure 4 from [20].*

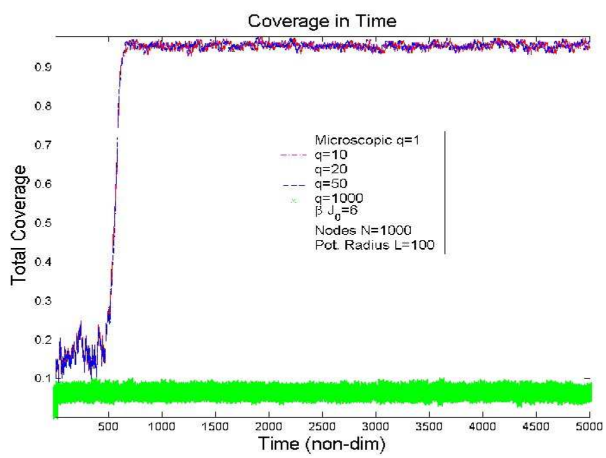

FIG. 7.3. *Time series of the coverage c_t^q. Simulations for different coarse-graining ratios are shown in the phase transition regime. The case $q = 1000, m = 1$ (mean-field approximation) shows significant discrepancy. Parameters used: potential radius length $L = 100$, $\beta J_0 = 6$, $d_0 = 1$, $c_0 = .072$*

examples we fix $\beta J_0 = 6 > \beta_c J_0$. The simulations become difficult when $\beta \simeq \beta_c$ and there is no external field h applied. We note that the coarse-graining algorithm will not perform well close to the critical point β_c when $h = 0$. In the numerical studies we first investigate approximation properties of the CGMC algorithms for certain global quantities.

Coverage: We define the coverage c_t to be the process computed as the spatial mean

$$c_t(\sigma_t) = \frac{1}{N}\sum_{x\in\Lambda_N}\sigma_t(x)\,, \qquad c_t^q(\eta_t) = \frac{1}{qM}\sum_{l\in\Lambda_M^c}\eta_t(k)\,.$$

We present time evolution of the coverage at the phase transition regime, $\beta J_0 = 6$. Note that the case $q = 1000$, $m = 1$ which corresponds to the mean-field approximation ("over coarse-grained" interactions) does not follow the phase transition path of the other simulations. On the other hand the agreement in the results is extremely good for the remaining

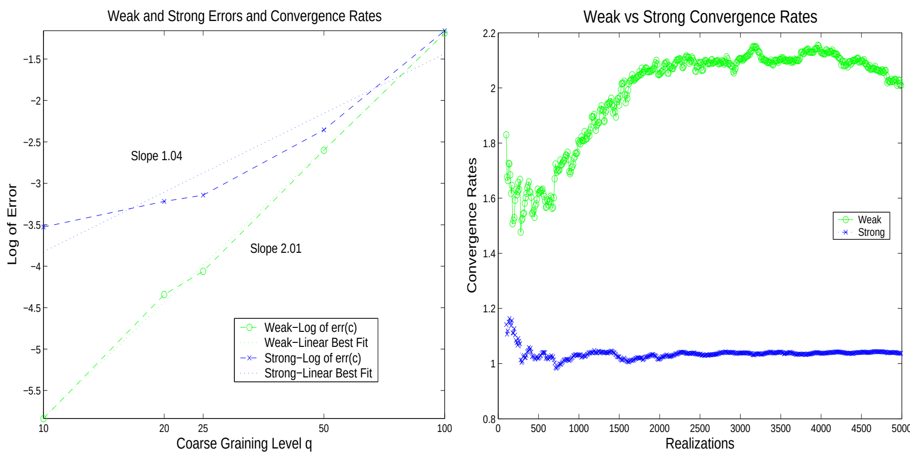

FIG. 7.4. *Estimated weak $e_w[c]$ and strong $e_s[c]$ errors. We compare the exact process c_t, $q = 1$ with coarse approximations c_t^q, $q = 10, 25, 50$ and 100. The simulation parameters were fixed at $L = 100$, $d_0 = 1$, $c_0 = .07$, $\beta J_0 = 6 > \beta_c J_0$ and the lattice size $N = 1000$. The convergence rates depicted are estimated by the linear best fit on the logarithmic scale. The statistical error or dependence of the estimates on the number of realisations is depicted in the right figure.*

values of q. Furthermore, these numerical results indicate path-wise (strong) approximation of the microscopic process by the coarse-grained process. This observation suggests a stronger error control than the relative entropy estimate provided by Proposition 5.2.

To quantify the error behaviour we calculate two errors between the exact stochastic process c_t and its coarse approximation c_t^q at the level of coarse-graining q. We define the weak error $e_w[c]$ and the strong error $e_s[c]$ respectively:

$$e_w[c] = \int_0^T |\mathbb{E}\,[c_t] - \mathbb{E}\,[c_t^q]|\,dt\,, \qquad e_s[c] = \int_0^T \mathbb{E}\,[|\mathbf{T}c_t - c_t^q|]dt\,.$$

The expected values are estimated by empirical means and the integral in time by the piecewise constant quadrature.

The simulations allow us to estimate the convergence rate for both errors. The rates in the case of fixed parameters $L = 100$, $d_0 = 1.0$, $c_0 = 0.07$ and $\beta J_0 = 6$ on the lattice of the size $N = 1000$ are depicted in Figure 7.4. Note that we need to eliminate the statistical error, arising from approximation of expected values by empirical means. However, as seen in Figure 7.4 the estimator of the rate converges as the number of realisations tends to infinity.

Since the coarse-grained Hamiltonian neglects higher order corrections arising from the fluctuations on fine scales one may expect that the approximation is poor if q/L is not very small. This is certainly true at the critical point (i.e., $\beta = \beta_c$ and $h = 0$) but further from the critical point the approximation properties are improved. This is demonstrated in the following table, where the simulations were performed in the presence of different (large) external fields. The relative error becomes small even for fairly crude coarse-graining $q = 20$ in the case of shorter interaction radii L.

Mean time to reach phase transition: One quantity of interest that is calculated from the simulations is the mean time $\bar{\tau}_T = \mathbb{E}\,[\tau_T]$ until the coverage reaches C^+ in its phase transition regime (see Figure 7.3). The random exit time is defined as $\tau_T = \inf\{t > 0 \,|\, c_t \geq C^+\}$. We estimate the probability distributions ρ_τ and ρ_τ^q from the simulations. We record a phase transition at the time $\bar{\tau}_T$ when the coverage exceeds the threshold value $C^+ = 0.9$.

TABLE 7.1
Relative strong error $e_s[c]$ in the presence of an external field defined by c_0. Comparisons are made for different values of the interaction radius L and different coarse-graining levels q. Size of the lattice fixed at $N = 1000$.

c_0	L	$q=5$	$q=10$	$q=20$
.07	100	.0591	.0733	.1134
	40	.0820	.0880	.1113
	20	.1508	.2214	.1832
.09	100	.0186	.0563	.0480
	40	.0678	.0749	.1064
	20	.1760	.1767	.1812
1	100	.0010	.0010	.0025
	40	.0036	.0040	.0054
	20	.0016	.0043	.0065

TABLE 7.2
Approximation of $\bar{\tau}_T, \mathcal{R}\left(\rho^q_\tau \,|\, \mathbf{T}_*\rho_\tau\right)$ and relative error.

L	q	$\bar{\tau}_T$	$\mathcal{R}\left(\rho^q_\tau \,\|\, \mathbf{T}_*\rho_\tau\right)$	Rel. Err.	CPU [s]
100	1	532	0.0	0	309647
100	2	532	0.003	0.01%	132143
100	4	530	0.001	0.22%	86449
100	5	534	0.003	0.38%	58412
100	10	536	0.004	0.82%	38344
100	20	550	0.007	3.42%	16215
100	25	558	0.010	4.91%	7574
100	50	626	0.009	17.69%	4577
100	100	945	0.087	77.73%	345

In Figure 7.5 we plot approximations of the Probability Density Functions (PDFs) of τ_T and compare them for different values of q.

The qualitative agreement observed in Figure 7.5 is quantified by using the information distance for error estimation, i.e., by estimating the relative entropy

$$\mathcal{R}\left(\rho_1 \,|\, \rho_2\right) = \sum_\lambda \rho_1(\lambda) \log\left(\frac{\rho_1(\lambda)}{\rho_2(\lambda)}\right) . \tag{7.1}$$

Nucleation: The nucleation of a new phase is a typical phenomenon in the regime where $\beta > \beta_c$. Essentially, there exist two equilibria (phases). Random fluctuations will induce transitions from one state to another by overcoming energy barriers that separate the equilibria. We investigate approximation of the path-wise behaviour on the configuration space for nucleation of a new phase. Two different initial configurations are used.

TEST CASE I: The initial state is at the metastable equilibrium where the coverage is zero. The fluctuations will cause the transition to the full coverage equilibrium which is stable due to the external applied field. We present only qualitative comparison in the series of snap-shots (Figure 7.8) of the phase transition from the uniform (zero) initial coverage to the full coverage. We observe a striking path-wise agreement on the configuration space for relatively large values of q compared to the interaction radius L. However, as the ratio q/L increases the corresponding coarse-grained process lags behind which is also demonstrated in the expected values of transition times. Such behaviour suggests that fluctuations at regions

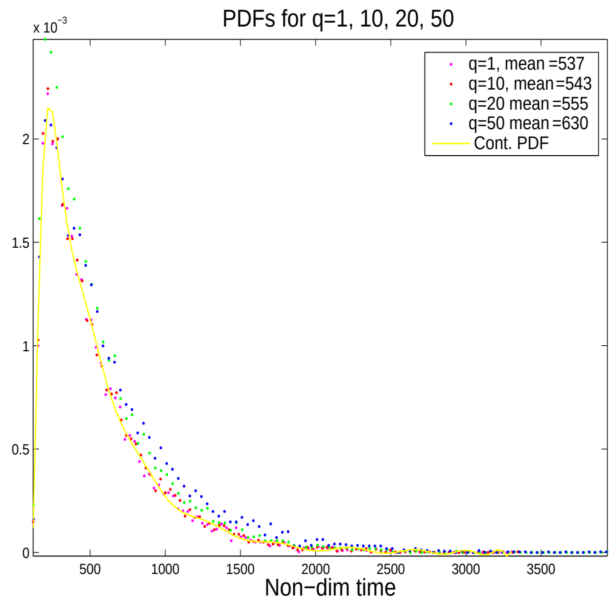

FIG. 7.5. *Probability Density Function (PDFs) comparisons between different coarse-graining levels q. The estimated mean times for each PDF are shown in the figures. All PDFs comprised of 10000 samples and the histogram is approximated by 100 bins.*

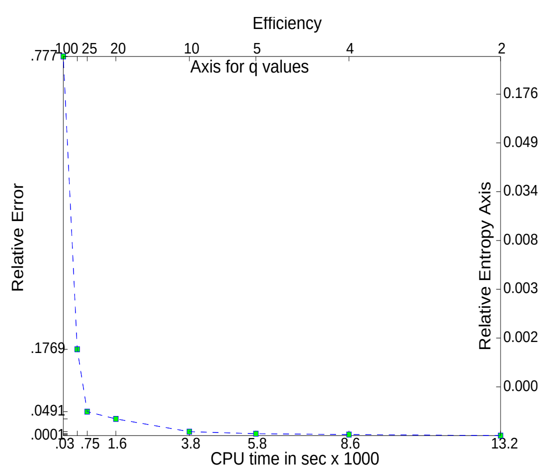

FIG. 7.6. *The dependence of the relative error and the relative entropy on the coarse-graining level q. The left scale on the vertical axis depicts the relative error and q on the log-log scale. Measurements based on averaging over* 10000 *realisations for each q.*

with uniform states are well-approximated by a highly coarse-grained process while finer resolution is necessary for resolving nucleation of new phases through islands.

TEST CASE II: We have already documented the path-wise agreement of the approximating dynamics under both transition and relaxation cases. In this example we examine the nucleation phenomenon at the critical parameter regime of phase transition $\beta J_0 = 6$. We chose the initial state to be at a saddle point of the energy surface, i.e., the mean coverage is set to 0.5. Snapshots of the spatial distribution of spins are presented in Figure 7.7. Under all four dynamics examined $q = 1, 5, 10$ and 20 we observe complete spatial path-wise agreement. Over time the total coverage may fall towards zero or rise towards one in which case it will remain there since we are at the phase transition regime where these represent stable equilibria. Furthermore such spatial examples of nucleation are shown below in Test Case III under the assumption of an "island-type" of initial state.

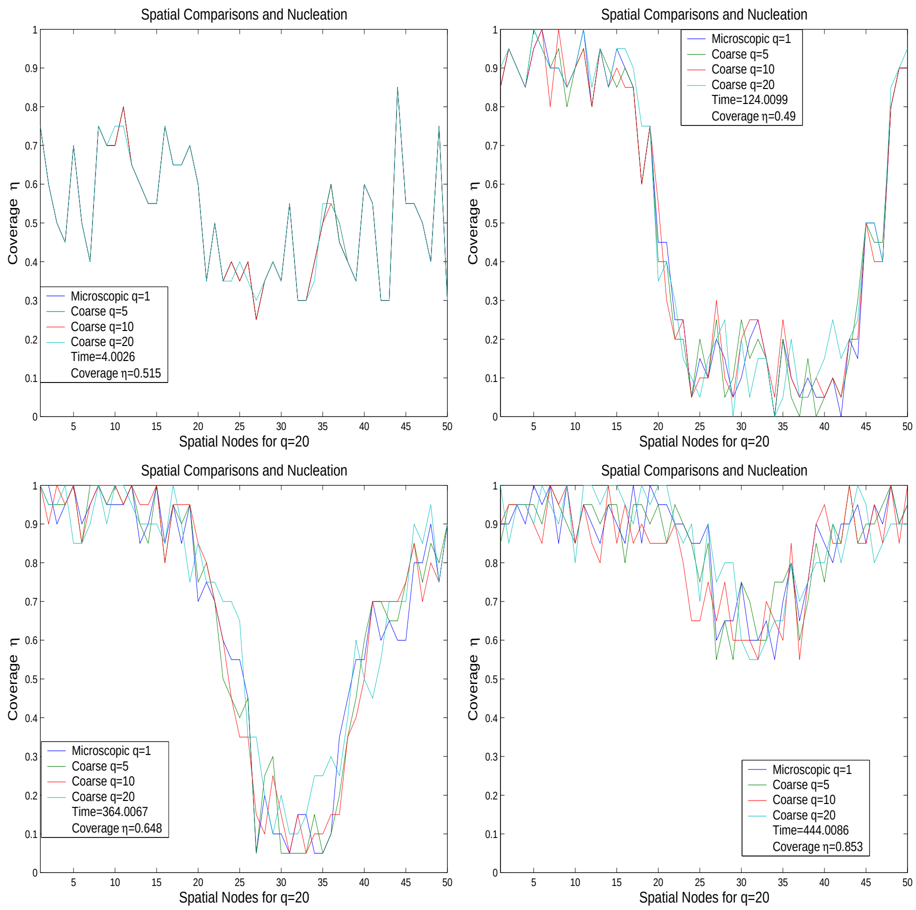

FIG. 7.7. *Snap-shots of the transition from the initial state with the mean coverage at* 0.5*. Comparisons between the microscopic* $q = 1$ *and coarse grained simulations* $q = 5, 10$ *and* $q = 20$*. The interaction radius is set to* $L = 100$*, the external field* $c_0 = 0.0492$*,* $d_0 = 1$ *and the total number of lattice sites* $N = 1000$*.*

TEST CASE III: The last set of simulations presents evolution from the non-uniform initial state, giving a qualitative comparison of nucleation from an island of a given size (Figure 7.9). In these simulations we observe spatial propagation of the interface in time for different initial size of the island.

Acknowledgements: The research of M.K. is partially supported by NSF-DMS-0413864 and NSF-ITR-0219211. The research of P.P. was partially supported by NSF-DMS-0303565. M.K. also acknowledges many valuable conversations with A. Szepessy. The authors would like to thank the Institute for Mathematics and its Applications where part of this work was carried out during the programme "Mathematics of Materials and Macromolecules: Multiple Scales, Disorder, and Singularities"

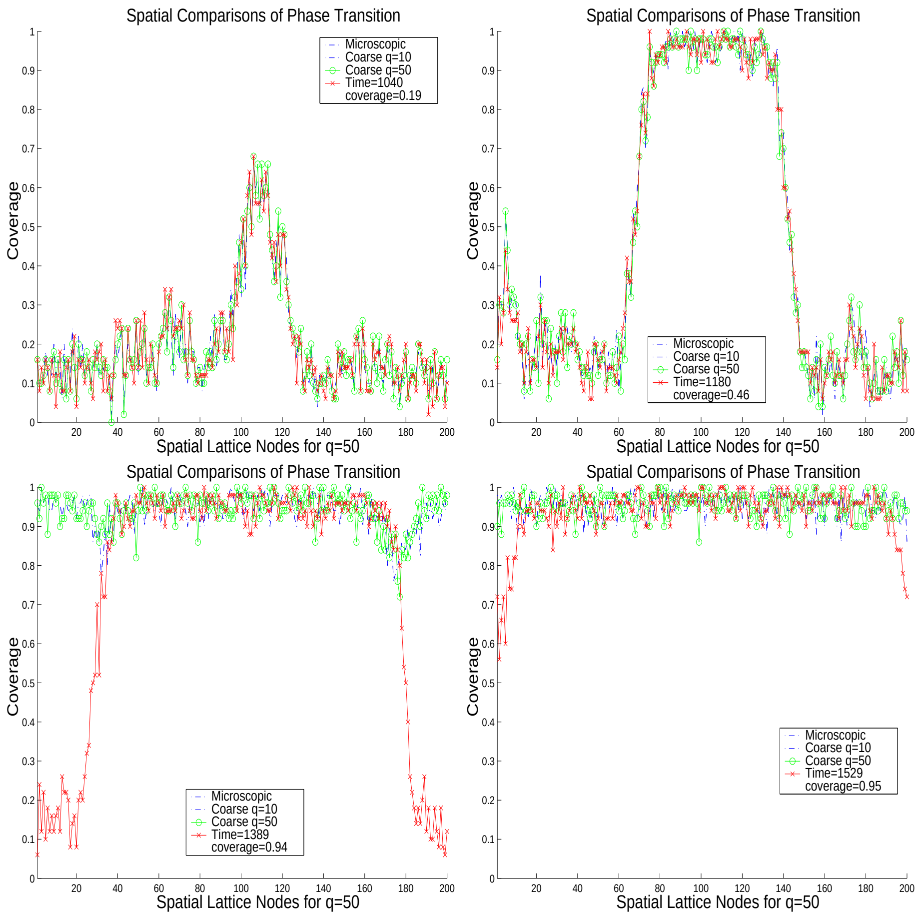

FIG. 7.8. *Snap-shots of the transition from zero initial spatial distribution. Comparisons between the microscopic* $q = 1$ *and two coarse grained simulations* $q = 10$ *and* $q = 50$*. The interaction radius is set to* $L = 200$ *while total nodes are* $N = 10000$*.*

REFERENCES

[1] D. R. A. BRANDT AND D. J. AMIT, *Multi-level approaches to discrete-state and stochastic problems*, in Multigrid Methods, II, W. Hackbusch and U. Trottenberg, eds., Springer-Verlag, 1986, pp. 66–99.

[2] D. BAI AND A. BRANDT, *Multiscale computation of polymer models.*, in Multiscale Computational Methods in Chemistry and Physics, J. B. A. Brandt and K. Binder, eds., vol. 177 of NATO Science Series: Computer and System Sciences, IOS Press, 2001, pp. 250–266.

[3] P. BERNARD, D. TALAY, AND L. TUBARO, *Rate of convergence of a stochastic particle method for the Kolmogorov equation with variable coefficients*, Math. Comp., 63 (1994), pp. 555–587, S11–S17.

[4] G. BEYLKIN, R. COIFMAN, AND V. ROKHLIN, *Fast wavelet transforms and numerical algorithms. I*, Comm. Pure Appl. Math., 44 (1991), pp. 141–183.

[5] A. BRANDT, *Multigrid methods in lattice field computations*, Nucl. Phys. B, 26 (1992), pp. 137–180.

[6] A. BRANDT AND V. ILYIN, *Multilevel Monte Carlo methods for studying large-scale phenomena in fluids.*, J. of Molecular Liquids, 105 (2003), pp. 253–256.

[7] A. BRANDT AND D. RON, *Renormalization multigrid: Statistically optimal renormalization group flow and coarse-to-fine Monte Carlo acceleration*, J. Stat. Phys., 102 (2001), pp. 231–257.

[8] T. M. COVER AND J. A. THOMAS, *Elements of Information Theory*, New York: Wiley, 1991.

[9] P. DUPUIS AND R. S. ELLIS, *A weak convergence approach to the theory of large deviations*, John Wiley & Sons Inc., New York, 1997. A Wiley-Interscience Publication.

[10] W. E AND B. ENGQUIST, *Multiscale modeling and computation*, Notices Amer. Math. Soc., 50 (2003),

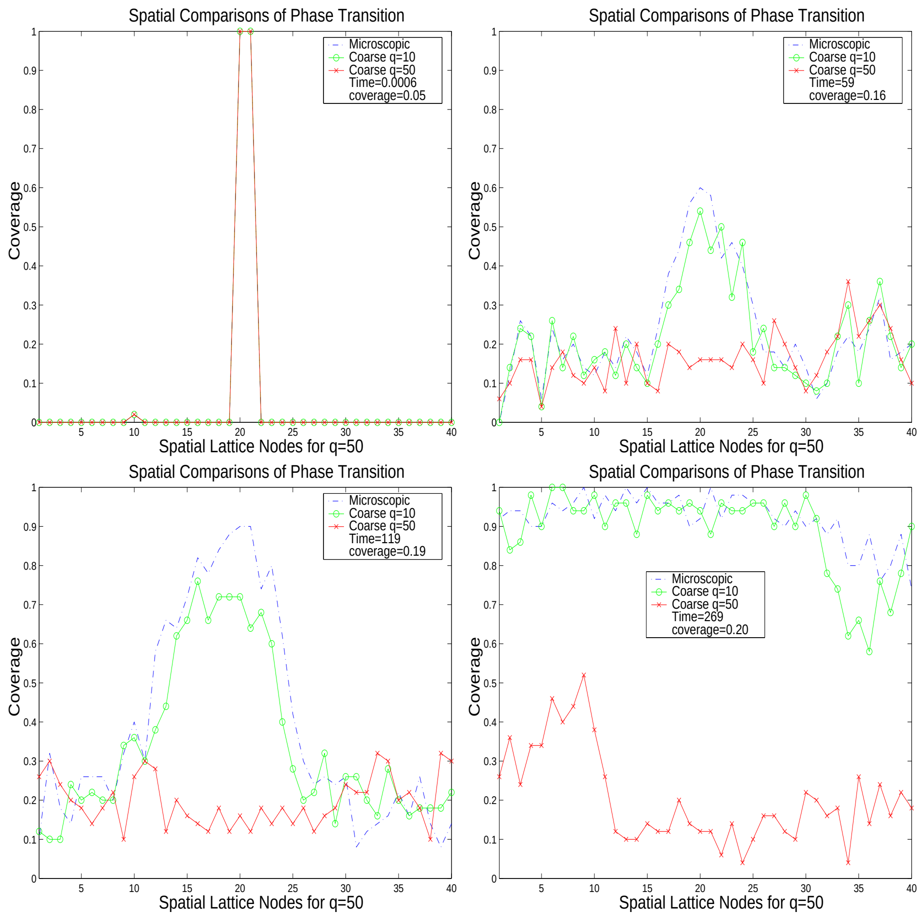

FIG. 7.9. *Snap-shots of the nucleation from a small-size initial seed. Comparisons between the microscopic* $q = 1$ *and two coarse grained simulations* $q = 10$ *and* $q = 50$. *Potential radius is set to* $L = 100$ *and the size of the lattice to* $N = 2000$.

pp. 1062–1070.

[11] K. ERIKSSON, D. ESTEP, P. HANSBO, AND C. JOHNSON, *Computational differential equations*, Cambridge University Press, 1996.

[12] J. GOODMAN AND A. D. SOKAL, *Multigrid Monte Carlo methods for lattice field theories.*, Phys. Rev. Lett., 56 (1986), pp. 1015–1018.

[13] Q. HOU, N. GOLDENFELD, AND A. MCKANE, *Renormalization group and perfect operators for stochastic differential equations*, Phys. Rev. E, 6303 (2001), p. 6125.

[14] T. Y. HOU AND X.-H. WU, *A multiscale finite element method for PDEs with oscillatory coefficients*, in Numerical treatment of multi-scale problems (Kiel, 1997), vol. 70 of Notes Numer. Fluid Mech., Vieweg, Braunschweig, 1999, pp. 58–69.

[15] A. E. ISMAIL, G. C. RUTLEDGE, , AND G. STEPHANOPOULOS, *Multiresolution analysis in statistical mechanics. I. using wavelets to calculate thermodynamic properties.*, J. Chem. Phys., 118 (2003), pp. 4414–4424.

[16] A. E. ISMAIL, G. C. RUTLEDGE, AND G. STEPHANOPOULOS, *Multiresolution analysis in statistical mechanics. II. wavelet transform as a basis for Monte Carlo simulations on lattices*, J. Chem. Phys., 118 (2003), p. 4424.

[17] L. P. KADANOFF, *Statistical physics : statics, dynamics and renormalization*, World Scientific, 1999.

[18] M. KATSOULAKIS, P. PLECHÁČ, L. REY-BELLET, AND D. TSAGKAROGIANNIS, *Higher order corrections to the coarse-grained Monte Carlo method.* in preparation, 2005.

[19] M. A. KATSOULAKIS, A. J. MAJDA, AND A. SOPASAKIS, *Multiscale couplings in prototype hybrid deter-*

ministic/stochastic systems: Part i, deterministic closures, Comm. Math. Sci., 2 (2004).
[20] 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 (2003), pp. 250–278.
[21] M. A. KATSOULAKIS, A. J. MAJDA, AND D. G. VLACHOS, *Coarse-grained stochastic processes for microscopic lattice systems*, Proc. Natl. Acad. Sci. USA, 100 (2003), pp. 782–787 (electronic).
[22] M. A. KATSOULAKIS AND A. SZEPESSY, *Stochastic hydrodynamical limits of stochastic spin systems*. in preparation, 2005.
[23] M. A. KATSOULAKIS AND J. TRASHORRAS, *Information loss in coarse-graining of stochastic particle dynamics*, tech. report, University of Massachussets, 2004. accepted in J. Stat. Phys.
[24] M. A. KATSOULAKIS AND D. G. VLACHOS, *Hierarchical kinetic Monte Carlo simulations for diffusion of interacting molecules*, J. Chem. Phys, 112 (2003).
[25] I. G. KEVREKIDIS, C. W. GEAR, AND G. HUMMER, *Equation-free: the computer aided analysis of complex multiscale systems*, AIChE Journal, 50 (2004), pp. 1346–1355.
[26] C. KIPNIS AND C. LANDIM, *Scaling Limits of Interacting Particle Systems*, Springer-Verlag, 1999.
[27] P. E. KLOEDEN AND E. PLATEN, *Numerical Solution of Stochastic Differential Equations*, vol. 23 of Applications of Mathematics, Springer Verlag, 3rd ed., 1999.
[28] D. P. LANDAU AND K. BINDER, *A Guide to Monte Carlo Simulations in Statistical Physics*, Cambridge University Press, 2000.
[29] A. J. MAJDA, I. TIMOFEYEV, AND E. V. EIJNDEN, *A mathematical framework for stochastic climate models*, Comm. Pure Appl. Math., 54 (2001), p. 891.
[30] F. MULLER-PLATHE, *Coarse-graining in polymer simulation: from the atomistic to the mesoscale and back*, Chem. Phys. Chem., 3 (2002), p. 754.
[31] C. SCHÜTTE, A. FISCHER, W. HUISINGA, AND P. DEUFLHARD, *A direct approach to conformational dynamics based on hybrid Monte Carlo*, J. Comput. Phys., 151 (1999), pp. 146–168. Computational molecular biophysics.
[32] A. SZEPESSY, R. TEMPONE, AND G. E. ZOURARIS, *Adaptive weak approximation of stochastic differential equations*, Comm. Pure Appl. Math., 54 (2001), pp. 1169–1214.
[33] D. TALAY AND L. TUBARO, *Expansion of the global error for numerical schemes solving stochastic differential equations*, Stochastic Anal. Appl., 8 (1990), pp. 483–509 (1991).
[34] D. G. VLACHOS, L. D. SCHMIDT, AND R. ARIS, *The effects of phase transitions, surface diffusion and defects on surface catalyzed reactions:fluctuations and oscillations*, J. Chem. Phys., 93 (1990), p. 8306.