# ERROR CONTROL AND ANALYSIS IN COARSE-GRAINING OF STOCHASTIC LATTICE DYNAMICS

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

**Abstract.** The coarse-grained Monte Carlo (CGMC) algorithm was originally proposed in the series of works [15, 16].

In this paper we further investigate the approximation properties of the coarse-graining procedure and relation between the coarse-grained and microscopic processes. We provide both analytical and numerical evidence that the hierarchy of the coarse models is built in a systematic way that allows for the error control of quantities that may also depend on the path. We also demonstrate that CGMC leads to a significant CPU speed-up of simulations of metastable phenomena, e.g., estimation of switching times or nucleation of new phases. Numerical evidence guided by analytical results suggests that CGMC probes the energy landscape in path-wise agreement to MC simulations at the microscopic level.

**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, or degrees of freedom and even larger number of interactions between them present the principal limitation for efficient simulations. A related limiting factor is represented by the essentially sequential nature of resolving the time evolution in many particle systems that yields a substantial slowdown in the resolution of dynamics especially in metastable regimes.

In [15, 16, 18] 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 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 [1], spanning length/time scales from the microscopic to the mesoscopic. The resulting *stochastic coarse-grained processes* involve Markovian birth-death and generalized exclusion processes and their combinations, and as demonstrated in [15, 16, 18], they share the same ergodic properties with their microscopic counterparts. More specifically, the full hierarchy of the coarse-grained stochastic dynamics satisfies detailed balance relations and as a result not only yields self-consistent random fluctuation mechanism, but also consistent with the underlying microscopic fluctuations and the unresolved degrees of freedom (DOFs). From the computational complexity perspective, a comparison of CGMC with conventional MC methods for the same real time shows, [15], that the CPU time can decrease approximately as $O(1/q^2)$ 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

*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

on a single processor, the computational savings of CGMC make it a suitable tool capable of capturing large scale features, while retaining microscopic information on intermolecular forces and particle fluctuations. A striking difference between CGMC for diffusion (spin exchange) and adsorption/desorption (spin flip) simulations is that in the case of diffusion we also have an additional coarse-graining in time by a factor $q^2$, improving the hydrodynamic slowdown effect in conservative MC, [18].

In the recent paper [17] the authors attempted to rigorously analyze 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 during coarse-graining from finer to coarser scales the information-theoretic concept of the *relative entropy* between probability measures, [4]. Such relative entropy estimates give a first mathematical reasoning for the parameter regimes, i.e., the degree of coarse-graining versus the potential range, for which CGMC is expected to give errors within a given tolerance. In related earlier work [18], information loss estimates between CGMC and microscopic processes were derived for the equilibrium canonical Gibbs states; we also refer to [9] for loss of information results at equilibrium for the coarse-graining of polymeric systems. More specifically, the suggested strategy, at least in the case of spin-flip dynamics, consists of the following steps: The interaction potential (or operator) is decomposed in terms of a truncated multi-resolution decomposition within a given tolerance. The CGMC algorithm is subsequently defined at the coarsening level specified by the truncation of the decomposition, since the error in the approximation is due only to the MRA decomposition and remains small in time. Spin exchange dynamics have a more complicated error structure, as the stochastic closure process in [18] suggests, and a rigorous error analysis is currently unavailable. Using the rigorous results in [17] as a starting point, in this paper we focus on carrying out a detailed numerical error analysis and error propagation for spin flip lattice dynamics. We first define a notion of *order of convergence* for CGMC algorithms using relative entropy, i.e., loss of information, as a measure of accuracy. 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, e.g., exit times in domain switching problems. Furthermore, we demonstrate the capabilities of CGMC for *accurate* large scale simulations. We comment on possible pitfalls due to improper coarse-graining, by concentrating on two rather demanding examples: nucleation and domain switching in spin flip stochastic lattice dynamics. We computationally demonstrate 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 certainly related to a number of methods involving coarse-graining at various levels, for instance fast summation techniques, renormalization group theory and simulation and multi-scale computational methods for stochastic systems. We next discuss some connections of CGMC with these approaches. 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 (refs). 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 rather than simulating every individual particle. The equilibrium set-up of CGMC is essentially given by the renormalized Hamiltonian after a single iteration in the renormalization group (RG) flow. It is not surprising that such an approach has many limitations, for instance in nearest neighbor Ising-type models close to critical temperatures, and this fact is manifested in the aforementioned error estimates and the comparative simulations in [15]. On the other hand the focus of CGMC is dynamic simulations usually coupled to a macroscopic system (see for instance the hybrid systems in [5, 14]), 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 RG flow given by RGMC and multigrid MC methods (refs) can improve the order of convergence of the CGMC. In that sense the CGMC method is order one accurate as explained in Section 4, where it is also demonstrated that RG corrections can give rise to higher order CGMC methods. This latter observation was partly inspired by the numerical analysis of finite element methods for PDEs, where smoother elements give rise to higher order accuracy, [8].

In recent years there has been a growing interest in developing coarse-graining methods for the purpose of modeling and simulation across scales. Such systems arise in a broad spectrum of scientific disciplines ranging from materials science to macromolecular dynamics, to epidemiology and to atmosphere/ocean science. These approaches may yield explicitly derived stochastic coarse models such as CGMC or [10, 12, 13, 22], or can be statistics-based [23] or may rely on on-fly simulations, e.g., equation-free [19], heterogeneous multiscale [7] or multiscale FEM methods [11]. For instance, the coarse-graining of atomistic models of polymer chains [23] is typically carried out by collecting a number of atoms (on the order of $10-20$) in a polymer chain into a "super-atom" and statistically fitting parameters to a known potential type, e.g., Lennard-Jones, to derive the coarse-grained potential for the super-atoms. Other coarse-graining techniques in the polymer science literature include the bond fluctuation model and its variants [21], in which an atomistic chain is mapped on a lattice, where a super-atom occupies a lattice cell. Such coarse-graining methodologies rely on parametrization, hence at different conditions (e.g., temperature, density, composition) coarse potentials need to be re-parametrized [23].

**2. Microscopic lattice models.** The presented analysis applies to the class of Ising-type lattice systems. For 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 discretization of the $d$-dimensional torus $\mathbb{T} = [0,1)^d$ and $d$ denotes the spatial dimension. 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 $\Sigma = \{0, 1, \dots s\}$ (Potts models). We shall study cases where the spin variable belongs to a compact Riemannien manifold, e.g., $\Sigma = \mathbb{S}^2$ (Heisenberg model), $\Sigma = \mathrm{SU}(2)$ (matrix model), elsewhere. We denote $\sigma = \{\sigma(x) \,|\, x \in \Lambda_N\}$ 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 $\sigma$ 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

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

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

$$\begin{aligned} &V \text{ is smooth on } \mathbb{R} \setminus \{0\}, \\ &\int_{\mathbb{R}} |V(r)|\, dr < \infty\,, \text{ and } \int_{\mathbb{R}} |\partial_r V(r)|\, dr < \infty\,. \end{aligned}$$

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

$$\mu_{N,\beta}(d\sigma) = \frac{1}{Z_{N,\beta}} e^{-\beta H(\sigma)} \mathrm{P}_N(d\sigma)\,, \qquad Z_{N,\beta} = \int_{\mathcal{S}_N} e^{-\beta H(\sigma)} \mathrm{P}_N(d\sigma)\,, \tag{2.2}$$

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 $\rho$. Typically for the Ising model the prior distribution on $\Sigma = \{0, 1\}$ would 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 $\sigma_t$, then the probability that over the time interval $[t, t+\Delta t]$ the spin at the site $x \in \Lambda_N$ spontaneously changes from $\sigma_t(x)$ to a new value in the state space $\xi \in \Sigma$ is $c(x, \sigma; \xi)\Delta t + O(\Delta t^2)$. We denote the resulting configuration $\sigma^{x,\xi}$. In the case of the Ising-type state space and spin-flip dynamics we omit $\xi$ 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 $f \in L^\infty(\mathcal{S}_N)$ defined on the space of configurations is given by

$$(\mathcal{L}f)(\sigma) = \sum_{x \in \Lambda_N} \int_\Sigma c(x, \sigma; \xi) \left(f(\sigma^{x,\xi}) - f(\sigma)\right) d\xi\,. \tag{2.3}$$

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

$$\frac{d}{dt} \mathbb{E}[f(\sigma_t)] = \mathbb{E}[\mathcal{L}f(\sigma_t)]\,, \tag{2.4}$$

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 (2.2). The sufficient condition is known as *Detailed Balance* (DB) and it imposes condition on the form of the rate

$$c(x,\sigma;\xi)e^{-\beta H(\sigma)} = c(x,\sigma^{x,\xi};\sigma(x))e^{-\beta H(\sigma^{x,\xi})}\,. \tag{2.5}$$

This condition has a simple interpretation: $c(x,\sigma;\xi)$ is the rate of converting $\sigma(x)$ to the value $\xi$ while $c(x,\sigma^{x,\xi};\sigma(x))$ is the rate of changing the spin with the value $\xi$ at the site $x$ back to $\sigma(x)$. The widely used class of Metropolis-type dynamics satisfies (2.5) and has the rate given by

$$c(x,\sigma;\xi) = G(\beta\Delta_{x,\xi}H(\sigma))\,,\ \text{where}\ \Delta_{x,\xi}H(\sigma) = H(\sigma^{x,\xi}) - H(\sigma), \tag{2.6}$$

and $G$ 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*, the rates are usually derived from transition state theory or obtained from molecular dynamics simulations. To avoid unnecessary generality from now we restrict the description to the Ising-type model with $\Sigma = \{0,1\}$ used for modeling adsorption/desorption processes. We also omit $\xi$ in the notation. The Arrhenius rate is defined as follows

$$c(x,\sigma) = \begin{cases} d_0 & \text{if } \sigma(x) = 0, \\ d_0 e^{-\beta U(x,\sigma)} & \text{if } \sigma(x) = 1, \end{cases} \tag{2.7}$$

where

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

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 (2.2)) 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 introducing the coarse-grained observables as block-spin variables. This follows the standard procedure of real-space renormalization as it was introduced by Kadanoff (refs). We remark that although we introduce block-spins our aim is not to approximate the renormalization group flow (either on the space of Gibbs measures or on the path space) rather we want to find an approximation that can be controlled with computable error estimates. Such task would be difficult for

problems where the renormalization flow defines non-trivial transformations in the space of Hamiltonians, for example, in the case of the Ising model with short range (nearest neighbor) interactions. Although, we study the case of finite- or long-range potentials we focus on approximating the coarse Hamiltonian in a simple and fast way. We obtain explicit error control of such procedure which can be combined with computationally more involved approximation of the RG flow, should higher accuracy be necessary.

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 $\Lambda^c_M$, and with the new state space $\Sigma^c$, 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 $\Sigma^c$ 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

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

*Ising-type spins.* To be more specific we analyze 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 $\Lambda_N$. 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 sites on the lattices in a 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 $\{\sigma_t\}_{t\geq 0}$ 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 $\{\eta_t\}_{t\geq 0}$ 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 $\Lambda_N$. Given the microscopic state $\sigma$ and corresponding coarse state $\eta = \mathbf{T}(\sigma)$

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

$$\sum_{k \in \Lambda^c_M} \left[ \sum_{x \in C_k} c(x,\sigma)\sigma(x) \right] [\psi(\eta - \delta_k) - \psi(\eta)]. \tag{3.3}$$

The configuration $\delta_k$ 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

$$\mathcal{L}^c\psi(\eta) = \sum_{k \in \Lambda^c_M} c_a(k) [\psi(\eta + \delta_k) - \psi(\eta)] + \sum_{k \in \Lambda^c_M} c_d(k) [\psi(\eta - \delta_k) - \psi(\eta)], \tag{3.4}$$

where the new rates

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

correspond to the adsorption and desorption processes. In this form the rates depend on the microscopic configuration $\sigma$ and not on the coarse random variable $\mathbf{T}(\sigma)$. Therefore, we 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 $\mathcal{S}^c_{M,q} = \{0,1,\dots q\}$. This process is defined by the generator $\bar{\mathcal{L}}^c$ of the form (2.3) where the rates $c_a$ and $c_d$ are replaced by approximate rates

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

For details we refer to [15]. 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 $U(\sigma)$

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

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

$$\bar{J}(k,l) = q^2 \int_{C_k}\int_{C_l} J(r-s)\,drds\,. \tag{3.7}$$

Note that this 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 renormalization 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 $O(q/L)$ as we discuss in the next section. The coarse interaction Hamiltonian is then given explicitly in terms of $\bar{J}$ and $\bar{h}$ as

$$\bar{H}(\eta) = -\frac{1}{2}\sum_{l\in\Lambda^c_M}\sum_{k\neq l}\bar{J}(k,l)\eta(k)\eta(l) - \frac{1}{2}\bar{J}(0,0)\sum_{l\in\Lambda^c_M}\eta(l)(\eta(l)-1) + \sum_{l\in\Lambda^c_M}\bar{h}(l)\eta(l)\,. \tag{3.8}$$

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 CG projections 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 including $q$ independent sites:

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

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

$$\begin{aligned} \bar{c}_a(k,\eta)\mu_{M,q,\beta}(\eta) &= \bar{c}_d(k,\eta+\delta_k)\mu_{M,q,\beta}(\eta+\delta_k), \\ \bar{c}_d(k,\eta)\mu_{M,q,\beta}(\eta) &= \bar{c}_a(k,\eta-\delta_k)\mu_{M,q,\beta}(\eta-\delta_k). \end{aligned}$$

We only verify the first relation; using that $\bar{H}(\eta+\delta_k) - \bar{H}(\eta) = -\bar{U}(k)$ and the definitions of the rates (3.5), we have (for $d_0 = 1$):

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

Since $(q-\lambda)\rho_q(\lambda) = (\lambda+1)\rho_q(\lambda+1)$, for all integers $0 \leq \lambda \leq q$, the last curly bracket is equal to zero, hence 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 (3.8) and hence the fluctuations from microscopic dynamics are properly included into the coarse-grained process.

In summary, the coarse-graining procedure described here has the following characteristics:

(i) the derived coarse-grained stochastic process $\{\eta_t\}_{t \geq 0}$ approximates a pre-specified observable, e.g. (3.1). In particular, time-dependent error estimates such as (5.6) can rigorously demonstrate that the process $\{\eta_t\}_{t \geq 0}$ keeps track of fluctuations from the microscopic level so expected values of certain path dependent (global) quantities can be properly estimated. More precisely, we can characterize 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) the invariant measure $\mu_{M,q,\beta}^c(d\eta)$ for the process $\{\eta_t\}_{t \geq 0}$ defined on $\mathcal{S}_{M,q}^c$ is close in a suitable probability metric to the projection of the microscopic Gibbs 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. Here we also note that the coarse-graining modifies the microscopic prior $\mathrm{P}_N(d\sigma)$ in (2.2), yielding the coarse prior $\mathrm{P}_{M,q}(d\eta)$.

**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 propose to 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 discrete state space $\mathcal{S}$ but analogous properties and definitions hold for the relative entropy of measures on general probability spaces (see [6]). 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 discrete state space $\mathcal{S}$. 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 $\pi_1$ and $\pi_2$, 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 $\pi_1$, $\pi_2$. 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 $\pi_1$ and $\pi_2$:

$$\mathcal{R}\left(\pi_1 \,|\, \pi_2\right) \geq \frac{1}{2} \left( \sum_{\sigma \in \mathcal{S}} |\pi_1(\sigma) - \pi_2(\sigma)| \right)^2 \equiv \frac{1}{2} ||\pi_1 - \pi_2||_1^2 \,. \tag{4.2}$$

Furthermore, for any observable $f = f(\sigma)$ we have the bound

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

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

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

The relative entropy (4.1) has the variational representation (see [20, pp. 338-339])

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

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 and information theory and use 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 $\sigma$ 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 $\sigma$. Since the information is carried in the random variable $\sigma$ we have to ask what is the *expected length* of the code required to capture the states of $\sigma$ provided we know the distribution of $\sigma$. 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 [4]) 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 $\sigma$ one needs to know the measure $\pi$. 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 or information about the random variable $\sigma$. 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 $\sigma$ are sampled by the Markov chain Monte Carlo algorithms and hence samples of a random variable $\sigma$ 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 hence assuming a wrong measure/distribution for the random variable $\sigma$. Using the relative entropy for estimating the approximation properties estimates 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 $\{\mathbf{T}\sigma_t\}_{t\geq 0}$, hence the approximation properties of such construction needs to be clarified.

As noted in the previous section we do not attempt to capture the effect of fine scales exactly and incorporate them into coarse model through the renormalization 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 need to be addressed is comparison and error estimate on 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 projection of fine to coarse spin variables. The principal idea proposed in [18] is to control *the specific loss of information* quantified by the relative entropy $\mathcal{R}\left(\mu^c_{M,q,\beta} \,|\, \mathbf{T}_*\mu_{N,\beta}\right)$.

PROPOSITION 5.1 ([18], *A priori estimate*).

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

This a priori estimate quantifies the dependence of the information distance (the specific relative entropy $\mathcal{R}(\mu\,|\,\nu)$) 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 parametrized 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 may be neglected. However, we also provide a posteriori estimate ([3, 2]) that allows to locally refine the hierarchy in adaptive way resembling the adaptive mesh refinement for approximation of PDEs.

PROPOSITION 5.2 (*A posteriori estimate*). *Given the coarse lattice* $\Lambda^c_M = \bigcup_k C_k$ *with* $q_k = |C_k|$

$$\mathcal{R}\left(\mu^c_{M,q,\beta} \,|\, \mathbf{T}_*\mu_{N,\beta}\right) \leq \mathbb{E}_{\mu^c}\left[R(\eta)\right] + \log\left(\mathbb{E}_{\mu^c}\left[e^{R(\eta)}\right]\right), \tag{5.2}$$

*where the residuum operator* $R(.)$ *is given by*

$$R(\eta) = 4\sum_{\substack{k,l\\k\neq l}} \frac{j_{kl}}{q_k q_l}\left(q_k^2\eta_l(q_l-\eta_l) - 2\eta_k\eta_l(q_k-\eta_k)(q_l-\eta_l)\right)$$
$$+4\sum_k \frac{j_{kk}}{q_k(q_k-1)}\eta_k(q_k-\eta_k)\left(\eta_k(\eta_k-1)+(q_k-\eta_k)(q_k-\eta_k+1)\right),$$

*and*

$$j_{kl} = \max_{\substack{x,x'\in D_k\\ y,y'\in D_l}} |J(x-y)-J(x'-y')|.$$

*The expected value* $\mathbb{E}_{\mu^c}[.]$ *is computed with respect to the Gibbs measure* $\mu^c_{M,q,\beta}$ *of the coarse-grained process* $\{\eta_t\}_{t\geq 0}$.

The error control is defined by the residuum operator $R(.)$ which can be computed a posteriori from sampling the measure $\mu^c_{M,q,\beta}$ (note that the expectations $\mathbb{E}_{\mu^c}[.]$ are computed with respect to the coarse-grained process).

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

PROPOSITION 5.3 (*Lower bound*).

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

*for all* $m' \leq m$ *and* $m'q' = mq = N$.

PROOF We first recall the variational formulation for the relative entropy (see for instance [20], Appendix 1):

$$\mathcal{R}(\mu \mid \nu) = \sup_f \left\{ \int f \, d\mu - \log \int e^f \, d\nu \right\}, \tag{5.4}$$

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

$$\mathcal{R}(\mu \mid \nu) \geq \sup_{f \circ \mathbf{T}} \left\{ \int f \circ \mathbf{T} \, d\mu - \log \int e^{f \circ \mathbf{T}} \, d\nu \right\} = \mathcal{R}(\mathbf{T}_* \mu \mid \mathbf{T}_* \nu) \tag{5.5}$$

where $\mathbf{T}$ is the projection operator (subscripts 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 $\mathbf{T}_*^{m',q'} \mu_{M,q,\beta}^c$ 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 noise) and the microscopic Gibbs states respectively. We characterize such an estimate as *a priori* since the bound depends on the exact microscopic process, in analogy to bounds for finite element approximations to PDE which depend on the Sobolev norm of the exact solution, [8]. 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, (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 (5.2). 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.

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 $\mathcal{D}(\Sigma^c)$, i.e., on the space of all right-continuous paths $\eta_t : [0, \infty) \to \Sigma^c$. Above we derived estimates for the exact coarse graining $\mathbf{T}_* \mu_{N,\beta}$ of the invariant measure $\mu_{N,\beta}$ and its approximation $\mu_{M,q,\beta}^c$ computed in terms of the coarse Hamiltonian. In the similar way we treat the measures on the path space: we denote $Q_{\sigma_0,[0,T]}$ the measure on $\mathcal{D}(\Sigma)$ for the process on the interval $[0, T]$, $\{\sigma_t\}_{t \in [0,T]}$ with the initial distribution $\sigma_0$. Similarly $Q_{\eta_0,[0,T]}^c$ denotes the measure on the coarse path space $\mathcal{D}(\Sigma^c)$. 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 full rigorous analysis on the path space is more involved and we refer to [17] and will only state the main a priori estimate here

PROPOSITION 5.4 ([17]). *Suppose the process $\{\eta_t\}_{t \in [0,T]}$, defined by the coarse generator $\mathcal{L}^c$ is the coarse approximation of the microscopic process $\{\sigma_t\}_{t \in [0,T]}$ then for any $q < L$ and $N$, $M$ the information loss as $q/L \to 0$ is*

$$\frac{1}{N} \mathcal{R}\left( \mathbf{T}_* Q_{\mathbf{T}_* \sigma_0,[0,T]} \mid Q_{\eta_0,[0,T]}^c \right) = T \, O\left( \frac{q}{L} \right) \tag{5.6}$$

REMARK The detailed proof of this information estimate (see [17]) 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 far as $q/L$, is small the approximation by the coarse-graining scheme yields small error independently of the potential $V$ or the initial distribution $\sigma_0$.

**6. Implementation of the coarse-grained Monte Carlo algorithms.** The hierarchy of coarse-grained Monte Carlo processes (CGMC) parametrized 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 $\Delta t$ 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 $c_d$ 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 $\Lambda_M^c$

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 $R = R_a + R_d$.

Step 3 Obtain two random numbers $\rho_1$ and $\rho_2$.

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,

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

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 realizations we initialize each new microscopic realization with a different seed. Once again, for comparison purposes, we initialize each subsequent coarse grained realization with the same seeds used in the respective

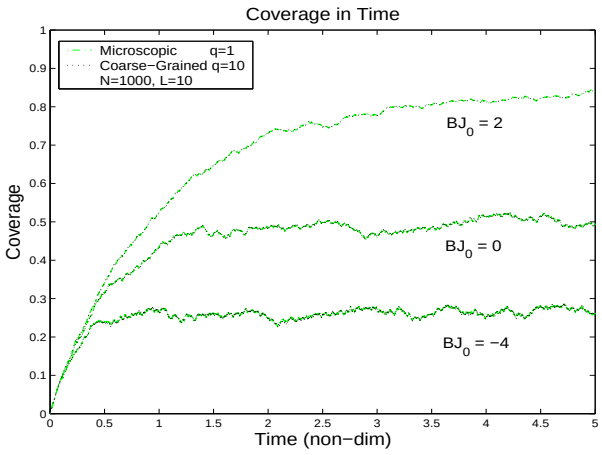


FIG. 7.1. *Relaxation dynamics. Comparison of microscopic ($q = 1$) and coarse grained ($q = 10$) simulations. The plot depicts a short time simulation in order to calibrate the code and compare to Figure 4 from [15].*

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.

**7. Numerical simulations.** We use the CGMC described and analyzed in the previous sections for efficient simulations in the systems that undergo phase transitions. Within the context of spin-flip dynamics a typical example will be, for example, nucleation of regions with a phase or transition from one phase (all spins equal to zero) to another (all spins equal to one). The emphasis is on the path-wise properties of the coarse-grained process. We compare simulations on the microscopic level $q = 1$ with those performed on different levels of coarse-graining hierarchy parametrized by $q$.

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 of $\beta_c J_0 = 4$. For the phase transition examples we fix $\beta J_0 = 6 > \beta_c J_0$

We investigate approximation of 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 values of $q$. Furthermore, these numerical results indicate path-wise (strong) approximation of the microscopic process by the coarse-grained

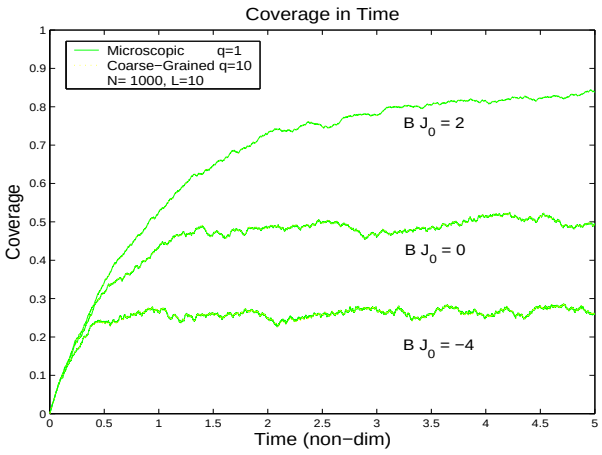


FIG. 7.2. *Time series of the coverage $c_t^q$. Simulations for different coarse-graining ratios are shown in the phase transition regime. The case $q = 1000, m = 1$ (mean-field approximation) shows significant discrepancy. Parameters used: potential radius length $L = 100$, $\beta J_0 = 6$. External potential is kept at the critical value of $1/.072$.*

TABLE 7.1
*Approximation of $\bar{\tau}_T, \mathcal{R}\left(\rho_\tau^q \,|\, \mathbf{T}_*\rho_\tau\right)$ and relative error.*

| $L$ | $q$ | $\bar{\tau}_T$ | $\mathcal{R}\left(\rho_\tau^q \,\|\, \mathbf{T}_*\rho_\tau\right)$ | 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 |

process. This observation suggests a stronger error control than the relative entropy estimate provided by Proposition 5.4.

*Mean time to reach phase transition:* One quantity of interest that can be 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 7.2). The random exit time is defined as $\tau_T = \inf\{t > 0 \,|\, c_t \geq C^+\}$. We estimate the probability distributions $\rho_\tau$ and $\rho_\tau^q$ from the simulations. We record a phase transition at the time $\bar{\tau}_T$ when the coverage exceeds the threshold value $C^+ = 0.9$.

In Figure 7.3 we plot approximations of the Probability Density Functions (PDFs) and compare them for different values of $q$.

The qualitative agreement observed in Figure 7.3 can be quantified using the information distance for error estimation, i.e., by estimating the relative entropy

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

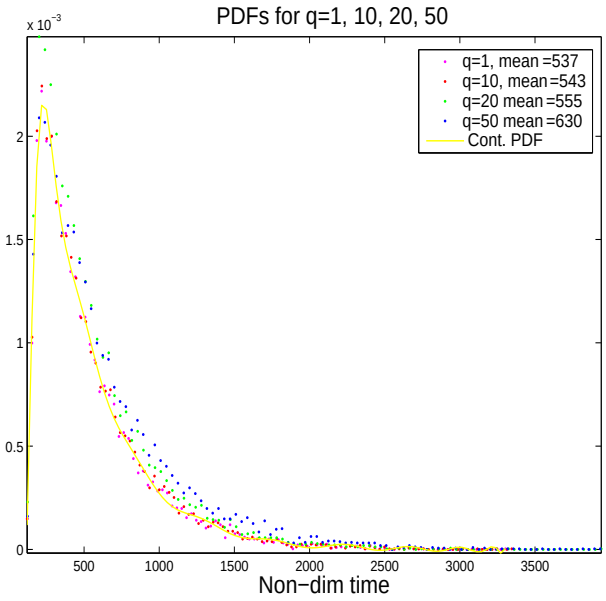


FIG. 7.3. *Probability Density Function (PDFs) comparisons between different coarse graining $q$. The mean times for each PDF are shown in the figures. All PDFs comprised of 10000 samples. The histogram is approximated by 100 bins.*

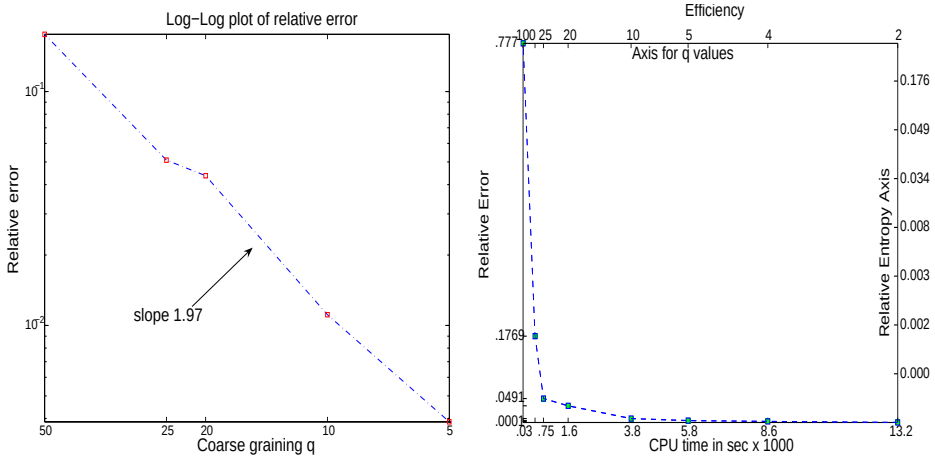


FIG. 7.4. *The dependence of the relative error and or the relative entropy on the coarse-graining level $q$. The left figure depicts the relative error and $q$ on the log-log scale. Measurements based on averaging over* 10000 *realizations for each $q$.*

*Nucleation:* The last set of simulations deals with the path-wise behaviour on the configuration space for nucleation of a new phase. We present only qualitative comparison in the series of snap-shots (Figure 7.5) 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 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. This point is also partly justified by the *a posteriori* estimate (5.2) where the residual term concentrates on transition regions.

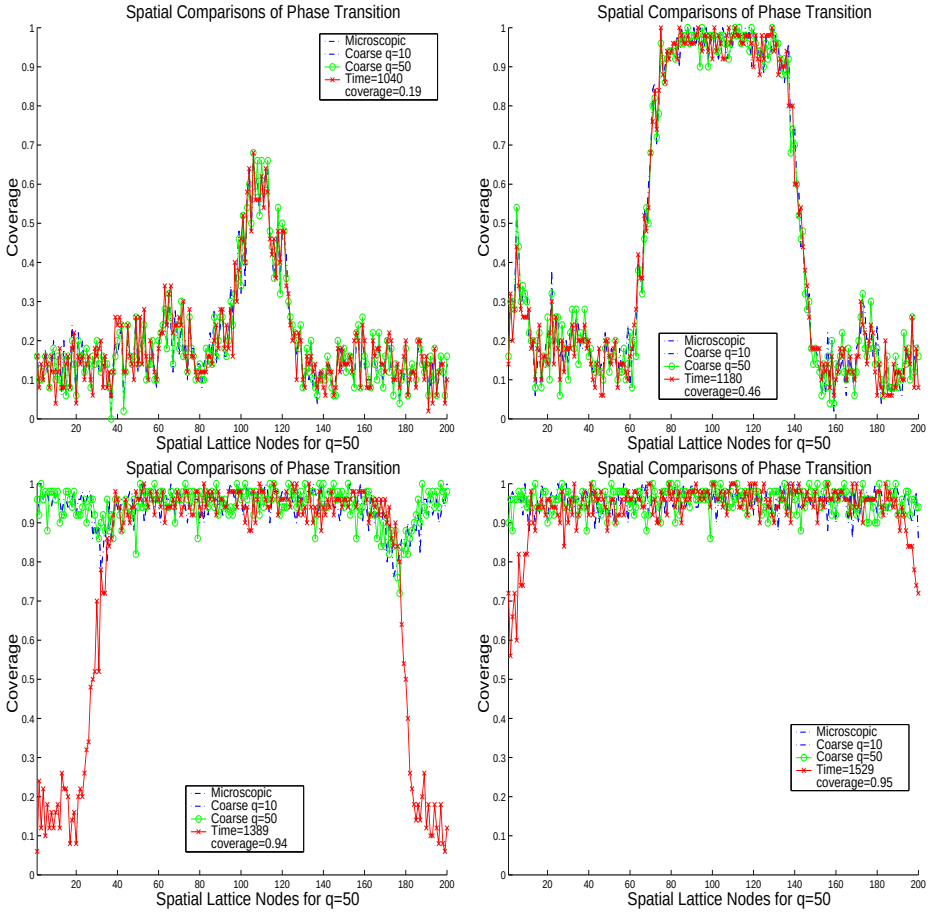


FIG. 7.5. *Snap-shots of the transition from zero initial spatial distribution. Comparisons between the microscopic* $q = 1$ *and two coarse grained simulations* $q = 10$ *and* $q = 50$. *The interaction radius is set to* $L = 200$ *while total nodes are* $N = 10000$.

In the next set of figures we present qualitative comparison of nucleation from an island of a given size (Figure 7.6–7.8). In these simulations we observe spatial propagation of the interface in time for different initial size of the island.

**Acknowledgments:** 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. The authors would like also 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"

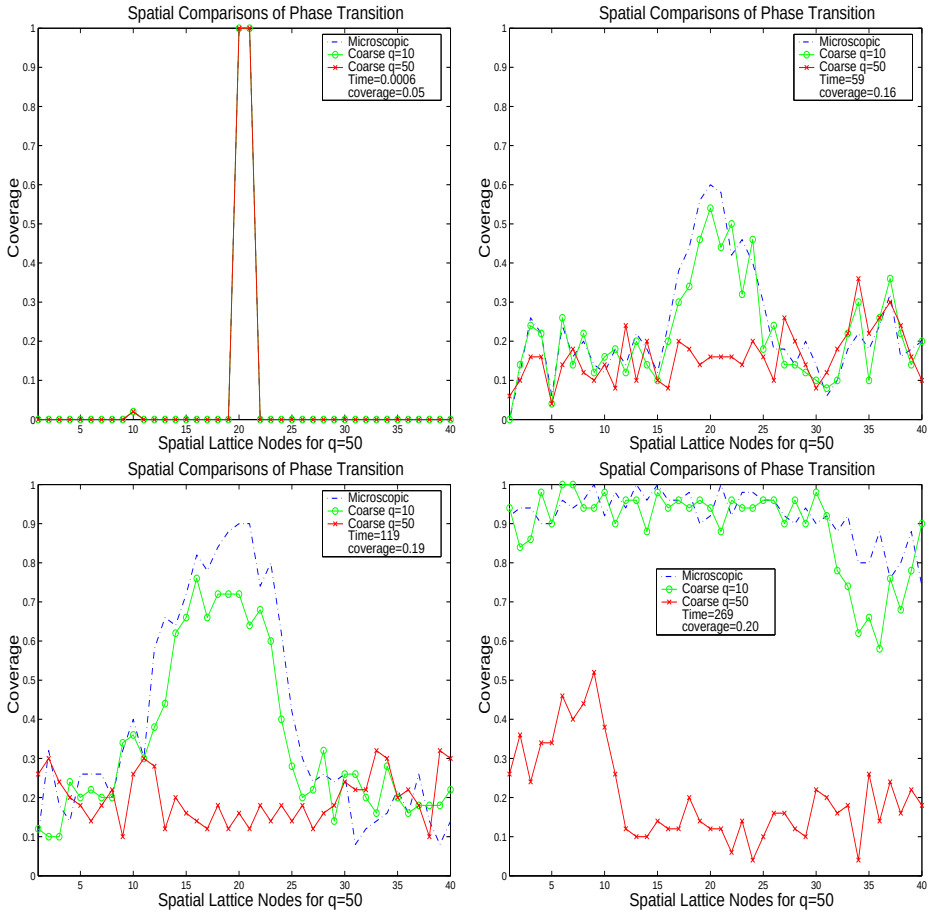


FIG. 7.6. *Snap-shots of the nucleation from a small-size initial seed. Comparisons between the microscopic* $q = 1$ *and two coarse grained simulations* $q = 10$ *and* $q = 50$. *Potential radius is set to* $L = 100$ *and the size of the lattice to* $N = 2000$.

## REFERENCES


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

[2] A. CHATTERJEE, M. A. KATSOULAKIS, P. PLECHÁČ, AND D. G. VLACHOS, *Higher-order correctors for adaptivity in lattice Monte Carlo simulations.* in preparation.

[3] A. CHATTERJEE, M. A. KATSOULAKIS, AND D. G. VLACHOS, *Spatially adaptive grand-canonical Monte Carlo simulations*, to appear Phys. Rev. E, (2004).

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

[5] L. D. S. D. G. VLACHOS 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.

[6] 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.

[7] W. E AND B. ENGQUIST, *Multiscale modeling and computation*, Notices Amer. Math. Soc., 50 (2003), pp. 1062–1070.

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

[9] P. L. FERRARI AND J. L. LEBOWITZ, *Information loss in coarse graining of polymer configurations via contact matrices*, J. Phys. A, 36 (2003), pp. 5719–5730.

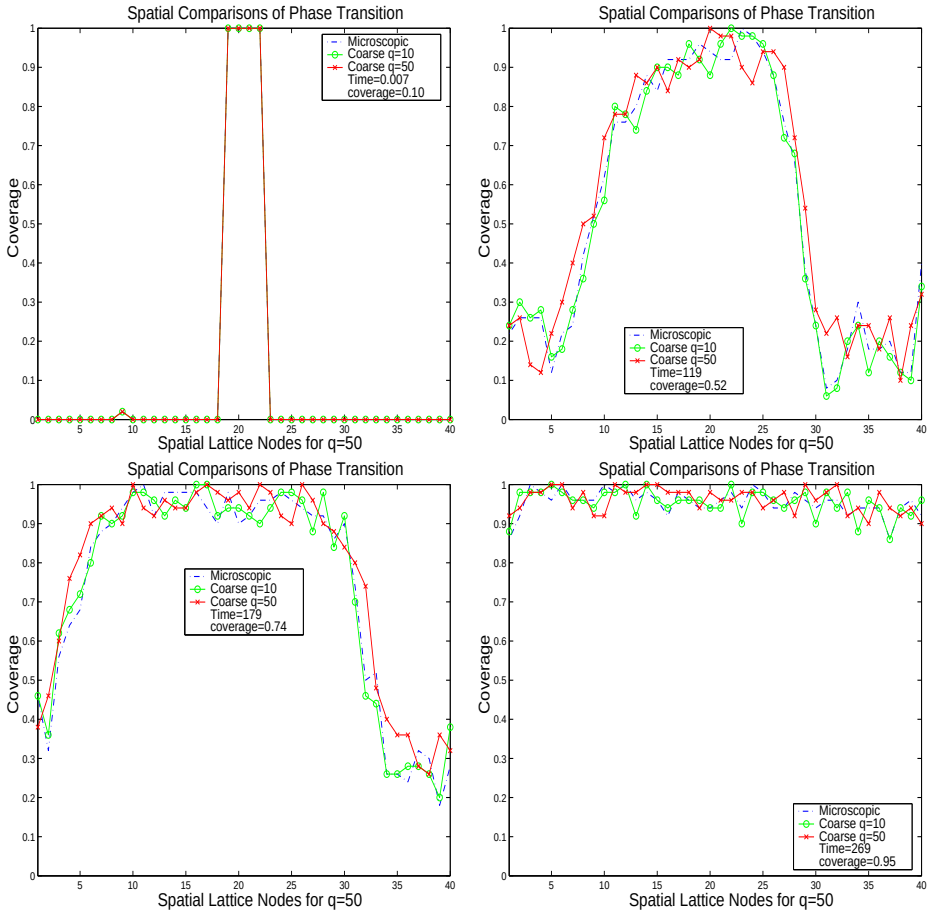


FIG. 7.7. *Snap-shots of the nucleation from a medium-size initial seed. Comparisons between the microscopic* $q = 1$ *and two coarse grained simulations* $q = 10$ *and* $q = 50$. *Potential radius is set to* $L = 100$ *and the size of the lattice to* $N = 2000$.


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

[11] 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.

[12] 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), p. 4414.

[13] 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.

[14] M. A. KATSOULAKIS, A. J. MAJDA, AND A. SOPASAKIS, *Multiscale couplings in prototype hybrid deterministic/stochastic systems: Part i, deterministic closures*, Comm. Math. Sci., 2 (2004).

[15] 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.

[16] 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).

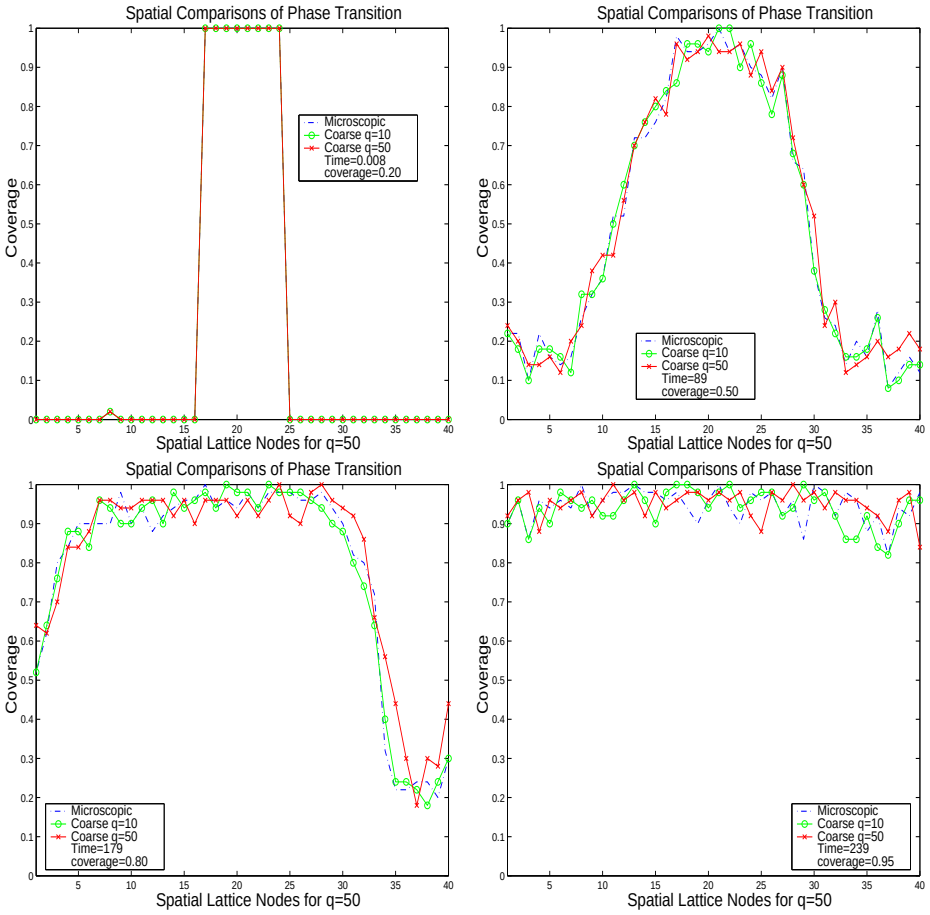


FIG. 7.8. *Snap-shots of the nucleation from a large-size initial seed. Comparisons between the microscopic* $q = 1$ *and two coarse grained simulations* $q = 10$ *and* $q = 50$. *Potential radius is set to* $L = 100$ *and the size of the lattice to* $N = 2000$.

[17] M. A. KATSOULAKIS AND J. TRASHORRAS, *Information loss in coarse-graining of stochastic particle dynamics*, tech. report, University of Massachussets, 2004. submitted to J. Stat. Phys.

[18] M. A. KATSOULAKIS AND D. G. VLACHOS, *Hierarchical kinetic Monte Carlo simulations for diffusion of interacting molecules*, J. Chem. Phys, 112 (2003).

[19] 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.

[20] C. KIPNIS AND C. LANDIM, *Scaling Limits of Interacting Particle Systems*, Springer-Verlag, 1999.

[21] D. P. LANDAU AND K. BINDER, *A Guide to Monte Carlo Simulations in Statistical Physics*, Cambridge University Press, 2000.

[22] A. J. MAJDA, I. TIMOFEYEV, AND E. V. EIJNDEN, *A mathematical framework for stochastic climate models*, Comm. Pure Appl. Math., 54 (2001), p. 891.

[23] F. MULLER-PLATHE, *Coarse-graining in polymer simulation: from the atomistic to the mesoscale and back*, Chem. Phys. Chem., 3 (2002), p. 754.